中国科学院指定考研参考书

線性代數

李炯生　查建国　编著

中国科学技术大学出版社
1989

内 容 提 要

本书是作者在中国科学技术大学数学系多年教学的基础上编写而成的. 它由多项式、行列式、矩阵、线性空间、线性变换、Jordan 标准形、Euclid 空间、酉空间和双线性函数等九章组成. 在内容的叙述上,力图做到矩阵方法与几何方法并重. 每章都配有丰富的典型例题和充足的习题可供读者选用.

附录中收录龚昇教授编著的《线性代数五讲》,从现代数学,尤其是模论的观点来重新审视与认识线性代数,讨论了向量空间、线性变换,着重研究了主理想整环上的模及其分解,并以此来重新理解向量空间在线性算子作用下的分解,可以使读者从高一个层次上来认识线性代数.

本书适合作为综合性大学理科数学专业的教材,也可以作为各类大专院校师生的教学参考书,以及关心线性代数与矩阵论的科技工作者和数学爱好者的自学读物或参考书.

图书在版编目 (CIP) 数据

线性代数 / 李炯生,查建国编著. —合肥: 中国科学技术大学出版社,1989 (2005.9 重印,2010.10 重排)
(中国科学院指定考研参考书)

ISBN 978-7-312-00110-9

I. 线··· II. ① 李··· ② 查··· III. 线性代数—高等学校—教材 IV. O 151.2

中国版本图书馆 CIP 数据核字(2003)第 078690 号

序　言

本书初稿完成于1983年. 当时中国科学技术大学数学系领导冯克勤教授委托编著者编写一本供数学系用的线性代数讲义. 接受这项任务后,我们专程到北京,拜访了中国科学院系统科学研究所万哲先研究员、中国科学院数学研究所许以超研究员、北京大学数学系聂灵沼教授和中国科学院研究生院曾肯成教授,请教他们对数学系线性代数教学的设想. 他们都热情地给予指导,从而为编写讲义提供了坚实的基础. 1984年春天,讲义便开始在数学系83级使用,并作为数学系线性代数教材一直使用到现在. 1985年,讲义曾获得中国科学技术大学首次颁发的优秀教材一等奖. 此后,在使用过程中对讲义又作了进一步的修改. 出版前编著者又作了全面的加工和充实.

线性代数研究的是线性空间以及线性空间的线性变换. 在线性空间取定一组基下,线性变换便和矩阵建立了一一对应关系. 这样,线性变换就和矩阵紧密联系起来. 于是,研究线性空间以及线性空间关于线性变换的分解即构成了线性代数的几何理论,而研究矩阵在各种关系下的分类问题则是线性代数的代数理论. 本书编写的一个着眼点是,着力于建立线性代数的这两大理论之间的联系,并从这种联系去阐述线性代数的理论.

当然,线性代数内容非常丰富,本书尽可能地按照1980年教育部颁发的综合性大学理科数学专业高等代数教学大纲进行选择,并在体系安排与叙述方式上尽量吸收中国科学技术大学数学系长期从事线性代数教学的老师与同事们,特别是曾肯成教授、许以超研究员的教学经验. 在处理行列式理论时,采用了曾肯成教授1965年首先在中国科学技术大学数学系使用的将n阶行列式视为数域$\mathbb{F}$上的n维向量空间的规范反对称n重线性函数的讲法; 在处理线性方程组理论时,利用了矩阵在相抵下的标准形理论; 在处理Jordan标准形时,先考虑了线性空间关于线性变换的分解,然后再用纯矩阵方法处理了Jordan标准形. 同时也着重于阐述已故著名数学家华罗庚教授的独具特色的矩阵方法.

为了便于读者掌握解题技巧,提高分析问题、解决问题的能力,本书几乎每一章都专门用一节讲述各种典型例题. 这部分内容是为习题课安排的. 每一节后面都附有大量习题,教师与读者可以根据具体情况选择使用. 这些习题除了在众多的线性代数、矩阵论教材以及习题集中选取之外,有一些是取自我国历届研究生试题,还有一些是自己编撰的. 在教学过程中,有些同学对线性代数的某些课题产生了兴趣,进行了一些研究. 有些结果即成为本书的习题. 这里应该提到的有中国科学技术大学数学系 81 级同学陈贵忠、黄瑜、窦昌柱,82 级同学陈秀雄等.

冯克勤教授对本书的编写自始至终都给予了热情的关心与帮助. 在编写过程中,得到万哲先、许以超、聂灵沼、曾肯成等研究员和教授的热心指导,编者谨致衷心感谢. 中国科学技术大学数学系陆洪文教授,杜锡录、李尚志副教授曾经使用本书的前身——《线性代数讲义》——作为教材,他们对讲义的修改提出许多有益的意见. 中国科学技术大学数学系讲师屈善坤、徐俊明协助编者仔细地审核了原稿,安徽大学数学系夏恩虎同志、中国科学技术大学 86 级硕士研究生黄道德审核了习题,在此一并致谢.

由于编著者水平所限,错误与缺点在所难免. 热忱欢迎同行们和广大读者批评指正.

李炯生　查建国

一九八八年元月于合肥

目　录

附 录 线 性 代 数 五 讲

第一章

多 项 式

✍ 本章将介绍数域上的多项式理论. 读者如果有机会学习抽象代数中的环论的话,将会对本章的内容有更深刻的理解.

✍ §1.1 从代数的观点定义了数环与数域,即具有加法与乘法两种运算且满足一定的运算规则的数的集合.

✍ §1.2 给出了一元多项式环的定义,以及多项式的加法与乘法的基本性质. 读者将会看到,多项式有许多性质与整数相类似.

✍ §1.3 讨论了多项式的整除性以及一组多项式的最大公因式,这里的关键是两个多项式的辗转相除法.

✍ §1.4 给出了本章的第一个主要定理——唯一析因定理,即每一个多项式都可以唯一地写成不可约多项式的乘积. 读者把它同整数中的算术基本定理进行比较,就可知道这一定理的重要意义.

✍ 根据唯一析因定理,不可约多项式的地位相当于整数中素数的地位. 因此,自然需要一些方法来判定多项式的不可约性. §1.5 说明了复系数不可约多项式只能是一次多项式,而实系数不可约多项式只能是一次或二次多项式.

✍ §1.6 给出了最有应用价值的判断整系数多项式不可约性的 Eisenstein 准则.

✍ §1.7 把一元多项式推广为多元多项式.

✍ §1.8 含本章的第二个主要定理——对称多项式基本定理,即每一个对称多项式都是基本对称多项式的多项式.

§1.1 整数环与数域

迄今为止,我们已经接触到的数系有自然数系,整数系,有理数系,实数系与复数系. 在这些数系中,都可以进行加法运算与乘法运算. 譬如,自然数系中的加法运算是指一个对应关系,即对于任意一对自然数 m 与 n,按照加法,可以确定唯一一个自然数与它们对应,这个自然数就是 m 与 n 的和 $m+n$; 而自然数系中的乘法运算也是一个对应关系,即对于任意一对自然数 m 与 n,按照乘法,可以确定唯一一个自然数与它们对应,这个自然数就是 m 与 n 的积 mn.

抽象地说,所谓集合 S 中的**代数运算**是指一个对应关系,即对于集合 S 中任意一对元素 a 与 b,按照这一对应关系,可以确定集合 S 中的唯一一个元素 c 与它们对应. 例如,复数的加,减,乘,除四则运算都是复数系中的代数运算.

一个集合引进了代数运算,这些代数运算往往具有某些性质. 例如,整数系的加法运算与乘法运算具有以下的性质:

(A1) 加法结合律

$$(a+b)+c=a+(b+c);$$

(A2) 加法交换律

$$a+b=b+a;$$

(A3) 有整数 0，它具有性质：

$$a+0=0+a=a;$$

(A4) 对每个整数 a，总有负数 $-a$，使得

$$a+(-a)=(-a)+a=0;$$

(M1) 乘法结合律

$$(ab)c=a(bc);$$

(M2) 乘法交换律

$$ab=ba;$$

(M3) 有整数 1，它具有性质，

$$a1=1a=a;$$

(D) 加乘分配律

$$a(b+c)=ab+ac,$$

其中 a, b 和 c 是任意整数.

集合 S 的每一种代数运算所适合的一些最基本的性质，以及不同代数运算之间最基本的联系便构成了界定这些代数运算的**公理**. 例如，上面提到的整数的加法与乘法就适合结合律，交换律以及加乘分配律等.

把整数系连同加法与乘法运算的特性抽象化，便引出以下的定义.

定义 1.1.1 在集合 R 中规定两种代数运算，一种称为**加法**运算，即对于集合 R 中任意一对元素 a 与 b，按照加法运算，集合 R 中有唯一一个元素 $a+b$ 与它们对应，元素 $a+b$ 称为 a 与 b 的**和**. 另一种称为**乘法**运算，即对于集合 R 中任意一对元素 a 与 b，按照乘法运算，集合 R 中有唯一一个元素 ab 与它们对应，元素 ab 称为 a 与 b 的**积**.

并且，加法运算与乘法运算适合下列**公理**：对于 R 中任意元素 a, b 与 c，有

(A1) 加法结合律

$$(a+b)+c=a+(b+c);$$

(A2) 加法交换律

$$a+b=b+a;$$

(A3) 存在零元素 R 中存在一个元素，它称为 R 的**零元素**，记作 0，使得

$$a+0=0+a=a;$$

(A4) 存在负元素 对于 R 中每个元素 a，存在元素 b，使得

$$a+b=b+a=0,$$

元素 b 称为元素 a 的**负元素**，记为 $-a$；

(M1) 乘法结合律

$$a(bc) = (ab)c;$$

(M2) 乘法交换律

$$ab = ba;$$

(M3) 存在单位元素 R 中存在一个元素,它称为**单位元素**,记为 1,使得

$$a1 = 1a = a;$$

(D) 加乘分配律

$$a(b + c) = ab + ac,$$

则集合 R 称为**交换环**.

容易验证,整数系是一个交换环,它称为整数环,记为 $\mathbb{Z}$. 另外,有理数系,实数系与复数系也都是交换环,它们都是复数系的子集合.

凡复数系的子集合,如果对复数的加法与乘法成为交换环,则称为**数环**.

应当指出,有理数系,实数系和复数系的乘法运算所具有的性质有些是和整数环的乘法性质不同的. 例如,在整数环中,对于非零整数 $a \neq \pm 1$,不存在整数 b,使得 $ab = ba = 1$; 但在实数环中,对于非零实数 a,一定存在实数 b,使得 $ab = ba = 1$. 为区别起见,引进以下的定义.

定义 1.1.2 设 $\mathbb{F}$ 是至少有两个元素的交换环. 如果对于 $\mathbb{F}$ 中每个非零元素 a,存在元素 $b \in \mathbb{F}$,使得 $ab = ba = 1$,则 b 称为 a 的**逆元素**,记作 a^{-1}. 这时交换环 $\mathbb{F}$ 称为**域**.

例如,有理数系,实数系与复数系都是域,它们依次称为有理数域,实数域与复数域,并依次记为 $\mathbb{Q}$,$\mathbb{R}$ 和 $\mathbb{C}$.

如果复数域 $\mathbb{C}$ 的子集合 $\mathbb{F}$ 对复数的加法与乘法成为一个域,则 $\mathbb{F}$ 称为**数域**.

可以验证,复数域的子集合

$$\mathbb{Q}\left[\sqrt{2}\right] = \left\{a + b\sqrt{2} \,\middle|\, a, b \in \mathbb{Q}\right\}$$

对复数的加法与乘法成为一个域,所以,$\mathbb{Q}\left[\sqrt{2}\right]$ 是一个数域.

习 题 1.1

1. 记 $\mathbb{Q}\left[\sqrt{2}\right] = \left\{a + b\sqrt{2} \,\middle|\, a, b \in \mathbb{Q}\right\}$. 验证 $\mathbb{Q}\left[\sqrt{2}\right]$ 是数域.

2. 记 $\mathbb{Z}\left[\sqrt{2}\right] = \left\{a + b\sqrt{2} \,\middle|\, a, b \in \mathbb{Z}\right\}$. 验证 $\mathbb{Z}\left[\sqrt{2}\right]$ 是数环. $\mathbb{Z}\left[\sqrt{2}\right]$ 是数域吗?

3. 设 $\mathbb{F}$ 是数域,a, b 和 c 是 $\mathbb{F}$ 中的任意三个元素,证明下列性质成立.

(1) 如果 $a + b = a + c$,则 $b = c$;

(2) 定义 $a - b = a + (-b)$,则 $a + (b - a) = b$;

(3) $a0 = 0a = 0$;

(4) $(-1)a = -a$;

(5) 如果 $ab = 0$,则 $a = 0$,或 $b = 0$.

4. 设 $\mathbb{F}$ 是所有有序实数对 (a, b) 的集合,其中 $a, b \in \mathbb{R}$.

(1) 如果集合 $\mathbb{F}$ 的加法与乘法分别定义为

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b)(c, d) = (ac, bd),$$

那么 $\mathbb{F}$ 是否成为域?

(2) 如果 $\mathbb{F}$ 的加法与乘法分别定义为

$$(a,b)+(c,d)=(a+c,b+d),\quad (a,b)(c,d)=(ac-bd,ad+bc),$$

那么 $\mathbb{F}$ 是否成为域?

(3) 如果 $\mathbb{F}$ 表示所有有序复数对的集合; 加法与乘法仍如 (1) 与 (2) 那样规定,结论又怎样?

5. 证明,在交换环的定义中,如果除加法交换律外,其它公理都假定成立,则可以推出加法交换律也成立. 换句话说,在交换环的定义中,加法交换律这一公理可以去掉.

§1.2 一元多项式环

在中学里,我们遇到过一次方程与二次方程,它们可以从两方面推广. 一方面从次数推广,即推广为 3 次,4 次以至 n 次的方程; 另一方面从系数所属的范围推广. 由 §1.1 可以看到,系数所属的实数域可以推广为其它的数域. 这就引出以下的定义.

定义 1.2.1 设 $\mathbb{F}$ 是数域,x 是未定元,$a_1,a_2,\ldots,a_n\in\mathbb{F}$,$a_n\neq 0$,$n$ 是非负整数. 则

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

称为**数域 $\mathbb{F}$ 上的一元多项式**,数域 $\mathbb{F}$ 上的一元多项式 $f(x)$ 的全体所成的集合记为 $\mathbb{F}[x]$. 其中 a_ix^i 称为多项式 $f(x)$ 的 ***i* 次项**,数 a_i 称为 $f(x)$ 的 ***i* 次项系数**.

特别地,a_0 称为 $f(x)$ 的**常数项**,a_nx^n 称为 $f(x)$ 的**首项**(或**最高次项**),a_n 称为 $f(x)$ 的**首项系数**. 如果 $a_n=1$,则 $f(x)$ 称为**首一多项式**.

非负整数 n 称为 $f(x)$ 的**次数**,记为 $\deg f(x)$.

如果多项式 $f(x)$ 的系数全为零,则 $f(x)$ 称为**零多项式**,这时仍记为 0. 约定零多项式的次数为 $-\infty$. 注意,零次多项式不是零多项式. 有时也称零次多项式为**纯量多项式**.

如果上述定义中,把数域 $\mathbb{F}$ 改成数环,则 $f(x)$ 称为**数环 $\mathbb{F}$ 上的一元多项式**,其它的规定是相同的.

设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\qquad a_n\neq 0;$$
$$g(x)=b_mx^m+b_{m-1}x^{n-1}+\cdots+b_1x+b_0,\quad b_m\neq 0$$

是数域 $\mathbb{F}$ 上的两个多项式. 如果 $f(x)$ 与 $g(x)$ 适合 $a_i=b_i$,$i=0,1,2,\ldots$,则 $f(x)$ 与 $g(x)$ 称为**相等**,记为 $f(x)=g(x)$.

多项式 $f(x)$ 与 $g(x)$ 的和 $\boldsymbol{f(x)+g(x)}$ 定义为多项式

$$(a_0+b_0)+(a_1+b_1)x+(a_2+b_2)x^2+\cdots,$$

即多项式 $f(x)+g(x)$ 的 i 次项系数为 a_i+b_i,$i=1,2,\ldots$. 其中当 $n\geqslant m$ 时,约定 $g(x)$ 的系数 $b_{m+1},b_{m+2},\ldots,b_n$ 都为零,而当 $n<m$ 时,约定 $f(x)$ 的系数 $a_{n+1},a_{n+2},\ldots,a_m$

都为零. 于是便定义了多项式的加法. 容易看出,

$$\deg\big(f(x)+g(x)\big)\leqslant \max\{\deg f(x),\ \deg g(x)\}.$$

容易验证,多项式的加法满足以下公理. 设多项式 $f(x), g(x), h(x) \in \mathbb{F}[x]$,则

(A1) 加法结合律

$$\big(f(x)+g(x)\big)+h(x)=f(x)+\big(g(x)+h(x)\big);$$

(A2) 加法交换律

$$f(x)+g(x)=g(x)+f(x);$$

(A3) 存在零元素 即存在零多项式 $0\in\mathbb{F}[x]$,使得

$$f(x)+0=0+f(x)=f(x);$$

(A4) 存在负元素 对每个多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$

都存在多项式

$$-f(x)=-a_nx^n-a_{n-1}x^{n-1}-\cdots-a_1x-a_0,$$

它称为多项式 $f(x)$ 的**负多项式**,使得

$$f(x)+\big(-f(x)\big)=\big(-f(x)\big)+f(x)=0.$$

对于 $\mathbb{F}[x]$ 中两个多项式 $f(x)$ 和 $g(x)$,其**乘积** $f(x)g(x)$ 定义为

$$f(x)g(x)=c_{n+m}x^{n+m}+c_{n+m-1}x^{n+m-1}+\cdots+c_1x+c_0,$$

其中

$$\begin{aligned} c_{n+m}&=a_nb_m,\\ c_{n+m-1}&=a_nb_{m-1}+a_{n-1}b_m,\\ &\cdots\cdots\\ c_i&=\sum_{j+k=i}a_jb_k,\\ &\cdots\cdots\\ c_0&=a_0b_0. \end{aligned}$$

于是规定了多项式的乘法. 因为 $a_n\neq 0, b_m\neq 0$,故 $a_nb_m\neq 0$. 所以,

$$\deg\big(f(x)g(x)\big)=\deg f(x)+\deg g(x).$$

容易验证,多项式的乘法适合以下的公理. 设 $f(x), g(x), h(x)\in\mathbb{F}[x]$,则

(M1) 乘法结合律

$$f(x)\big(g(x)h(x)\big)=\big(f(x)g(x)\big)h(x);$$

(M2) 乘法交换律

$$f(x)g(x)=g(x)f(x);$$

(M3) 存在单位元素 即存在纯量多项式 $e(x)=1$,使得

$$f(x)e(x)=e(x)f(x)=f(x);$$

(D) 加乘分配律

$$f(x)\big(g(x)+h(x)\big)=f(x)g(x)+f(x)h(x).$$

于是根据定义,$\mathbb{F}[x]$ 是交换环,它称为数域 $\mathbb{F}$ 上的**一元多项式环**.

习 题 1.2

1. 设 $f(x),g(x)\in\mathbb{F}[x]$. 证明,当且仅当 $f(x)=0$,或 $g(x)=0$ 时,$f(x)g(x)=0$.

2. 设 $f(x),g(x),h(x)\in\mathbb{F}[x]$,且 $f(x)\neq 0$. 证明,若 $f(x)g(x)=f(x)h(x)$,则 $g(x)=h(x)$.

3. 设非零的实系数多项式 $f(x)$ (即系数都是实数的多项式)满足 $f\big(f(x)\big)=f^k(x)$,其中 k 是给定的正整数. 求多项式 $f(x)$.

4. 设非零的实系数多项式 $f(x)$ 满足 $f(x^2)=f^2(x)$. 求多项式 $f(x)$.

5. 设实系数多项式 $f(x)=ax^2+bx+c$,$a\neq 0$. 证明,对任意给定的正整数 n,不存在 n 次实系数多项式 $g(x)$,使得 $f\big(g(x)\big)=g\big(f(x)\big)$. (本题似有误)

6. 设实系数多项式 $P(x)=a_0+a_1x+\cdots+a_nx^n$ 满足

$$0\leqslant a_0=a_n\leqslant a_1=a_{n-1}\leqslant\cdots\leqslant a_{[\frac{n}{2}]}=a_{[\frac{n+1}{2}]},$$

所有这样的多项式 $P(x)$ 的集合记作 $A(n)$. 证明,如果 $P(x)\in A(n)$,$Q(x)\in A(m)$,则乘积

$$P(x)Q(x)\in A(n+m).$$

§1.3 整除性与最大公因式

数域 $\mathbb{F}$ 上的一元多项式环 $\mathbb{F}[x]$ 是我们遇到的第一个不是由数构成的交换环. 它的性质是否与数环,特别是与整数环 $\mathbb{Z}$ 相同? 譬如,在整数环 $\mathbb{Z}$ 中,对于任意整数 $a,b\in\mathbb{Z}$,$b\neq 0$,总存在唯一一对整数 q 和 r,$0\leqslant r<|b|$,使得 $a=qb+r$. 整数环 $\mathbb{Z}$ 的这一性质,多项式环 $\mathbb{F}[x]$ 是否也具有? 对此,有

定理 1.3.1 (带余除法) 设多项式 $f(x),g(x)\in\mathbb{F}[x]$,$g(x)\neq 0$. 则存在唯一一对多项式 $q(x),r(x)\in\mathbb{F}[x]$,$\deg r(x)<\deg g(x)$,使得

$$f(x)=q(x)g(x)+r(x). \tag{1.3.1}$$

证明 存在性 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\qquad a_n\neq 0,$$
$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0,\qquad b_m\neq 0.$$

显然,当 $n<m$ 时,取 $q(x)=0$,$r(x)=f(x)$,则式 (1.3.1) 成立. 当 $n\geqslant m$ 时,记

$$f(x)-\frac{a_n}{b_m}x^{n-m}g(x)=f_1(x),$$

显然,$\deg f_1(x)<\deg f(x)$.

于是对 $\deg f(x)=n$ 用归纳法,则存在多项式 $q_1(x),r(x)\in\mathbb{F}[x]$,$\deg r(x)<$

$\deg g(x)$，使得

$$f(x)-\frac{a_n}{b_m}x^{n-m}g(x)=f_1(x)=q_1(x)g(x)+r(x).$$

因此，

$$f(x)=\Big(\frac{a_n}{b_m}x^{n-m}+q_1(x)\Big)g(x)+r(x).$$

这表明，如果记 $q(x)=a_nb_m^{-1}x^{n-m}+q_1(x)$，则式 (1.3.1) 成立.

唯一性 设 $q_1(x), r_1(x)\in\mathbb{F}[x]$，$\deg r_1(x)<\deg g(x)$，使得式 (1.3.1) 成立. 则

$$\big(q(x)-q_1(x)\big)g(x)=r_1(x)-r(x).$$

因此，如果 $q(x)\neq q_1(x)$，则由上式，

$$\deg g(x)<\deg\big(q(x)-q_1(x)\big)+\deg g(x)=\deg\big(r_1(x)-r(x)\big).$$

但因 $\deg r(x)<\deg g(x)$，$\deg r_1(x)<\deg g(x)$，因此，

$$\deg\big(r_1(x)-r(x)\big)<\deg g(x)$$

不可能. 所以，$q(x)=q_1(x)$，从而 $r(x)=r_1(x)$. ■

和整数环 $\mathbb{Z}$ 相仿，定理 1.3.1 中多项式 $q(x)$ 与 $r(x)$ 分别称为多项式 $f(x)$ 除以 $g(x)$ 的**商式**与**余式**.

应当指出，定理 1.3.1 关于商式 $q(x)$ 与余式 $r(x)$ 的存在性证明是构造性的. 换句话说，给定多项式 $f(x)$ 与 $g(x)$，$g(x)\neq 0$，可以按照定理 1.3.1 的证明方法求出商式 $g(x)$ 与余式 $r(x)$，其过程如下：

当 $\deg f(x)<\deg g(x)$ 时，$q(x)=0$，$r(x)=f(x)$.

当 $\deg f(x)\geqslant\deg g(x)$ 时，记

$$f(x)-\frac{a_n}{b_m}x^{n-m}g(x)=f_1(x).$$

如果 $\deg f_1(x)<\deg g(x)$，则

$$f(x)=\frac{a_n}{b_m}x^{n-m}g(x)+f_1(x),$$

此时 $q(x)=a_nb_m^{-1}x^{n-m}$，$r(x)=f_1(x)$.

如果 $\deg f_1(x)\geqslant\deg g(x)$，且 $f_1(x)$ 的首项系数为 $c_t\neq 0$，$t<\deg f(x)$，则记

$$f_1(x)-\frac{c_t}{b_m}x^{t-m}g(x)=f_2(x).$$

如果 $\deg f_2(x)<\deg g(x)$，则

$$f(x)=\Big(\frac{a_n}{b_m}x^{n-m}+\frac{c_t}{b_m}x^{t-m}\Big)g(x)+f_2(x),$$

此时 $q(x)=a_nb_m^{-1}x^{n-m}+c_tb_m^{-1}x^{t-m}$，$r(x)=f_2(x)$.

如果 $\deg f_2(x)\geqslant\deg g(x)$，则重复上述过程. 于是得到多项式序列 $f(x), f_1(x), \ldots, f_k(x), \ldots$，它们适合

$$\deg f(x)>\deg f_1(x)>\deg f_2(x)>\cdots>\deg f_k(x)>\cdots\geqslant\deg g(x).$$

由于 $\deg f(x)-\deg g(x)$ 是有限的，因此，经有限步后，必有 ℓ，使得

$$f_{\ell-1}(x) = \frac{d_s}{b_m}x^{s-m}g(x) + f_\ell(x), \qquad f_\ell(x) = \frac{e_h}{b_m}x^{h-m}g(x) + f_{\ell+1}(x),$$

其中 d_s 与 e_h 分别是多项式 $f_{\ell-1}(x)$ 与 $f_\ell(x)$ 的首项系数，$s > h \geqslant m$，并且

$$\deg f_{\ell+1}(x) < \deg g(x).$$

于是

$$f(x) = \left(\frac{a_n}{b_m}x^{n-m} + \frac{c_t}{b_m}x^{t-m} + \cdots + \frac{d_s}{b_m}x^{s-m} + \frac{e_h}{b_m}x^{h-m}\right)g(x) + f_{\ell+1}(x).$$

因此，

$$q(x) = \frac{a_n}{b_m}x^{n-m} + \frac{c_t}{b_m}x^{t-m} + \cdots + \frac{d_s}{b_m}x^{s-m} + \frac{e_h}{b_m}x^{h-m},$$
$$r(x) = f_{\ell+1}(x).$$

上述过程可写成表 1.3.1 的形式，这种算法称为 **Euclid 长除法**.

$$\begin{array}{r|l|l}
 & b_m^{-1}a_nx^{n-m} + b_m^{-1}c_tx^{t-m} + \cdots + b_m^{-1}d_sx^{s-m} + b_m^{-1}e_hx^{h-m} & \\
\hline
g(x) = b_mx^m + b_{m-1}x^{m-1} + \cdots + b_1x + b_0 & a_nx^n + a_{n-1}x^{n-1} + \cdots + a_{n-m}x^{n-m} + \cdots + a_1x + a_0 & f(x) \\
 & -)\ a_nx^n + b_m^{-1}a_nb_{m-1}x^{n-1} + \cdots + b_m^{-1}a_nb_0x^{n-m} & b_m^{-1}a_nx^{n-m}g(x) \\
\cline{2-2}
 & c_tx^t + c_{t-1}x^{t-1} + \cdots + c_{t-m}t^{t-m} + \cdots + c_0 & f_1(x) \\
 & -)\ c_tx^t + b_m^{-1}c_tb_{m-1}x^{t-1} + \cdots + b_m^{-1}c_tb_0x^{t-m} & b_m^{-1}c_tx^{t-m}g(x) \\
\cline{2-2}
 & \cdots\cdots\cdots\cdots\cdots\cdots & \cdots\ \cdots \\
\cline{2-2}
 & d_sx^s + \cdots\cdots + d_{s-m}x^{s-m} + \cdots + d_0 & f_{\ell-1}(x) \\
 & -)\ d_sx^s + \cdots\cdots + b_m^{-1}d_sb_0x^{s-m} & b_m^{-1}d_sx^{s-m}g(x) \\
\cline{2-2}
 & e_hx^h + \cdots + e_{h-m}x^{h-m} + \cdots + e_0 & f_\ell(x) \\
 & -)\ e_hx^h + \cdots + b_m^{-1}e_hb_0x^{h-m} & b_m^{-1}e_hx^{h-m}g(x) \\
\cline{2-2}
 & p_qx^q + \cdots + p_1x + p_0\ (q < m) & r(x) = f_{\ell+1}(x)
\end{array}$$

表 1.3.1 Euclid 长除法

$$\begin{array}{r|l}
 & \frac{1}{3}x - \frac{1}{9} \\
\hline
3x^3 + 10x^2 + 2x - 3 & x^4 + 3x^3 - x^2 - 4x - 3 \\
 & -x^4 - \frac{10}{3}x^3 - \frac{2}{3}x^2 + x \\
\cline{2-2}
 & -\frac{1}{3}x^3 - \frac{5}{3}x^2 - 3x - 3 \\
 & \frac{1}{3}x^3 + \frac{10}{9}x^2 + \frac{2}{9}x - \frac{1}{3} \\
\cline{2-2}
 & -\frac{5}{9}x^2 - \frac{25}{9}x - \frac{10}{3}
\end{array}$$

表 1.3.2 Euclid 长除法(例 1.3.1)

设多项式 $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{F}[x]$，$a \in \mathbb{F}$. 记

$$f(a) = a_na^n + a_{n-1}a^{n-1} + \cdots + a_1a + a_0.$$

$f(a)$ 称为**多项式 $f(x)$ 在 $x=a$ 处的值**. 若 $f(a)=0$,则 a 称为**多项式 $f(x)$ 的根**.

定理 1.3.1 的一个重要特例是:

推论 1.3.1(剩余定理) 设 $f(x)\in\mathbb{F}[x]$,$a\in\mathbb{F}$,则存在唯一 $q(x)\in\mathbb{F}[x]$,使得

$$f(x)=(x-a)q(x)+f(a). \tag{1.3.2}$$

证明 由定理 1.3.1,存在唯一一对多项式 $q(x),r(x)\in\mathbb{F}[x]$,使得

$$f(x)=(x-a)q(x)+r(x),$$

其中 $\deg r(x)<1$. 将 $x=a$ 代入,得到 $f(a)=r(x)$. ■

例 1.3.1 设 $f(x)=x^4+3x^3-x^2-4x-3$,$g(x)=3x^3+10x^2+2x-3$,求 $f(x)$ 除以 $g(x)$ 的商式 $q(x)$ 和余式 $r(x)$.

解 如表 1.3.2 作 Euclid 长除法. 由此得到

$$q(x)=\frac{1}{3}x-\frac{1}{9},\qquad r(x)=-\frac{5}{9}x^2-\frac{25}{9}x-\frac{10}{3}.$$ ■

定义 1.3.1 设非零多项式 $g(x)\in\mathbb{F}[x]$. 如果存在 $q(x)\in\mathbb{F}[x]$,使得多项式

$$f(x)=g(x)q(x),$$

则多项式 $g(x)$ 称为多项式 $f(x)$ 的一个**因式**,而多项式 $f(x)$ 称为多项式 $g(x)$ 的一个**倍式**. 这时说多项式 $g(x)$ **整除**多项式 $f(x)$,记为 $g(x)\mid f(x)$. 否则就说多项式 $g(x)$ **不能整除**多项式 $f(x)$,记作 $g(x)\nmid f(x)$.

关于多项式的整除性,有以下性质.

性质 1.3.1 如果 $g(x)\mid f(x)$,$h(x)\mid g(x)$,则 $h(x)\mid f(x)$.

性质 1.3.2 如果 $g(x)\mid f_1(x)$,$g(x)\mid f_2(x)$,则对任意 $h_1(x),h_2(x)\in\mathbb{F}[x]$,

$$g(x)\mid\big(h_1(x)f_1(x)+h_2(x)f_2(x)\big).$$

性质 1.3.3 如果 $g(x)\mid f(x)$,$f(x)\mid g(x)$,则存在非零的数 $\lambda\in\mathbb{F}$,使得

$$f(x)=\lambda g(x).$$

上述性质请读者自证之.

推论 1.3.2(因式定理) 设 $f(x)\in\mathbb{F}[x]$,$a\in\mathbb{F}$. 则当且仅当 $f(a)=0$ 时,

$$(x-a)\mid f(x).$$

定义 1.3.2 设多项式 $f_1(x),f_2(x),h(x)\in\mathbb{F}[x]$. 如果 $h(x)$ 是 $f_1(x)$ 与 $f_2(x)$ 的因式,则 $h(x)$ 称为 $f_1(x)$ 与 $f_2(x)$ 的一个**公因式**. 设 $f_1(x)$ 与 $f_2(x)$ 不全为零. 如果首一多项式 $d(x)$ 是 $f_1(x)$ 与 $f_2(x)$ 的公因式,而且 $f_1(x)$ 与 $f_2(x)$ 的每个公因式都是 $d(x)$ 的一个因式,则 $d(x)$ 称为 $f_1(x)$ 与 $f_2(x)$ 的**最大公因式**,记为

$$d(x)=\big(f_1(x),f_2(x)\big).$$

关于两个不全为零的多项式 $f_1(x)$ 与 $f_2(x)$ 的最大公因式,有

定理 1.3.2　任意两个不全为零的多项式 $f_1(x)$ 与 $f_2(x)$ 的最大公因式 $d(x)$ 存在而且唯一.

证明　**唯一性**　设 $d_1(x)$ 与 $d_2(x)$ 都是 $f_1(x)$ 与 $f_2(x)$ 的最大公因式. 由最大公因式的定义, $d_1(x)$ 是 $f_1(x)$ 与 $f_2(x)$ 的一个公因式, 而 $d_2(x)$ 是 $f_1(x)$ 与 $f_2(x)$ 的最大公因式, 因此, $d_1(x) \mid d_2(x)$.

反之, 同样有 $d_2(x) \mid d_1(x)$. 由性质 1.3.3, 存在非零的数 $\lambda \in \mathbb{F}$, 使 $d_1(x) = \lambda d_2(x)$. 由于 $d_1(x)$ 与 $d_2(x)$ 都是首一多项式, 比较上式两边的首项系数, 得到 $\lambda = 1$, 即 $d_1(x) = d_2(x)$.

存在性　不妨设 $f_2(x) \neq 0$. 由定理 1.3.1, 存在多项式 $q_1(x)$ 与 $r_1(x)$, 使得

$$f_1(x) = q_1(x) f_2(x) + r_1(x),$$

其中 $\deg r_1(x) < \deg f_2(x)$.

如果 $\deg r_1(x) = -\infty$ 为零多项式, 则停止. 如果 $r_1(x)$ 是非零多项式, 则由定理 1.3.1, 存在多项式 $q_2(x)$ 与 $r_2(x)$, 使得

$$f_2(x) = q_2(x) r_1(x) + r_2(x),$$

其中 $\deg r_2(x) < \deg r_1(x)$.

如果 $r_2(x)$ 是零多项式, 则停止. 如果 $r_2(x)$ 是非零多项式, 则重复上述过程. 于是得一连串的等式:

$$\begin{aligned}
f_1(x) &= q_1(x) f_2(x) + r_1(x), && (1)\\
f_2(x) &= q_2(x) r_1(x) + r_2(x), && (2)\\
r_1(x) &= q_3(x) r_2(x) + r_3(x), && (3)\\
&\cdots\cdots\cdots\cdots\cdots\cdots &&\\
r_{k-2}(x) &= q_k(x) r_{k-1}(x) + r_k(x), && (k)\\
&\cdots\cdots\cdots\cdots\cdots\cdots &&
\end{aligned}$$

其中

$$\deg f_2(x) > \deg r_1(x) > \deg r_2(x) > \cdots > \deg r_k(x) > \cdots.$$

由于 $\deg f_2(x)$ 是一个给定的非负整数, 因此, 存在某个正整数 ℓ, 使得 $r_\ell(x) \neq 0$, 而 $r_{\ell+1}(x) = 0$, 即最后必有

$$\begin{aligned}
r_{\ell-2}(x) &= q_\ell(x) r_{\ell-1}(x) + r_\ell(x), && (\ell)\\
r_{\ell-1}(x) &= q_{\ell+1}(x) r_\ell(x). && (\ell+1)
\end{aligned}$$

$(\ell+1)$ 式表明, $r_\ell(x) \mid r_{\ell-1}(x)$.

因此, 由 (ℓ) 式, $r_\ell(x) \mid r_{\ell-2}(x)$; 再由 $(\ell-1)$ 式, $r_\ell(x) \mid r_{\ell-3}(x)$, 等等. 于是 $r_\ell(x)$ 整除

$$r_{\ell-1}(x),\quad r_{\ell-2}(x),\quad \ldots,\quad r_2(x),\quad r_1(x).$$

因此, 由 (2), $r_\ell(x) \mid f_2(x)$. 最后, 由 (1) 式 $r_\ell(x) \mid f_1(x)$. 所以 $r_\ell(x)$ 是 $f_1(x)$ 与 $f_2(x)$ 的一个公因式.

反之，设 $h(x)$ 是 $f_1(x)$ 与 $f_2(x)$ 的一个公因式. 由 (1) 式，$h(x)$ 是 $r_1(x)$ 的因式. 再由 (2) 式，$h(x)$ 是 $r_2(x)$ 的因式，等等. 于是 $h(x)$ 是 $r_{\ell-2}(x)$ 及 $r_{\ell-1}(x)$ 的因式. 由 (ℓ) 式，$h(x)$ 是 $r_\ell(x)$ 的一个因式.

因此，如果 $r_\ell(x)$ 的首项系数为 a，则由最大公因式的定义，$d(x) = a^{-1}r_\ell(x)$ 是 $f_1(x)$ 与 $f_2(x)$ 的最大公因式. ■

定理 1.3.2 关于多项式 $f_1(x)$ 与 $f_2(x)$ 的最大公因式的存在性证明给出了求最大公因式的一个方法，即所谓**辗转相除法**.

其过程如下：设 $f_1(x), f_2(x) \in \mathbb{F}[x]$，$\deg f_1(x) \geqslant \deg f_2(x)$. 先对 $f_1(x)$ 与 $f_2(x)$ 用 Euclid 长除法，得到商式 $q_1(x)$ 与余式 $r_1(x)$，$\deg r_1(x) < \deg f_2(x)$；在进行长除法时，将商式 $q_1(x)$ 记入表 1.3.3 的右边.

$f_2(x)$		$f_1(x)$	$q_1(x)$
	$-)$	$q_1(x)f_1(x)$	
		$r_1(x)$	

表 1.3.3

$q_2(x)$		$f_2(x)$		$f_1(x)$	$q_1(x)$
	$-)$	$q_2(x)r_1(x)$	$-)$	$q_1(x)f_2(x)$	
		$r_2(x)$		$r_1(x)$	

表 1.3.4

如果 $r_1(x)$ 是零多项式，则停止. 如果 $r_1(x)$ 是非零多项式，则对 $f_2(x)$ 和 $r_1(x)$ 用长除法，得到商式 $q_2(x)$ 与余式 $r_2(x)$，$\deg r_2(x) < \deg r_1(x)$，在进行长除法时，将商式 $q_2(x)$ 记入表 1.3.4 的左边.

$q_2(x)$		$f_2(x)$		$f_1(x)$	$q_1(x)$
	$-)$	$q_2(x)r_1(x)$	$-)$	$q_1(x)f_2(x)$	
$q_4(x)$		$r_2(x)$		$r_1(x)$	$q_3(x)$
	$-)$	$q_4(x)r_3(x)$	$-)$	$q_3(x)r_2(x)$	
$q_6(x)$		$r_4(x)$		$r_3(x)$	$q_5(x)$
$\vdots$		$\vdots$		$\vdots$	$\vdots$

表 1.3.5

如果 $r_2(x)$ 是零多项式，则停止. 如果 $r_2(x)$ 是非零多项式，则对 $r_1(x)$ 和 $r_2(x)$ 用长除法，得到商式 $q_3(x)$ 和余式 $r_3(x)$，并把商式 $q_3(x)$ 记入表 1.3.5 的右边.

如此继续，即可求得 $f_1(x)$ 与 $f_2(x)$ 的最大公因式.

例 1.3.2 求多项式 $f_1(x) = x^4 + 3x^3 - x^2 - 4x - 3$ 与 $f_2(x) = 3x^3 + 10x^2 + 2x - 3$ 的最大公因式.

解 如表 1.3.6 所示，对多项式 $f_1(x)$ 和 $f_2(x)$ 作辗转相除法.

因此，$r_2(x) = 9x + 27$，$r_3(x) = 0$. 于是，$\big(f_1(x), f_2(x)\big) = x + 3$. ■

定理 1.3.3 设不全为零的多项式 $f(x), g(x) \in \mathbb{F}[x]$，$d(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因式. 则存在多项式 $u(x), v(x) \in \mathbb{F}[x]$，使得

	$3x^3+10x^2+2x-3=f_2(x)$	$x^4+3x^3-x^2-4x-3=f_1(x)$	$\frac{1}{3}x$
		$-)\ x^4+\frac{10}{3}x^3+\frac{2}{3}x^2-x$	
$-\frac{27}{5}x$	$3x^3+10x^2+2x-3=f_2(x)$	$-\frac{1}{3}x^3-\frac{5}{3}x^2-3x-3$	$-\frac{1}{9}$
	$-)\ 3x^3+15x^2+18x$	$-)\ -\frac{1}{3}x^3-\frac{10}{9}x^2-\frac{2}{9}x+\frac{1}{3}$	
9	$-5x^2-16x-3$	$-\frac{5}{9}x^2-\frac{25}{9}x-\frac{10}{3}=r_1(x)$	
	$-)\ -5x^2-25x-30$		
	$9x+27=r_2(x)$	$-\frac{5}{9}x^2-\frac{25}{9}x-\frac{10}{3}=r_1(x)$	$-\frac{5}{81}x$
		$-)\ -\frac{5}{9}x^2-\frac{5}{3}x$	
	$9x+27=r_2(x)$	$-\frac{10}{9}x-\frac{10}{3}$	$-\frac{10}{81}$
		$-)\ -\frac{10}{9}x-\frac{10}{3}$	
		$0=r_3(x)$	

表 1.3.6　辗转相除法(例 1.3.2)

$$f(x)u(x)+g(x)v(x)=d(x). \tag{1.3.3}$$

证明　记 $f(x)=f_1(x)$, $g(x)=f_2(x)$. 由定理 1.3.2 的证明,存在多项式 $q_1(x)$, $q_2(x),\ldots,q_{k+1}(x)$ 和 $r_1(x),r_2(x),\ldots,r_k(x)$,使得

$$f_1(x)=q_1(x)f_2(x)+r_1(x), \tag{1}$$
$$f_2(x)=q_2(x)r_1(x)+r_2(x), \tag{2}$$
$$r_1(x)=q_3(x)r_2(x)+r_3(x),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$r_{k-3}(x)=q_{k-1}(x)r_{k-2}(x)+r_{k-1}(x), \tag{$k-1$}$$
$$r_{k-2}(x)=q_k(x)r_{k-1}(x)+r_k(x), \tag{k}$$
$$r_{k-1}(x)=q_{k+1}(x)r_k(x), \tag{$k+1$}$$

其中

$$\deg f_2(x)>\deg r_1(x)>\deg r_2(x)>\cdots>\deg r_k(x)\geqslant 0;$$

而且 $r_k(x)=\lambda d(x)$,λ 为 $r_k(x)$ 的首项系数. 由式 (k) 得到,

$$\lambda d(x)=r_{k-2}(x)-r_{k-1}(x)q_k(x).$$

设 $u_1(x)=1$, $v_1(x)=-q_k(x)$,上式为

$$\lambda d(x)=r_{k-2}u_1(x)+r_{k-1}(x)v_1(x).$$

把式 $(k-1)$ 代入上式,得到

$$\lambda d(x)=r_{k-3}(x)v_1(x)+r_{k-2}(x)\big(u_1(x)-q_{k-1}(x)v_1(x)\big).$$

记 $u_2(x)=v_1(x)$, $v_2(x)=u_2(x)-q_{k-1}(x)v_1(x)$,上式即为

$$\lambda d(x)=r_{k-3}(x)u_2(x)+r_{k-2}(x)v_2(x).$$

逐个地把以上等式代入. 假设进行了 $k-2$ 次,得到

$$\lambda d(x) = f_2(x)u_{k-1}(x) + r_1(x)v_{k-1}(x).$$

最后把式 (1) 代入上式,得到

$$\lambda d(x) = f_1(x)v_{k-1}(x) + f_2(x)\big(u_{k-1}(x) - q_1(x)v_{k-1}(x)\big).$$

取 $$u(x) = \lambda^{-1}v_{k-1}(x), \qquad v(x) = \lambda^{-1}\big(u_{k-1}(x) - q_1(x)v_{k-1}(x)\big),$$

即得定理 1.3.3. ■

容易看出,对两个不全为零的多项式 $f(x)$ 与 $g(x)$,所有零次多项式都是它们的公因式. 如果 $f(x)$ 与 $g(x)$ 除零次多项式外不含其它的公因式,则 $f(x)$ 与 $g(x)$ 称为**互素**.

根据最大公因式的定义可以看出,两个多项式互素的充分必要条件是它们的最大公因式为 1. 因此,由定理 1.3.3 直接得到

推论 1.3.3 多项式 $f(x)$ 与 $g(x)$ 互素当且仅当存在 $u(x)$ 与 $v(x)$,使得

$$f(x)u(x) + g(x)v(x) = 1. \tag{1.3.4}$$

关于两个多项式的互素,有以下性质.

性质 1.3.4 设 $f(x), g(x), h(x)$ 是多项式,$\big(f(x), g(x)\big) = 1$,$\big(f(x), h(x)\big) = 1$,则

$$\big(f(x), g(x)h(x)\big) = 1.$$

证明 因为 $\big(f(x), g(x)\big) = 1$,故存在多项式 $u(x)$ 和 $v(x)$,使得

$$f(x)u(x) + g(x)v(x) = 1.$$

上式两端同乘以 $h(x)$,得到

$$f(x)\big(u(x)h(x)\big) + \big(g(x)h(x)\big)v(x) = h(x).$$

由此可知,如果多项式 $w(x)$ 是 $f(x)$ 与 $g(x)h(x)$ 的公因式,则 $w(x)$ 也是 $f(x)$ 与 $h(x)$ 的公因式. 由于 $\big(f(x), h(x)\big) = 1$,因此 $w(x)$ 是零次多项式. 这表明 $f(x)$ 与 $g(x)h(x)$ 除零次多项式外不含其它公因式,即 $f(x)$ 与 $g(x)h(x)$ 互素. ■

性质 1.3.5 设 $f(x), g(x)$ 与 $h(x)$ 是多项式,其中 $\big(g(x), h(x)\big) = 1$,并且 $h(x) \mid f(x)g(x)$,则

$$h(x) \mid f(x).$$

证明 因为 $\big(g(x), h(x)\big) = 1$,故存在多项式 $u(x)$ 与 $v(x)$,使得

$$h(x)u(x) + g(x)v(x) = 1.$$

上式两端同乘以 $f(x)$,得到

$$h(x)\big(u(x)f(x)\big) + \big(g(x)f(x)\big)v(x) = f(x).$$

由此可知,$h(x) \mid f(x)$. ■

性质 1.3.6 设 $f(x),g(x)$ 与 $h(x)$ 是多项式，$f(x)\mid h(x)$，$g(x)\mid h(x)$，且

$$\big(f(x),g(x)\big)=1,$$

则 $f(x)g(x)\mid h(x)$.

例 1.3.3 求多项式 $u(x)$ 和 $v(x)$，使得

$$x^m u(x)+(1-x)^n v(x)=1. \tag{1.3.5}$$

解 显然，多项式 x^m 与 $(1-x)^n$ 互素. 所以由定理 1.3.3，适合式 (1.3.5) 的多项式 $u(x)$ 与 $v(x)$ 是存在的.

如果 $\deg u(x)\geqslant n$，$\deg v(x)\geqslant m$，则由定理 1.3.1，存在多项式 $q_1(x),u_1(x),q_2(x)$ 与 $v_1(x)$，$\deg u_1(x)<n$，$\deg v_1(x)<m$，使得

$$u(x)=(1-x)^n q_1(x)+u_1(x),\quad v(x)=x^m q_2(x)+v_1(x).$$

代入式 (1.3.5) 得到

$$x^m(1-x)^n\big(q_1(x)+q_2(x)\big)=1-x^m u_1(x)-(1-x)^n v_1(x).$$

比较两端多项式的次数，得到

$$x^m u_1(x)+(1-x)^n v_1(x)=1.$$

因此可设 $\deg u(x)<n$，$\deg v(x)<m$.

设

$$u(x)=\sum_{i=0}^{n-1}a_i(1-x)^i,\qquad v(x)=\sum_{j=0}^{m-1}b_jx^j.$$

显然，

$$a_i=\frac{(-1)^i}{i!}u^{(i)}(1),\qquad b_j=\frac{1}{j!}v^{(j)}(0).$$

将 $u(x)$ 与 $v(x)$ 代入式 (1.3.5). 令 $x=0$，则得到 $b_0=v(0)=1$. 令 $x=1$，则得到 $a_0=u(1)=1$. 对式 (1.3.5) 求 k 阶微商，得到

$$\sum_{i=0}^{k}C_k^i\frac{m!}{(m-k+i)!}x^{m-k+i}u^{(i)}(x)+\sum_{i=0}^{k}(-1)^{k-i}C_k^i\frac{n!}{(n-k+i)!}(1-x)^{n-k+i}v^{(i)}(x)=0. \tag{1.3.6}$$

设 $1\leqslant k\leqslant m-1$，则 $m-k+i\geqslant 1$，其中 $i=0,1,\ldots,k$. 在式 (1.3.6) 中令 $x=0$，得到

$$\sum_{i=0}^{k}(-1)^{k-i}C_k^i\frac{n!}{(n-k+i)!}v^{(i)}(0)=0.$$

由此知

$$\sum_{i=0}^{k}(-1)^iC_n^{k-i}b_i=0.$$

进而解得 $b_j=C_{n+j-1}^j$. 于是，

$$v(x)=\sum_{i=0}^{m-1}C_{n+j-1}^jx^j.$$

同样,设 $1\leqslant k\leqslant n-1$,则 $n-k+i\geqslant 1$,其中 $i=0,1,\ldots,k$. 在式 (1.3.6) 中令 $x=1$,得到

$$\sum_{i=0}^{k}\mathrm{C}_k^i\frac{m!}{(m-k+i)!}u^{(i)}(1)=0.$$

由此知

$$\sum_{i=0}^{k}(-1)^i\mathrm{C}_m^{k-i}a_i=0.$$

解得 $a_i=\mathrm{C}_{m+i-1}^i$. 因此,

$$u(x)=\sum_{i=0}^{n-1}\mathrm{C}_{m+i-1}^i(1-x)^i.$$

■

注 对于给定多项式 $f(x)$ 与 $g(x)$,求多项式 $u(x)$ 与 $v(x)$,使得

$$f(x)u(x)+g(x)v(x)=d(x),$$

这里 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因式,其方法很多的.

方法之一是,用辗转相除法,求出定理 1.3.3 的证明中的等式 (1),(2),...,(k);然后,如同定理 1.3.3 的证明,把这里些等式逐个地由后往前代入,即可求出 $u(x)$ 与 $v(x)$.

方法之二是,由于定理 1.3.3 保证了 $u(x)$ 与 $v(x)$ 的存在性,因此可以用**待定系数法**. 例 1.3.3 采用的就是待定系数法.

最大公因式的概念可以推广到有限多个不全为零的多项式的情形.

定义 1.3.3 设不全为零的多项式 $f_1(x),f_2(x),\ldots,f_s(x)\in\mathbb{F}[x]$,$h(x)\in\mathbb{F}[x]$. 如果对任意 $i=1,2,\ldots,s$,都有

$$h(x)\mid f_i(x),$$

则 $h(x)$ 称为 $f_1(x),f_2(x),\ldots,f_s(x)$ 的**公因式**. 如果首一多项式 $d(x)\in\mathbb{F}[x]$ 是 $f_1(x)$, $f_2(x),\ldots,f_s(x)$ 的公因式,而 $f_1(x),f_2(x),\ldots,f_s(x)$ 的每个公因式都是 $d(x)$ 的因式,则 $d(x)$ 称为 $f_1(x),f_2(x),\ldots,f_s(x)$ 的**最大公因式**,记为

$$d(x)=\big(f_1(x),f_2(x),\ldots,f_s(x)\big).$$

定理 1.3.4 不全为零的多项式 $f_1(x),f_2(x),\ldots,f_s(x)$ 的最大公因式存在且唯一,而且

$$d(x)=\big(f_1(x),f_2(x),\ldots,f_s(x)\big)=\Big(\big(f_1(x),f_2(x),\ldots,f_{s-1}(x)\big),f_s(x)\Big).\tag{1.3.7}$$

证明 先证明最大公因式的存在性与式 (1.3.7) 成立. 对多项式的个数 s 用归纳法. 当 $s=2$ 时,显然最大公因式存在且式 (1.3.7) 成立.

假设当 $s=k-1$ 时最大公因式存在且式 (1.3.7) 成立. 记 $k-1$ 个不全为零的多项式 $f_1(x),f_2(x),\ldots,f_{k-1}(x)$ 的最大公因式为 $\widetilde{d}(x)$. 由定理 1.3.2,多项式 $\widetilde{d}(x)$ 与 $f_k(x)$ 的最大公因式存在,记为 $d(x)$.

显然,$d(x)$ 是 $\widetilde{d}(x)$ 与 $f_k(x)$ 的公因式. 由于 $\widetilde{d}(x)$ 是 $f_1(x),f_2(x),\ldots,f_{k-1}(x)$ 的公因式,因此,$d(x)$ 是 $f_1(x),f_2(x),\ldots,f_k(x)$ 的公因式.

另一方面,设 $h(x)$ 是 $f_1(x),f_2(x),\ldots,f_k(x)$ 的公因式,则 $h(x)$ 是 $f_1(x),f_2(x)$,

$\ldots, f_{k-1}(x)$ 的公因式，从而 $h(x)$ 是 $\widetilde{d}(x)$ 的因式. 即 $h(x)$ 是 $\widetilde{d}(x)$ 和 $f_k(x)$ 的公因式. 因此，$h(x)$ 也是 $d(x)$ 的因式.

这表明，$d(x)$ 是 $f_1(x), f_2(x), \ldots, f_k(x)$ 的最大公因式. 即

$$d(x) = \left(\widetilde{d}(x), f_k(x)\right).$$

这就证明了 $f_1(x), f_2(x), \ldots, f_k(x)$ 的最大公因式存在而且式 (1.3.7) 成立.

至于唯一性的证明和 $s=2$ 的情形是类似的. 从略. ■

由式 (1.3.7) 与定理 1.3.3 直接得到，

推论 1.3.4　设 $d(x)$ 是多项式 $f_1(x), f_2(x), \ldots, f_s(x) \in \mathbb{F}[x]$ 的最大公因式，则存在多项式 $u_1(x), u_2(x), \ldots, u_s(x) \in \mathbb{F}[x]$，使得

$$f_1(x)u_1(x) + f_2(x)u_2(x) + \cdots + f_s(x)u_s(x) = 0. \tag{1.3.8}$$

如果多项式 $f_1(x), f_2(x), \ldots, f_s(x)$ 的公因式只能是零次多项式，则称 $f_1(x), f_2(x), \ldots, f_s(x)$ 是互素的.

容易看出，多项式 $f_1(x), f_2(x), \ldots, f_s(x)$ 互素的充分必要条件是它们的最大公因式为 1. 注意，当 $s>2$ 时，如果 $f_1(x), f_2(x), \ldots, f_s(x)$ 互素，这些多项式并不一定两两互素.

习　题　1.3

1. 设多项式 $g(x) = x^2 - 2ax + 2$ 整除多项式 $f(x) = x^4 + 3x^2 + ax + b$，求 a 和 b. 这里 $a, b \in \mathbb{R}$.

2. 设 m, n 和 p 为正整数. 证明，多项式 $g(x) = x^2 + x + 1$ 整除多项式 $f(x) = x^{3m} + x^{3n+1} + x^{3p+2}$.

3. 证明，当 $n = 6m + 5$ 时，多项式 $x^2 + xy + y^2$ 整除多项式 $(x+y)^n - x^n - y^n$；当 $n = 6m + 1$ 时，多项式 $(x^2 + xy + y^2)^2$ 整除多项式 $(x+y)^n - x^n - y^n$. 这里 m 是使 $n > 0$ 的整数，而 x 与 y 是实数.

4. 求多项式 $f(x)$ 与 $g(x)$ 的最大公因式.

(1) $f(x) = x^4 + x^3 - 3x^2 - 4x - 1, g(x) = x^3 + x^2 - x - 1$;

(2) $f(x) = x^6 + 2x^4 - 4x^3 - 3x^2 + 8x - 5, g(x) = x^5 + x^2 - x + 1$;

(3) $f(x) = 3x^6 - x^5 - 9x^4 - 14x^3 - 11x^2 - 3x - 1, g(x) = 3x^5 + 8x^4 + 9x^3 + 15x^2 + 10x + 9$.

5. 确定多项式 $u(x)$ 与 $v(x)$，使得 $f(x)u(x) + g(x)v(x) = d(x)$，其中 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因式.

(1) $f(x) = x^4 + 2x^3 - x^2 - 4x - 2, g(x) = x^4 + x^3 - x^2 - 2x - 2$;

(2) $f(x) = 3x^5 + 5x^4 - 16x^3 - 6x^2 - 5x - 6, g(x) = 3x^4 - 4x^3 - x^2 - x - 2$;

(3) $f(x) = 3x^3 - 2x^2 + x + 2, g(x) = x^2 - x + 1$;

(4) $f(x) = x^4 - x^3 - 4x^2 + 4x + 1, g(x) = x^2 - x - 1$.

6. 用待定系数法确定多项式 $u(x)$ 与 $v(x)$，使得 $f(x)u(x) + g(x)v(x) = 1$，其中 $f(x)$ 与 $g(x)$ 如下：

(1) $f(x) = x^3, g(x) = (1-x)^2$;　　(2) $f(x) = x^4, g(x) = (1-x)^4$;

(3) $f(x) = x^4 - 4x^3 + 1, g(x) = x^3 - 3x^2 + 1$.

7. 求次数最低的多项式 $u(x)$ 与 $v(x)$，使得

(1) $(x^4 - 2x^3 - 4x^2 + 6x + 1)u(x) + (x^3 - 5x - 3)v(x) = x^4$;

(2) $(x^4+2x^3+x+1)u(x)+(x^4+x^3-2x^2+2x-1)v(x)=x^3-2x$.

8. 求次数最低的多项式 $f(x)$,使得 $f(x)$ 被多项式 $(x-1)^2$ 除时余式为 $2x$,被多项式 $(x-2)^3$ 除时余式为 $3x$.

9. 求次数最低的多项式 $f(x)$,使得 $f(x)$ 被多项式 $x^4-2x^3-2x^2+10x-7$ 除时余式为 x^2+x+1,被多项式 $x^4-2x^3-3x^2+13x-10$ 除时余式为 $2x^2-3$.

10. 设 $f(x)$ 是 $2n+1$ 次多项式,n 为正整数,$f(x)+1$ 被 $(x-1)^n$ 整除,而 $f(x)-1$ 被 $(x+1)^n$ 整除. 求 $f(x)$.

§1.4 唯一析因定理

大家知道,在整数环 $\mathbb{Z}$ 中素数起着重要的作用. 所谓素数是指,除 ±1 和自身外不含其它因子的整数. 整数环 $\mathbb{Z}$ 中每个非零整数都可以分解为若干个素数的乘积,而且不计素因子的正负号和顺序,这种分解是唯一的.

对数域 $\mathbb{F}$ 上的一元多项式环 $\mathbb{F}[x]$,也有类似的结论. 为了介绍多项式环的唯一析因定理,先引述以下的定义.

定义 1.4.1 设 $f(x)$ 是数域 $\mathbb{F}$ 上的 n 次多项式,$n\geqslant1$. 如果存在次数小于 n 的多项式 $g(x),h(x)\in\mathbb{F}[x]$,使得 $f(x)=g(x)h(x)$,则多项式 $f(x)$ 称为**在 $\mathbb{F}$ 上可约**. 如果多项式 $f(x)$ 不是在 $\mathbb{F}$ 上可约,则 $f(x)$ 称为**在 $\mathbb{F}$ 上不可约**.

应当注意,一个多项式在数域 $\mathbb{F}$ 上不可约,在包含数域 $\mathbb{F}$ 的数域 $\mathbb{K}$ 上这个多项式有可能是可约的. 例如,多项式 x^2+1 在实数域 $\mathbb{R}$ 上不可约,但是,由于 $x^2+1=(x-\mathrm{i})(x+\mathrm{i})$,这里 $\mathrm{i}^2=-1$. 因此多项式 x^2+1 在复数域上是可约的. 所以,多项式的不可约性是相对给定的数域而言的.

关于不可约多项式,有以下简单性质.

性质 1.4.1 设多项式 $p(x)$ 在 $\mathbb{F}$ 上不可约,且 a 是 $\mathbb{F}$ 中非零的数,则多项式 $ap(x)$ 在 $\mathbb{F}$ 上不可约.

性质 1.4.2 设多项式 $f(x)\in\mathbb{F}[x]$,且 $p(x)$ 是 F 上的不可约多项式,则 $p(x)\mid f(x)$,或者 $p(x)$ 与 $f(x)$ 互素.

证明 设 $f(x)$ 与 $p(x)$ 不互素,则它们的最大公因式 $d(x)\neq1$,即 $d(x)$ 是 $p(x)$ 的因式. 因为 $p(x)$ 在 $\mathbb{F}$ 上不可约,所以 $p(x)=ad(x)$,$a\in\mathbb{F}$. 而 $d(x)$ 是 $f(x)$ 的因式,故 $p(x)$ 也是 $f(x)$ 的因式,即 $p(x)\mid f(x)$. ■

性质 1.4.3 设多项式 $f(x),g(x)\in\mathbb{F}[x]$,$p(x)$ 是数域 $\mathbb{F}$ 上的不可约多项式. 如果 $p(x)\mid f(x)g(x)$,则 $p(x)\mid f(x)$,或者 $p(x)\mid g(x)$.

证明 如果 $p(x)\nmid f(x)$,则 $\big(p(x),f(x)\big)=1$. 因此,存在多项式 $u(x),v(x)\in\mathbb{F}[x]$,使得

$$p(x)u(x)+f(x)v(x)=1.$$

因此,

$$p(x)\big(u(x)g(x)\big)+\big(f(x)g(x)\big)v(x)=g(x).$$

由此即知, $p(x)\mid g(x)$. ■

下面是本节的主要定理.

定理 1.4.1 (唯一析因定理) 设 n 次多项式 $f(x)\in\mathbb{F}[x]$, 则存在数域 $\mathbb{F}$ 上的不可约多项式 $p_1(x),p_2(x),\ldots,p_s(x)\in\mathbb{F}[x]$, 使得

$$f(x)=p_1(x)p_2(x)\cdots p_s(x).$$

如果另有不可约多项式 $q_1(x),q_2(x),\ldots,q_t(x)\in\mathbb{F}[x]$, 使得

$$f(x)=q_1(x)q_2(x)\cdots q_t(x),$$

则 $s=t$, 并且可以适当调动因式的次序, 使得

$$q_i(x)=a_ip_i(x),$$

其中 $a_i\in\mathbb{F}$, $i=1,2,\ldots,s$.

如果不可约多项式 $p(x)$ 整除多项式 $f(x)$, 则 $p(x)$ 称为 $f(x)$ 的**不可约因式**. 把多项式 $f(x)$ 分解为若干个不可约因式的乘积, 称为**对 $f(x)$ 施行不可约分解**. 于是, 定理 1.4.1 可以简单叙述为:

每个多项式都可以分解为不可约因式的乘积, 而且如果不计不可约因式的次序和零次因式, 这种不可约分解是唯一的.

证明　存在性　对多项式的次数 n 用归纳法.

显然, 一次多项式在数域 $\mathbb{F}$ 上都是不可约的, 因此结论对 $n=1$ 成立. 假设结论对次数小于 n 的多项式都成立, 下面证明结论对 n 次多项式 $f(x)$ 成立.

如果 $f(x)$ 本身在 $\mathbb{F}$ 上不可约, 则 $f(x)$ 的不可约分解由自身组成; 如果 $f(x)$ 在 $\mathbb{F}$ 上可约, 则存在次数小于 n 的多项式 $g(x),h(x)\in\mathbb{F}[x]$, 使得

$$f(x)=g(x)h(x).$$

由于 $g(x)$ 和 $h(x)$ 的次数都小于 n, 故由归纳假设, 存在不可约多项式 $p_1(x)$, $p_2(x),\ldots,p_k(x)$ 和 $p_{k+1}(x),\ldots,p_s(x)\in\mathbb{F}[x]$, 使得

$$g(x)=p_1(x)p_2(x)\cdots p_k(x),\qquad h(x)=p_{k+1}(x)p_{k+2}(x)\cdots p_s(x).$$

于是,

$$f(x)=p_1(x)p_2(x)\cdots p_k(x)p_{k+1}(x)\cdots p_s(x).$$

唯一性　现在设多项式 $f(x)$ 具有两个不可约分解, 即设

$$f(x)=p_1(x)p_2(x)\cdots p_s(x)=q_1(x)q_2(x)\cdots q_t(x).\tag{1.4.1}$$

因为 $q_1(x)$ 在 $\mathbb{F}$ 上不可约, 并且 $q_1(x)\mid p_1(x)p_2(x)\cdots p_s(x)$, 因此, 由性质 1.4.3, 存在某 $1\leqslant i\leqslant s$ 使得 $q_1(x)\mid p_i(x)$. 适当地调整不可约因式 $p_1(x),p_2(x),\ldots,p_s(x)$ 的次序, 可设 $q_1(x)\mid p_1(x)$. 由于 $p_1(x)$ 在 $\mathbb{F}$ 上不可约, 因此, 存在 $a_1\in\mathbb{F}$ 使得 $p_1(x)=a_1q_1(x)$. 于是由式 (1.4.1) 得到:

$$\big(a_1p_2(x)\big)p_3(x)\cdots p_s(x)=q_2(x)q_3(x)\cdots q_t(x)=g(x). \tag{1.4.2}$$

由性质 1.4.1, $a_1p_2(x)$ 在 $\mathbb{F}$ 上不可约. 因此式 (1.4.2) 是次数小于 n 的多项式 $g(x)$ 的两个不可约分解,根据归纳假设得到, $s-1=t-1$,即 $s=t$. 并且可适当调整不可约因式 $a_1p_2(x),p_3(x),\ldots,p_s(x)$ 的次序,使得

$$q_2(x)=a_2'a_1p_2(x),q_3(x)=a_3p_3(x),\ldots,q_s(x)=a_sp_s(x),$$

其中 $a_2',a_3,\ldots,a_s\in\mathbb{F}$. 记 $a_2'a_1=a_2$,则得到 $q_i(x)=a_ip_i(x)$, $i=1,2,\ldots,s$. ■

应当指出,如果要求数域 $\mathbb{F}$ 上的不可约多项式是首一的,则由定理 1.4.1 直接得到,存在首一不可约多项式 $p_1(x),p_2(x),\ldots,p_s(x)\in\mathbb{F}[x]$,使得多项式 $f(x)$ 可以表为

$$f(x)=a_0p_1(x)p_2(x)\cdots p_s(x),$$

其中 a_0 为 $f(x)$ 的首项系数. 对于多项式 $f(x)$ 的这种不可约分解,除了不可约因式的次序外是唯一的.

一般地说,出现在多项式 $f(x)$ 的一个不可约分解中的不可约因式不一定都不相同. 如果不可约因式 $p(x)$ 不只出现一次,则 $p(x)$ 称为 $f(x)$ 的**重因式**,否则称为**单因式**. 如果 $p(x)$ 恰好出现 k 次,则 $p(x)$ 称为 $f(x)$ 的 $\boldsymbol{k}$ **重因式**.

设多项式 $f(x)$ 具有不可约分解

$$f(x)=a_0p_1(x)p_2(x)\cdots p_s(x), \tag{1.4.3}$$

其中 $p_1(x),p_2(x),\ldots,p_s(x)$ 是不可约的首一多项式, a_0 是 $f(x)$ 的首项系数. 又设分解式 (1.4.3) 中所有不同的不可约因式为 $p_1(x),p_2(x),\ldots,p_\ell(x)$,它们分别是 $f(x)$ 的 $k_1,k_2,\ldots,k_\ell$ 重因式,则式 (1.4.3) 可以写成

$$f(x)=a_0p_1^{k_1}(x)p_2^{k_2}(x)\cdots p_\ell^{k_\ell}(x). \tag{1.4.4}$$

设多项式 $f(x)$ 和 $g(x)$ 的所有不同的首一不可约因式分别为 $h_1(x),h_2(x),\ldots,h_s(x)$ 和 $q_1(x),q_2(x),\ldots,q_s(x)$. 它们的并集记为 $\{p_1(x),p_2(x),\ldots,p_\ell(x)\}$. 则 $f(x)$ 与 $g(x)$ 的不可约分解可以表为

$$f(x)=a_0p_1^{k_1}(x)p_2^{k_2}(x)\cdots p_\ell^{k_\ell}(x),$$
$$g(x)=b_0p_1^{e_1}(x)p_2^{e_2}(x)\cdots p_\ell^{e_\ell}(x),$$

其中 a_0 与 b_0 分别是 $f(x)$ 和 $g(x)$ 的首项系数,而 k_i 与 e_i 是非负整数, $i=1,2,\ldots,\ell$. 于是, $f(x)$ 与 $g(x)$ 的最大公因式为

$$\big(f(x),g(x)\big)=p_1^{m_1}(x)p_2^{m_2}(x)\cdots p_\ell^{m_\ell}(x),$$

其中 $m_i=\min\{k_i,e_i\}$, $i=1,2,\ldots,\ell$.

§1.5　实系数与复系数多项式

系数都是实数或者都是复数的多项式分别称为**实系数**或**复系数多项式**. 这节讨论实系数多项式与复系数多项式的唯一析因理论. 先证明以下的定理.

定理 1.5.1　数域 $\mathbb{F}$ 上的 n 次多项式 $f(x)$ 在 $\mathbb{F}$ 上至多有 n 个不同的根, 其中 $n>0$.

证明　设 $a_1,a_2,\ldots,a_r$ 是 $f(x)$ 的不同的根, $a_1,a_2,\ldots,a_r\in\mathbb{F}$. 下面对 r 用归纳法证明

$$(x-a_1)(x-a_2)\cdots(x-a_r)\,\big|\,f(x).$$

事实上, 当 $r=1$ 时. 因为 a_1 是 $f(x)$ 的根, 故由因式定理, $(x-a_1)\,\big|\,f(x)$.

假设结论对 $r-1$ 成立, 现在证明结论对 r 成立. 因为 $a_1,a_2,\ldots,a_{r-1}$ 是 $f(x)$ 的根, 故由归纳假设, $(x-a_1)(x-a_2)\cdots(x-a_{r-1})\,\big|\,f(x)$,

$$f(x)=(x-a_1)(x-a_2)\cdots(x-a_{r-1})h(x).$$

其中 $h(x)\in\mathbb{F}[x]$. 由于 a_r 为 $f(x)$ 的根, 故

$$f(a_r)=(a_r-a_1)(a_r-a_2)\cdots(a_r-a_{r-1})h(a_r)=0.$$

因为 $a_1,a_2,\ldots,a_r$ 是 $f(x)$ 的不同的根, 因此 $a_r-a_i\neq0$, $i=1,2,\ldots,r-1$. 所以 $h(a_r)=0$. 由因式定理, $h(x)=(x-a_r)g(x)$, 于是,

$$f(x)=(x-a_1)(x-a_2)\cdots(x-a_{r-1})(x-a_r)g(x).\qquad\blacksquare$$

定理 1.5.1 并没有告诉我们, n 次多项式 $f(x)\in\mathbb{F}[x]$ 一定在数域 $\mathbb{F}$ 上有根. 例如, 多项式 x^2+1 在实数域 $\mathbb{R}$ 上就没有根. 但是, 当数域 $\mathbb{F}$ 为复数域 $\mathbb{C}$ 时, 有

定理 1.5.2 (代数基本定理) 任意一个 n 次复系数多项式一定有复数根, 其中 $n\geqslant1$.

这个定理是人们早就知道的. 直到 1797 年, 二十岁的德国大数学家 Gauss 才第一个给出证明. 后来 Gauss 又给出三个证明. 由于十九世纪以前的代数是以研究代数方程为中心的, 而这个定理对代数方程论又具有基本重要性, 所以人们称它为代数基本定理. 这个定理的证明有的涉及复变函数论知识, 而初等证明的篇幅又嫌太长, 这里就不给出了.

利用代数基本定理容易证明,

定理 1.5.3　设 $f(x)$ 是任意一个 n 次复系数多项式, $n>0$, 则 $f(x)$ 恰有 n 个复数根 $c_1,c_2,\ldots,c_n$, 而且

$$f(x)=a_0(x-c_1)(x-c_2)\cdots(x-c_n),\tag{1.5.1}$$

其中 a_0 是 $f(x)$ 的首项系数.

证明 对多项式 $f(x)$ 的次数 n 用归纳法. 当 $n=1$ 时定理显然成立.

假设定理对次数为 $n-1$ 的多项式成立, $f(x)$ 是 n 次复系数多项式. 由代数基本定理, $f(x)$ 具有复数根 c_1, 由因式定理, $f(x)=(x-c_1)g(x)$, 其中 $g(x)$ 为 $n-1$ 次复系数多项式.

由归纳假设, $g(x)$ 恰有 $n-1$ 个复数根 $c_2,c_3,\dots,c_n$, 并且

$$g(x)=a_0(x-c_2)(x-c_3)\cdots(x-c_n).$$

于是, $f(x)=a_0(x-c_1)(x-c_2)\cdots(x-c_n)$. 显然, $c_1,c_2,\dots,c_n$ 是 $f(x)$ 的 n 个复数根, 而且 a_0 是 $f(x)$ 的首项系数. ■

应当说明, n 次多项式 $f(x)$ 的 n 个根 $c_1,c_2,\dots,c_n$ 不一定都不相同. 如果 $f(x)$ 的根 c 在 $c_1,c_2,\dots,c_n$ 中出现 k 次, 则 c 称为 $f(x)$ 的 **k 重根**. 1 重根称为**单根**.

设 n 次多项式 $f(x)$ 的所有不同的根为 $c_1,c_2,\dots,c_s$, 它们的重数分别为 $k_1,k_2,\dots,k_s$, 则 $f(x)$ 的分解式 (1.5.1) 可以写为

$$f(x)=a_0(x-c_1)^{k_1}(x-c_2)^{k_2}\cdots(x-c_s)^{k_s},$$

其中正整数 $k_1,k_2,\dots,k_s$ 适合 $k_1+k_2+\cdots+k_s=n$.

我们知道, 复系数一次多项式一定是不可约的. 定理 1.5.2 表明, 任何 n 次复系数多项式 $f(x)$ 在复数域上都是可约的, 其中 $n\geqslant 2$. 因此, 复系数多项式 $p(x)$ 在复数域 $\mathbb{C}$ 上不可约的充分必要条件是 $\deg p(x)=1$. 利用这一事实和 §1.4 证明的唯一析因定理, 也可以直接得到定理 1.5.3. 所以, 定理 1.5.3 是复系数多项式的唯一析因定理.

下面讨论实系数多项式的不可约分解.

定理 1.5.4 实系数多项式 $f(x)$ 的复数根共轭成对出现.

证明 设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, 其中 $a_0,a_1,\dots,a_n\in\mathbb{R}$, 且设 c 是 $f(x)$ 的复数根, 则

$$f(c)=a_nc^n+a_{n-1}c^{n-1}+\cdots+a_1c+a_0=0.$$

上式两端取共轭, 并注意 a_i 是实数, $i=1,2,\dots,n$, 则得到,

$$a_n\overline{c}^n+a_{n-1}\overline{c}^{n-1}+\cdots+a_1\overline{c}+a_0=0.$$

其中 $\overline{c}$ 是 c 的共轭复数, 即 $f(\overline{c})=0$. 因此, $\overline{c}$ 也是 $f(x)$ 的根. ■

对复系数多项式, 定理 1.5.4 并不成立. 例如, 复系数多项式 $x^2-\mathrm{i}x=x(x-\mathrm{i})$ 的根为 0 和 i, 它们并不共轭.

由定理 1.5.4 可以知道, 奇次实系数多项式一定有实数根.

定理 1.5.5 实系数多项式 $p(x)$ 在实数域上不可约, 则 $p(x)$ 的次数为 1 或 2.

证明 反证法. 设 $\deg p(x)=n\geqslant 3$. 根据定理 1.5.2, 作为复系数多项式, $p(x)$ 具有复数根. 如果 $p(x)$ 具有实数根 a, 则由因式定理,

$$p(x)=(x-a)f(x),$$

其中 $f(x)$ 是实系数多项式. 这表明, $p(x)$ 在实数域上可约, 与假设矛盾.

如果 $p(x)$ 的根都是复数(不能是实数), 则由定理 1.5.4, n 为偶数. 设 $n=2k$, $k\geqslant 2$, 且设 $c_1,\bar{c}_1,c_2,\bar{c}_2,\ldots,c_k,\bar{c}_k$ 是 $p(x)$ 的根. 因此,

$$p(x)=a_0(x-c_1)(x-\bar{c}_1)\cdots(x-c_k)(x-\bar{c}_k).$$

记

$$p_i(x)=(x-c_i)(x-\bar{c}_i)=x^2-(c_i+\bar{c}_i)x+c_i\bar{c}_i,$$

其中 $i=1,2,\ldots,k$. 显然, $p_i(x)$ 是实系数的. 因此, $p(x)$ 在实数域上可约, 与假设矛盾. 这就证明, $1\leqslant\deg(x)\leqslant 2$. ■

利用二次方程的判别式, 容易知道, 实二次多项式 x^2+px+q 在实数域上不可约的充分必要条件是, 它的判别式 $p^2-4q<0$.

定理 1.5.6 n 次实系数多项式 $f(x)$ 可分解为一次因式和二次不可约因式的乘积, 即

$$\begin{aligned}f(x)=a_0&(x-c_1)^{k_1}(x-c_2)^{k_2}\cdots(x-c_s)^{k_s}\\&\times(x^2+p_1x+q_1)^{e_1}(x^2+p_2x+q_2)^{e_2}\cdots(x2+p_tx+q_t)^{e_t},\end{aligned}\tag{1.5.2}$$

其中 k_i 和 e_j 是正整数, $1\leqslant i\leqslant s$, $1\leqslant j\leqslant t$, 且

$$k_1+k_2+\cdots+k_s+2(e_1+e_2+\cdots+e_t)=n,$$

而 a_0 是 $f(x)$ 的首项系数, $c_1,c_2,\ldots,c_s$, $p_1,p_2,\ldots,p_t$ 和 $q_1,q_2,\ldots,q_t$ 都是实数, 并且

$$p_j^2-4q_j<0.$$

当然, 如果 $f(x)$ 没有实根, 则式 (1.5.2) 中的一次因式不出现; 如果 $f(x)$ 的根都是实数, 则二次因式不出现.

证明 这是 §1.4 中唯一析因定理和定理 1.5.5 的直接推论. ■

习 题 1.5

1. 把下列复系数多项式分解为一次因式的乘积.

(1) $(x+\cos\theta+\mathrm{i}\sin\theta)^n+(x+\cos\theta-\mathrm{i}\sin\theta)^n$;

(2) $(x+1)^n-(x-1)^n$;

(3) $x^n-\mathrm{C}_{2n}^2x^{n-1}+\mathrm{C}_{2n}^4x^{n-2}+\cdots+(-1)^n\mathrm{C}_{2n}^{2n}$;

(4) $x^{2n}+\mathrm{C}_{2n}^2x^{2n-2}(x^2-1)+\mathrm{C}_{2n}^4x^{2n-4}(x^2-1)^2+\cdots+(x^2-1)^n$;

(5) $x^{2n+1}+\mathrm{C}_{2n+1}^2x^{2n-1}(x^2-1)+\mathrm{C}_{2n+1}^4x^{2n-3}(x^2-1)^2+\cdots+x(x^2-1)^n$.

2. 把下列实系数多项式分解为实的不可约因式的乘积.

(1) x^4+1;

(2) x^6+27;

(3) $x^4+4x^3+4x^2+1$;

(4) $x^{2n}-2x^n+2$;

(5) $x^4-ax^2+1, -2<a<2$;

(6) $x^{2n}+x^n+1$.

3. 证明, 复系数多项式 $f(x)$ 对所有实数 x 恒取正值的充分必要条件是, 存在没有实数根的复系数多项式 $\varphi(x)$, 使得 $f(x)=|\varphi(x)|^2$.

4. 证明, 实系数多项式 $f(x)$ 对所有实数 x 恒取非负实数值的充分必要条件是, 存在实系数多项式 $\varphi(x)$ 和 $\psi(x)$, 使得 $f(x)=\varphi^2(x)+\psi^2(x)$.

§1.6 整系数与有理系数多项式

系数都是整数或者都是有理数的多项式称为**整系数多项式**或**有理系数多项式**. 根据定理 1.4.1,有理系数多项式可以分解为有理系数不可约多项式的乘积,而且不计不可约因式的次序与零次因式,不可约分解是唯一的. 问题是,如何判定一个有理系数多项式是否在有理数域 $\mathbb{Q}$ 上不可约? 另外,定理 1.4.1 (即多项式的唯一析因定理)是对数域 $\mathbb{F}$ 而言的,对整数环 $\mathbb{Z}$,唯一析因定理是否仍成立,这是本节所要讨论的.

和数域 $\mathbb{F}$ 上不可约多项式的定义相仿,可以给出不可约整系数多项式的定义.

设 n 是正整数,如果 n 次整系数多项式 $f(x)$ 可以表为两个次数小于 n 的整系数多项式的乘积,则 $f(x)$ 称为在整数环 $\mathbb{Z}$ 上**不可约**. 否则,$f(x)$ 称为在整数环 $\mathbb{Z}$ 上**可约**.

容易看出,如果整系数多项式 $f(x)$ 在 $\mathbb{Z}$ 上可约,则作为有理系数多项式,$f(x)$ 在有理数域 $\mathbb{Q}$ 上也可约. 反之,如果整系数多项式 $f(x)$ 在 $\mathbb{Q}$ 上可约,$f(x)$ 是否在 $\mathbb{Z}$ 上也可约?

为了回答这个问题,先引进以下的概念.

定义 1.6.1 如果整系数多项式 $f(x) = a_0 + a_1x + \cdots + a_nx^n$ 的系数 $a_0, a_1, \ldots, a_n$ 的最大公因子为 1,则 $f(x)$ 称为**本原多项式**.

设整系数多项式 $f(x) = a_0 + a_1x + \cdots + a_nx^n$ 的系数的最大公因子为

$$d = (a_0, a_1, \ldots, a_n),$$

则 $a_i = da_i'$,其中 $a_i' \in \mathbb{Z}, i = 0, 1, \ldots, n$,并且 $(a_0', a_1', \ldots, a_n') = 1$. 因此,

$$f(x) = d(a_0' + a_1'x + \cdots + a_n'x^n).$$

记 $f_1(x) = a_0' + a_1'x + \cdots + a_n'x^n$. 显然,$f_1(x)$ 是本原多项式,并且

$$f(x) = df_1(x).$$

这说明,每个整系数多项式都可以表成系数的最大公因子和本原多项式的乘积.

Gauss 引理 任意两个本原多项式的乘积是本原多项式.

证明 设 $f(x) = a_0 + a_1x + \cdots + a_nx^n$ 与 $g(x) = b_0 + b_1x + \cdots + b_mx^m$ 是本原多项式. 设 $f(x)g(x) = c_0 + c_1x + \cdots + c_{n+m}x^{n+m}$,其中对于 $k = 0, 1, \ldots, n+m$,

$$c_k = a_0b_k + a_1b_{k-1} + \cdots + a_{k-1}b_1 + a_kb_0,$$

这里约定,当 $i > m$ 时,$b_i = 0$,而当 $j > n$ 时,$a_j = 0$.

如果 $f(x)g(x)$ 不是本原的,则 $(c_0, c_1, \ldots, c_{n+m}) \neq 1$. 设素数 p 是 $c_0, c_1, \ldots, c_{n+m}$ 的公因子. 由于 $f(x)$ 是本原的,故 p 不是 $a_0, a_1, \ldots, a_n$ 的公因子. 因此,可设 a_i 是

$a_0, a_1, \ldots, a_n$ 中第一个不被 p 整除的系数. 同理可设 b_j 是 $b_0, b_1, \ldots, b_m$ 第一个不被 p 整除的系数.

现在考察 $f(x)g(x)$ 的系数

$$c_{i+j} = a_0 b_{i+j} + \cdots + a_{i-1} b_{j+1} + a_i b_j + a_{i+1} b_{j-1} + \cdots + a_{i+j} b_0.$$

由于素数 p 整除 $a_0, a_1, \ldots, a_{i-1}, b_0, b_1, \ldots, b_{j-1}$ 和 c_{i+j}, 因此, p 整除 $a_i b_j$. 因为 p 是素数, 故 p 整除 a_i, 或者整除 b_j, 不可能. 因此,

$$(c_0, c_1, \ldots, c_{n+m}) = 1. \qquad \blacksquare$$

定理 1.6.1　设 n 次整系数多项式 $f(x)$ 在 $\mathbb{Z}$ 上不可约, 则 $f(x)$ 在 $\mathbb{Q}$ 不可约.

证明　设 $f(x)$ 在 $\mathbb{Q}$ 上可约, 则存在次数小于 n 的 $g(x), h(x) \in \mathbb{Q}[x]$, 使得

$$f(x) = g(x)h(x).$$

将多项式 $g(x)$ 的系数通分, 得到 $g(x) = b_1 g_1(x)$, $b_1 \in \mathbb{Q}$, $g_1(x) \in \mathbb{Z}[x]$. 而 $g_1(x)$ 可以表为系数最大公因子 d_1 和本原多项式 $\widetilde{g}(x)$ 的乘积. 因此, $g(x) = b\widetilde{g}(x)$, 其中 $b = b_1 d_1 \in \mathbb{Q}$.

同理, $h(x) = c\widetilde{h}(x)$, $c \in \mathbb{Q}$, $\widetilde{h}(x)$ 为本原多项式. 于是,

$$f(x) = bc\widetilde{g}(x)\widetilde{h}(x).$$

由 Gauss 引理, $\widetilde{g}(x)\widetilde{h}(x)$ 是本原多项式, 记 $bc = uv^{-1}$, $u, v \in \mathbb{Z}$. 由于 $f(x)$ 是整系数多项式, 且 $\widetilde{g}(x)\widetilde{h}(x)$ 是本原的, 因此, $v \mid u$. 所以, $bc \in \mathbb{Z}$. 这就说明, $f(x)$ 在 $\mathbb{Z}$ 上可约, 但这是一个矛盾. $\blacksquare$

定理 1.6.1 说, 如果整系数多项式 $f(x)$ 在 $\mathbb{Q}$ 上可约, 则 $f(x)$ 在 $\mathbb{Z}$ 上可约; 反之, 如果整系数多项式 $f(x)$ 在 $\mathbb{Z}$ 上可约, $f(x)$ 当然在 $\mathbb{Q}$ 上可约. 因此, 整系数多项式 $f(x)$ 相对于整数环 $\mathbb{Z}$ 和有理数域 $\mathbb{Q}$ 的不可约性是相同的.

定理 1.6.2　n 次整系数多项式 $f(x)$ 可以分解为一个整数和若干个本原不可约多项式的乘积, 而且不计因式的次序和符号, 这种分解是唯一的.

证明　根据数域 $\mathbb{F}$ 上的多项式的唯一析因定理, 作为有理系数多项式, $f(x)$ 可以表为

$$f(x) = a_0 p_1(x) \cdots p_s(x),$$

其中 $a_0 \in \mathbb{Q}$ 是 $f(x)$ 的首项系数, 诸 $p_i(x)$ 是首一有理系数多项式, 并且都在 $\mathbb{Q}$ 上不可约.

如同定理 1.6.1 的证明, $p_i(x) = b_i q_i(x)$, 其中 $b_i \in \mathbb{Q}$, 而 $q_i(x)$ 是本原多项式.

如果 $q_i(x)$ 在 $\mathbb{Z}$ 上可约, 则 $q_i(x)$ 在 $\mathbb{Q}$ 上可约, 从而 $p_i(x)$ 在 $\mathbb{Q}$ 可约, 不可能. 因此, $q_i(x)$ 在 $\mathbb{Z}$ 上不可约, $i = 1, 2, \ldots, s$. 于是,

$$f(x) = a_0 b_1 \cdots b_s q_1(x) q_2(x) \cdots q_s(x).$$

和定理 1.6.1 的证明相同, 可以证明, $a_0 b_1 \cdots b_s \in \mathbb{Z}$. 因此, $f(x)$ 可以表示为一个整数和若干个本原多项式的乘积.

现在设

$$f(x)=a_0p_1(x)p_2(x)\cdots p_s(x)=b_0q_1(x)q_2(x)\cdots q_s(x),$$

其中 $a_0,b_0\in\mathbb{Z}$, $p_1(x),p_2(x),\ldots,p_s(x)$ 和 $q_1(x),q_2(x),\ldots,q_s(x)$ 是本原不可约多项式. 根据定理 1.6.1, $p_1(x),\ldots,p_s(x)$ 和 $q_1(x),\ldots,q_s(x)$ 在 $\mathbb{Q}$ 上不可约. 显然, $a_0p_1(x),b_0q_1(x)$ 在 $\mathbb{Q}$ 上也不可约. 把 $f(x)$ 视为有理系数多项式,根据有理数域 $\mathbb{Q}$ 上的多项式的唯一析因定理,$s=t$,并且可以适当地调整不可约因式的次序,使得相应的有理系数不可约因式只相差一个有理数因子. 为简单计,设

$$a_0p_1(x)=c_1b_0q_1(x),p_2(x)=c_2q_2(x),\ldots,p_s(x)=c_sq_s(x),$$

其中 $c_1,c_2,\ldots,c_s\in\mathbb{Q}$.

当 $2\leqslant i\leqslant s$ 时,由于 $p_i(x)$ 与 $q_i(x)$ 是本原的,因此,$c_i=\pm1$,即 $p_i(x)=\pm q_i(x)$. 由于 $p_1(x)=a_0^{-1}c_1b_0q_1(x)$,且 $p_1(x)$ 与 $q_1(x)$ 是本原的,因此,$a_0^{-1}c_1b_0=\pm1$,即 $a_0=\pm c_1b_0$,所以,$p_1(x)=\pm q_1(x)$. 从而 $a_0=\pm b_0$. ■

定理 1.6.3(Eisenstein 判别准则) 设 $f(x)=a_0+a_1x+\cdots+a_nx^n\in\mathbb{Z}[x]$. 如果存在素数 p,使得 $p\mid a_i,i=0,\ldots,n-1$,但 $p\nmid a_n$ 且 $p^2\nmid a_0$,则 $f(x)$ 在 $\mathbb{Z}$ 上不可约.

证明 反证法. 设 $f(x)$ 在 $\mathbb{Z}$ 可约,则

$$f(x)=(b_0+b_1x+\cdots+b_kx^k)(c_0+c_1x+\cdots+c_lx^\ell),$$

其中 $b_i,c_j\in\mathbb{Z}$,并且 $k<n,\ell<n,k+\ell=n$. 于是得到,

$$a_r=b_0c_r+b_1c_{r-1}+\cdots+b_{r-1}c_1+b_rc_0,$$

其中 $r=0,1,\ldots,n$; 并且当 $i\geqslant k+1$ 时,约定 $b_k=0$; 当 $j\geqslant\ell+1$ 时,约定 $c_j=0$.

由于 $p\mid a_0,a_0=b_0c_0$,故 $p\mid b_0$,或者 $p\mid c_0$. 由于 $p^2\nmid a_0$,故 p 不同时整除 b_0 与 c_0. 因此可设 $p\mid b_0$,但 $p\nmid c_0$.

又因为 $p\nmid a_n$,故 $p\nmid b_k$. 所以必有某个 $1\leqslant i_0\leqslant k$,使得 $p\mid b_i,i=0,1,\ldots,i_0-1$,但 $p\nmid b_{i_0}$. 由于 $p\nmid a_{i_0}$, $p\mid b_i,i=0,1,\ldots,i_0-1$,并且

$$a_{i_0}=b_0c_{i_0}+b_1c_{i_0-1}+\cdots+b_{i_0-1}c_1+b_0c_0,$$

故 $p\mid b_{i_0}c_0$. 因为 $p\nmid b_i$,故 $p\mid c_0$. 与 $p\nmid c_0$ 的假设相矛盾. ■

利用 Eisenstein 判别准则容易看出,对每个整数 $n\geqslant2$,都存在 n 次多项式 $f(x)\in\mathbb{Q}[x]$,使得,$f(x)$ 在 $\mathbb{Q}$ 不可约. 例如,多项式 $f(x)=x^n+2\in\mathbb{Z}[x]$,取 $p=2$,则 $f(x)$ 适合 Eisenstein 判别准则的条件,因此,$f(x)$ 在 $\mathbb{Z}$ 上不可约. 根据定理 1.6.1,作为有理系数多项式,$f(x)$ 在 $\mathbb{Q}$ 上不可约.

例 1.6.1 设 p 是素数. 多项式

$$\Phi_p(x)=x^{p-1}+x^{p-2}+\cdots+x+1$$

称为**分圆多项式**. 证明,分圆多项式 $\Phi_p(x)$ 在 $\mathbb{Z}$ 上(当然也在 $\mathbb{Q}$ 上)不可约.

证明 令 $x=y+1$. 则

$$
\begin{aligned}
f(y) = \Phi_p(y+1) &= \frac{(y+1)^p - 1}{(y+1)-1} \\
&= y^{p-1} + py^{p-2} + \cdots + C_p^{k-1} y^{p-k} + \cdots + C_p^{p-2} y + p.
\end{aligned}
$$

显然，p 不能整除 $f(y)$ 的首项系数，p^2 不能整除 $f(y)$ 的常数项，但 p 整除 $f(y)$ 中首项系数外的其它各项. 根据 Eisenstein 判别准则，$f(y)$ 在 $\mathbb{Z}$ 上不可约.

如果 $\Phi_p(x)$ 在 $\mathbb{Z}$ 上可约，则

$$\Phi_p(x) = g(x)h(x),$$

其中 $g(x), h(x) \in \mathbb{Z}[x]$，且 $\deg g(x) < p-1, \deg h(x) < p-1$. 于是，

$$f(y) = g(y+1)h(y+1).$$

显然，$g(y+1), h(y+1) \in \mathbb{Z}[y]$. 从而 $f(y)$ 在 $\mathbb{Z}$ 上可约，矛盾. ■

例 1.6.2　设 $a_1, a_2, \ldots, a_n$ 是 n 个不同的整数. 证明，多项式

$$f(x) = (x-a_1)(x-a_2)\cdots(x-a_n) - 1$$

在 $\mathbb{Q}$ 上不可约.

证明　设 $f(x)$ 在 $\mathbb{Q}$ 上可约，则 $f(x)$ 在 $\mathbb{Z}$ 上可约，因此，

$$f(x) = g(x)h(x),$$

其中 $g(x), h(x) \in \mathbb{Z}[x]$，且 $\deg g(x) < \deg f(x), \deg h(x) < \deg f(x)$.

由于 $f(a_i) = g(a_i)h(a_i) = -1$，故 $|g(a_i)| = |h(a_i)| = 1$，且 $g(a_i) + h(a_i) = 0$. 这表明，多项式 $g(x) + h(x)$ 至少有 n 个不同的根.

由于 $\deg g(x) < n, \deg h(x) < n$，因此，$\deg(g(x)+h(x)) < n$. 因此若 $g(x)+h(x)$ 是非零多项式，则 $g(x)+h(x)$ 的根的个数小于 n，不可能. 因此，$g(x)+h(x)$ 是零多项式，从而 $f(x) = -g^2(x)$. 这和 $f(x)$ 的首项系数为 1 相矛盾. ■

习　题　1.6

1. 利用 Eisenstein 判别准则判定下述整系数多项式的不可约性.

(1) $x^4 - 8x^3 + 12x^2 - 6x + 2$;　　(2) $x^4 - x^3 + 2x + 1$;

(3) $x^4 + 1$;　　(4) $x^6 + x^3 + 1$;

(5) $\sum\limits_{i=0}^{p-1}(x+1)^i$，其中 p 是素数.

2. 设 $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ 是整系数多项式，且素数 p 适合：$p \nmid a_0, p \nmid a_1, \ldots, p \nmid a_k, p \mid a_i, i = k+1, k+2, \ldots, n$，而 $p^2 \nmid a_n$. 证明，$f(x)$ 具有次数不低于 $n-k$ 的整系数不可约因式.

3. 设

$$f(x) = a_0x^{2n+1} + \cdots + a_nx^{n+1} + a_{n+1}x^n + \cdots + a_{2n}x + a_{2n+1}$$

是整系数多项式，且素数 p 适合：$p \nmid a_0, p \mid a_i, i = 1, \ldots, n, p^2 \mid a_i, i = n+1, \ldots, 2n+1$，但 $p^3 \nmid a_{2n+1}$. 证明 $f(x)$ 在 $\mathbb{Q}$ 上不可约.

4. 设 $a_1, a_2, \ldots, a_n$ 是 n 个不同的整数. 证明,多项式

$$f(x) = (x - a_1)^2 (x - a_2)^2 \cdots (x - a_n)^2 + 1$$

在 $\mathbb{Q}$ 上不可约.

5. 试给出有理系数多项式 $f(x) = x^4 + px^2 + q$ 在 $\mathbb{Q}$ 上不可约的充分必要条件.

6. 设整系数多项式 $f(x)$ 在 x 的 4 个不同整数值上都取值为 1,则 $f(x)$ 在 x 的其它整数值上的值不可能是 -1.

7. 证明,设正整数 $n \geqslant 12$,并且 n 次整系数多项式 $f(x)$ 在 x 的 $\left[\frac{n}{2}\right] + 1$ 个以上的整数值上取值为 ± 1,则 $f(x)$ 在 $\mathbb{Q}$ 不可约. 次数 n 的下界 12 是否还可缩小?

8. 设整系数多项式 $ax^2 + bx + 1$ 在有理数域 $\mathbb{Q}$ 上不可约,并且设

$$\varphi(x) = (x - a_1)(x - a_2) \cdots (x - a_n),$$

其中 $a_1, a_2, \ldots, a_n$ 是 n 个不同的整数,$n \geqslant 7$. 证明,多项式

$$f(x) = a\varphi^2(x) + b\varphi(x) + 1$$

在 $\mathbb{Q}$ 上不可约. 并问次数 n 的下界 7 是否还可缩小?

§1.7 多元多项式环

设 $\mathbb{F}$ 是数域,$x_1, x_2, \ldots, x_n$ 是 n 个未定元. 设 $\mathbb{N}$ 是所有非负整数的集合. 记

$$\mathbb{N}^n = \left\{ (k_1, k_2, \ldots, k_n) \,\middle|\, k_1, k_2, \ldots, k_n \in \mathbb{N} \right\}.$$

设 $(k_1, k_2, \ldots, k_n) \in \mathbb{N}$, $a_{k_1 k_2 \cdots k_n} \in \mathbb{F}$,则

$$a_{k_1 k_2 \cdots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

称为**数域 $\mathbb{F}$ 上的 n 元单项式**,$k_1 + k_2 + \cdots + k_n$ 称为它的**次数**,$a_{k_1 k_2 \cdots k_n}$ 称为它的**系数**.

设 M 是集合 $\mathbb{N}^n$ 的有限子集合,则

$$f(x_1, x_2, \ldots, x_n) = \sum_{(k_1, k_2, \ldots, k_n) \in M} a_{k_1 k_2 \cdots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

称为**数域 $\mathbb{F}$ 上的 n 元多项式**,其中 $a_{k_1 k_2 \cdots k_n} \in \mathbb{F}$. n 元多项式 $f(x_1, x_2, \ldots, x_n)$ 的所有单项式的最高次数称为 $f(x_1, x_2, \cdots, x_n)$ 的**次数**,记为 $\deg f(x_1, x_2, \ldots, x_n)$,或简记为 $\deg f$; $a_{k_1 k_2 \cdots k_n}$ 称为项 $a_{k_1 k_2 \cdots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ 的**系数**.

例如,

$$f(x_1, x_2, x_3) = 4x_1 x_2 x_3 + \sqrt{3}\, x_2^3 x_3 + \pi x_1^3 x_2^2$$

是实数域 $\mathbb{R}$ 上的 5 次 3 元多项式.

所有数域 $\mathbb{F}$ 上的 n 元多项式的集合记为 $\mathbb{F}[x_1, x_2, \ldots, x_n]$.

给定数域 $\mathbb{F}$ 上的 n 元多项式 $f(x_1, x_2, \ldots, x_n)$,可以接照**字典排列法**把它所有的项逐一写出来. 设

$$a_{k_1 k_2 \cdots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} \text{ 和 } a_{\ell_1 \ell_2 \cdots \ell_n} x_1^{\ell_1} x_2^{\ell_2} \cdots x_n^{\ell_n}$$

是 $f(x_1, x_2, \ldots, x_n)$ 的两个项. 如果存在正整数 i, $1 \leqslant i \leqslant n$,使得 $k_1 = \ell_1, k_2 = \ell_2, \ldots,$

$k_{i-1}=\ell_{i-1}$，而 $k_i>\ell_i$，则将项 $a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 写在项 $a_{\ell_1\ell_2\cdots\ell_n}x_1^{\ell_1}x_2^{\ell_2}\cdots x_n^{\ell_n}$ 之前.

例如，3 元多项式

$$f(x_1,x_2,x_3)=x_1x_2^2x_3^2+x_1^3x_2+x_1x_3^3+x_1^2x_2x_3^2$$

可以按照字典排列写成

$$f(x_1,x_2,x_3)=x_1^3x_2+x_1^2x_2x_3^2+x_1x_2^2x_3^2+x_2x_3^3.$$

按照字典排列法写在多项式 $f(x_1,x_2,\ldots,x_n)$ 的和式中最前面的项称为 $f(x_1,x_2,\ldots,x_n)$ 的**首项**，相应的系数称为 $f(x_1,x_2,\ldots,x_n)$ 的**首项系数**.

注意，$f(x_1,x_2,\ldots,x_n)$ 的首项不一定是最高次项.

设 $f(x_1,x_2,\ldots,x_n)$ 与 $g(x_1,x_2,\ldots,x_n)$ 是数域 $\mathbb{F}$ 上的 n 元多项式. 如果它们的相应项的系数都相等. 则 $f(x_1,x_2,\ldots,x_n)$ 与 $g(x_1,x_2,\ldots,x_n)$ 称为**相等**. 两个 n 元多项式的**和**是以它们相应项的系数之和作为相应项的系数的多项式，记为

$$f(x_1,x_2,\ldots,x_n)+g(x_1,x_2,\ldots,x_n).$$

具体地说，设

$$f(x_1,x_2,\ldots,x_n)=\sum_{(k_1,k_2,\ldots,k_n)\in M_1}a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n},$$
$$g(x_1,x_2,\ldots,x_n)=\sum_{(k_1,k_2,\ldots,k_n)\in M_2}b_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n},$$

其中 $M_1,M_2\subseteq\mathbb{N}^n$. 记 $M=M_1\cup M_2$，当 $(k_1,k_2,\ldots,k_n)\in M_1$ 但 $(k_1,k_2,\ldots,k_n)\notin M_2$ 时，约定 $b_{k_1k_2\cdots k_n}=0$；而当 $(k_1,k_2,\ldots,k_n)\notin M_1$，$(k_1,k_2,\ldots,k_n)\in M_2$ 时，约定 $a_{k_1k_2\cdots k_n}=0$. 则 $f(x_1,\ldots,x_n)$ 与 $g(x_1,\ldots,x_n)$ 的和为

$$f(x_1,x_2,\ldots,x_n)+g(x_1,x_2,\ldots,x_n)=\sum_{(k_1,k_2,\ldots,k_n)\in M}\left(a_{k_1k_2\cdots k_n}+b_{k_1k_2\cdots k_n}\right)x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}.$$

容易看出，

$$\deg\big(f(x_1,\ldots,x_n)+g(x_1,\ldots,x_n)\big)\leqslant\max\big\{\deg f(x_1,\ldots,x_n),\ \deg g(x_1,\ldots,x_n)\big\}.$$

对 n 元多项式的加法，容易验证下述公理成立：

(A1) 结合律　对任意 $f(x_1,x_2,\ldots,x_n),g(x_1,x_2,\ldots,x_n)$ 和 $h(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$，

$$\begin{aligned}&\big(f(x_1,x_2,\ldots,x_n)+g(x_1,x_2,\ldots,x_n)\big)+h(x_1,x_2,\ldots,x_n)\\&\qquad=f(x_1,x_2,\ldots,x_n)+\big(g(x_1,x_2,\ldots,x_n)+h(x_1,x_2,\ldots,x_n)\big);\end{aligned}$$

(A2) 交换律　对任意 $f(x_1,x_2,\ldots,x_n),g(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$，

$$f(x_1,x_2,\ldots,x_n)+g(x_1,x_2,\ldots,x_n)=g(x_1,x_2,\ldots,x_n)+f(x_1,x_2,\ldots,x_n);$$

(A3) 存在零多项式　每个系数都为零的 n 元多项式称为**零多项式**，记为 0，同时零多项式的次数约定为 $-\infty$. 显然，$0\in\mathbb{F}[x_1,x_2,\ldots,x_n]$，并且对任意 $f(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x_1,x_2,\ldots,x_n]$，

$$f(x_1,x_2,\dots,x_n)+0=0+f(x_1,x_2,\dots,x_n)=f(x_1,x_2,\dots,x_n);$$

(A4) 存在负多项式 设

$$f(x_1,x_2,\dots,x_n)=\sum_{(k_1,k_2,\dots,k_n)\in M} a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}\in\mathbb{F}[x_1,x_2,\dots,x_n].$$

多项式

$$\sum_{(k_1,k_2,\dots,k_n)\in M}\left(-a_{k_1k_2\cdots k_n}\right)x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}\in\mathbb{F}[x_1,x_2,\dots,x_n]$$

称为 $f(x_1,x_2,\dots,x_n)$ 的**负多项式**. 记为 $-f(x_1,x_2,\dots,x_n)$. 显然,

$$f(x_1,x_2,\dots,x_n)+\big(-f(x_1,x_2,\dots,x_n)\big)=0=\big(-f(x_1,x_2,\dots,x_n)\big)+f(x_1,x_2,\dots,x_n).$$

设

$$f(x_1,x_2,\dots,x_n)=\sum_{(k_1,k_2,\dots,k_n)\in M_1} a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n},$$
$$g(x_1,x_2,\dots,x_n)=\sum_{(\ell_1,\ell_2,\dots,\ell_n)\in M_2} b_{\ell_1\ell_2\cdots \ell_n}x_1^{\ell_1}x_2^{\ell_2}\cdots x_n^{\ell_n},$$

是数域 $\mathbb{F}$ 上的 n 元多项式. 又令 $m_i=k_i+\ell_i,\ i=1,2,\dots,n$, 且记

$$M=\left\{(m_1,m_2,\dots,m_n)\;\middle|\;(k_1,k_2,\dots,k_n)\in M_1,(\ell_1,\ell_2,\dots,\ell_m)\in M_2\right\}.$$

则 $f(x_1,x_2,\dots,x_n)$ 与 $g(x_1,x_2,\dots,x_n)$ 的**乘积** $f(x_1,x_2,\dots,x_n)g(x_1,x_2,\dots,x_n)$ 规定为

$$f(x_1,x_2,\dots,x_n)g(x_1,x_2,\dots,x_n)=\sum_{(m_1,m_2,\dots,m_n)\in M} c_{m_1m_2\cdots m_n}x_1^{m_1}x_2^{m_2}\cdots x_n^{m_n},$$

其中

$$c_{m_1m_2\cdots m_n}=\sum_{\substack{1\leqslant j\leqslant n\\ k_j+\ell_j=m_j}} a_{k_1k_2\cdots k_n}b_{\ell_1\ell_2\cdots\ell_n}.$$

显然, $f(x_1,x_2,\dots,x_n)g(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$, 并且其首项系数等于 $f(x_1,x_2,\dots,x_n)$ 与 $g(x_1,x_2,\dots,x_n)$ 的首项系数的乘积, 即

$$\deg\big(f(x_1,x_2,\dots,x_n)g(x_1,x_2,\dots,x_n)\big)=\deg f(x_1,x_2,\dots,x_n)+\deg g(x_1,x_2,\dots,x_n).$$

此外, 对 n 元多项式的乘法, 乘法结合律、乘法交换律以及乘法对加法的分配律成立. 同时, 对任意 $f(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$, 均有

$$1\cdot f(x_1,x_2,\dots,x_n)=f(x_1,x_2,\dots,x_n)=f(x_1,x_2,\dots,x_n)\cdot 1,$$

其中 1 是数域 $\mathbb{F}$ 上的零次多项式.

于是, $\mathbb{F}[x_1,x_2,\dots,x_n]$ 在上述多项式的加法与乘法下构成一个交换环. 它称为数域 $\mathbb{F}$ 上的 ***n* 元多项式环**.

习 题 1.7

1. 设 $f(x_1,x_2,\dots,x_n), g(x_1,x_2,\dots,x_n)\in\mathbb{F}[x_1,x_2,\dots,x_n]$. 证明, 如果

$$f(x_1,x_2,\dots,x_n)g(x_1,x_2,\dots,x_n)$$

为零多项式, 则 $f(x_1,x_2,\dots,x_n)$ 与 $g(x_1,x_2,\dots,x_n)$ 至少有一个是零多项式.

2. 设 $f(x_1,x_2,\ldots,x_n), g(x_1,x_2,\ldots,x_n), h(x_1,x_2,\ldots,x_n) \in \mathbb{F}[x_1,x_2,\ldots,x_n]$. 证明,如果

$$f(x_1,x_2,\ldots,x_n)g(x_1,x_2,\ldots,x_n) = f(x_1,x_2,\ldots,x_n)h(x_1,x_2,\ldots,x_n),$$

则 $g(x_1,x_2,\ldots,x_n) = h(x_1,x_2,\ldots,x_n)$.

3. 验证 $\mathbb{F}[x_1,x_2,\ldots,x_n]$ 在 n 元多项式的加法与乘法下成为一个交换环.

§1.8　对称多项式

设 $\mathbb{F}[x_1,x_2,\ldots,x_n]$ 是数域 $\mathbb{F}$ 上的 n 元多项式环. 在 n 元多项式中,经常遇到的是所谓对称多项式. 其定义如下:

定义 1.8.1　设 $f(x_1,x_2,\ldots,x_n) \in \mathbb{F}[x_1,x_2,\ldots,x_n]$. 如果对自然数 $1,2,\ldots,n$ 的任意一个排列 $i_1i_2\cdots i_n$,都有

$$f(x_{i_1},x_{i_2},\ldots,x_{i_n}) = f(x_1,x_2,\ldots,x_n),$$

则 $f(x_1,x_2,\ldots,x_n)$ 称为 n 元**对称多项式**.

例如,容易看出,

$$\begin{aligned}
\sigma_1 &= x_1+x_2+\cdots+x_n = \sum_{1\leqslant i_1\leqslant n} x_{i_1},\\
\sigma_2 &= x_1x_2+\cdots+x_1x_n+x_2x_3+\cdots+x_2x_n+\cdots+x_{n-1}x_n = \sum_{1\leqslant i_1<i_2\leqslant n} x_{i_1}x_{i_2},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\sigma_k &= \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} x_{i_1}x_{i_2}\cdots x_{i_k},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\sigma_n &= x_1x_2\cdots x_n,
\end{aligned}$$

都是 n 元对称多项式. 它们称为 n 元**基本对称多项式**.

可验证,两个 n 元对称多项式之和、差与积仍是 n 元对称多项式.

此外,对任意多项式 $f(x_1,x_2,\ldots,x_n) \in \mathbb{F}[x_1,x_2,\ldots,x_n]$,如果用基本对称多项式 $\sigma_1,\sigma_2,\ldots,\sigma_n$ 分别替换 $f(x_1,x_2,\ldots,x_n)$ 中的未定元 $x_1,x_2,\ldots,x_n$,得到 $f(\sigma_1,\sigma_2,\ldots,\sigma_n)$,则 $f(\sigma_1,\sigma_2,\ldots,\sigma_n)$ 是一个关于未定元 $x_1,x_2,\ldots,x_n$ 的对称多项式. 例如取 $n=3$, $f(x_1,x_2,x_3) = x_1x_2+2x_3 \in \mathbb{R}[x_1,x_2,x_3]$. 用 $\sigma_1 = x_1+x_2+x_3$, $\sigma_2 = x_1x_2+x_1x_3+x_2x_3$ 与 $\sigma_3 = x_1x_2x_3$ 分别替换 $f(x_1,x_2,x_3)$ 的未定元 x_1,x_2,x_3,得到

$$f(\sigma_1,\sigma_2,\sigma_3) = \sigma_1\sigma_2+2\sigma_3 = x_1^2x_2+x_1^2x_3+x_1x_2^2+x_1x_3^2+x_2^2x_3+x_2x_3^2+5x_1x_2x_3,$$

显然,它是一个三元对称多项式. 反之,有

定理 1.8.1(对称多项式基本定理) 设 $f(x_1,x_2,\ldots,x_n)$ 是数域 $\mathbb{F}$ 上的 n 元对称多项式. 则存在唯一的多项式 $g(x_1,x_2,\ldots,x_n) \in \mathbb{F}[x_1,x_2,\ldots,x_n]$,使得

$$f(x_1,x_2,\ldots,x_n) = g(\sigma_1,\sigma_2,\ldots,\sigma_n),$$

其中 $\sigma_1,\sigma_2,\ldots,\sigma_n$ 是数域 $\mathbb{F}$ 上的 n 元基本对称多项式.

证明　存在性　设对称多项式 $f(x_1,x_2,\ldots,x_n)$ 的首项为

$$a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n},$$

则一定有 $k_1\geqslant k_2\geqslant\cdots\geqslant k_n$. 否则,将有某个 j,使得 $k_j<k_{j+1}$. 由于 $f(x_1,x_2,\ldots,x_n)$ 是对称的,则通过对换未定元 x_j 和 x_{j+1} 便可看出,$f(x_1,x_2,\ldots,x_n)$ 含有项

$$a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_{j-1}^{k_{j-1}}x_j^{k_{j+1}}x_{j+1}^{k_j}x_{j+1}^{k_{j+1}}\cdots x_n^{k_n}.$$

显然,按字典排列法它应排在项 $a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 之前,这和 $a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 为 $f(x_1,x_2,\ldots,x_n)$ 的首项的假设相矛盾.

取 n 元对称多项式 $\varphi_1(x_1,x_2,\ldots,x_n)$ 为

$$\varphi_1(x_1,x_2,\ldots,x_n)=a_{k_1k_2\cdots k_n}\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\cdots\sigma_{n-1}^{k_{n-1}-k_n}\sigma_n^{k_n}.$$

容易看出,$\sigma_1,\sigma_2,\ldots,\sigma_n$ 的首项依次是 $x_1,x_1x_2,\ldots,x_1x_2\cdots x_n$. 因此,$\varphi_1(x_1,x_2,\ldots,x_n)$ 的首项为 $a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$. 于是,

$$f_1(x_1,x_2,\ldots,x_n)=f(x_1,x_2,\ldots,x_n)-\varphi_1(x_1,x_2,\ldots,x_n)$$

是数域 $\mathbb{F}$ 上的 n 元对称多项式.

如果 $f_1(x_1,x_2,\ldots,x_n)$ 是零多项式,则 $f(x_1,x_2,\ldots,x_n)=\varphi_1(x_1,x_2,\ldots,x_n)$ 是关于 $\sigma_1,\sigma_2,\ldots,\sigma_n$ 的多项式.

如果 $f_1(x_1,x_2,\ldots,x_n)$ 是非零多项式,则 $f(x_1,x_2,\ldots,x_n)$ 的首项 $a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 应在 $f_1(x_1,x_2,\ldots,x_n)$ 的首项 $b_{\ell_1\ell_2\cdots\ell_n}x_1^{\ell_1}x_2^{\ell_2}\cdots x_n^{\ell_n}$ 之前,其中 $\ell_1\geqslant\ell_2\geqslant\cdots\geqslant\ell_n$. 记

$$\varphi_2(x_1,x_2,\ldots,x_n)=b_{\ell_1\ell_2\cdots\ell_n}\sigma_1^{\ell_1-\ell_2}\sigma_2^{\ell_2-\ell_3}\cdots\sigma_{n-1}^{\ell_{n-1}-\ell_n}\sigma_n^{\ell_n}.$$

则数域 $\mathbb{F}$ 上的 n 元对称多项式

$$f_2(x_1,x_2,\ldots,x_n)=f_1(x_1,x_2,\ldots,x_n)-\varphi_2(x_1,x_2,\ldots,x_n)$$

为零多项式,或为非零多项式,并且 $f_1(x_1,x_2,\ldots,x_n)$ 的首项在 $f_2(x_1,x_2,\ldots,x_n)$ 的首项之前.

设 $f_2(x_1,x_2,\ldots,x_n)$ 为非零多项式. 重复上述过程,得到对称多项式序列:

$$\begin{aligned}
f_0(x_1,x_2,\ldots,x_n)&=f(x_1,x_2,\ldots,x_n),\\
f_1(x_1,x_2,\ldots,x_n)&=f_0(x_1,x_2,\ldots,x_n)-\varphi_1(x_1,x_2,\ldots,x_n),\\
&\cdots\cdots\cdots\cdots\\
f_{i+1}(x_1,x_2,\ldots,x_n)&=f_i(x_1,x_2,\ldots,x_n)-\varphi_{i+1}(x_1,x_2,\ldots,x_n),
\end{aligned}$$

其中 $\varphi_{i+1}(x_1,x_2,\ldots,x_n)$ 是关于 $\sigma_1,\sigma_2,\ldots,\sigma_n$ 的多项式,而且 $f_i(x_1,x_2,\ldots,x_n)$ 的首项在 $f_{i+1}(x_1,x_2,\ldots,x_n)$ 的首项之前,$i=0,1,\ldots$.

设 $f_i(x_1,x_2,\ldots,x_n)$ 的首项为 $c_{m_1m_2\cdots m_n}x_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}$,则 $m_1\geqslant m_2\geqslant\cdots\geqslant m_n$,并且由于 $f(x_1,x_2,\ldots,x_n)$ 的首项在 $f_i(x_1,x_2,\ldots,x_n)$ 的首项之前,故 $k_1\geqslant m_1$.

因为适合 $k_1\geqslant m_1$ 的非负整数 m_1 只有有限多个,而且对每个适合 $k_1\geqslant m_1$ 的非

负整数 m_1，适合 $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_n$ 的 n 元非负整数组 $(m_1, m_2, \ldots, m_n)$ 也只有有限多个，因此存在某个 j，使得 $f_j(x_1, x_2, \ldots, x_n) = 0$. 于是，

$$f(x_1, \ldots, x_n) = \varphi_1(x_1, \ldots, x_n) + \varphi_2(x_1, \ldots, x_n) + \cdots + \varphi_j(x_1, \ldots, x_n).$$

这就证明，$f(x_1, x_2, \ldots, x_n)$ 可以表为系数在 $\mathbb{F}$ 中的关于 $\sigma_1, \sigma_2, \ldots, \sigma_n$ 的多项式.

唯一性　设存在 $g(x_1, x_2, \ldots, x_n), h(x_1, x_2, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$，使得

$$f(x_1, x_2, \ldots, x_n) = g(\sigma_1, \sigma_2, \ldots, \sigma_n) = h(\sigma_1, \sigma_2, \ldots, \sigma_n). \tag{1.8.1}$$

设 $g(x_1, x_2, \ldots, x_n)$ 与 $h(x_1, x_2, \ldots, x_n)$ 的首项分别为

$$a_{k_1 k_2 \cdots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} \text{ 与 } b_{\ell_1 \ell_2 \cdots \ell_n} x_1^{\ell_1} x_2^{\ell_2} \cdots x_n^{\ell_n}.$$

于是，

$$g\big(\sigma_1(x_1, x_2, \ldots, x_n), \sigma_2(x_1, x_2, \ldots, x_n), \ldots, \sigma_n(x_1, x_2, \ldots, x_n)\big)$$

的首项为

$$a_{k_1 k_2 \cdots k_n} x_1^{k_1} (x_1 x_2)^{k_2} \cdots (x_1 x_2 \cdots x_n)^{k_n} = a_{k_1 k_2 \cdots k_n} x_1^{k_1 + k_2 + \cdots + k_n} x_2^{k_2 + \cdots + k_n} \cdots x_n^{k_n},$$

而

$$h\big(\sigma_1(x_1, x_2, \ldots, x_n), \sigma_2(x_1, x_2, \ldots, x_n), \ldots, \sigma_n(x_1, x_2, \ldots, x_n)\big)$$

的首项为

$$b_{\ell_1 \ell_2 \cdots \ell_n} x_1^{\ell_1 + \ell_2 + \cdots + \ell_n} x_2^{\ell_2 + \cdots + \ell_n} \cdots x_n^{\ell_n}.$$

由式 (1.8.1)，$a_{k_1 k_2 \cdots k_n} = b_{\ell_1 \ell_2 \cdots \ell_n}$，并且

$$\begin{aligned} k_1 + k_2 + \cdots + k_n &= \ell_1 + \ell_2 + \cdots + \ell_n, \\ k_1 + k_2 + \cdots + k_{n-1} &= \ell_1 + \ell_2 + \cdots + \ell_{n-1}, \\ &\cdots\cdots \\ k_{n-1} + k_n &= \ell_{n-1} + \ell_n, \\ k_n &= \ell_n. \end{aligned}$$

由此得到 $k_1 = \ell_1, k_2 = \ell_2, \ldots, k_n = \ell_n$. 即 $g(x_1, x_2, \ldots, x_n)$ 与 $h(x_1, x_2, \ldots, x_n)$ 具有相同的首项. 从 $g(x_1, x_2, \ldots, x_n)$ 与 $h(x_1, x_2, \ldots, x_n)$ 中各减去首项，分别记为 $g_1(x_1, x_2, \ldots, x_n)$ 与 $h_1(x_1, x_2, \ldots, x_n)$. 显然，

$$\begin{aligned} & g_1\big(\sigma_1(x_1, x_2, \ldots, x_n), \sigma_2(x_1, x_2, \ldots, x_n), \ldots, \sigma_n(x_1, x_2, \ldots, x_n)\big) \\ &\qquad = h_1\big(\sigma_1(x_1, x_2, \ldots, x_n), \sigma_2(x_1, x_2, \ldots, x_n), \ldots, \sigma_n(x_1, x_2, \ldots, x_n)\big), \end{aligned}$$

上式是关于 $x_1, x_2, \ldots, x_n$ 的对称多项式，记为 $f_1(x_1, x_2, \ldots, x_n)$.

再对 $f_1(x_1, x_2, \ldots, x_n)$，$g_1(x_1, x_2, \ldots, x_n)$ 与 $h_1(x_1, x_2, \ldots, x_n)$ 用上述证明，如此继续，即可证明

$$g(x_1, x_2, \ldots, x_n) = h(x_1, x_2, \ldots, x_n) \qquad \blacksquare$$

定理 1.8.1 中关于存在性部分的证明是构造性的，它给出了对称多项式表为关于基本对称多项式 $\sigma_1, \sigma_2, \ldots, \sigma_n$ 的多项式的具体方法.

各个项次数相等的多项式称为**齐次多项式**，否则称为**非齐次多项式**.

例如,多项式

$$g(x_1,x_2,\ldots,x_n)=\sum_{1\leqslant i<j\leqslant n}x_i^2x_j$$

是一个三次齐次多项式.

对于给定的 m 次齐次多项式 $f(x_1,x_2,\ldots,x_n)$,可以把它的同次项归并在一起,得到

$$f(x_1,x_2,\ldots,x_n)=\sum_{j=0}^{m}f_j(x_1,x_2,\ldots,x_n),$$

其中 $f_j(x_1,x_2,\ldots,x_n)$ 是 j 次齐次多项式,即 $f(x_1,x_2,\ldots,x_n)$ 可以表为若干个齐次多项式之和.

把对称多项式 $f(x_1,x_2,\ldots,x_n)$ 表为关于基本对称多项式 $\sigma_1,\sigma_2,\ldots,\sigma_n$ 的多项式,除了定理 1.8.1 所给出的方法外,还可以采用待定系数法,其步骤如下:

(1) 把多项式 $f(x_1,x_2,\ldots,x_n)$ 分解为齐次多项式之和. 容易看出,由于 $f(x_1,x_2,\ldots,x_n)$ 是对称的,因此,这些齐次多项式也是对称的;

(2) 设 m 次齐次对称多项式 $f_m(x_1,x_2,\ldots,x_n)$ 的首项是

$$a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n},$$

其中 $k_1\geqslant k_2\geqslant\cdots\geqslant k_n$,并且 $k_1+k_2+\cdots+k_n=m$. 写出所有可能排在 $f_m(x_1,x_2,\ldots,x_n)$ 的首项之后的项 $a_{\ell_1\ell_2\cdots\ell_n}x_1^{\ell_1}x_2^{\ell_2}\cdots x_n^{\ell_n}$,其中 $\ell_1\geqslant\ell_2\geqslant\cdots\geqslant\ell_n$,$\ell_1+\ell_2+\cdots+\ell_n=m$,并且存在某个 j,使得 $k_1=\ell_1,\ldots,k_{j-1}=\ell_{j-1},k_j>\ell_j,1\leqslant j\leqslant n$,这里 j 与项 $a_{\ell_1\ell_2\cdots\ell_n}x_1^{\ell_1}x_2^{\ell_2}\cdots x_n^{\ell_n}$ 有关. 所有适合这些条件的 $(\ell_1,\ell_2,\ldots,\ell_n)\in\mathbb{N}^n$ 的集合记为 M;

(3) 对每个可能出现的项 $a_{\ell_1\ell_2\cdots\ell_n}x_1^{\ell_1}x_2^{\ell_2}\cdots x_n^{\ell_n}$ 构作一个项

$$A_{\ell_1\ell_2\cdots\ell_n}\sigma_1^{\ell_1-\ell_2}\sigma_2^{\ell_2-\ell_3}\cdots\sigma_{n-1}^{\ell_{n-1}-\ell_n}\sigma_n^{\ell_n},$$

并令

$$f(x_1,x_2,\ldots,x_n)=\sum_{(\ell_1,\ell_2,\ldots,\ell_n)\in M}A_{\ell_1\ell_2\cdots\ell_n}\sigma_1^{\ell_1-\ell_2}\sigma_2^{\ell_2-\ell_3}\cdots\sigma_{n-1}^{\ell_{n-1}-\ell_n}\sigma_n^{\ell_n},$$

其中 $A_{\ell_1\ell_2\cdots\ell_n}$ 为待定常数;

(4) 取 $x_1,x_2,\ldots,x_n$ 的一些特殊值,代入上式,便得到一组关于 $A_{\ell_1\ell_2\cdots\ell_n}$ 的方程,解之即得 $A_{\ell_1\ell_2\cdots\ell_n}$. 通常特殊值可以取为 $x_1=x_2=\cdots=x_k=1,x_{k+1}=\cdots=x_n=0$. 把它们代入 $\sigma_j(x_1,x_2,\ldots,x_n)$,得到

$$\sigma_j(\underbrace{1,1,\ldots,1}_{k\text{个}},0,\ldots,0)=\begin{cases}\mathrm{C}_k^j, & 1\leqslant j\leqslant k,\\ 0, & k<j\leqslant n.\end{cases}$$

例 1.8.1 把 n 元对称多项式

$$f(x_1,x_2,\ldots,x_n)=\sum_{1\leqslant j_1<j_2<j_3\leqslant n}\left(x_{j_1}^2x_{j_2}^2x_{j_3}+x_{j_1}^2x_{j_2}x_{j_3}^2+x_{j_1}x_{j_2}^2x_{j_3}^2\right)$$

表为关于基本对称多项式的多项式,其中 $n\geqslant 5$.

解 $f(x_1,x_2,\ldots,x_n)$ 本身是 5 次齐次对称多项式,它的首项是 $x_1^2x_2^2x_3$. 可能出现在它后面的项有 $x_1^2x_2x_3x_4$,$x_1x_2x_3x_4x_5$. 令

$$\begin{aligned} f(x_1,x_2,\ldots,x_n) &= \sigma_1^{2-1}\sigma_2^{2-1}\sigma_3^{1} + A\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4 + B\sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4^{1-1}\sigma_5^{1} \\ &= \sigma_2\sigma_3 + A\sigma_1\sigma_4 + B\sigma_5. \end{aligned}$$

取 $x_1 = x_2 = x_3 = x_4 = 1, x_5 = 0$. 则 $\sigma_1 = \mathrm{C}_4^1 = 4, \sigma_2 = \mathrm{C}_4^2 = 6, \sigma_3 = \mathrm{C}_4^3 = 4, \sigma_4 = \mathrm{C}_4^4 = 1$, $\sigma_5 = 0$,并且 $f(1,1,1,1,0,\ldots,0) = 12$. 因此,$12 = 24 + 4A$. 所以,$A = -3$.

再取 $x_1 = x_2 = \cdots = x_5 = 1, x_6 = x_7 = \cdots = x_n = 0$. 则 $\sigma_1 = \mathrm{C}_5^1 = 5, \sigma_2 = \mathrm{C}_5^2 = 10, \sigma_3 = \mathrm{C}_5^3 = 10$, $\sigma_4 = \mathrm{C}_5^4 = 5, \sigma_5 = 1$,并且 $f(1,1,1,1,1,0,\ldots,0) = 3\mathrm{C}_5^3 = 30$. 因此,$30 = 100 - 3 \times 25 + B$. 所以,$B = 5$.

由此得到,

$$f(x_1,x_2,\ldots,x_n) = \sigma_2\sigma_3 - 3\sigma_1\sigma_4 + 5\sigma_5. \qquad \blacksquare$$

利用基本对称多项式,可以得到关于一元多项式的根与系数的定理.

定理 1.8.2(Vièta 定理) 设数域 $\mathbb{F}$ 上的 n 次多项式

$$f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$$

的根为 $c_1, c_2, \ldots, c_n$,则对于 $k = 0, 1, \ldots, n$,

$$a_k = (-1)^{n-k}\sigma_{n-k}(c_1, c_2, \ldots, c_n),$$

其中 $a_n = 1$,且 $\sigma_0(c_1, c_2, \ldots, c_n) = 1$.

证明 因为 $c_1, c_2, \ldots, c_n$ 是首一多项式 $f(x)$ 的根,所以

$$f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n = (x - c_1)(x - c_2)\cdots(x - c_n).$$

上式右端乘开,并比较上式两端同次项系数,得到

$$\begin{aligned} a_{n-1} &= -(c_1 + c_2 + \cdots + c_n) \\ &= -\sigma_1(c_1, c_2, \ldots, c_n), \\ a_{n-2} &= c_1c_2 + c_1c_3 + \cdots + c_1c_n + c_2c_3 + \cdots + c_2c_n + \cdots + c_{n-1}c_n \\ &= \sigma_2(c_1, c_2, \ldots, c_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ a_k &= (-1)^{n-k}(c_1\cdots c_{n-k-1}c_{n-k} + c_1\cdots c_{n-k-1}c_{n-k+1} + \cdots + c_1\cdots c_{n-k-1}c_n \\ &\qquad + c_1\cdots c_{n-k-2}c_{n-k}c_{n-k+1} + \cdots + c_1\cdots c_{n-k-2}c_{n-k}c_n + \cdots + c_{k+1}c_{k+2}\cdots c_n) \\ &= (-1)^{n-k}\sigma_{n-k}(c_1, c_2, \ldots, c_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ a_1 &= (-1)^{n-1}(c_1c_2\cdots c_{n-1} + c_1c_2\cdots c_{n-2}c_n + c_1c_2\cdots c_{n-3}c_{n-1}c_n + \cdots + c_2c_3\cdots c_n) \\ &= (-1)^{n-1}\sigma_{n-1}(c_1, c_2, \ldots, c_n), \\ a_0 &= (-1)^n(c_1c_2\cdots c_n) \\ &= (-1)^n\sigma_n(c_1, c_2, \ldots, c_n). \end{aligned} \qquad \blacksquare$$

例 1.8.2 证明多项式 $f(x) = x^9 + x^7 + x^5 + x^2 + x - 1$ 的根的平方和为 -2.

证明 设 $c_1, c_2, \ldots, c_n$ 是 $f(x)$ 的根. 把 $f(x)$ 的根的平方和表为基本对称多项式的多项式,得到

$$\sum_{j=1}^{9} c_j^2 = \sigma_1^2(c_1, c_2, \ldots, c_n) - 2\sigma_2(c_1, c_2, \ldots, c_n).$$

由定理 1.8.2, $\sigma_1(c_1, c_2, \ldots, c_n) = 0$, $\sigma_2(c_1, c_2, \ldots, c_n) = 1$, 于是,

$$\sum_{j=1}^{9} c_j^2 = -2.$$ ■

在对称多项式中,除基本对称多项式外,还有一组重要的对称多项式,即等幂和,其定义为,对于 $k = 0, 1, 2, \ldots$,

$$s_k(x_1, x_2, \ldots, x_n) = x_1^k + x_2^k + \cdots + x_n^k.$$

关于基本对称多项式与等幂和,有

定理 1.8.3 (Newton 恒等式) 当 $k \geqslant n$ 时,

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^{n-1}\sigma_{n-1}s_{k-n+1} + (-1)^n \sigma_n s_{k-n} = 0;$$

当 $k < n$ 时,

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0.$$

证明 将 $x_1, x_2, \ldots, x_n$ 视为常数,考虑多项式

$$f(x) = (x - x_1)(x - x_2)\cdots(x - x_n).$$

由 Vièta 定理,

$$f(x) = \sum_{j=0}^{n} (-1)^{n-j}\sigma_{n-j}(x_1, x_2, \ldots, x_n)x^j. \tag{1.8.2}$$

因为 $f(x) = (x - x_1)(x - x_2)\cdots(x - x_n)$, 所以

$$f'(x) = f(x)\sum_{i=1}^{n} \frac{1}{x - x_i}.$$

记

$$g(x) = f(x)\sum_{i=1}^{n} \frac{x_i^{k+1}}{x - x_i},$$

显然, $\deg g(x) < n$. 于是,对任意正整数 k,

$$\begin{aligned} x^{k+1}f'(x) - g(x) &= f(x)\sum_{i=1}^{n} \frac{x^{k+1} - x_i^{k+1}}{x - x_i} \\ &= f(x)\sum_{i=1}^{n} (x^k + x^{k-1}x_i + \cdots + x_i^k) \\ &= f(x)\sum_{i=1}^{n} (nx^k + s_1 x^{k-1} + \cdots + s_{k-1}x + s_k). \end{aligned}$$

即

$$x^{k+1}f'(x) = f(x)\sum_{i=1}^{n} (nx^k + s_1 x^{k-1} + \cdots + s_{k-1}x + s_k) + g(x),$$

其中 $\deg g(x) < n$. 将 (1.8.2) 代入上式,得到

$$\sum_{i=1}^{n} (-1)^{n-i} i\sigma_{n-i}x^{k+i} = \left(\sum_{j=0}^{n} (-1)^{n-j}\sigma_{n-j}x^j\right)\left(\sum_{i=0}^{k} s_i x^{k-i}\right) + g(x). \tag{1.8.3}$$

当 $k<n$ 时，比较上式两端 n 次项系数，得到

$$(-1)^k(n-k)\sigma_k=s_k-\sigma_1 s_{k-1}+\cdots+(-1)^{k-1}\sigma_{k-1}s_1+(-1)^k n\sigma_k.$$

因此，当 $k<n$ 时，

$$s_k-\sigma_1 s_{k-1}+\cdots+(-1)^{k-1}\sigma_{n-1}s_1+(-1)^k k\sigma_k=0.$$

当 $k\geqslant n$ 时，比较式 (1.8.3) 两端 n 次项系数，得到

$$s_k-\sigma_1 s_{k-1}+\cdots+(-1)^{n-1}\sigma_{n-1}s_{k-n+1}+(-1)^n\sigma_n s_{k-n}=0.$$ ■

利用 Newton 恒等式，可以把基本对称多项式表为等幂和的多项式. 例如，

$$\sigma_1=s_1,\qquad \sigma_2=\frac{s_1^2-s_2}{2},\qquad \sigma_3=\frac{s_1^3-3s_1s_2+2s_3}{6},$$

等等. 于是再利用定理 1.8.1，即可得到

推论 1.8.1　数域 $\mathbb{F}$ 上的 n 元对称多项式 $f(x_1,x_2,\ldots,x_n)$ 可以唯一地表为关于等幂和 $s_1,s_2,\ldots,s_n$ 的多项式，即存在唯一的 $g(x_1,x_2,\ldots,x_n)\in\mathbb{F}[x]$，使得

$$f(x_1,x_2,\ldots,x_n)=g\big(s_1(x_1,x_2,\ldots,x_n),s_2(x_1,x_2,\ldots,x_n),\ldots,s_n(x_1,x_2,\ldots,x_n)\big).$$

习　题　1.8

1. 把下列对称多项式表为关于基本对称多项式的多项式.

(1) $(x_1^2+x_2^2)(x_2^2+x_3^2)(x_3^2+x_1^2)$；

(2) $(2x_1-x_2-x_3)(2x_2-x_3-x_1)(2x_3-x_1-x_2)$；

(3) $(-x_1+x_2+\cdots+x_n)(x_1-x_2+\cdots+x_n)\cdots(x_1+x_2+\cdots+x_{n-1}-x_n)$；

(4) $\displaystyle\sum_{1\leqslant i<j\leqslant n}(x_i-x_j)^2$；　　(5) $\displaystyle\sum_{\substack{1\leqslant i<j\leqslant n\\ k\neq i,j}}(x_i+x_j-x_k)^2$；

(6) $$\sum_{\left(\begin{smallmatrix}1&2&\cdots&n\\ i_1&i_2&\cdots&i_n\end{smallmatrix}\right)}\left(a_1x_{i_1}+a_2x_{i_2}+\cdots+a_nx_{i_n}\right)^2,$$

这里的求和号表示对遍历自然数 $1,2,\ldots,n$ 的所有排列 $i_1,i_2,\ldots,i_n$ 求和.

2. 证明，三次实系数方程 $x^3+ax^2+bx+c=0$ 的每个根的实部都是负数的充分必要条件为

$$a>0,\qquad ab-c>0,\qquad c>0.$$

3. 设三次方程 $x^3+ax^2+bx+c=0$ 的三个根是某个三角形的内角的正弦. 证明，

$$a(4ab-a^3-8c)=4c^2.$$

4. 设 $x_1,x_2,\ldots,x_n$ 是多项式 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+x^n$ 的 n 个根. 证明，关于 $x_2,x_3,\ldots,x_n$ 的对称多项式可以表为关于 x_1 的多项式.

5. 求

$$\sum_{i=1}^n\frac{\partial\sigma_k}{\partial x_i},$$

其中求和号后的偏导数表示 $\sigma_k(x_1,x_2,\ldots,x_n)$ 关于 x_i 的偏导数.

6. 设对称多项式 $f(x_1,x_2,\ldots,x_n)$ 满足

$$f(x_1+a,x_2+a,\ldots,x_n+a)=f(x_1,x_2,\ldots,x_n),$$

其中 a 是任意常数. 设 $f(x_1,x_2,\ldots,x_n)=g(\sigma_1,\sigma_2,\ldots,\sigma_n)$. 证明

$$n\frac{\partial g}{\partial\sigma_1}+(n-1)\sigma_1\frac{\partial g}{\partial\sigma_2}+\cdots+\sigma_{n-1}\frac{\partial g}{\partial\sigma_n}=0.$$

7. 把 n 元等幂和 $s_1, s_2, \ldots, s_6$ 表为关于 n 元基本对称多项式的多项式，其中 $n \geqslant 6$.

8. 把 n 元基本对称多项式 $\sigma_4, \sigma_5, \sigma_6$ 表为 n 元等幂和 $s_1, s_2, \ldots$ 的多项式.

9. 把下列 n 元对称多项式表为 n 元等幂和的多项式，其中 k 是正整数.

(1) $\sum\limits_{1\leqslant i<j\leqslant n} x_i^k x_j^k$;　　(2) $\sum\limits_{1\leqslant i<j\leqslant n} (x_i + x_j)^k$;　　(3) $\sum\limits_{1\leqslant i<j\leqslant n} (x_i - x_j)^{2k}$.

10. 求多项式

$$f(x) = x^n + (a+b)x^{n-1} + (a^2+b^2)x^{n-2} + \cdots + (a^{n-1}+b^{n-1})x + (a^n+b^n)$$

的根的等幂和 $s_1, s_2, \ldots, s_n$.

11. 自然数 $1, 2, \ldots, n$ 的**循环排列**是指排列 $12\cdots n$ 与 $j(j+1)\cdots n12\cdots(j-1)$，$j = 2, 3, \ldots, n$. 如果对于 $1, 2, \ldots, n$ 的每个循环排列 $j(j+1)\cdots n12\cdots(j-1)$，多项式 $f(x_1, \ldots, x_n)$ 适合

$$f\big(x_j, x_{j+1}, \ldots, x_n, x_1, x_2, \ldots, x_{j-1}\big) = f(x_1, x_2, \ldots, x_n),$$

其中 $j = 2, 3, \ldots, n$，则 $f(x_1, x_2, \ldots, x_n)$ 称为在未定元 $x_1, x_2, \ldots, x_n$ 的**循环变换下不变**. 证明，循环变换下不变的多项式 $f(x_1, x_2, \ldots, x_n)$ 可表为多项式

$$g_j(x_1, x_2, \ldots, x_n) = x_1\omega^j + x_2\omega^{2j} + \cdots + x_n\omega^{nj}$$

的多项式，其中 $j = 0, 1, \ldots, n-1, \omega, \omega^2, \ldots, \omega^n$ 分别是方程 $x^n = 1$ 的 n 个相异复数根.

行 列 式

- 尽管行列式理论并不是线性代数的主体，但它无疑是处理各类线性代数问题的不可缺少的工具.
- 作为三维向量空间的直接推广，§2.1 定义了数域上的 n 元数组的空间，即数域上的 n 维向量空间，并把行列式定义为 n 元数组空间上的规范反对称 n 重线性函数.
- 在 §2.2 中证明了行列式的存在性和唯一性以及行列式的基本性质. 这一节的内容是行列式理论的最基本部分.
- 如何具体计算行列式，是行列式理论中的一个重要问题. §2.3 给出了行列式的 Laplace 展开定理. 它说明，一个高阶行列式可以归结为一些低阶行列式的和.
- 作为行列式理论的一个应用，§2.4 导出了解线性方程组的 Cramer 法则.
- 最后，在 §2.5 中用一些典型例子阐明计算行列式的各种常用方法，以帮助读者提高计算行列式的能力.

§2.1　数域 $\mathbb{F}$ 上的 n 维向量空间

通过解析几何的学习，我们知道，在通常空间中建立坐标系（例如直角坐标系）之后，空间中给定的向量 α 便唯一地确定一个坐标，即有序三元实数组 (a_1, a_2, a_3). 反之，给定有序三元实数组，利用所建立的坐标系，便可在空间中唯一地确定一个向量. 换句话说，如果记所有的有序三元实数组 (a_1, a_2, a_3) 的集合为 $\mathbb{R}^3$，即

$$\mathbb{R}^3 = \left\{(a_1, a_2, a_3) \,\middle|\, a_i \in \mathbb{R}, i = 1, 2, 3\right\}.$$

则空间中所有向量的集合与 $\mathbb{R}^3$ 之间存在一一对应. 而且如果空间中向量 α 与 β 分别对应于有序三元实数组 (a_1, a_2, a_3) 与 (b_1, b_2, b_3)，则向量 α 与 β 的和 $\alpha + \beta$ 便对应于有序三元实数组

$$(a_1 + b_1, a_2 + b_2, a_3 + b_3);$$

而纯量 λ 与向量 α 的乘积 $\lambda\alpha$ 便对应于有序三元实数组

$$(\lambda\alpha_1, \lambda\alpha_2, \lambda\alpha_3).$$

因此，可以把通常空间与 $\mathbb{R}^3$ 等同起来. 所以 $\mathbb{R}^3$ 称为三维实向量空间；而 $\mathbb{R}^3$ 中的元素，即有序三元实数组称为三维实向量. 根据这一点，可以把通常空间的概念加以推广.

设 n 是正整数，由 n 个实数 $a_1, a_2, \ldots, a_n$ 组成的有序 n 元数组

$$\alpha=(a_1,a_2,\ldots,a_n)$$

称为 **n 维实向量**(简称为**向量**),实数 a_i 称为 n 维实向量 α 的**第 i 个坐标**,$1\leqslant i\leqslant n$. 所有 n 维实向量的集合记为 $\mathbb{R}^n$,即

$$\mathbb{R}^n=\left\{(a_1,a_2,\ldots,a_n)\;\middle|\;a_i\in\mathbb{R},1\leqslant i\leqslant n\right\}.$$

设 $\alpha=(a_1,a_2,\ldots,a_n)$ 和 $\beta=(b_1,b_2,\ldots,b_n)\in\mathbb{R}^n$. 如果 $a_i=b_i$, $i=1,2,\ldots,n$,则称 n 维实向量 α 与 β **相等**,记为 $\alpha=\beta$.

定义

$$\alpha+\beta=(a_1+b_1,a_2+b_2,\ldots,a_n+b_n),$$

它称为向量 α 与 β 的**和**.

显然,$\alpha+\beta\in\mathbb{R}^n$. 这样便规定了 $\mathbb{R}^n$ 中向量的加法运算. 容易验证,$\mathbb{R}^n$ 中向量的加法满足以下的公理.

(A1) 加法结合律 设 $\alpha,\beta,\gamma\in\mathbb{R}^n$,则

$$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma);$$

(A2) 加法交换律 设 $\alpha,\beta\in\mathbb{R}^n$,则

$$\alpha+\beta=\beta+\alpha;$$

(A3) 存在零向量 即 $\mathbb{R}^n$ 中坐标全为零的向量 $0=(0,0,\ldots,0)$,使得对每个 $\alpha\in\mathbb{R}^n$,均有

$$\alpha+0=\alpha=0+\alpha;$$

(A4) 设 $\alpha\in\mathbb{R}^n$,则存在 $\beta\in\mathbb{R}^n$,使得

$$\alpha+\beta=0=\beta+\alpha.$$

向量 β 称为向量 α 的**负向量**,记为 $-\alpha$. 容易看出,对于 $(a_1,a_2,\ldots,a_n)=\alpha$,

$$-\alpha=(-a_1,a_2,\ldots,-a_n).$$

利用负向量的概念,可以在 $\mathbb{R}^n$ 中引进减法运算:设 $\alpha,\beta\in\mathbb{R}^n$,则规定向量 α 与 β 之差为 $\alpha+(-\beta)$,并记为 $\alpha-\beta$. 显然,如果 $\alpha=(a_1,a_2,\ldots,a_n)$,$\beta=(b_1,b_2,\ldots,b_n)$,则

$$\alpha-\beta=(a_1-b_1,a_2-b_2,\ldots,a_n-b_n).$$

纯量与 $\mathbb{R}^n$ 中向量的乘法规定如下:设 $\lambda\in\mathbb{R}$,$\alpha=(a_1,a_2,\ldots,a_n)\in\mathbb{R}^n$,则

$$\lambda\alpha=(\lambda a_1,\lambda a_2,\ldots,\lambda a_n)$$

称为**纯量 λ 与向量 α 的乘积**.

显然,$\lambda\alpha\in\mathbb{R}^n$. 并且容易验证,纯量与向量的乘法满足以下公理.

(M1) 设 $\lambda,\mu\in\mathbb{R}$,$\alpha\in\mathbb{R}^n$,则

$$(\lambda\mu)\alpha=\lambda(\mu\alpha);$$

(M2) 对任意 $\alpha\in\mathbb{R}^n$,均有

$$1\cdot\alpha=\alpha.$$

向量的加法和纯量与向量的乘法之间有以下关系：

(D1)　设 $\lambda \in \mathbb{R}, \alpha, \beta \in \mathbb{R}^n$，则

$$\lambda(\alpha+\beta) = \lambda\alpha + \lambda\beta;$$

(D2)　设 $\lambda, \mu \in \mathbb{R}, \alpha \in \mathbb{R}^n$，则

$$(\lambda+\mu)\alpha = \lambda\alpha + \mu\alpha.$$

$\mathbb{R}^n$ 连同向量的加法以及纯量与向量的乘法一起称为 **n 维实向量空间**. 有时也径称 $\mathbb{R}^n$ 为 n 维实向量空间. 同样可以定义数域 $\mathbb{F}$ 上的 n 维向量空间：设 $a_1, a_2, \ldots, a_n \in \mathbb{F}$，则有序 n 元数组 $(a_1, a_2, \ldots, a_n)$ 称为数域 $\mathbb{F}$ 上的 **n 维向量**.

数域 $\mathbb{F}$ 上的所有 n 维向量的集合记为 $\mathbb{F}^n$. 数域 $\mathbb{F}$ 上的 n 维向量的相等，加法以及纯量(即 $\mathbb{F}$ 的元素)与数域 $\mathbb{F}$ 上的 n 维向量的乘法的定义如同 $\mathbb{R}^n$.

数域 $\mathbb{F}$ 上的 n 维向量的加法，纯量与数域 $\mathbb{F}$ 上的 n 维向量的乘法以及这两种代数运算之间的关系所满足的公理也都与 $\mathbb{R}^n$ 相同.

$\mathbb{F}^n$ 连同向量的加法以及纯量与向量的乘法一起称为**数域 $\mathbb{F}$ 上的 n 维向量空间**. 特别，复数域 $\mathbb{C}$ 上的 n 维向量空间 $\mathbb{C}^n$ 称为 **n 维复向量空间**，$\mathbb{C}^n$ 中的向量称为 n 维复向量.

与通常空间一样，可以引进数域 $\mathbb{F}$ 上的 n 维向量空间 $\mathbb{F}^n$ 的函数概念.

所谓定义在 n 维向量空间 $\mathbb{F}^n$ 上而取值在数域 $\mathbb{F}$ 上的**一元函数**是指 $\mathbb{F}^n$ 到 $\mathbb{F}$ 的一种对应规律 f，使得对 $\mathbb{F}^n$ 中每个向量 ξ，依照对应规律 f，可以唯一地确定 $\mathbb{F}$ 中一个元素 $f(\xi)$ 与之对应. 例如，设 $\xi = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$，则

$$f(\xi) = x_1 + x_2 + \cdots + x_n$$

就是一个定义在 $\mathbb{F}^n$ 上而取值在 $\mathbb{F}$ 上的一元函数.

同样，所谓 n 维向量空间上的 **k 元函数**是指一种对应规律 f，使得对每一个由 k 个向量 $\xi_1, \xi_2, \ldots, \xi_k$ 组成的有序 k 元向量组 $(\xi_1, \xi_2, \ldots, \xi_k)$，依照对应规律，可以唯一地确定 $\mathbb{F}^n$ 中一个元素与之对应. 例如，设 $\xi_i = (x_{i1}, x_{i2}, \ldots, x_{in}) \in \mathbb{F}^n$，则

$$f(\xi_1, \xi_2, \ldots, \xi_k) = x_{11}x_{22}\cdots x_{kk} \tag{$*$}$$

就是 $\mathbb{F}^n$ 上的一个 k 元函数.

设 $f(\xi_1, \xi_2, \ldots, \xi_k)$ 是 $\mathbb{F}^n$ 上的一个 k 元函数. 如果对每个 i，$1 \leqslant i \leqslant k$，均有

$$\begin{aligned} &f(\xi_1, \ldots, \xi_{i-1}, \lambda\eta + \mu\zeta, \xi_{i+1}, \ldots, \xi_k) \\ &\qquad = \lambda f(\xi_1, \ldots, \xi_{i-1}, \eta, \xi_{i+1}, \ldots, \xi_k) + \mu f(\xi_1, \ldots, \xi_{i-1}, \zeta, \xi_{i+1}, \ldots, \xi_k), \end{aligned}$$

其中 $\lambda, \mu, \eta, \zeta \in \mathbb{F}$，则 $f(\xi_1, \xi_2, \ldots, \xi_k)$ 称为 **k 重线性函数**. 1 重线性函数常简称为线性函数. 例如，上述定义的 k 元函数 $(*)$ 就是一个 k 重线性函数.

例 2.1.1　确定 $\mathbb{F}^2$ 上的所有二重线性函数.

解　设向量 $\xi_1 = (a_{11}, a_{12})$，$\xi_2 = (a_{21}, a_{22}) \in \mathbb{F}^2$，且设 $\varepsilon_1 = (1, 0)$，$\varepsilon_2 = (0, 1)$. 则 $\xi_1 = a_{11}\varepsilon_1 + a_{12}\varepsilon_2$，$\xi_2 = a_{21}\varepsilon_1 + a_{22}\varepsilon_2$.

设 $f(\xi_1,\xi_2)$ 是一个二重线性函数. 则由定义

$$\begin{aligned}f(\xi_1,\xi_2)&=f(a_{11}\varepsilon_1+a_{12}\varepsilon_2,a_{21}\varepsilon_1+a_{22}\varepsilon_2)\\&=a_{11}f(\varepsilon_1,a_{21}\varepsilon_1+a_{22}\varepsilon_2)+a_{12}f(\varepsilon_2,a_{21}\varepsilon_1+a_{22}\varepsilon_2)\\&=a_{11}a_{21}f(\varepsilon_1,\varepsilon_1)+a_{11}a_{22}f(\varepsilon_1,\varepsilon_2)+a_{12}a_{21}f(\varepsilon_2,\varepsilon_1)+a_{12}a_{22}f(\varepsilon_2,\varepsilon_2).\end{aligned}$$

这表明,$f(\xi_1,\xi_2)$ 完全由 4 个值 $f(\varepsilon_1,\varepsilon_1),f(\varepsilon_1,\varepsilon_2),f(\varepsilon_2,\varepsilon_1),f(\varepsilon_2,\varepsilon_2)$ 所确定. 现在设 a,b,c,d 是 $\mathbb{F}$ 中任意四个数,记

$$f(\xi_1,\xi_2)=a_{11}a_{21}a+a_{11}a_{22}b+a_{12}a_{21}c+a_{12}a_{22}d.$$

容易验证,$f(\xi_1,\xi_2)$ 是 $\mathbb{F}^2$ 上的二重线性函数. ■

设 $f(\xi_1,\xi_2,\dots,\xi_k)$ 是 $\mathbb{F}^n$ 上的 k 重线性函数. 如果对任意 $i,j,1\leqslant i\neq j\leqslant k$,当向量 ξ_i 与 ξ_j 相同时,$f(\xi_1,\xi_2,\dots,\xi_k)$ 的值为 0,则 $f(\xi_1,\xi_2,\dots,\xi_k)$ 称为**反对称的**.

例如,设 $\xi_1=(a_{11},a_{12}),\xi_2=(a_{21},a_{22})\in\mathbb{F}^2$,则

$$f(\xi_1,\xi_2)=\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}=a_{11}a_{22}-a_{21}a_{12}$$

是 $\mathbb{F}^2$ 上的反对称二重线性函数.

关于 $\mathbb{F}^n$ 上的反对称 k 重线性函数,有

定理 2.1.1 设 $f(\xi_1,\xi_2,\dots,\xi_k)$ 是 $\mathbb{F}^n$ 上的 k 重线性函数. 则 $f(\xi_1,\xi_2,\dots,\xi_k)$ 是反对称的充分必要条件为,对任意 $i,j,1\leqslant i\neq j\leqslant k$,均有

$$f(\xi_1,\dots,\xi_j,\dots,\xi_i,\dots,\xi_k)=-f(\xi_1,\dots,\xi_i,\dots,\xi_j,\dots,\xi_k),\tag{2.1.1}$$

即任意对调 $f(\xi_1,\xi_2,\dots,\xi_k)$ 中两个向量元的位置,$f(\xi_1,\xi_2,\dots,\xi_k)$ 的值变号.

证明 设 $f(\xi_1,\xi_2,\dots,\xi_k)$ 是反对称的,则对任意 $i,j,1\leqslant i\neq j\leqslant k$,

$$f(\xi_1,\dots,\xi_i+\xi_j,\dots,\xi_i+\xi_j,\dots,\xi_k)=0.$$

因为 $f(\xi_1,\xi_2,\dots,\xi_k)$ 是 k 重线性的,所以,

$$\begin{aligned}&f(\xi_1,\dots,\xi_i+\xi_j,\dots,\xi_i+\xi_j,\dots,\xi_k)\\&\quad=f(\xi_1,\dots,\xi_i,\dots,\xi_i,\dots,\xi_k)+f(\xi_1,\dots,\xi_i,\dots,\xi_j,\dots,\xi_k)\\&\qquad+f(\xi_1,\dots,\xi_j,\dots,\xi_i,\dots,\xi_k)+f(\xi_1,\dots,\xi_j,\dots,\xi_j,\dots,\xi_k)\\&\quad=0.\end{aligned}$$

再用一次 $f(\xi_1,\xi_2,\dots,\xi_k)$ 的反对称性,便得到式 (2.1.1).

反之,设式 (2.1.1) 成立. 则对任意 $i,j,1\leqslant i\neq j\leqslant k$,

$$f(\xi_1,\dots,\xi,\dots,\xi,\dots,\xi_k)=-f(\xi_1,\dots,\xi,\dots,\xi,\dots,\xi_k).$$

因此 $f(\xi_1,\dots,\xi,\dots,\xi,\dots,\xi_k)=0$. 所以 $f(\xi_1,\xi_2,\dots,\xi_k)$ 是反对称的. ■

设 $f(\xi_1,\xi_2,\dots,\xi_n)$ 是 $\mathbb{F}$ 上的 n 元函数,且设 ε_i 是 $\mathbb{F}^n$ 中第 i 个分量为 1 其它分量为 0 的向量. 如果 $f(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)=1$,则 $f(\xi_1,\xi_2,\dots,\xi_k)$ 称为**规范的**.

例如,设 $\xi_i=(a_{i1},a_{i2},a_{i3})\in\mathbb{F}^3,i=1,2,3$. 且设

$$f(\xi_1,\xi_2,\xi_3)=\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix},$$

其中右端为三阶行列式. 则 $f(\xi_1,\xi_2,\xi_3)$ 是 $\mathbb{F}^3$ 上的一个规范反对称三重线性函数.

反之,有

定理 2.1.2　设 $f(\xi_1,\xi_2,\xi_3)$ 是三维向量空间 $\mathbb{F}^3$ 上的规范反对称三重线性函数,且 $\xi_i=(a_{i1},a_{i2},a_{i3})$, $i=1,2,3$. 则

$$f(\xi_1,\xi_2,\xi_3)=\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}. \tag{2.1.2}$$

证明　记 $\varepsilon_1=(1,0,0)$, $\varepsilon_2=(0,1,0)$, $\varepsilon_3=(0,0,1)$. 则对于 $i=1,2,3$,

$$\xi_i=(a_{i1},a_{i2},a_{i3})=\sum_{j=1}^{3}a_{ij}\varepsilon_j.$$

因为 $f(\xi_1,\xi_2,\xi_3)$ 是三重线性的,所以

$$f(\xi_1,\xi_2,\xi_3)=\sum_{1\leqslant r,s,t\leqslant 3}a_{1r}a_{2s}a_{3t}f(\varepsilon_r,\varepsilon_s,\varepsilon_t). \tag{2.1.3}$$

因为 $f(\xi_1,\xi_2,\xi_3)$ 是反对称的,所以当 $f(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ 的下标 r,s,t 中有两个相同时,

$$f(\varepsilon_r,\varepsilon_s,\varepsilon_t)=0.$$

于是式 (2.1.3) 化为

$$f(\xi_1,\xi_2,\xi_3)=\sum_{\substack{1\leqslant r,s,t\leqslant 3\\ r,s,t\text{ 两两不等}}}a_{1r}a_{2s}a_{3t}f(\varepsilon_r,\varepsilon_s,\varepsilon_t). \tag{2.1.4}$$

由于 $f(\xi_1,\xi_2,\xi_3)$ 是规范的,因此 $f(\varepsilon_1,\varepsilon_2,\varepsilon_3)=1$.

由定理 2.1.1 得到,

$$\begin{aligned} &f(\varepsilon_1,\varepsilon_3,\varepsilon_2)=f(\varepsilon_3,\varepsilon_2,\varepsilon_1)=f(\varepsilon_2,\varepsilon_1,\varepsilon_3)=-f(\varepsilon_1,\varepsilon_2,\varepsilon_3)=-1,\\ &f(\varepsilon_3,\varepsilon_1,\varepsilon_2)=f(\varepsilon_2,\varepsilon_3,\varepsilon_1)=f(\varepsilon_1,\varepsilon_2,\varepsilon_3)=1.\end{aligned}$$

由式 (2.1.4) 得到,

$$f(\xi_1,\xi_2,\xi_3)=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}.\ \blacksquare$$

习　题　2.1

1. 设 $\xi_1=(x_1,x_2)$ 与 $\xi_2=(y_1,y_2)$ 是二维实向量. 指出下列二维实向量空间 $\mathbb{R}^2$ 上的二元实函数中哪些是二重线性函数?

(1) $f(\xi_1,\xi_2)=1$;　　(2) $f(\xi_1,\xi_2)=(x_1-y_1)^2+x_2y_2$;

(3) $f(\xi_1,\xi_2)=(x_1+y_1)^2-(x_1-y_1)^2$;　　(4) $f(\xi_1,\xi_2)=(x_1x_2)^2-x_2y_1$;

(5) $f(\xi_1,\xi_2)=x_1y_1+x_2y_2$.

2. 设 $\xi=(x_1,x_2,x_3)$ 与 $\xi_2=(y_1,y_2,y_3)$ 是三维实向量,而 $\varepsilon_1=(1,0,0)$,$\varepsilon_2=(0,1,0)$,$\varepsilon_3=(0,0,1)$. 证明,三维实向量空间 $\mathbb{R}^3$ 上的每个反对称二重线性函数都可以表为

$$f(\xi_1,\xi_2)=(x_1y_2-x_2y_1)f(\varepsilon_1,\varepsilon_2)+(x_1y_3-x_3y_1)f(\varepsilon_1,\varepsilon_3)+(x_2y_3-x_3y_2)f(\varepsilon_2,\varepsilon_3).$$

3. 设 $\xi_1=(x_1,x_2,\ldots,x_n)$ 与 $\xi_2=(y_1,y_2,\ldots,y_3)$ 是 n 维实向量. 定义

$$f(\xi_1,\xi_2)=x_1y_1+x_2y_2+\cdots+x_ny_n.$$

验证 $f(\xi_1,\xi_2)$ 是 n 维实向量空间上的二重线性函数.

§2.2 n 阶行列式的定义与性质

定理 2.1.2 揭示了三阶行列式的本质. 它说明,如果将三阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

中每个元素 a_{ij} 视为变量,而将每个行 $\xi_i=(a_{i1},a_{i2},a_{i3})$ 视为三维向量,则三阶行列式即等价于三维向量空同上的规范反对称三重线性函数. 这就为推广三阶行列式为 n 阶行列式提供了根据.

定义 2.2.1 n 维向量空间 $\mathbb{F}^n$ 上的规范反对称 n 重线性函数称为数域 $\mathbb{F}$ 上的 ***n* 阶行列式函数**,简称为 ***n* 阶行列式**,并记为

$$\det(\xi_1,\xi_2,\ldots,\xi_n).$$

对于给定的 $\xi_i^0=(a_{i1},a_{i2},\ldots,a_{in})\in\mathbb{F}^n$,可以将 n^2 个数 $a_{ij}\in\mathbb{F}$,$1\leqslant i,j\leqslant n$,排成如下的正方形表:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

其中第 i 行由向量 ξ_i 的 n 个分量组成. A 称为数域 $\mathbb{F}$ 上的 ***n* 阶方阵**. $\mathbb{F}^n$ 上的 n 阶行列式函数 $\det(\xi_1,\xi_2,\ldots,\xi_n)$ 在 $\xi_i=\xi_i^0$ $(i=1,2,\ldots,n)$ 的取值 $\det(\xi_1^0,\xi_2^0,\ldots,\xi_n^0)$ 记作

$$\det A=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{2.2.1}$$

根据定义,可以得出如下定理.

定理 2.2.1 **(1)** 对换 n 阶行列式的某两行,其值变号;

(2) n 阶行列式的某一行遍乘以某个数,加到另一行,其值不变;

(3) n 阶行列式的某一行遍乘以数 λ,其值为原行列式的 λ 倍.

具体地说,设 n 阶行列式为 (2.2.1),则

(1)

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix};$$

(2)

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1}+\lambda a_{j1} & a_{i2}+\lambda a_{j2} & \cdots & a_{in}+\lambda a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix};$$

(3)

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{i1} & \lambda a_{i2} & \cdots & \lambda a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证明 记 $\xi_i = (a_{i1}, a_{i2}, \ldots, a_{in}), i = 1, 2, \ldots, n$. 由定理 2.1.1,

$$\det(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_n) = -\det(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_n).$$

此即为 (1). 由于 $\det(\xi_1, \xi_2, \ldots, \xi_n)$ 是 n 重线性的,所以,

$$\begin{aligned} &\det\big(\xi_1, \ldots, \underset{i}{\xi_i + \lambda\xi_j}, \ldots, \underset{j}{\xi_j}, \ldots, \xi_n\big) \\ &\qquad = \det\big(\xi_1, \ldots, \underset{i}{\xi_i}, \ldots, \underset{j}{\xi_j}, \ldots, \xi_n\big) + \lambda \det\big(\xi_1, \ldots, \underset{i}{\xi_j}, \ldots, \underset{j}{\xi_j}, \ldots, \xi_n\big). \end{aligned}$$

因为 $\det(\xi_1, \xi_2, \ldots, \xi_n)$ 是反对称的,故得 (2). 因为 $\det(\xi_1, \xi_2, \ldots, \xi_n)$ 是 n 重线性的,所以,

$$\det\big(\xi_1, \ldots, \underset{i}{\lambda\xi_i}, \ldots, \xi_n\big) = \lambda \det\big(\xi_1, \ldots, \underset{i}{\xi_i}, \ldots, \xi_n\big).$$

此即为 (3). ■

推论 2.2.1 **(1)** n 阶行列式的某一行的元素全为零,其值为 0;

(2) n 阶行列式的某两行的相应元素相同,其值为 0;

(3) n 阶行列式的某两行的相应元素成比例,其值为 0.

证明 由定理 2.2.1 的性质 (3) 即得 (1).

由定理 2.2.1 的性质 (1) 或者 n 阶行列式的反对称性定义即得 (2).

由定理 2.2.1 的性质 (3) 与本推论之 (2) 即得 (3). ■

定理 2.2.2 如果 n 阶行列式的某一行是两组数的和,则其值等于两个 n 阶行列式的值之和,这两个行列式除这一行外其它各行与原行列式相应的行相同. 具体地说,有

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证明 记 $\xi_k = (a_{k1}, a_{k2}, \ldots, a_{kn})$, $k = 1, 2, \ldots, n$, $\eta_i = (b_{i1}, b_{i2}, \ldots, b_{in})$. 由于行列式函数 $\det(\xi_1, \xi_2, \ldots, \xi_n)$ 是 n 重线性的,所以,

$$\det\big(\xi_1, \ldots, \underset{i}{\xi_i + \eta_i}, \ldots, \xi_n\big) = \det\big(\xi_1, \ldots, \underset{i}{\xi_i}, \ldots, \xi_n\big) + \det\big(\xi_1, \ldots, \underset{i}{\eta_i}, \ldots, \xi_n\big). \quad ■$$

问题是,n 阶行列式是否存在? 如果存在,是否唯一?

为了讨论这一问题,先引进如下记号. 对于 $\xi_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbb{F}^n$, $i = 1, 2, \ldots, n$,可以确定 n 阶方阵 $A = (a_{ij})$. 划掉 A 的第 i 行与第 1 列,得到 $n-1$ 阶方阵为

$$\begin{pmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

记 $\eta_{ik} = (a_{k2}, a_{k3}, \ldots, a_{kn})$, $1 \leqslant k \leqslant i-1$, $\eta_{ik} = (a_{k+1,2}, a_{k+1,3}, \ldots, a_{k+1,n})$, $i \leqslant k \leqslant n-1$, 则 $\eta_{ik} \in \mathbb{F}^{n-1}$. 设 $\det(\eta_1, \eta_2, \ldots, \eta_{n-1})$ 为 $n-1$ 阶行列式函数,记

$$A_{i1} = (-1)^{i+1} \det(\eta_{i1}, \eta_{i2}, \ldots, \eta_{i,n-1}) = (-1)^{i+1} \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

定理 2.2.3 设 $n > 1$,且设 $\det(\eta_1, \eta_2, \ldots, \eta_{n-1})$ 是 $n-1$ 阶行列式函数,则

$$f(\xi_1, \xi_2, \ldots, \xi_n) = \sum_{i=1}^{n} a_{i1} A_{i1}$$

是 n 阶行列式函数.

证明　设 $\lambda,\mu\in\mathbb{F},\alpha_k=(\widetilde{a}_{i1},\widetilde{a}_{i2},\dots,\widetilde{a}_i),\beta_k=(\widetilde{b}_{i1},\widetilde{b}_{i2},\dots,\widetilde{b}_i)$,且设 $\xi_k=\lambda\alpha_k+\mu\beta_k,\xi_i=(a_{i1},a_{i2},\dots,a_{in}),i\neq k$. 记

$$\widetilde{A}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ \lambda\widetilde{a}_{k1}+\mu\widetilde{b}_{k1}&\cdots&\lambda\widetilde{a}_{kn}+\mu\widetilde{b}_{kn}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix},\quad \widetilde{B}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ \widetilde{a}_{k1}&\cdots&\widetilde{a}_{kn}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix},\quad \widetilde{C}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ \widetilde{b}_{k1}&\cdots&\widetilde{b}_{kn}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}.$$

则当 $i\neq k$ 时,

$$\widetilde{A}_{i1}=(-1)^{i+1}\begin{vmatrix}a_{12}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{i-1,2}&\cdots&a_{i-1,n}\\ a_{i+1,2}&\cdots&a_{i+1,n}\\ \vdots&&\vdots\\ \lambda\widetilde{a}_{k2}+\mu\widetilde{b}_{k2}&\cdots&\lambda\widetilde{a}_{kn}+\mu\widetilde{b}_{kn}\\ \vdots&&\vdots\\ a_{n2}&\cdots&a_{nn}\end{vmatrix}$$

由于 $n-1$ 阶行列式函数是 $n-1$ 重线性的,因此对于 $i\neq k$

$$\widetilde{A}_{i1}=\lambda\widetilde{B}_{i1}+\mu\widetilde{C}_{i1}.$$

而

$$\widetilde{A}_{k1}=(-1)^{k+1}\begin{vmatrix}a_{12}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{k-1,2}&\cdots&a_{k-1,n}\\ a_{k+1,2}&\cdots&a_{k+1,n}\\ \vdots&&\vdots\\ a_{n2}&\cdots&a_{nn}\end{vmatrix}=\widetilde{B}_{k1}=\widetilde{C}_{k1},$$

因此

$$\begin{aligned}f(\xi_1,\dots,\lambda\alpha_k+\mu\beta_k,\dots,\xi_n)&=\left(\lambda\widetilde{a}_{k1}+\mu\widetilde{b}_{k1}\right)\widetilde{A}_{k1}+\sum_{i\neq k}a_{i1}\widetilde{A}_{i1}\\&=\lambda\widetilde{a}_{k1}\widetilde{B}_{k1}+\mu\widetilde{b}_{k1}\widetilde{C}_{k1}+\sum_{i\neq k}a_{i1}\left(\lambda\widetilde{B}_{i1}+\mu\widetilde{C}_{i1}\right)\\&=\left(\lambda\widetilde{a}_{k1}\widetilde{B}_{k1}+\sum_{i\neq k}\lambda a_{i1}\widetilde{B}_{i1}\right)+\left(\mu\widetilde{b}_{k1}\widetilde{C}_{k1}+\sum_{i\neq k}\mu a_{i1}\widetilde{C}_{i1}\right)\\&=\lambda f(\xi_1,\dots,\alpha_k,\dots,\xi_n)+\mu f(\xi_1,\dots,\beta_k,\dots,\xi_n).\end{aligned}$$

即 $f(\xi_1,\xi_2,\dots,\xi_n)$ 是 n 重线性的.

设 $\xi_i=\xi_j=(b_1,\dots,b_n)=\alpha,i\neq j,\xi_k=(a_{k1},\dots,a_{kn}),k\neq i,j$. 记

$$A=\begin{array}{c}\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_1 & \cdots & b_n \\ \vdots & & \vdots \\ b_1 & \cdots & b_n \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\end{array}\begin{array}{l} \\ \\ \text{第 } i \text{ 行} \\ \\ \text{第 } j \text{ 行} \\ \\ \\ \end{array}, \quad \text{则当 } k\neq i,j \text{ 时}, A_{k1}=(-1)^{k+1}\begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_2 & \cdots & b_n \\ \vdots & & \vdots \\ b_2 & \cdots & b_n \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}=0,$$

而

$$A_{i1}=(-1)^{i+1}\begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \\ b_2 & \cdots & b_n \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}, \quad A_{j1}=(-1)^{j+1}\begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_2 & \cdots & b_n \\ \vdots & & \vdots \\ a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

由定理 2.2.1，$A_{j1}=-A_{i1}$. 因此，

$$f(\xi_1,\ldots,\alpha,\ldots,\alpha,\ldots,\xi_n)=\sum_{k\neq i,j}a_{k1}A_{k1}+b_1A_{i1}+b_1A_{j1}=0.$$

即 $f(\xi_1,\xi_2,\ldots,\xi_n)$ 是反对称的.

设 $\xi_i=\varepsilon_i$，$i=1,2,\ldots,n$. 记 n 阶方阵 E 为

$$E=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

则

$$E_{11}=\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{vmatrix}=1, E_{i1}=0, i\neq 1.$$

因此

$$f(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)=1\cdot E_{11}=1.$$

所以 $f(\xi_1,\xi_2,\ldots,\xi_n)$ 是规范的. 于是，$f(\xi_1,\xi_2,\ldots,\xi_n)$ 是 n 阶行列式函数. ■

定理 2.2.4 n 阶行列式函数确实存在.

证明 对 n 用归纳法.

当 $n=1$ 时，$\xi=(a)\in\mathbb{F}^1$，其中 $a\in\mathbb{F}$. 则 $\det(\xi)=a$ 显然是 1 阶行列式函数.

假设 $n-1$ 阶行列式函数 $\det(\eta_1,\eta_2,\ldots,\eta_{n-1})$ 存在，则由定理 2.2.3，定理 2.2.3 中定义的函数 $f(\xi_1,\xi_2,\ldots,\xi_n)$ 是 n 阶行列式函数. ■

为了证明 n 阶行列式函数的唯一性，需要一些有关自然数 $1,2,\ldots,n$ 的排列的基本性质. 设 $i_1i_2\cdots i_n$ 是自然数 $1,2,\ldots,n$ 的一个排列. 如果当 $p<q$ 时，$i_p>i_q$，则 i_p,i_q 称为排列 $i_1i_2\cdots i_n$ 的一个**逆序**. 排列 $i_1i_2\cdots i_n$ 中逆序的总数称为 $i_1i_2\cdots i_n$ 的**逆序数**. 逆序数为偶数的排列称为**偶排列**，否则称为**奇排列**. 排列 $i_1i_2\cdots i_n$ 的奇偶性符号 $\operatorname{sgn}\begin{pmatrix}1&2&\cdots&n\\ i_1&i_2&\cdots&i_n\end{pmatrix}$ 规定如下：

$$\operatorname{sgn}\begin{pmatrix}1&2&\cdots&n\\ i_1&i_2&\cdots&i_n\end{pmatrix}=\begin{cases}1, & \text{当 } i_1i_2\cdots i_n \text{ 为偶排列时；}\\ -1, & \text{当 } i_1i_2\cdots i_n \text{ 为奇排列时.}\end{cases}$$

对调排列 $i_1i_2\cdots i_n$ 中两个数 i_p 与 i_q，其它的数保持不动，得到一个新排列 $j_1j_2\cdots j_n$. 由 $i_1i_2\cdots i_n$ 得到 $j_1j_2\cdots j_n$ 的过程称为对 $i_1i_2\cdots i_n$ 实施一次**对换**. 如果对换的是 $i_1i_2\cdots i_n$ 中两个相邻的数 i_p 和 i_{p+1}，则这样的对换称为**相邻对换**.

命题 2.2.1　一次对换必改变排列的奇偶性.

证明　分两种情形.

(1) 排列 $i_1i_2\cdots i_pi_{p+1}\cdots i_n$ 经过一次相邻对换变为 $i_1i_2\cdots i_{p+1}i_p\cdots i_n$.

此时，当 $i_p<i_{p+1}$ 时，排列 $i_1i_2\cdots i_{p+1}i_p\cdots i_n$ 的逆序数比 $i_1i_2\cdots i_pi_{p+1}\cdots i_n$ 的增加 1；当 $i_p>i_{p+1}$ 时，排列 $i_1i_2\cdots i_{p+1}i_p\cdots i_n$ 的逆序数比 $i_1i_2\cdots i_pi_{p+1}\cdots i_n$ 的减少 1. 因此一次相邻对换改变排列的奇偶性.

(2) 排列 $i_1\cdots i_p\cdots i_q\cdots i_n$ 经过一次对换改变为 $i_1\cdots i_q\cdots i_p\cdots i_n$.

此时，这种对换可以通过有限次相邻对换实现.

例如，先对换 i_p 和 i_{p+1}，再对换 i_p 和 i_{p+2}，等等，最后再对换 i_p 和 i_q，共 $q-p$ 次相邻对换，得到排列 $i_1\cdots i_{p+1}\cdots i_qi_p\cdots i_n$. 然后对换 i_{q-1} 和 i_q，再对换 i_{q-2} 和 i_q，等等，最后再对换 i_{p+1} 和 i_q，共 $q-(p+1)$ 次相邻对换，得到 $i_1\cdots i_qi_{p+1}\cdots i_{q-1}i_p\cdots i_n$.

因此 $i_1\cdots i_p\cdots i_q\cdots i_n$ 可以通过 $2(q-p)-1$ 次相邻对换变为 $i_1\cdots i_q\cdots i_p\cdots i_n$. 由于一次相邻对换改变排列的奇偶性，而 $2(q-p)-1$ 是一个奇数，因此排列 $i_1\cdots i_p\cdots i_q\cdots i_n$ 与 $i_1\cdots i_q\cdots i_p\cdots i_n$ 必是一奇一偶的. ■

命题 2.2.2　任意排列 $i_1i_2\cdots i_n$ 都可经过有限次对换变为标准排列 $12\cdots n$，并且对换方式不唯一.

证明　对 n 用归纳法. 当 $n=2$ 时结论显然成立.

假设结论对 $n-1$ 成立. 因为 $i_1i_2\cdots i_n$ 是自然数 $1,2,\ldots,n$ 的一个排列，所以必有某个 $i_p=n$. 对换 i_p,i_n，得到排列 $i_1\cdots i_{p-1}i_ni_{p+1}\cdots i_{n-1}n$. 显然，$i_1\cdots i_{p-1}i_ni_{p+1}\cdots i_{n-1}$ 是自然数 $1,2,\ldots,n-1$ 的一个排列. 由归纳假设，可经有限次对换，将 $i_1\cdots i_{p-1}i_ni_{p+1}\cdots i_{n-1}$ 变为标准排列 $12\cdots(n-1)$. 从而 $i_1i_2\cdots i_n$ 可经有限次对换变为标准排列.

在上面证明中，将 $i_1i_2\cdots i_n$ 变为 $i_1\cdots i_{p-1}i_ni_{p+1}\cdots i_{n-1}n$ 时，既可直接对换 i_p,i_n，也可通过相邻对换实现. 因此对换方式不唯一. ■

命题 2.2.3 设排列 $i_1i_2\cdots i_n$ 经过 s 次对换变为标准排列，则 s 的奇偶性与 $i_1i_2\cdots i_n$ 的奇偶性相同，并且

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} = (-1)^s.$$

证明 由于标准排列 $12\cdots n$ 的逆序数为 0，所以 $12\cdots n$ 是偶排列.

由命题 2.2.1，排列 $i_1i_2\cdots i_n$ 经一次对换改变奇偶性，而排列 $i_1i_2\cdots i_n$ 经 s 次对换变为偶排列 $12\cdots n$，因此，当 $i_1i_2\cdots i_n$ 为奇排列时，s 为奇数；当 $i_1i_2\cdots i_n$ 为偶排列时，s 为偶数. 即 s 的奇偶性与 $i_1i_2\cdots i_n$ 的相同.

由 $\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$ 的定义，即得 $\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} = (-1)^s$. ■

现在来证明 n 阶行列式函数的唯一性.

定理 2.2.5 n 阶行列式函数 $\det(\xi_1, \xi_2, \ldots, \xi_n)$ 是唯一的. 且当 $\xi_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ 时，

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}} \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} a_{1i_1}a_{2i_2}\cdots a_{ni_n},$$

其中 $i_1i_2\cdots i_n$ 是自然数 $1, 2, \ldots, n$ 的排列，上述和号表示对 $i_1i_2\cdots i_n$ 遍历 $1, 2, \ldots, n$ 的所有排列求和.

证明 设 $\varepsilon_j = (0, \ldots, 0, \underset{j}{1}, 0, \ldots, 0)$，记

$$\xi_i = (a_{i1}, a_{i2}, \ldots, a_{in}) = \sum_{j=1}^{n} a_{ij}\varepsilon_j;$$

由于 $\det(\xi_1, \xi_2, \ldots, \xi_n)$ 是 n 重线性的，因此，

$$\begin{aligned} \det(\xi_1, \xi_2, \ldots, \xi_n) &= \det\big(a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + \cdots + a_{1n}\varepsilon_n, \xi_2, \ldots, \xi_n\big) \\ &= \sum_{i_1=1}^{n} a_{1i_1} \det(\varepsilon_{i_1}, \xi_2, \ldots, \xi_n), \end{aligned}$$

对 $\xi_2, \ldots, \xi_n$ 分别作同样考察，得到

$$\det(\xi_1, \xi_2, \ldots, \xi_n) = \sum_{1\leqslant i_1, i_2, \ldots, i_n \leqslant n} a_{1i_1}a_{2i_2}\cdots a_{ni_n} \cdot \det(\varepsilon_{i_1}, \varepsilon_{i_2}, \ldots, \varepsilon_{i_n}). \tag{2.2.2}$$

因为 $\det(\xi_1, \xi_2, \ldots, \xi_n)$ 是反对称的，所以当 $\varepsilon_{i_p} = \varepsilon_{i_q}$ 时，

$$\det(\varepsilon_{i_1}, \ldots, \varepsilon_{i_p}, \ldots, \varepsilon_{i_q}, \ldots, \varepsilon_n) = 0.$$

因此，

$$\det(\xi_1, \xi_2, \ldots, \xi_n) = \sum_{\substack{1\leqslant i_1, i_2, \ldots, i_n \leqslant n \\ i_\ell \neq i_m, \ell \neq m}} a_{1i_1}a_{2i_2}\cdots a_{ni_n} \cdot \det(\varepsilon_{i_1}, \varepsilon_{i_2}, \ldots, \varepsilon_{i_n}).$$

当 $i_1, i_2, \ldots, i_n$ 两两不等时，$i_1i_2\cdots i_n$ 是自然数 $1, 2, \ldots, n$ 的排列. 因此，

$$\det\big(\xi_1, \xi_2, \ldots, \xi_n\big) = \sum_{\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}} a_{1i_1}a_{2i_2}\cdots a_{ni_n} \cdot \det\big(\varepsilon_{i_1}, \varepsilon_{i_2}, \ldots, \varepsilon_{i_n}\big). \tag{2.2.3}$$

因为 $\det(\xi_1,\xi_2,\ldots,\xi_n)$ 是规范的，所以 $\det(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)=1$.

现在计算 $\det(\varepsilon_{i_1},\varepsilon_{i_2},\ldots,\varepsilon_{i_n})$ 的值，其中 $i_1i_2\cdots i_n$ 是 $1,2,\ldots,n$ 的排列.

由命题 2.2.2，排列 $i_1i_2\cdots i_n$ 可经有限次（设为 s 次）对换变为标准排列 $12\cdots n$. 因此，$\det(\varepsilon_{i_1},\varepsilon_{i_2},\ldots,\varepsilon_{i_n})$ 也可经 s 次向量的对换变为 $\det(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$.

由定理 2.2.1，$\det(\varepsilon_{i_1},\varepsilon_{i_2},\ldots,\varepsilon_{i_n})$ 中向量 $\varepsilon_{i_1},\varepsilon_{i_2},\ldots,\varepsilon_{i_n}$ 每对换一次，向量 ε_{i_p} 与 ε_{i_q} 的位置，其值改变一次符号，因此，由命题 2.2.3，

$$\det(\varepsilon_{i_1},\varepsilon_{i_2},\ldots,\varepsilon_{i_n})=(-1)^s\det(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)=\operatorname{sgn}\begin{pmatrix}1&2&\cdots&n\\ i_1&i_2&\cdots&i_n\end{pmatrix}.$$

将它代入式 (2.2.3)，即得定理 2.2.5. ■

定理 2.2.5 给出了 n 阶行列式函数的明显表达式.

给定 n 个向量 $\xi_i=(a_{i1},a_{i2},\ldots,a_{in})$，$i=1,2,\ldots,n$. 由 $\xi_1,\xi_2,\ldots,\xi_n$ 可以确定一个 n 阶方阵 A：

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}.$$

方阵 A 简记为 $A=(a_{ij})$，其中 a_{ij} 表示 A 的第 i 行第 i 列上的元素，$1\leqslant i,j\leqslant n$. n 阶行列式函数 $\det(\xi_1,\xi_2,\ldots,\xi_n)$ 在 $\xi_i=(a_{i1},a_{i2},\ldots,a_{in})$，$i=1,2,\ldots,n$ 处的取值也称为 n 阶方阵 A 的行列式，记为 $\det A$，即

$$\det A=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}.$$

定理 2.2.5 说明，$\det A$ 是 $n!$ 个形如

$$\operatorname{sgn}\begin{pmatrix}1&2&\cdots&n\\ i_1&i_2&\cdots&i_n\end{pmatrix}a_{1i_1}a_{2i_2}\cdots a_{ni_n}$$

的项之和，每一项是方阵 A 中 n 个元素 $a_{1i_1},a_{2i_2},\ldots,a_{ni_n}$ 以及奇偶性符号的连乘积. 这 n 个元素 $a_{1i_1},a_{2i_2},\ldots,a_{ni_n}$ 是依次从方阵 A 的第 $1,2,\ldots,n$ 行各取一个元素构成的. 因为 $i_1i_2\cdots i_n$ 是 $1,2,\ldots,n$ 的一个排列，所以这 n 个元素也是依次从 A 的第 $i_1,i_2,\ldots,i_n$ 列各取一个元素构成的.

由定理 2.2.3，定理 2.2.4 与定理 2.2.5 立即得，

推论 2.2.2

$$\det A=\sum_{i=1}^n a_{i1}A_{i1}.$$

例 2.2.1 证明下面的等式成立：

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{2.2.4}$$

证明 仿照定理 2.2.3 中关于 n 阶行列式的存在性的证明. 记

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

对每个 i，划掉 A 的第 i 行与第 1 列，得到一个 $n-1$ 阶行列式. 记这个 $n-1$ 阶行列式与 $(-1)^{i+1}$ 之乘积为 A_{i1}，$i=1,2,\ldots,n$. 则由定理 2.2.3 的证明得到

$$\det A = a_{11}A_{11}. \qquad \blacksquare$$

给定 n 阶方阵 $A=(a_{ij})$. 将方阵 A 的行排成列，将列排成行，得到的方阵称为方阵 A 的**转置方阵**，记为 A^{T}，即

$$A^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}.$$

记 $\eta_j = (a_{1j}, a_{2j}, \ldots, a_{nj})$，$j=1,2,\ldots,n$. 则 $\det(\eta_1,\eta_2,\ldots,\eta_n) = \det A^{\mathrm{T}}$.

定理 2.2.6 行列式经转置后其值不变，即 $\det A^{\mathrm{T}} = \det A$.

证明 记 $A^{\mathrm{T}} = (b_{ij})$，其中 $b_{ij} = a_{ji}$，$1 \leqslant i,j \leqslant n$. 由定理 2.2.5，

$$\begin{aligned} \det A^{\mathrm{T}} &= \sum_{\left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{smallmatrix}\right)} \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} b_{1i_1}b_{2i_2}\cdots b_{ni_n} \\ &= \sum_{\left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{smallmatrix}\right)} \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} a_{i_1 1}a_{i_2 2}\cdots a_{i_n n}. \end{aligned}$$

对于上式右端的项 $\operatorname{sgn}\left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{smallmatrix}\right) a_{i_1 1}a_{i_2 2}\cdots a_{i_n n}$，由于 $i_1 i_2 \cdots i_n$ 是 $1,2,\ldots,n$ 的排列，所以在 $i_1 i_2\cdots i_n$ 中必有 $i_{j_1}=1, i_{j_2}=2,\ldots,i_{j_n}=n$，其中 $j_1 j_2\cdots j_n$ 是 $1,2,\ldots,n$ 的排列. 于是

$$a_{i_1 1}a_{i_2 2}\cdots a_{i_n n} = a_{i_{j_1} j_1}a_{i_{j_2} j_2}\cdots a_{i_{j_n} j_n} = a_{1j_1}a_{2j_2}\cdots a_{nj_n}. \tag{2.2.5}$$

另一方面，由于 $i_{j_1}i_{j_2}\cdots i_{j_n} = 12\cdots n$，因此当 $i_1 i_2\cdots i_n$ 经 s 次对换便成为标准排列 $12\cdots n$ 时，排列 $12\cdots n$ 也经相同的对换方式变为 $j_1 j_2\cdots j_n$. 所以

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} = \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}.$$

于是，

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} b_{1i_1}b_{2i_2}\cdots b_{ni_n} = \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix} a_{1j_1}a_{2j_2}a_{nj_n}.$$

最后记 $1,2,\ldots,n$ 的所有排列的集合为 P_n. 利用等式 (2.2.5) 可以建立 P_n 到自

身的一个映射 φ:

$$\varphi(i_1i_2\cdots i_n)=j_1j_2\cdots j_n.$$

显然映射 φ 是双射. 因此当 $i_1i_2\cdots i_n$ 遍历 P_n 中所有排列时, $j_1j_2\cdots j_n$ 也遍历 P_n 中所有排列. 这就证明了

$$\det A^{\mathrm{T}}=\sum_{\left(\begin{smallmatrix}1&2&\cdots&n\\ j_1&j_2&\cdots&j_n\end{smallmatrix}\right)}\operatorname{sgn}\begin{pmatrix}1&2&\cdots&n\\ j_1&j_2&\cdots&j_n\end{pmatrix}a_{1j_1}a_{2j_2}\cdots a_{nj_n}.$$

由定理 2.2.5 即得, $\det A^{\mathrm{T}}=\det A$. ■

定理 2.2.6 说明, 在 n 阶行列式中, 行与列的地位是平等的. 也就是说, n 阶行列式的每一个关于行的性质对列也必定成立. 反之亦然. 因此定理 2.2.1、推论 2.2.1 以及定理 2.2.2 中的行换成列, 结论仍成立.

最后对阶行列式的定义作一点说明. 设 $A=(a_{ij})$ 是 n 阶方阵, 则 $\det A=\det(\xi_1,\xi_2,\ldots,\xi_n)$, 其中 $\xi_i=(a_{i1},a_{i2},\ldots,a_{in})\in\mathbb{F}^n$. 由于 $\mathbb{F}^n$ 中的向量是写成行的形式的, 因此 $\mathbb{F}^n$ 也称为 n 维行向量空间. $\mathbb{F}^n$ 中的向量也称为 n 维行向量.

自然想到的是, 数域 $\mathbb{F}$ 上的 n 元数组 $\eta=(a_1,a_2,\ldots,a_n)^{\mathrm{T}}$ 即是 $\mathbb{F}$ 上的 n 维列向量. $\mathbb{F}$ 上的所有 n 维列向量的集合仍记为 $\mathbb{F}^n$. n 维列向量加法规定为

$$(a_1,a_2,\ldots,a_n)^{\mathrm{T}}+(b_1,b_2,\ldots,b_n)^{\mathrm{T}}=(a_1+b_1,a_2+b_2,\ldots,a_n+b_n)^{\mathrm{T}},$$

纯量 $\lambda\in\mathbb{F}$ 与 n 维列向量的乘法规定为

$$\lambda(a_1,a_2,\ldots,a_n)^{\mathrm{T}}=(\lambda a_1,\lambda a_2,\ldots,\lambda a_n).$$

容易验证, 列向量空间中向量的加法以及纯量与向量的乘法满足 n 维行向量空间所具有的性质 (A1)—(A4), (M1), (M2) 与 (D1), (D2). 因此 $\mathbb{F}^n$ 也称为 $\mathbb{F}$ 上的 n 维列向量空间. 同样可以定义 $\mathbb{F}^n$ 上的函数、线性函数、反对称函数以及规范函数等概念. 记

$$\eta_j=(a_{1j},a_{2j},\ldots,a_{nj})^{\mathrm{T}}\in\mathbb{F}^n.$$

则列向量空间 $\mathbb{F}^n$ 上的规范反对称 n 重线性函数即称为 n 阶行列式函数, 也将之记为 $\det(\eta_1,\eta_2,\ldots,\eta_n)$. 将 $\eta_1,\eta_2,\ldots,\eta_n$ 排成一个 n 阶方阵 A,

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix},$$

其中 η_j 排在 A 的第 j 列, 记 $\xi_i=(a_{i1},a_{i2},\ldots,a_{in})$.

不难证明, 列向量空间 $\mathbb{F}$ 上的 n 阶行列式函数 $\det(\eta_1,\eta_2,\ldots,\eta_n)$ 在 η_j 时的取值等于行向量空间 $\mathbb{F}$ 上的 n 阶行列式函数 $\det(\xi_1,\xi_2,\ldots,\xi_n)$ 在 ξ_i 时的取值, 即

$$\det(\eta_1,\eta_2,\ldots,\eta_n)=\det A=\det(\xi_1,\xi_2,\ldots,\xi_n).$$

因此方阵 A 的行列式 $\det A$ 既可以看成 A 的行向量的规范反对称 n 重线性函数, 也可以看成 A 的列向量的规范反对称 n 重线性函数.

习 题 2.2

1. 在自然数 $1,2,\ldots,n$ 的所有排列中哪个排列的逆序数最大?

2. 自然数 $1,2,\ldots,n$ 的任意一个排列 $i_1i_2\cdots i_n$ 的逆序数与正序数之和等于多少?

3. 证明,自然数 $1,2,\ldots,n$ 的任意一个排列都可以经过至多 $n-1$ 次对换变为标准排列 $12\cdots n$.

4. 证明,在自然数 $1,2,\ldots,n$ 的所有排列中,一定存在这样的排列,它不能经过小于 $n-1$ 次对换变为标准排列 $12\cdots n$.

5. 确定以下的奇偶性符号.

(1) $\operatorname{sgn}\begin{pmatrix}1 & 2 & \cdots & (n-1) & n\\ n & (n-1) & \cdots & 2 & 1\end{pmatrix}$;　　(2) $\operatorname{sgn}\begin{pmatrix}1 & 2 & 3 & \cdots & (n-1) & n\\ 1 & n & (n-1) & \cdots & 3 & 1\end{pmatrix}$;

(3) $\operatorname{sgn}\begin{pmatrix}1 & 2 & \cdots & (n-1) & n & (n+1) & (n+2) & \cdots & (2n-1) & 2n\\ 2 & 4 & \cdots & (2n-2) & 2n & 1 & 3 & \cdots & (2n-3) & (2n-1)\end{pmatrix}$;

(4) $\operatorname{sgn}\begin{pmatrix}1 & 2 & \cdots & (n-1) & n & (n+1) & (n+2) & \cdots & (2n-1) & 2n\\ 1 & 3 & \cdots & (2n-3) & (2n-1) & 2 & 4 & \cdots & (2n-2) & 2n\end{pmatrix}$.

6. 确定正整数 i 和 j 的值,使得 7 阶行列式含有以下的项.

(1) $-a_{62}a_{i5}a_{33}a_{j4}a_{46}a_{21}a_{77}$;　　(2) $a_{1i}a_{24}a_{31}a_{47}a_{55}a_{63}a_{7j}$.

7. 写出 4 阶行列式

$$\begin{vmatrix}5x & 1 & 2 & 3\\ x & x & 1 & 2\\ 1 & 2 & x & 3\\ x & 1 & 2 & 2x\end{vmatrix}$$

中含 x^3 与 x^4 的项.

8. 求下列 n 阶行列式的和:

$$\sum_{j_1=1}\sum_{j_2=1}\cdots\sum_{j_n=1}\begin{vmatrix}a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n}\\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n}\end{vmatrix}.$$

9. 计算以下的行列式.

(1) $\begin{vmatrix}1 & 2 & 3 & 4\\ 2 & 3 & 4 & 1\\ 3 & 4 & 1 & 2\\ 4 & 1 & 2 & 3\end{vmatrix}$;　　(2) $\begin{vmatrix}\frac{1}{3} & -\frac{5}{2} & \frac{2}{5} & \frac{3}{2}\\ 3 & -12 & \frac{21}{5} & 15\\ \frac{2}{3} & -\frac{9}{2} & \frac{4}{5} & \frac{5}{2}\\ -\frac{1}{7} & \frac{2}{7} & -\frac{1}{7} & \frac{3}{7}\end{vmatrix}$;

(3) $\begin{vmatrix}0 & a & b & c\\ -a & 0 & d & e\\ -b & -d & 0 & f\\ -c & -e & -f & 0\end{vmatrix}$;　　(4) $\begin{vmatrix}a & b & c & d\\ -b & a & d & -c\\ -c & -d & a & b\\ -d & c & -b & a\end{vmatrix}$.

10. 证明以下等式.

(1)

$$\begin{vmatrix}a & b & c & d\\ a & a+b & a+b+c & a+b+c+d\\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d\\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d\end{vmatrix}=a^4;$$

(2) 设 $a_{ij}=a_{i,j-1}+a_{i-1,j}$, $i,j=1,2,\ldots,n$,则

$$\begin{vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 1 & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = 1.$$

11. 设 $f(\xi_1, \xi_2, \ldots, \xi_k)$ 是 $\mathbb{F}^n$ 上的 k 元函数. 如果对任意整数 $i, j, 1 \leqslant i, j \leqslant n$ 均有

$$f(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_k) = f(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_k),$$

则 $f(\xi_1, \xi_2, \ldots, \xi_k)$ 称为**对称的**. 数域 $\mathbb{F}^n$ 上的规范对称 n 重线性函数称为 n **阶积和式 (Permanent)**, 记为 $\mathrm{Per}(\xi_1, \xi_2, \ldots, \xi_n)$. 记 $\xi_i = (a_{i1}, a_{i2}, \ldots, a_{in}), i = 1, 2, \ldots, n$, 并记 n 阶方阵 A 为

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

则 n 阶积和式 $\mathrm{Per}(\xi_1, \xi_2, \ldots, \xi_n)$ 也记为 $\mathrm{Per}\,A$. 证明,

$$\mathrm{Per}\,A = \sum_{\left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{smallmatrix}\right)} a_{1i_1} a_{2i_2} \cdots a_{ni_n}.$$

§2.3 Laplace 展开定理

给定 n 阶方阵 $A = (a_{ij})_{n\times n}$, 从它的行列式 $\det A$ 中取出第 $i_1, i_2, \ldots, i_p$ 行与第 $j_1, j_2, \ldots, j_p$ 列的交叉位置上的元素, $1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n$, $1 \leqslant j_1 < j_2 < \cdots < j_p \leqslant n$, 并按这些元素在原来行列式 $\det A$ 中的次序排成一个 p 阶行列式, 它称为行列式 $\det A$ 的 $\boldsymbol{p}$ **阶子式**, 记为

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix},$$

上指标 $i_1, i_2, \ldots, i_p$ 是行列式 $\det A$ 的行指标, 下指标 $j_1, j_2, \ldots, j_p$ 是列指标. 即

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} = \begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_p} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_pj_1} & a_{i_pj_2} & \cdots & a_{i_pj_p} \end{vmatrix}.$$

从行列式 $\det A$ 中删去第 $i_1, i_2, \ldots, i_p$ 行和第 $j_1, j_2, \ldots, j_p$ 列上的所有元素, 把余下的元素按它们在原来行列式中的次序排成一个 $n-p$ 阶子式, 它称为 p 阶子式 $A\left(\begin{smallmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{smallmatrix}\right)$ 的**余子式**.

设从行列式 $\det A$ 的 n 个行中删去第 $i_1, i_2, \ldots, i_p$ 行后余下的行是第 $i_{p+1}, \ldots, i_n$ 行, 其中 $1 \leqslant i_{p+1} < \cdots < i_n \leqslant n$, 从行列式 $\det A$ 的 n 个列中删去第 $j_1, j_2, \ldots, j_p$ 列后余下的列是第 $j_{p+1}, \ldots, j_n$ 列, 其中 $1 \leqslant j_{p+1} < \cdots < j_n \leqslant n$, 则 p 阶子式 $A\left(\begin{smallmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{smallmatrix}\right)$ 的余子式就是 $n-p$ 阶子式 $A\left(\begin{smallmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{smallmatrix}\right)$.

特别地, 一阶子式 $A\left(\begin{smallmatrix} i \\ j \end{smallmatrix}\right)$ 是由行列式 $\det A$ 的第 i 行和第 j 列的交叉元素 a_{ij} 构

成的一阶行列式,即元素 a_{ij} 自身,即 $A\begin{pmatrix} i \\ j \end{pmatrix} = a_{ij}$,而它的余子式为

$$A\begin{pmatrix} 1 & 2 & \cdots & (i-1) & (i+1) & \cdots & n \\ 1 & 2 & \cdots & (j-1) & (j+1) & \cdots & n \end{pmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

子式 $A\begin{pmatrix} i \\ j \end{pmatrix}$ 的余子式有时也称为元素 a_{ij} 的余子式.

子式 $A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}$ 的余子式 $A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}$ 与符号 $(-1)^{i_1+i_2+\cdots+i_p+j_1+j_2+\cdots+j_p}$ 的乘积

$$(-1)^{i_1+i_2+\cdots+i_p+j_1+j_2+\cdots+j_p} A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}$$

称为子式 $A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}$ 的**代数余子式.**

特别地,元素 a_{ij} 的代数余子式记为 A_{ij},即

$$A_{ij} = (-1)^{i+j} A\begin{pmatrix} 1 & 2 & \cdots & (i-1) & (i+1) & \cdots & n \\ 1 & 2 & \cdots & (j-1) & (j+1) & \cdots & n \end{pmatrix}.$$

为了下面推导的需要,先证明一个简单事实.

命题 2.3.1 下述行列式等式成立:

$$\begin{vmatrix} a_{11} & \cdots & a_{1p} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} & 0 & \cdots & 0 \\ a_{p+1,1} & \cdots & a_{p+1,p} & a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} & a_{n,p+1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pp} \end{vmatrix} \begin{vmatrix} a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,p+1} & \cdots & a_{nn} \end{vmatrix}. \tag{2.3.1}$$

证明 对 p 用归纳法. 当 $p=1$ 时,式 (2.3.1) 即化为例 2.2.1. 因此式 (2.3.1) 对 $p=1$ 成立. 假设式 (2.3.1) 对 $p-1$ 成立,下面证明式 (2.3.1) 对 p 成立. 记

$$A = \begin{vmatrix} a_{11} & \cdots & a_{1p} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} & 0 & \cdots & 0 \\ a_{p+1,1} & \cdots & a_{p+1,p} & a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} & a_{n,p+1} & \cdots & a_{nn} \end{vmatrix}.$$

由推论 2.2.2,

$$\det A = a_{11}A_{11} + \cdots + a_{p1}A_{p1} + a_{p+1,1}A_{p+1,1} + \cdots + a_{n1}A_{n1}.$$

由归纳假设,当 $1 \leqslant i \leqslant p$ 时,

$$a_{i1}A_{i1} = (-1)^{i+1}a_{i1}\begin{vmatrix} a_{12} & \cdots & a_{1p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,p} & 0 & \cdots & 0 \\ a_{i+1,2} & \cdots & a_{i+1,p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p2} & \cdots & a_{pp} & 0 & \cdots & 0 \\ a_{p+1,2} & \cdots & a_{p+1,p} & a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{np} & a_{n,p+1} & \cdots & a_{nn} \end{vmatrix}$$

$$= (-1)^{i+1}a_{i1}\begin{vmatrix} a_{12} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,p} \\ a_{i+1,2} & \cdots & a_{i+1,p} \\ \vdots & & \vdots \\ a_{p2} & \cdots & a_{pp} \end{vmatrix}\begin{vmatrix} a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,p+1} & \cdots & a_{nn} \end{vmatrix}.$$

而当 $p+1 \leqslant i \leqslant n$ 时,

$$a_{i1}A_{i1} = (-1)^{i+1}a_{i1}\begin{vmatrix} a_{12} & \cdots & a_{1p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p-1,2} & \cdots & a_{p-1,p} & 0 & \cdots & 0 \\ a_{p2} & \cdots & a_{pp} & 0 & \cdots & 0 \\ a_{p+1,2} & \cdots & a_{p+1,p} & a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,p} & a_{i-1,p+1} & \cdots & a_{i-1,n} \\ a_{i+1,2} & \cdots & a_{i+1,p} & a_{i+1,p+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n2} & \cdots & a_{np} & a_{n,p+1} & \cdots & a_{nn} \end{vmatrix}$$

$$= (-1)^{i+1}a_{i1}\begin{vmatrix} a_{12} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p-1,2} & \cdots & a_{p-1,p} \end{vmatrix}\begin{vmatrix} 0 & \cdots & 0 \\ a_{p+1,2} & \cdots & a_{p+1,n} \\ \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n,2} & \cdots & a_{nn} \end{vmatrix} = 0.$$

因此,

$$\begin{aligned}\det A &= a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{p1}A_{p1} \\ &= \left(\sum_{i=1}^{p}(-1)^{i+1}a_{i1}\begin{vmatrix} a_{12} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,p} \\ a_{i+1,2} & \cdots & a_{i+1,p} \\ \vdots & & \vdots \\ a_{p2} & \cdots & a_{pp}\end{vmatrix}\right)\begin{vmatrix} a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,p+1} & \cdots & a_{nn}\end{vmatrix}.\end{aligned}$$

再由推论 2.2.2,

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp}\end{vmatrix} = \sum_{i=1}^{p}(-1)^{i+1}a_{i1}\begin{vmatrix} a_{12} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,p} \\ a_{i+1,2} & \cdots & a_{i+1,p} \\ \vdots & & \vdots \\ a_{p2} & \cdots & a_{pp}\end{vmatrix}.$$

由此即得式 (2.3.1). ■

命题 2.3.2 给定 n 阶方阵 $A = (a_{ij})$. 设 $\varepsilon_j = (0, \ldots, 0, \underset{j}{1}, 0, \ldots, 0)$,

$$\xi_i = \sum_{j=1}^{n} a_{ij}\varepsilon_j,$$

且 $1 \leqslant k_1 < k_2 < \cdots < k_p \leqslant n$, 则

$$\begin{aligned}&\sum_{\left(\begin{smallmatrix} k_1 & k_2 & \cdots & k_p \\ j_1 & j_2 & \cdots & j_p\end{smallmatrix}\right)} a_{1j_1}a_{2j_2}\cdots a_{pj_p}\det(\varepsilon_{j_1}, \varepsilon_{j_2}, \ldots, \varepsilon_{j_p}, \xi_{p+1}, \ldots, \xi_n) \\ &\qquad = A\left(\begin{smallmatrix} 1 & 2 & \cdots & p \\ k_1 & k_2 & \cdots & k_p\end{smallmatrix}\right)(-1)^{1+\cdots+p+k_1+\cdots+k_p}A\left(\begin{smallmatrix} p+1 & \cdots & n \\ k_{p+1} & \cdots & k_n\end{smallmatrix}\right),\end{aligned}$$

其中 $k_1k_2\cdots k_p\cdots k_n$ 是 $1, 2, \ldots, n$ 的排列, 且 $1 \leqslant k_{p+1} < \cdots < k_n \leqslant n$, 而上述求和号表示对 $j_1j_2\cdots j_p$ 遍历 $k_1, k_2, \ldots, k_p$ 的所有排列求和.

证明 先证明

$$\begin{aligned}&\sum_{\left(\begin{smallmatrix} k_1 & k_2 & \cdots & k_p \\ j_1 & j_2 & \cdots & j_p\end{smallmatrix}\right)} a_{1j_1}a_{2j_2}\cdots a_{pj_p}\det(\varepsilon_{j_1}, \varepsilon_{j_2}, \ldots, \varepsilon_{j_p}, \xi_{p+1}, \ldots, \xi_n) \\ &= \left|\begin{matrix} 0 & \cdots & 0 & a_{1k_1} & 0 & \cdots \\ \vdots & & \vdots & \vdots & \vdots & \\ 0 & \cdots & 0 & a_{pk_1} & 0 & \cdots \\ a_{p+1,1} & \cdots & a_{p+1,k_1-1} & a_{p+1,k_1} & a_{p+1,k_1+1} & \cdots \\ \vdots & & \vdots & \vdots & \vdots & \\ a_{n1} & \cdots & a_{n,k_1-1} & a_{n,k_1} & a_{n,k_1+1} & \cdots \end{matrix}\right.\end{aligned}$$

$$\left.\begin{array}{cccccc} \cdots & 0 & a_{1k_p} & 0 & \cdots & 0 \\ & \vdots & \vdots & \vdots & & \vdots \\ \cdots & 0 & a_{pk_p} & 0 & \cdots & 0 \\ \cdots & a_{p+1,k_p-1} & a_{p+1,k_p} & a_{p+1,k_p+1} & \cdots & a_{p+1,n} \\ & \vdots & \vdots & \vdots & & \vdots \\ \cdots & a_{n,k_p-1} & a_{n,k_p} & a_{n,k_p+1} & \cdots & a_{nn} \end{array}\right| . \tag{2.3.2}$$

事实上,记上式右端为 Δ,并且对任意 $1 \leqslant i \leqslant p$,

$$\widetilde{\xi}_i = \sum_{j_i}{}^{*} a_{ij_i}\varepsilon_{j_i},$$

其中上述求和号表示对 j_i 遍历 $k_1, k_2, \ldots, k_p$ 求和.

于是 $\Delta = \det\left(\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots, \widetilde{\xi}_p, \xi_{p+1} \ldots, \xi_n\right)$. 因为行列式是 n 重线性函数,所以

$$\Delta = \sum_{j_1, j_2, \ldots, j_p}{}^{*} a_{1j_1}a_{2j_2}\cdots a_{pj_p}\det\left(\varepsilon_{j_1}, \varepsilon_{j_2}, \ldots, \varepsilon_{j_p}, \xi_{p+1} \ldots, \varepsilon_n\right),$$

其中上述求和号表示对 $j_1, j_2, \ldots, j_p$ 分别遍历 $k_1, k_2, \ldots, k_p$ 求和.

因为行列式是反对称函数. 所以,当 $\varepsilon_{j_1}, \varepsilon_{j_2}, \ldots, \varepsilon_{j_p}$ 的某两个下指标相同时,

$$\det\left(\varepsilon_{j_1}, \varepsilon_{j_2}, \ldots, \varepsilon_{j_p}, \xi_{p+1}, \ldots, \varepsilon_n\right) = 0.$$

因此,$j_1j_2\cdots j_p$ 应是 $k_1, k_2, \ldots, k_p$ 的排列. 于是得到

$$\Delta = \sum_{\left(\begin{smallmatrix} k_1 & k_2 & \cdots & k_p \\ j_1 & j_2 & \cdots & j_p \end{smallmatrix}\right)} a_{1j_1}a_{2j_2}\cdots a_{pj_p}\det\left(\varepsilon_{j1}, \varepsilon_{j2}, \ldots, \varepsilon_{jp}, \xi_{p+1}, \ldots, \xi_n\right).$$

把式 (2.3.2) 右端行列式 Δ 中第 k_1 列依次和第 $k_1 - 1$ 列,第 $k_1 - 2$ 列,$\ldots$,第 1 列对换,于是 Δ 的第 k_1 列经过 $k_1 - 1$ 次对换调到第 1 列,再把第 k_2 列依次和第 $k_2 - 1$ 列,第 $k_2 - 2$ 列,$\ldots$,第 2 列对换,于是 Δ 的第 k_2 列经过 $k_2 - 2$ 次对换调到第 2 列,等等. 最后 Δ 的第 k_p 列经过 $k_p - p$ 次对换调到第 p 列. 由定理 2.2.1 之 (1),

$$\Delta = (-1)^{1+\cdots+p+k_1+\cdots+k_p} \cdot \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{pk_1} & \cdots & a_{pk_p} & 0 & \cdots & 0 \\ a_{p+1,k_1} & \cdots & a_{p+1,k_p} & a_{p+1,k_{p+1}} & \cdots & a_{p+1,k_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,k_1} & \cdots & a_{n,k_p} & a_{n,k_{p+1}} & \cdots & a_{nk_n} \end{vmatrix}.$$

于是由命题 2.3.1 即得命题 2.3.2. ■

现在证明本节的主要定理.

定理 2.3.1 (Laplace 展开定理) 取定行指标 $i_1, i_2, \ldots, i_p$,$1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n$. 遍取行列式 $\det A$ 中第 $i_1, i_2, \ldots, i_p$ 行上的 p 阶子式,并分别乘以相应的代数余子式,其和即为 $\det A$. 具体地说,有

$$\det A = \sum_{1 \leqslant j_1 < j_2 < \cdots < j_p \leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}\left((-1)^{i_1+i_2+\cdots+i_p+j_1+j_2+\cdots+j_p} A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}\right),$$

其中 $i_1i_2\cdots i_p i_{p+1}\cdots i_n$ 和 $j_1j_2\cdots j_p j_{p+1}\cdots j_n$ 都是 $1,2,\dots,n$ 的排列，并且 $1\leqslant i_{p+1}<\cdots<i_n\leqslant n$，$1\leqslant j_{p+1}<\cdots<j_n\leqslant n$.

证明 采用命题 2.3.1 和命题 2.3.1 的记号. 先证明特殊情形：$i_1=1,i_2=2,\dots,i_p=p$. 此时 $i_{p+1}=p+1,\dots,i_n=n$. 因为 n 阶行列式是 n 重线性函数，因此，

$$\det A=\det(\xi_1,\dots,\xi_n)=\sum_{1\leqslant k_1,k_2,\dots,k_p\leqslant n} a_{1k_1}a_{2k_2}\cdots a_{pk_p}\det\big(\varepsilon_{k_1},\dots,\varepsilon_{k_p},\xi_{p+1},\xi_{p+2},\dots,\xi_n\big),$$

因为 n 阶行列式是反对称函数，所以，当下指标 $k_1,k_2,\dots,k_p$ 有两个相同时，

$$\det\big(\varepsilon_{k_1},\varepsilon_{k_2},\dots,\varepsilon_{k_p},\xi_{p+1},\xi_{p+2},\dots,\xi_n\big)=0.$$

因此

$$\det A=\sum_{\substack{1\leqslant k_1,k_2,\dots,k_p\leqslant n\\ k_1,k_2,\dots,k_p\ \text{两两不等}}} a_{1k_1}a_{2k_2}\cdots a_{pk_p}\det(\varepsilon_{k_1},\varepsilon_{k_2},\dots,\varepsilon_{k_p},\xi_{p+1},\xi_{p+2},\dots,\xi_n).$$

不难证明（这里留给读者作为练习）

$$\sum_{\substack{1\leqslant k_1,k_2,\dots,k_p\leqslant n\\ k_1,k_2,\dots,k_p\ \text{两两不等}}} b_{k_1k_2\cdots k_p}=\sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n}\ \sum_{\begin{pmatrix} j_1 & j_2 & \cdots & j_p\\ k_1 & k_2 & \cdots & k_p\end{pmatrix}} b_{k_1k_2\cdots k_p},$$

因此得到

$$\det A=\sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n}\ \sum_{\begin{pmatrix} j_1 & j_2 & \cdots & j_p\\ k_1 & k_2 & \cdots & k_p\end{pmatrix}} a_{1k_1}a_{2k_2}\cdots a_{pk_p}\det\big(\varepsilon_{k_1},\varepsilon_{k_2},\dots,\varepsilon_{k_p},\xi_{k_{p+1}},\xi_{p+1},\dots,\xi_n\big).$$

由命题 2.3.2，上式即为

$$\det A=\sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n} A\begin{pmatrix} 1 & 2 & \cdots & p\\ j_1 & j_2 & \cdots & j_p\end{pmatrix}(-1)^{1+2+\cdots+p+j_1+j_2+\cdots+j_p}A\begin{pmatrix} p+1 & \cdots & n\\ j_{p+1} & \cdots & j_n\end{pmatrix}.$$

这就证明了当 $i_1=1,i_2=2,\dots,i_p=p$ 时，定理 2.3.1成立.

现在转到一般情形. 由于对换行列式的某两行，行列式的值变号，而行列式 $\det A$ 的第 i_1 行可以经过 i_1-1 次相邻两行的对换调到第 1 行，第 i_2 行可以经过 $i-2$ 次相邻两行的对换调到第 2 行，等等，最后，第 i_p 行经过 i_p-p 次相邻两行的对换调到第 p 行. 因此得到，

$$\det A=(-1)^{i_1+i_2+\cdots+i_p+1+2+\cdots+p}\begin{vmatrix} a_{i_11} & a_{i_12} & \cdots & a_{i_1n}\\ \vdots & \vdots & & \vdots\\ a_{i_p1} & a_{i_p2} & \cdots & a_{i_pn}\\ a_{i_p+1,1} & a_{i_p+1,2} & \cdots & a_{i_p+1,n}\\ \vdots & \vdots & & \vdots\\ a_{i_n1} & a_{i_n2} & \cdots & a_{i_nn}\end{vmatrix}.$$

记 $a_{i_\ell k}=b_{\ell k}$，$B=(b_{\ell k})$，并且由上一段结论

$$\det A=(-1)^{i_1+i_2+\cdots+i_p+1+2+\cdots+p}\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \\ b_{p+1,1} & b_{p+1,2} & \cdots & b_{p+1,n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

$$\begin{aligned} &=(-1)^{i_1+i_2+\cdots+i_p+1+2+\cdots+p} \\ &\quad\times\sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n} B\begin{pmatrix} 1 & 2 & \cdots & p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}\left((-1)^{1+2+\cdots+p+j_1+j_2+\cdots+j_p}B\begin{pmatrix} p+1 & \cdots & n \\ j_{p+1} & \cdots & j_n \end{pmatrix}\right). \end{aligned}$$

由于 $a_{i_\ell k}=b_{\ell k}$，所以

$$B\begin{pmatrix} 1 & 2 & \cdots & p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}=A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix},\quad B\begin{pmatrix} p+1 & \cdots & n \\ j_{p+1} & \cdots & j_n \end{pmatrix}=A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}.$$

代入上式即得定理 2.3.1. ■

应当说明的是，在定理 2.3.1 的等式右端和式中共有 C_n^p 个项. 其次，定理 2.3.1 是对行列式 $\det A$ 给定的行 $i_1,i_2,\ldots,i_p$ 讲的，所以，它也称为对行列式 $\det A$ 按照第 $i_1,i_2,\ldots,i_p$ 行作 **Laplalce 展开**. 由于 $\det A=\det A^{\mathrm{T}}$，所以，对行列式 $\det A^{\mathrm{T}}$ 按第 $j_1,j_2,\ldots,j_p$ 行作 Laplace 展开，便得到 $\det A$ 按第 $j_1,j_2,\ldots,j_p$ 列的 Laplace 展开式：

$$\det A=\sum_{1\leqslant i_1<i_2<\cdots<i_p\leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}(-1)^{i_1+i_2+\cdots+i_p+j_1+j_2+\cdots+j_p}A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}.$$

定理 2.3.1 的特殊情形是 $p=1$，即行列式 $\det A$ 按第 i 行（或第 j 列）作 Laplace 展开：

$$\det A=\sum_{j=1}^n a_{ij}A_{ij},\quad 或者\quad \det A=\sum_{i=1}^n a_{ij}A_{ij},$$

其中 A_{ij} 是一阶子式 $A\binom{i}{j}=a_{ij}$ 的代数余子式.

定理 2.3.2　任给 n 阶行列式 $\det A=\det(a_{ij})$，则对任意 $1\leqslant i,k\leqslant n$，

$$\sum_{j=1}^n a_{ij}A_{kj}=\delta_{ik}\det A,\quad 且\quad \sum_{j=1}^n a_{ji}A_{jk}=\delta_{ik}\det A,$$

其中 δ_{ik} 是 Kronecker 符号，即当 $i=k$ 时，$\delta_{ik}=1$，否则 $\delta_{ik}=0$，$1\leqslant i,k\leqslant n$.

证明　当 $i=k$ 时，前一个等式即是行列式 $\det A$ 按第 i 行作 Laplace 展式，等式成立. 而当 $i\neq k$ 时，考虑行列式

$$\Delta=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,1} & \cdots & a_{k-1,n} \\ a_{i1} & \cdots & a_{in} \\ a_{k+1,1} & \cdots & a_{k+1,n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}\ 第\ k\ 行.$$

由于行列式 Δ 的第 i 行和第 k 行都是 $a_{i1}, a_{i2}, \ldots, a_{in}$，并且 $i \neq k$，因此，$\Delta = 0$；另一方面，对行列式 Δ 按第 k 行作 Laplace 展开，得到

$$\Delta = \sum_{j=1}^{n} a_{ij} A_{kj} = 0.$$

这就证明了前一个等式对 $i \neq k$ 也成立.

由于行列式的行和列的地位是平等的，故第二个等式也成立 ■

对行列式的行或列作 Laplace 展开，提供了一种可以将高阶行列式化为低阶行列式的计算方法.

例 2.3.1 计算 5 阶行列式

$$\Delta = \begin{vmatrix} -4 & 1 & 2 & -2 & 1 \\ 0 & 3 & 0 & 1 & -5 \\ 2 & -3 & 1 & -3 & 1 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 4 & 0 & 2 & 5 \end{vmatrix}.$$

解 由于行列式 Δ 的第 1 列和第 3 列上零的个数最多，所以将行列式 Δ 按第 1 列、第 3 列作 Laplace 展开，得到

$$\begin{aligned}
\Delta &= \sum_{1\leqslant i_1<i_2\leqslant 5} A\begin{pmatrix} i_1 & i_2 \\ 1 & 3 \end{pmatrix}\left((-1)^{i_1+i_2+1+3} A\begin{pmatrix} i_3 & i_4 & i_5 \\ 2 & 4 & 5 \end{pmatrix}\right) \\
&= (-1)^{1+2+1+3}\begin{vmatrix} -4 & 2 \\ 0 & 0 \end{vmatrix}\begin{vmatrix} -3 & -3 & 1 \\ -1 & -1 & 0 \\ 4 & 2 & 5 \end{vmatrix} + (-1)^{1+3+1+3}\begin{vmatrix} -4 & 2 \\ 2 & 1 \end{vmatrix}\begin{vmatrix} 3 & 1 & -5 \\ -1 & -1 & 0 \\ 4 & 2 & 5 \end{vmatrix} \\
&\quad + (-1)^{1+4+1+3}\begin{vmatrix} -4 & 2 \\ -1 & 3 \end{vmatrix}\begin{vmatrix} 3 & 1 & 5 \\ -3 & -3 & 1 \\ 4 & 2 & 5 \end{vmatrix} + (-1)^{1+5+1+3}\begin{vmatrix} -4 & 2 \\ 0 & 0 \end{vmatrix}\begin{vmatrix} 3 & 1 & 5 \\ -3 & -3 & 1 \\ -1 & -1 & 0 \end{vmatrix} \\
&\quad + (-1)^{2+3+1+3}\begin{vmatrix} 0 & 0 \\ 2 & 1 \end{vmatrix}\begin{vmatrix} 1 & -2 & 1 \\ -1 & -1 & 0 \\ 4 & 2 & 5 \end{vmatrix} + (-1)^{2+4+1+3}\begin{vmatrix} 0 & 0 \\ -1 & 3 \end{vmatrix}\begin{vmatrix} 1 & -2 & 1 \\ -3 & -3 & 1 \\ 4 & 2 & 5 \end{vmatrix} \\
&\quad + (-1)^{2+5+1+3}\begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}\begin{vmatrix} 1 & -2 & 1 \\ -3 & -3 & 1 \\ -1 & -1 & 0 \end{vmatrix} + (-1)^{3+4+1+3}\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix}\begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -5 \\ 4 & 2 & 5 \end{vmatrix} \\
&\quad + (-1)^{3+5+1+3}\begin{vmatrix} 2 & 1 \\ 0 & 0 \end{vmatrix}\begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -5 \\ -1 & -1 & 0 \end{vmatrix} + (-1)^{4+5+1+3}\begin{vmatrix} -1 & 3 \\ 0 & 0 \end{vmatrix}\begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -5 \\ -3 & -3 & 1 \end{vmatrix}
\end{aligned}$$

$$= \begin{vmatrix} -4 & 2 \\ 2 & 1 \end{vmatrix} \begin{vmatrix} 3 & 1 & -5 \\ -1 & -1 & 0 \\ 4 & 2 & 5 \end{vmatrix} - \begin{vmatrix} -4 & 2 \\ -1 & 3 \end{vmatrix} \begin{vmatrix} 3 & 1 & 5 \\ -3 & -3 & 1 \\ 4 & 2 & 5 \end{vmatrix} - \begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix} \begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -5 \\ 4 & 2 & 5 \end{vmatrix}$$

$$= (-8)(-20) - (-10)(-62) - 7 \cdot 87 = -1069. \quad \blacksquare$$

注意,在把 Δ 按第 1 和第 3 列作 Laplace 展开时共有 $C_5^2 = 10$ 个项. 在展开时务必不要漏掉一些项.

习 题 2.3

1. 利用 Laplace 展开定理计算下列行列式.

(1) $\begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 \end{vmatrix}$; (2) $\begin{vmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix}$;

(3) $\begin{vmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ x_1 & x_2 & 0 & 0 & 0 & x_3 \\ a_1 & b_1 & 1 & 1 & 1 & c_1 \\ a_2 & b_2 & x_1 & x_2 & x_3 & c_2 \\ a_3 & b_3 & x_1^2 & x_2^2 & x_3^2 & c_3 \\ x_1^2 & x_2^2 & 0 & 0 & 0 & x_3^2 \end{vmatrix}$; (4) $\begin{vmatrix} \lambda & 0 & 0 & \cdots & 0 & a \\ x_1 & c & b & \cdots & b & y_1 \\ x_2 & b & c & \cdots & b & y_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & b & b & \cdots & c & y_n \\ a & 0 & 0 & \cdots & 0 & \lambda \end{vmatrix}$.

2. 设 A, B, C 和 D 依次是由下表

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix}$$

中删去第 1, 2, 3 和第 4 列而得到的三阶行列式. 证明,

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 & 0 & 0 \\ a_2 & b_2 & c_2 & d_2 & 0 & 0 \\ a_3 & b_3 & c_3 & d_3 & 0 & 0 \\ 0 & 0 & a_1 & b_1 & c_1 & d_1 \\ 0 & 0 & a_2 & b_2 & c_2 & d_2 \\ 0 & 0 & a_3 & b_3 & c_3 & d_3 \end{vmatrix} = AD - BC.$$

3. 证明 $D = \delta^3\Delta^2$,其中

$$D = \begin{vmatrix} a_1x_1 & b_1x_1 & a_1x_2 & b_1x_2 & a_1x_3 & b_1x_3 \\ a_2x_1 & b_2x_1 & a_2x_2 & b_2x_2 & a_2x_3 & b_2x_3 \\ a_1y_1 & b_1y_1 & a_1y_2 & b_1y_2 & a_1y_3 & b_1y_3 \\ a_2y_1 & b_2y_1 & a_2y_2 & b_2y_2 & a_2y_3 & b_2y_3 \\ a_1z_1 & b_1z_1 & a_1z_2 & b_1z_2 & a_1z_3 & b_1z_3 \\ a_2z_1 & b_2z_1 & a_2z_2 & b_2z_2 & a_2z_3 & b_2z_3 \end{vmatrix}, \quad \delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \Delta = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

4. 设 n 阶方阵 $A = (a_{ij})$ 的元素 a_{ij} 都是变量 x 的可微函数,$1 \leqslant i, j \leqslant n$. 证明,

$$\frac{\mathrm{d}(\det A)}{\mathrm{d}x} = \sum_{1 \leqslant i, j \leqslant n} \frac{\mathrm{d}a_{ij}}{\mathrm{d}x} A_{ij},$$

其中 A_{ij} 是元素 a_{ij} 的代数余子式，$1 \leqslant i, j \leqslant n$.

5. 给定 n 阶方阵 $A = (a_{ij})$. 证明，

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_{21}-a_{11} & a_{22}-a_{12} & \cdots & a_{2n}-a_{1n} \\ a_{31}-a_{11} & a_{32}-a_{12} & \cdots & a_{3n}-a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1}-a_{11} & a_{n2}-a_{12} & \cdots & a_{nn}-a_{1n} \end{vmatrix} = \sum_{1 \leqslant i,j \leqslant n} A_{ij},$$

其中 A_{ij} 是行列式 $\det A$ 中元素 a_{ij} 的代数余子式，$1 \leqslant i, j \leqslant n$.

6. 证明，如果 n 阶行列式 Δ 的某一行(列)的所有元素都是等于 1，则 Δ 的所有元素的代数余子式之和等于 Δ.

7. 用 $x_1, x_2, \ldots, x_{n-1}, 1$ 替换 n 阶行列式

$$\Delta = \begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & 1 \\ a_{21} & \cdots & a_{2,n-1} & 1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & 1 \end{vmatrix}$$

的第 i 行，得到的行列式记为 Δ_i，$i = 1, 2, \ldots, n$. 证明

$$\Delta = \Delta_1 + \Delta_2 + \cdots + \Delta_n.$$

8. 给定 n 阶方阵 $A = (a_{ij})$. 证明

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ y_1 & y_2 & \cdots & y_n & z \end{vmatrix} = z \det A - \sum_{1 \leqslant i,j \leqslant n} A_{ij} x_i y_j,$$

其中 A_{ij} 是行列式 $\det A$ 的元素 a_{ij} 的代数余子式，$1 \leqslant i, j \leqslant n$.

9. 设 $b_{ij} = (a_{i1} + a_{i2} + \cdots + a_{in}) - a_{ij}$，$1 \leqslant i, j \leqslant n$. 证明

$$\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix} = (-1)^{n-1}(n-1) \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

如果 $b_{ij} = (a_{i1} + a_{i2} + \cdots + a_{in}) - k a_{ij}$，其中 $1 \leqslant k \leqslant n$，$1 \leqslant i, j \leqslant n$，结论又怎样?

10. 计算 $2n$ 阶行列式 Δ_{2n}，其中未写出的元素都是零：

$$\Delta_{2n} = \begin{vmatrix} a_1 & & & & & & & b_{2n} \\ & a_2 & & & & & b_{2n-1} & \\ & & \ddots & & & \iddots & & \\ & & & a_n & b_{n+1} & & & \\ & & & b_n & a_{n+1} & & & \\ & & \iddots & & & \ddots & & \\ & b_2 & & & & & a_{2n-1} & \\ b_1 & & & & & & & a_{2n} \end{vmatrix}.$$

§2.4　Cramer 法则

作为 Laplace 展开定理的应用,本节介绍关于线性方程组解的 Cramer 法则.

设 $x_1, x_2, \ldots, x_n$ 是 n 个未知量,它们满足以下 n 个线性方程构成的线性方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n, \end{cases} \tag{2.4.1}$$

其中系数 a_{ij} 和常数项 b_k 都属于数域 $\mathbb{F}$,并且都是已知的,$1 \leqslant i, j, k \leqslant n$. n^2 个系数 a_{ij} 构成的 n 阶方阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

称为线性方程组 (2.4.1) 的**系数矩阵**. 方阵 A 的行列式 $\det A$ 称为线性方程组 (2.4.1) 的**系数行列式**.

如果数域 $\mathbb{F}$ 上有序 n 元数组 $(x_1^0, x_2^0, \ldots, x_n^0)$ 满足线性方程组 (2.4.1) 中所有的方程,即

$$\begin{cases} a_{11}x_1^0 + a_{12}x_2^0 + \cdots + a_{1n}x_n^0 = b_1, \\ a_{21}x_1^0 + a_{22}x_2^0 + \cdots + a_{2n}x_n^0 = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1^0 + a_{n2}x_2^0 + \cdots + a_{nn}x_n^0 = b_n, \end{cases}$$

则有序 n 元数组 $(x_1^0, x_2^0, \ldots, x_n^0)$ 称为线性方程组 (2.4.1) 的**解**.

记线性方程组 (2.4.1) 的系数行列式 $\det A$ 为 Δ. 依次用线性方程组 (2.4.1) 的常数项 $b_1, b_2, \ldots, b_n$ 替换行列式 Δ 的第 j 列上的元素,得到的行列式记为 Δ_j,即

$$\Delta_j = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

于是,有

定理 2.4.1(Cramer 法则) 设线性方程组 (2.4.1) 的系数行列式 $\Delta \neq 0$,则线性方程组 (2.4.1) 具有唯一解

$$\left(\frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \ldots, \frac{\Delta_n}{\Delta}\right).$$

证明　记

$$\beta = \left(b_1, b_2, \ldots, b_n\right)^{\mathrm{T}}, \qquad \eta_j = \left(a_{1j}, a_{2j}, \ldots, a_{nj}\right)^{\mathrm{T}},$$

则线性方程组 (2.4.1) 可以改写为 n 维列向量之和的形式：

$$\sum_{j=1}^{n} x_j \eta_j = \beta.$$

而系数行列式 Δ 为 $\Delta = \det\left(\eta_1, \eta_2, \ldots, \eta_n\right)$，并且

$$\Delta_j = \det\left(\eta_1, \ldots, \eta_{j-1}, \beta, \eta_{j+1}, \ldots, \eta_n\right).$$

首先证明线性方程组 (2.4.1) 的解的存在性．取

$$x_j^0 = \frac{\Delta_j}{\Delta},$$

并把它们代入方程组 (2.4.1) 的第 i 个方程的左端，得到

$$\sum_{j=1}^{n} a_{ij} x_j^0 = \frac{1}{\Delta} \sum_{j=1}^{n} a_{ij} \Delta_j = \frac{1}{\Delta} \sum_{j=1}^{n} a_{ij} \det\left(\eta_1, \ldots, \eta_{j-1}, \beta, \eta_{j+1}, \ldots, \eta_n\right).$$

记

$$\beta = \sum_{k=1}^{n} b_k \varepsilon_k, \qquad \varepsilon_k = \left(0, \ldots, 0, \underset{k}{1}, 0, \ldots, 0\right)^{\mathrm{T}},$$

则由于行列式是线性函数，故

$$\sum_{j=1}^{n} a_{ij} x_j^0 = \frac{1}{\Delta} \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{n} b_k \det\left(\eta_1, \ldots, \eta_{j-1}, \varepsilon_k, \eta_{j+1}, \ldots, \eta_n\right).$$

显然由例 2.2.1，

$$\det\left(\eta_1, \ldots, \eta_{j-1}, \varepsilon_k, \eta_{j+1}, \ldots, \eta_n\right) = (-1)^{k+j} A\begin{pmatrix} 1 & \cdots & (k-1) & (k+1) & \cdots & n \\ 1 & \cdots & (j-1) & (j+1) & \cdots & n \end{pmatrix} = A_{kj}.$$

因此，

$$\sum_{j=1}^{n} a_{ij} x_j^0 = \frac{1}{\Delta} \sum_{k=1}^{n} b_k \left(\sum_{j=1}^{n} a_{ij} A_{kj}\right).$$

由定理 2.3.2，

$$\sum_{j=1}^{n} a_{ij} x_j^0 = \frac{1}{\Delta} \sum_{k=1}^{n} b_k \delta_{ik} \det A = b_i,$$

其中 $i = 1, 2, \ldots, n$．这就证明了

$$\left(\frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \ldots, \frac{\Delta_n}{\Delta}\right).$$

是方程组 (2.4.1) 的解.

其次证明线性方程组 (2.4.1) 的解的唯一性．设方程组 (2.4.1) 另有解 $(x_1^0, x_2^0, \ldots, x_n^0)$，则

$$\sum_{\ell=1}^{n} x_\ell^0 \eta_\ell = \beta.$$

于是，

$$\Delta_j = \det\left(\eta_1, \ldots, \eta_{j-1}, \sum_{\ell=1}^{n} x_\ell^0 \eta_\ell, \eta_{j+1}, \ldots, \eta_n\right).$$

由于 n 阶行列式是它的列向量的 n 重线性函数，所以，

$$\Delta_j = \sum_{\ell=1}^{n} x_\ell^0 \det\left(\eta_1, \ldots, \eta_{j-1}, \eta_\ell, \eta_{j+1}, \ldots, \eta_n\right).$$

由于行列式是列向量的反对称函数,所以,当 $\ell \neq j$ 时,

$$\det(\eta_1, \ldots, \eta_{j-1}, \eta_\ell, \eta_{j+1}, \ldots, \eta_n) = 0.$$

因此,

$$\Delta_j = x_j^0 \det(\eta_1, \ldots, \eta_{j-1}, \eta_j, \eta_{j+1}, \ldots, \eta_n) = x_j^0 \Delta.$$

因为 $\Delta \neq 0$,故

$$x_j^0 = \frac{\Delta_j}{\Delta},$$

其中 $j = 1, 2, \ldots, n$. 这就证明了解的唯一性. ■

对于线性方程组 (2.4.1),如果它的常数项 $b_1, b_2, \ldots, b_n$ 都是零,则方程组 (2.4.1) 称为**齐次线性方程组**,否则称为**非齐次线性方程组**.

显然,齐次线性方程组恒有解

$$(\underbrace{0, 0, 0, \ldots, 0, 0}_{n}),$$

它称为**零解**.

如果齐次方程组具有解 $(x_1^0, x_2^0, \ldots, x_n^0)$,其中 $x_1^0, x_2^0, \ldots, x_n^0$ 不全为零,则它称为**非零解**.

由 Cramer 法则直接得到,

推论 2.4.1 如果齐次方程组具有非零解,则它的系数行列式为零.

上述推论的逆命题也是成立的,即如果齐次方程组的系数行列式为零,则它具有非零解. 这一事实留待下一章证明.

Cramer 法则是 Cramer 于 1750 年发现的. 实际上,最早发现的是 Leibnitz,他比 Cramer 要早 50 年.

应当指出,Cramer 法则并没有完全解决线性方程组求解问题. 例如,当线性方程组的系数行列式为零时,线性方程组是否恒有解? 假定有解,又如何求解,等等. 这些问题是 Cramer 法则无法解决的. 即便在线性方程组的系数行列式不为零时,利用 Cramer 法则求解也是不方便的. 因为这时需要计算 $n+1$ 个 n 阶行列式. 而当 n 很大时,计算 n 阶行列式的工作量是相当大的. 因此,必需寻求实际可行的求解方法. 所有这些,将在下一章中深入讨论.

习　题　2.4

1. 解下列线性方程组.

(1) $\begin{cases} x_1 + x_2 + 2x_3 + 3x_4 = 1, \\ 3x_1 - x_2 - x_3 - 2x_4 = -4, \\ 2x_1 + 3x_2 - x_3 - x_4 = -6, \\ x_1 + 2x_2 + 3x_3 - x_4 = -4; \end{cases}$　　(2) $\begin{cases} 2x_1 - x_2 + 3x_3 + 2x_4 = 4, \\ 3x_1 + 3x_2 + 3x_3 + 2x_4 = 6, \\ 3x_1 - 2x_2 - x_3 + 2x_4 = 6, \\ 3x_1 - x_2 + 3x_3 - x_4 = 6. \end{cases}$

2. 设 a, b, c 和 d 是不全为零的实数. 证明,线性方程组

$$\begin{cases} ax + by + cz + dw = 0, \\ bx - ay + dz - cw = 0, \\ cx - dy - az + bw = 0, \\ dx + cy - bz - aw = 0, \end{cases}$$

具有唯一解,其中 x, y, z 和 w 是未知数.

3. 求下列线性方程组的解.

$$(1) \begin{cases} \mathrm{C}_0^0 \ x_1 + \mathrm{C}_1^0 \ x_2 + \mathrm{C}_2^0 \ x_3 + \cdots + \mathrm{C}_{n-1}^0 \ x_n = \mathrm{C}_n^1, \\ \mathrm{C}_1^1 \ x_1 + \mathrm{C}_2^1 \ x_2 + \mathrm{C}_3^1 \ x_3 + \cdots + \mathrm{C}_n^1 \ x_n = \mathrm{C}_{n+1}^2, \\ \mathrm{C}_2^2 \ x_1 + \mathrm{C}_3^2 \ x_2 + \mathrm{C}_4^2 \ x_3 + \cdots + \mathrm{C}_{n+1}^2 \ x_n = \mathrm{C}_{n+2}^3, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathrm{C}_{n-1}^{n-1} x_1 + \mathrm{C}_n^{n-1} x_2 + \mathrm{C}_{n+1}^{n-1} x_3 + \cdots + \mathrm{C}_{2n-2}^{n-1} x_n = \mathrm{C}_{2n-1}^n. \end{cases}$$

$$(2) \begin{cases} x_1 + x_2 + x_3 + \cdots + x_{n-1} + x_n = 1, \\ x_1 + \ 0 \ + x_3 + \cdots + x_{n-1} + x_n = 2, \\ x_1 + x_2 + \ 0 \ + \cdots + x_{n-1} + x_n = 3, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_1 + x_2 + x_3 + \cdots + x_{n-1} + \ 0 \ = n. \end{cases}$$

4. 设三次多项式 $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ 满足 $f(-1) = 0, f(1) = 4, f(2) = 3, f(3) = 16$. 求多项式 $f(x)$.

5. 设 $y_0, y_1, y_2, \ldots, y_n$ 是数域 $\mathbb{F}$ 中 $n+1$ 个数, $x_0, x_1, x_2, \ldots, x_n$ 是数域 $\mathbb{F}$ 中两两不同的 $n+1$ 个数. 证明,存在唯一一个多项式 $f(x)$, $\deg f(x) \leqslant n$, 使得 $f(x_i) = y_i, i = 0, 1, \ldots, n$.

6. 设线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{2.4.2}$$

的系数行列式 $\Delta \neq 0$. 利用 $n+1$ 阶行列式 $(i = 1, 2, \ldots, n)$

$$\begin{vmatrix} b_i & a_{i1} & a_{i2} & \cdots & a_{in} \\ b_1 & a_{11} & a_{12} & \cdots & a_{1n} \\ b_2 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0,$$

证明

$$\left(\frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \ldots, \frac{\Delta_n}{\Delta} \right)$$

是方程组 (2.4.2) 的解.

§2.5　行列式的计算

给定 n 阶行列式 $\det A$，要计算出它的值，如果采用定理 2.2.5 所给的行列式表达式，就必须先计算它的 $n!$ 个项

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} a_{1i_1}a_{2i_2}\cdots a_{ni_n}.$$

然后相加，才能得到行列式的值. 由数学分析中著名的 Stirling 公式①，随着行列式阶数 n 的增加，行列式表达式中项数 $n!$ 将以指数形式增加. 因此，当阶数 n 很大时，计算量相当大. 所以，在计算行列式的值时，往往不用行列式的表达式. 而是针对所给的具体行列式的特点，利用行列式的基本性质，将行列式的值求出来.

在计算行列式时，把高阶行列式化为低阶行列式，是经常采用的途径. 把高阶行列式化为低阶行列式的一个基本方法是对行列式实施行或列的**初等变换**.

所谓对行列式实施行(列)的初等变换是指

(1) 对换行列式的某两行(列)，其它的行(列)保持不动；

(2) 行列式的某一行(列)的元素遍乘以某个非零的数再加到另一行(列)；

(3) 行列式的某一行(列)的元素遍乘以某个非零的数.

对行列式实施行或列的初等变换的目的是把行列式化成特殊形式的行列式，使之便于计算. 下面通过一些例子来说明计算行列式的基本方法.

一、化为三角形

行列式 $\det A$ 中从西北角到东南角的对角线叫做**主对角线**. 方阵 $A=(a_{ij})$ 的元素 a_{ii} 位于主对角线上；而当 $i<j$ 时，元素 a_{ij} 位于主对角线的上侧；当 $i>j$ 时，元素 a_{ij} 位于主对角线的下侧. 主对角线的一侧的元素全为零的行列式称为**三角形的**. 对于给定的行列式，可以通过行或列的初等变换化为三角形，然后根据例 2.2.1，三角形的行列式等于主对角元素的乘积. 这样便可以求出原行列式的值.

例 2.5.1　求五阶行列式 Δ 的值，其中

$$\Delta=\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 7 & 10 & 13 \\ 3 & 5 & 11 & 16 & 21 \\ 2 & -7 & 7 & 7 & 2 \\ 1 & 4 & 5 & 3 & 10 \end{vmatrix}$$

解　Δ 的第 1 列分别乘以 $-2,-3,-4,-5$，然后分别加到第 2, 3, 4, 5 列，得到

---

① **Stirling 公式**　$n! \sim \sqrt{2n\pi}\left(\dfrac{n}{\mathrm{e}}\right)^n \ (n\to\infty)$.

$$\Delta \xlongequal[\substack{-2(1)+(2),\ -3(1)+(3),\\ -4(1)+(4),\ -5(1)+(5)}]{} \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & 2 & 3 \\ 3 & -1 & 2 & 4 & 6 \\ 2 & -11 & 1 & -1 & -8 \\ 1 & 2 & 2 & -1 & -5 \end{vmatrix}.$$

其中为了便于验算,标明了所作的初等变换,行(列)的初等变换记在等号上(下)方. 右端的行列式记为 $\widetilde{\Delta}$. 对调行列式 $\widetilde{\Delta}$ 的第 2,3 列,得到,

$$\Delta = \widetilde{\Delta} \xlongequal[(2,\ 3)]{} -\begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & -1 & 2 & 3 \\ 3 & 2 & -1 & 4 & 6 \\ 2 & 1 & -11 & -1 & -8 \\ 1 & 2 & 2 & -1 & -5 \end{vmatrix} \xlongequal[\substack{1(2)+(3),\\ -2(1)+(4),\\ -3(2)+(5)}]{} -\begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 \\ 2 & 1 & -10 & -3 & -11 \\ 1 & 2 & 4 & -5 & -1 \end{vmatrix}$$

$$\xlongequal{-11(5)+(4)} -\begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 \\ -9 & -21 & -54 & 52 & 0 \\ 1 & 2 & 4 & -5 & -1 \end{vmatrix} = 52.$$ ■

行列式 $\det A$ 的第 i 行(列)上所有元素之和称为 $\det A$ 的第 i **个行(列)和**. 有些行列式的 n 个行(列)和都相等. 这时就可把行列式的各个行(列)都加到第 1 行,然后再三角化.

例 2.5.2 计算 n 阶行列式 Δ,其中

$$\Delta = \begin{vmatrix} a & a+d & \cdots & \cdots & a+(n-2)d & a+(n-1)d \\ a+d & a+2d & \cdots & \cdots & a+(n-1)d & a \\ \vdots & \vdots & & \iddots & a & a+d \\ \vdots & \vdots & \iddots & \iddots & \iddots & \vdots \\ a+(n-2)d & a+(n-1)d & a & a+d & \cdots & a+(n-3)d \\ a+(n-1)d & a & a+d & a+2d & \cdots & a+(n-2)d \end{vmatrix}.$$

解

$$\Delta \xlongequal[1(j)+(1)]{2\leqslant j\leqslant n} \begin{vmatrix} na+\dfrac{n(n-1)}{2}d & a+d & \cdots & \cdots & a+(n-2)d & a+(n-1)d \\ na+\dfrac{n(n-1)}{2}d & a+2d & \cdots & \cdots & a+(n-1)d & a \\ \vdots & \vdots & & \iddots & a & a+d \\ \vdots & \vdots & \iddots & \iddots & \iddots & \vdots \\ na+\dfrac{n(n-1)}{2}d & a+(n-1)d & a & a+d & \cdots & a+(n-3)d \\ na+\dfrac{n(n-1)}{2}d & a & a+d & a+2d & \cdots & a+(n-2)d \end{vmatrix}$$

$$
=\left(na+\frac{n(n-1)}{2}d\right)\begin{vmatrix}
1 & a+d & \cdots & \cdots & a+(n-2)d & a+(n-1)d\\
1 & a+2d & \cdots & \cdots & a+(n-1)d & a\\
\vdots & \vdots & & \iddots & a & a+d\\
\vdots & \vdots & \iddots & \iddots & \iddots & \vdots\\
1 & a+(n-1)d & a & a+d & \cdots & a+(n-3)d\\
1 & a & a+d & a+2d & \cdots & a+(n-2)d
\end{vmatrix}
$$

$$
\xlongequal[1\leqslant i\leqslant n-1]{-1(i)+(i+1)} n\left(a+\frac{n-1}{2}d\right)\begin{vmatrix}
1 & a+d & \cdots & a+(n-2)d & a+(n-1)d\\
0 & d & \cdots & d & -(n-1)d\\
\vdots & \vdots & \iddots & \iddots & d\\
0 & d & \iddots & \iddots & \vdots\\
0 & -(n-1)d & d & \cdots & d
\end{vmatrix}
$$

$$
=nd^{n-1}\left(a+\frac{n-1}{2}d\right)\begin{vmatrix}
1 & 1 & \cdots & 1 & -(n-1)\\
\vdots & \vdots & \iddots & -(n-1) & 1\\
\vdots & 1 & \iddots & \iddots & \vdots\\
1 & -(n-1) & 1 & \cdots & 1\\
-(n-1) & 1 & 1 & \cdots & 1
\end{vmatrix}
$$

$$
\xlongequal[1(j)+(1]{2\leqslant j\leqslant n-1} nd^{n-1}\left(a+\frac{n-1}{2}d\right)\begin{vmatrix}
-1 & 1 & \cdots & 1 & -(n-1)\\
-1 & \vdots & \iddots & -(n-1) & 1\\
\vdots & 1 & \iddots & \iddots & \vdots\\
-1 & -(n-1) & 1 & \cdots & 1\\
-1 & 1 & 1 & \cdots & 1
\end{vmatrix}
$$

$$
\xlongequal[2\leqslant i\leqslant n-2]{-1(i)+(i+1)} nd^{n-1}\left(a+\frac{n-1}{2}d\right)\begin{vmatrix}
-1 & 1 & \cdots & 1 & 1 & -(n-1)\\
0 & 0 & \cdots & 0 & -n & n\\
\vdots & \vdots & \iddots & \iddots & \iddots & 0\\
0 & 0 & \iddots & \iddots & \iddots & \vdots\\
0 & -n & \iddots & \iddots & & \vdots\\
0 & n & 0 & \cdots & \cdots & 0
\end{vmatrix}
$$

$$
=-nd^{n-1}\left(a+\frac{n-1}{2}d\right)\begin{vmatrix}
0 & \cdots & 0 & -n & n\\
\vdots & \iddots & \iddots & n & 0\\
0 & \iddots & \iddots & \iddots & \vdots\\
-n & \iddots & \iddots & & \vdots\\
n & 0 & \cdots & \cdots & 0
\end{vmatrix}.
$$

最后这个 $n-2$ 阶行列式的第 $n-2$ 行经过 $n-3$ 次相邻行的对换调到第一行，第 $n-3$ 行经过 $n-4$ 次相邻行的对换调至到第二行，等等，得到，

$$\Delta = -(-1)^{1+2+\cdots+(n-4)+(n-3)} nd^{n-1}\left(a+\frac{n-1}{2}d\right)\begin{vmatrix} n & 0 & \cdots & \cdots & 0 \\ -n & n & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -n & n \end{vmatrix}$$

$$= (-1)^{\frac{n(n-1)}{2}}(nd)^{n-1}\left(a+\frac{n-1}{2}d\right). \qquad \blacksquare$$

在例 2.5.2 中,除第一步使用将行列式的各个列都加到第 1 列的技巧外,还采用了下一行减去前一行、逐行相减的技巧. 这也是常用技巧之一,应予重视.

二、建立递推公式

把 n 阶行列式通过行或列的初等变换或其它方法化为同种形式的 $n-1$ 阶行列式,或者阶数更低的行列式,从而建立递推公式. 再利用递推公式求出原来行列式的值.

例 2.5.3 求 n 阶 Vandermonde 行列式 $\Delta_n(x_1,x_2,\ldots,x_n)$,

$$\Delta_n(x_1,x_2,\ldots,x_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}.$$

解

$$\Delta_n(x_1,x_2,\ldots,x_n) \xlongequal[1\leqslant i\leqslant n-1]{-x_1(i)+(i+1)} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & x_2-x_1 & \cdots & x_n-x_1 \\ 0 & (x_2-x_1)x_2 & \cdots & (x_n-x_1)x_n \\ \vdots & \vdots & & \vdots \\ 0 & (x_2-x_1)x_2^{n-2} & \cdots & (x_n-x_1)x_n^{n-2} \end{vmatrix}$$

$$= (x_2-x_1)(x_3-x_1)\cdots(x_n-x_1)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}.$$

由此得到递推公式,

$$\Delta_n(x_1,x_2,\ldots,x_n) = \left(\prod_{i=2}^{n}(x_i-x_1)\right)\Delta_{n-1}(x_2,x_3,\ldots,x_n).$$

因此得到,

$$\Delta_n(x_1,x_2,\ldots,x_n) = \prod_{1\leqslant i<j\leqslant n}(x_j-x_i). \qquad \blacksquare$$

例 2.5.4　求 n 阶三对角行列式 Δ_n，其中未写出的元素都是零，$bc \neq 0$：

$$\Delta_n = \begin{vmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{vmatrix}_n;$$

解　将行列式 Δ_n 按第一行作 Laplace 展开，得到

$$\Delta_n = a\begin{vmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{vmatrix}_{n-1} - b\begin{vmatrix} c & b & & & & \\ 0 & a & b & & & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & & & & c & a \end{vmatrix}_{n-1} = a\Delta_{n-1} - bc\Delta_{n-2}.$$

由此得到递推方程

$$\Delta_n - a\Delta_{n-1} + bc\Delta_{n-2} = 0, \tag{2.5.1}$$

其初始条件为 $\Delta_0 = 1, \Delta_1 = a$．递推方程 (2.5.1) 的特征方程为 $x^2 - ax + bc = 0$，它的根记为 α, β，于是，$\alpha + \beta = a, \alpha\beta = bc$．

由 (2.5.1) 得到，$\Delta_n - (\alpha + \beta)\Delta_{n-1} + \alpha\beta\Delta_{n-2} = 0$．因此，

$$\left.\begin{aligned} \Delta_n - \alpha\Delta_{n-1} &= \beta(\Delta_{n-1} - \alpha\Delta_{n-2}) = \cdots = \beta^{n-1}(\Delta_1 - \alpha\Delta_0) = \beta^n, \\ \Delta_n - \beta\Delta_{n-1} &= \alpha(\Delta_{n-1} - \beta\Delta_{n-2}) = \cdots = \alpha^{n-1}(\Delta_1 - \beta\Delta_0) = \alpha^n, \end{aligned}\right\} \tag{2.5.2}$$

当 $\alpha \neq \beta$ 时，由式 (2.5.2) 得到

$$\Delta_{n-1} = \frac{\beta^n - \alpha^n}{\beta - \alpha} \implies \Delta_n = \frac{\beta^{n+1} - \alpha^{n+1}}{\beta - \alpha}.$$

当 $\alpha = \beta$ 时，由 $2\alpha = a, \alpha^2 = bc \neq 0$ 可知 $\alpha = \dfrac{a}{2} \neq 0$．再由 (2.5.2) 可得

$$\frac{\Delta_n}{\alpha^n} - \frac{\Delta_{n-1}}{\alpha^{n-1}} = 1 \implies \frac{\Delta_n}{\alpha^n} = n + \Delta_0;$$

因此

$$\Delta_n = (n+1)\alpha^n = (n+1)\left(\frac{a}{2}\right)^n. \qquad \blacksquare$$

关于递推方程求解理论，由于不属于本书范围，这里不拟介招．有兴趣的读者可参阅黄国勋和李炯生著《计数》(上海教育出版社，1983 年出版)．

三、Laplace 展开

Laplace 展开是把高阶行列式化为低阶行列式的一种常见方法．

例 2.5.5　求 n 阶行列式 Δ，其中未写出的元素都为零，并且 $k + \ell = n, bc \neq 0$：

$$\Delta = \begin{vmatrix} a & b & & & & & & \\ c & a & b & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & c & a & b & & & \\ & & & 2c & 2a & 2b & & \\ & & & & \ddots & \ddots & \ddots & \\ & & & & & 2c & 2a & 2b \\ & & & & & & 2c & 2a \end{vmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} a\\ a\\ a\\ a \end{matrix}}\right\} k\text{ 行} \\ \left.\vphantom{\begin{matrix} a\\ a\\ a\\ a \end{matrix}}\right\} \ell\text{ 行} \end{array}$$

解 按前 k 行作 Laplace 展开，得到

$$\Delta = (-1)^{1+2+\cdots+k+1+2+\cdots+k} \begin{vmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{vmatrix}_k \begin{vmatrix} 2a & 2b & & & \\ 2c & 2a & 2b & & \\ & \ddots & \ddots & \ddots & \\ & & 2c & 2a & 2b \\ & & & 2c & 2a \end{vmatrix}_\ell$$

$$+ (-1)^{1+\cdots+k+1+\cdots+(k-1)+(k+1)} \begin{vmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a & 0 \\ & & & & c & b \end{vmatrix}_k \begin{vmatrix} 2c & 2b & & & & \\ 0 & 2a & 2b & & & \\ & 2c & 2a & 2b & & \\ & & \ddots & \ddots & \ddots & \\ & & & 2c & 2a & 2b \\ & & & & 2c & 2a \end{vmatrix}_\ell .$$

记 k 阶行列式

$$\Delta_k = \begin{vmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{vmatrix},$$

将 Δ 的展开式的第二项中的第一个因子按第 k 列作 Laplace 展开，第二个因子按第一列作 Laplace 展开，得到

$$\Delta = 2^\ell \Delta_k \Delta_\ell - 2^\ell bc \Delta_{k-1} \Delta_{\ell-1} = 2^\ell \big(\Delta_k \Delta_\ell - bc \Delta_{k-1} \Delta_{\ell-1} \big).$$

由例 2.5.4 得到，

$$\Delta = \begin{cases} \dfrac{a^n}{2^k}(n+1), & \text{当 } a^2 = 4bc \text{ 时；} \\ \dfrac{2^\ell \big((\beta^{k+1} - \alpha^{k+1})(\beta^{\ell+1} - \alpha^{\ell+1}) - bc(\beta^k - \alpha^k)(\beta^\ell - \alpha^\ell) \big)}{(\beta - \alpha)^2}, & \text{其它情形，} \end{cases}$$

其中 α, β 是方程 $x^2 - ax + bc = 0$ 的根. ■

上面介绍的计算行列式的方法，其实质是把高阶行列式化为低阶行列式．尽管是可行的，但不一定最有效．有时，把低阶行列式化为高阶行列式来计算反而更为简便．其方法有

四、加边

由于

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} 1 & b_1 & b_2 & \cdots & b_n \\ 0 & a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

所以可以针对 n 阶行列式 $\det(a_{ij})$ 的特点，添加上1行（列）和1列（行），得到一个与原行列式值相等的 $n+1$ 阶行列式．然后对新的 $n+1$ 阶行列式进行计算，以求得原行列式的值．

例 2.5.6　求 n 阶行列式

$$\Delta_n = \begin{vmatrix} c_1 & a_2 & a_3 & \cdots & a_n \\ a_1 & c_2 & a_3 & \cdots & a_n \\ a_1 & a_2 & c_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_n \\ a_1 & a_2 & \cdots & a_{n-1} & c_n \end{vmatrix}.$$

解　添加一行一列到行列式 Δ_n：

$$\Delta_n = \begin{vmatrix} 1 & -a_1 & -a_2 & -a_3 & \cdots & -a_n \\ 0 & c_1 & a_2 & a_3 & \cdots & a_n \\ 0 & a_1 & c_2 & a_3 & \cdots & a_n \\ 0 & a_1 & a_2 & c_3 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & a_n \\ 0 & a_1 & a_2 & \cdots & a_{n-1} & c_n \end{vmatrix}.$$

再将第1行加到其它各行，得到，

$$\Delta_n = \begin{vmatrix} 1 & -a_1 & -a_2 & -a_3 & \cdots & -a_n \\ 1 & c_1-a_1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & c_2-a_2 & 0 & \cdots & 0 \\ 1 & 0 & 0 & c_3-a_3 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & 0 & \cdots & 0 & c_n-a_n \end{vmatrix}. \tag{2.5.3}$$

如果存在某 $c_i = a_i$，则对行列式 (2.5.3) 的第 $i+1$ 列作 Laplace 展开，得到

$$\Delta_n = (-1)^{2+i}(-a_i)\begin{vmatrix} 1 & c_1-a_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & c_{i-1}-a_{i-1} & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 0 & c_{i+1}-a_{i+1} & \ddots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & c_n-a_n \end{vmatrix}$$

再对上一行列式按第 i 行作 Laplace 展开，得到

$$\Delta_n = (-1)^{2+i+i+1}(-a_i)\begin{vmatrix} c_1-a_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & c_{i-1}-a_{i-1} & \ddots & & \vdots \\ \vdots & & \ddots & c_{i+1}-a_{i+1} & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & c_n-a_n \end{vmatrix}$$

$$= (c_1-a_1)(c_2-a_2)\cdots(c_{i-1}-a_{i-1})a_i(c_{i+1}-a_{i+1})\cdots(c_n-a_n).$$

当 $c_i \neq a_i, i = 1, 2, \ldots, n$ 时，由 (2.5.3) 得到

$$\Delta_n \xlongequal[(c_j-a_j)^{-1}*(j+1)]{1\leqslant j\leqslant n} \prod_{i=1}^{n}(c_i-a_i)\begin{vmatrix} 1 & -\dfrac{a_1}{c_1-a_1} & -\dfrac{a_2}{c_2-a_2} & \cdots & -\dfrac{a_n}{c_n-a_n} \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

$$\xlongequal[-1(j)+(1)]{2\leqslant j\leqslant n+1} \prod_{i=1}^{n}(c_i-a_i)\begin{vmatrix} 1+\sum\limits_{i=1}^{n}\dfrac{a_i}{c_i-a_i} & -\dfrac{a_1}{c_1-a_1} & -\dfrac{a_2}{c_2-a_2} & \cdots & -\dfrac{a_n}{c_n-a_n} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

$$= \prod_{i=1}^{n}(c_i-a_i)\left(1+\sum_{i=1}^{n}\frac{a_i}{c_i-a_i}\right)$$

$$= \prod_{i=1}^{n}(c_i-a_i) + \sum_{j=1}^{n}(c_1-a_1)\cdots(c_{j-1}-a_{j-1})a_j(c_{j+1}-a_{j+1})\cdots(c_n-a_n). \qquad \blacksquare$$

注 在上式中，如果令 $c_i = a_i$，所得到的值与前面 $c_i = a_i$ 的情形所算得的相同. 因此上式包含 $c_i = a_i, i = 1, 2, \ldots, n$ 的情形.

有时行列式的阶数不用升高或降低，也可以计算出行列式. 下面就是其中一种方法.

五、拆行(列)

所谓拆行(列)是指,按照行列式的某一行(列)将行列式拆成两个行列式之和.其根据是定理 2.2.2.

例 2.5.7　计算 n 阶行列式,其中 $x \neq y$:

$$\Delta_n(x, y;\ a_1, a_2, \ldots, a_n) = \begin{vmatrix} a_1 & x & \cdots & x \\ y & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x \\ y & \cdots & y & a_n \end{vmatrix}.$$

解　由定理 2.2.2,

$$\Delta_n(x, y;\ a_1, a_2, \ldots, a_n) = \begin{vmatrix} a_1 & x & \cdots & x & x+0 \\ y & a_2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & x & x+0 \\ y & \cdots & y & a_{n-1} & x+0 \\ y & \cdots & y & y & x+(a_n-x) \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & x & \cdots & x & x \\ y & a_2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & x & x \\ y & \cdots & y & a_{n-1} & x \\ y & \cdots & y & y & x \end{vmatrix} + \begin{vmatrix} a_1 & x & \cdots & x & 0 \\ y & a_2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & x & 0 \\ y & \cdots & y & a_{n-1} & 0 \\ y & \cdots & y & y & (a_n-x) \end{vmatrix}.$$

对上式右端第一个行列式,用 -1 乘以它的第 n 行,然后加到其它各行,对第二个行列式的第 n 列作 Laplace 展开,得到

$$\Delta_n = \begin{vmatrix} a_1-y & x-y & \cdots & x-y & 0 \\ 0 & a_2-y & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & x-y & 0 \\ 0 & \cdots & 0 & a_{n-1}-y & 0 \\ y & \cdots & y & y & x \end{vmatrix} + (a_n-x)\begin{vmatrix} a_1 & x & \cdots & x \\ y & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x \\ y & \cdots & y & a_{n-1} \end{vmatrix}$$

$$= x(a_1-y)(a_2-y)\cdots(a_{n-1}-y) + (a_n-x)\Delta_{n-1}(x, y;\ a_1, a_2, \ldots, a_{n-1}).$$

考虑 Δ_n 的转置行列式,得到,

$$\Delta_n = y(a_1-x)(a_2-x)\cdots(a_{n-1}-x) + (a_n-y)\Delta_{n-1}(x, y;\ a_1, a_2, \ldots, a_{n-1}).$$

于是得到,

$$\Delta_{n-1}(x, y;\ a_1, a_2, \ldots, a_{n-1}) = \frac{y(a_1-x)\cdots(a_{n-1}-x) - x(a_1-y)\cdots(a_{n-1}-y)}{y-x}.$$

因此,

$$\Delta_n(x, y;\ a_1, a_2, \ldots, a_n) = \frac{y(a_1-x)\cdots(a_n-x) - x(a_1-y)\cdots(a_n-y)}{y-x}.$$ ■

六、视行列式为某些元素的多项式

把行列式 $\det A$ 中某个元素 x 看成未定元,其它元素看成常数. 由理 定理 2.2.1, 行列式 $\det A$ 是 $n!$ 个项之和,而每个项是 $\det A$ 的某些元素之积,因此是关于 x 的多项式. 从而 $\det A$ 也是关于 x 的多项式,记为 $f(x)$. 于是计算行列式 $\det A$ 就转化为确定多项式 $f(x)$. 由多项式理论可知,为了确定多项式 $f(x)$,可以先确定它的次数 s,再求出它的 s 个根 $a_1, a_2, \ldots, a_s$,于是,

$$f(x) = c(x-a_1)(x-a_2)\cdots(x-a_s),$$

其中 c 是多项式 $f(x)$ 的首项系数. 最后求出 c,多项式 $f(x)$ 也就定出来了.

例 2.5.8 求 n 阶行列式

$$\Delta_n = \begin{vmatrix} x & a_1 & a_2 & \cdots & a_{n-1} \\ a_1 & x & a_2 & \cdots & a_{n-1} \\ a_1 & a_2 & x & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{n-1} \\ a_1 & a_2 & \cdots & a_{n-1} & x \end{vmatrix}.$$

解 把 Δ_n 中元素 x 看成未定元,元素 $a_1, a_2, \ldots, a_{n-1}$ 看成常数,则 Δ_n 是关于 x 的多项式 $f(x)$. 利用定理 2.2.1,容易看出,多项式 $f(x)$ 的次数为 n,首项系数为 1. 把其它各列加到第一列,得到,

$$\Delta_n = f(x) = \begin{vmatrix} x+a_1+a_2+\cdots+a_{n-1} & a_1 & a_2 & \cdots & a_{n-1} \\ x+a_1+a_2+\cdots+a_{n-1} & x & a_2 & \cdots & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ x+a_1+a_2+\cdots+a_{n-1} & a_2 & a_3 & \cdots & x \end{vmatrix}.$$

由此看出,多项式 $f(x)$ 有一根为 $-(a_1+a_2+\cdots+a_{n-1})$.

把 $x = a_i$ 代入 $f(x)$,得到

$$f(a_i) = \begin{vmatrix} a_i & a_1 & a_2 & \cdots & a_{i-2} & a_{i-1} & a_i & a_{i+1} & \cdots & a_{n-1} \\ a_1 & a_i & a_2 & \cdots & a_{i-2} & a_{i-1} & a_i & a_{i+1} & \cdots & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_{i-1} & a_i & a_i & a_{i+1} & \cdots & a_{n-1} \\ a_1 & a_2 & a_3 & \cdots & a_{i-1} & a_i & a_i & a_{i+1} & \cdots & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_{i-1} & a_i & a_{i+1} & a_{i+2} & \cdots & a_i \end{vmatrix}.$$

其中第 i 行和第 $i+1$ 行相同,所以 $f(a_i) = 0$,即 a_i 是 $f(x)$ 的根. 因此,

$$\Delta_n = f(x) = c\big(x+a_1+a_2+\cdots+a_{n-1}\big)(x-a_1)(x-a_2)\cdots(x-a_{n-1}).$$

其中 c 是待定常数. 因为多项式 $f(x)$ 的首项系数为 1,所以 $c=1$. 于是,

$$\Delta_n = \big(x+a_1+a_2+\cdots+a_{n-1}\big)(x-a_1)(x-a_2)\cdots(x-a_{n-1}). \qquad \blacksquare$$

例 2.5.9　求 n 阶行列式 Δ_n：

$$\Delta_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{k-1} & x_2^{k-1} & \cdots & x_k^{k-1} \\ x_1^{k+1} & x_2^{k+1} & \cdots & x_n^{k+1} \\ \vdots & \vdots & & \vdots \\ x_1^n & x_2^n & \cdots & x_n^n \end{vmatrix}. \qquad \Delta = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & \cdots & x_n & x \\ x_1^2 & x_2^2 & \cdots & x_n^2 & x^2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^{k-1} & x_2^{k-1} & \cdots & x_k^{k-1} & x^{k-1} \\ x_1^k & x_2^k & \cdots & x_k^k & x^k \\ x_1^{k+1} & x_2^{k+1} & \cdots & x_n^{k+1} & x^{k+1} \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^n & x_2^n & \cdots & x_n^n & x^n \end{vmatrix}.$$

解　在 Δ_n 中添加一行一列成为上述右边的行列式 Δ. 此时 Δ 是一个 $n+1$ 阶 Vandermonde 行列式. 由例 2.5.3，

$$\Delta = (x - x_1)(x - x_2)\cdots(x - x_n) \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i). \tag{2.5.4}$$

把 x 看成未定元，则 Δ 是关于 x 的多项式. 比较 Δ_n 和 Δ，对 Δ 的第 $n+1$ 行作 Laplace 展开，可以看出，Δ_n 是多项式 Δ 的 k 次项系数，只是相差符号

$$(-1)^{k+1+n+1}.$$

由 (2.5.4)，Δ 的 k 次项系数为

$$\prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)\Big((-1)^{n-k} \sum_{1 \leqslant i_1 < i_2 < \cdots < i_{n-k} \leqslant n} x_{i_1} x_{i_2} \cdots x_{i_{n-k}}\Big).$$

因此，

$$\Delta_n = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i) \sum_{1 \leqslant i_1 < i_2 < \cdots < i_{n-k} \leqslant n} x_{i_1} x_{i_2} \cdots x_{i_{n-k}}. \qquad \blacksquare$$

最后，应当指出，这里所给的方法只是基本的. 如何巧妙而迅速地求出行列式的值，还要靠自己通过做题悉心领会. 另外还有一些重要的计算方法，如利用行列式的乘法公式，只好留待以后介绍了.

习　题　2.5

1. 计算下列 n 阶行列式.

(1) $\begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & 3 & 3 & \cdots & n-1 & n \\ 1 & 2 & 5 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & n-1 & n \\ 1 & 2 & \cdots & n-2 & 2n-3 & n \\ 1 & 2 & \cdots & \cdots & n-1 & 2n-1 \end{vmatrix}$;　　(2) $\begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & n+1 \\ 3 & 4 & 5 & \cdots & n+1 & n+2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n & n+1 & n+2 & \cdots & 2n & 2n-1 \end{vmatrix}$;

(3) $\begin{vmatrix} a_0 & -1 & 0 & 0 & \cdots & 0 \\ a_1 & x & -1 & 0 & \cdots & 0 \\ a_2 & 0 & x & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ a_{n-2} & 0 & \cdots & 0 & x & -1 \\ a_{n-1} & 0 & \cdots & 0 & 0 & x \end{vmatrix}$;　(4) $\begin{vmatrix} x & a & a & a & \cdots & a \\ -a & x & a & a & \cdots & a \\ -a & -a & x & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & a \\ -a & -a & \cdots & -a & x & a \\ -a & -a & \cdots & -a & -a & x \end{vmatrix}$;

(5) $\begin{vmatrix} a_1 & -a_2 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & -a_3 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_{n-1} & -a_n \\ 1 & 1 & \cdots & 1 & 1 & 1+a_n \end{vmatrix}$;　(6) $\begin{vmatrix} x & y & 0 & \cdots & \cdots & 0 \\ 0 & x & y & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & & & \ddots & x & y \\ y & 0 & \cdots & \cdots & 0 & x \end{vmatrix}$;

(7) $\begin{vmatrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & x & \cdots & x & x \\ 1 & x & 0 & \ddots & & x \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & x & & \ddots & 0 & x \\ 1 & x & x & \cdots & x & 0 \end{vmatrix}$;　(8) $\begin{vmatrix} 1 & a & a^2 & \cdots & a^{n-1} \\ x_{11} & 1 & a & \cdots & a^{n-2} \\ x_{21} & x_{22} & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a \\ x_{n-1,1} & x_{n-1,2} & \cdots & x_{n-1,n-1} & 1 \end{vmatrix}$;

(9) $\begin{vmatrix} a & a+h & a+2h & \cdots & a+(n-2)h & a+(n-1)h \\ -a & a & 0 & 0 & \cdots & 0 \\ 0 & -a & a & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -a & a & 0 \\ 0 & \cdots & 0 & 0 & -a & a \end{vmatrix}$;

(10) $\begin{vmatrix} 1+a_1+x_1 & a_1+x_2 & \cdots & a_1+x_n \\ a_2+x_1 & 1+a_2+x_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1}+x_n \\ a_n+x_1 & \cdots & a_n+x_{n-1} & 1+a_n+x_n \end{vmatrix}$;

(11) $\begin{vmatrix} 2\cos\theta & 1 & & & \\ 1 & 2\cos\theta & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 2\cos\theta & 1 \\ & & & 1 & 2\cos\theta \end{vmatrix}_n$;　(12) $\begin{vmatrix} 1 & \mathrm{C}_{m_1}^1 & \cdots & \mathrm{C}_{m_1}^{n-1} \\ 1 & \mathrm{C}_{m_2}^1 & \cdots & \mathrm{C}_{m_2}^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mathrm{C}_{m_n}^1 & \cdots & \mathrm{C}_{m_n}^{n-1} \end{vmatrix}$;

(13) $\begin{vmatrix} \mathrm{C}_m^k & \mathrm{C}_m^{k+1} & \cdots & \mathrm{C}_m^{k+n-1} \\ \mathrm{C}_{m+1}^k & \mathrm{C}_{m+1}^{k+1} & \cdots & \mathrm{C}_{m+1}^{k+n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{C}_{m+n-1}^k & \mathrm{C}_{m+n-1}^{k+1} & \cdots & \mathrm{C}_{m+n-1}^{k+n-1} \end{vmatrix}$;　(14) $\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & \mathrm{C}_1^1 & 0 & \cdots & 0 & x \\ 1 & \mathrm{C}_2^1 & \mathrm{C}_2^2 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 & x^{n-3} \\ 1 & \mathrm{C}_{n-2}^1 & \mathrm{C}_{n-2}^2 & \cdots & \mathrm{C}_{n-2}^{n-2} & x^{n-2} \\ 1 & \mathrm{C}_{n-1}^1 & \mathrm{C}_{n-1}^2 & \cdots & \mathrm{C}_{n-1}^{n-2} & x^{n-1} \end{vmatrix}$;

(15) $\begin{vmatrix} \dfrac{1}{a_1+b_1} & \dfrac{1}{a_1+b_2} & \cdots & \dfrac{1}{a_1+b_n} \\ \dfrac{1}{a_2+b_1} & \dfrac{1}{a_2+b_2} & \cdots & \dfrac{1}{a_2+b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{1}{a_n+b_1} & \dfrac{1}{a_n+b_2} & \cdots & \dfrac{1}{a_n+b_n} \end{vmatrix}$; (16) $\begin{vmatrix} 1 & \cos\theta_1 & \cos 2\theta_1 & \cdots & \cos(n-1)\theta_1 \\ 1 & \cos\theta_2 & \cos 2\theta_2 & \cdots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \cos\theta_n & \cos 2\theta_n & \cdots & \cos(n-1)\theta_n \end{vmatrix}$;

(17) $\begin{vmatrix} \sin\theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1 \\ \sin\theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sin\theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n \end{vmatrix}$; (18) $\begin{vmatrix} 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_1^n \end{vmatrix}$;

(19) $\begin{vmatrix} x & 1 & & & & \\ -n & x-2 & 2 & & & \\ & -(n-1) & x-4 & 3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -2 & x-2n+2 & n \\ & & & & -1 & x-2n \end{vmatrix}$.

(20) 计算 $2n$ 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_{11} & \cdots & b_{1,n-1} & b_{1n} \\ & a_{22} & \cdots & a_{2n} & b_{21} & \cdots & b_{2,n-1} & b_{2n} \\ & & \ddots & \vdots & \vdots & \iddots & & \\ & & & a_{nn} & b_{n1} & & & \\ & & & c_{1n} & d_{11} & & & \\ & & \iddots & \vdots & \vdots & \ddots & & \\ & c_{n-1,2} & \cdots & c_{n-1,n} & d_{n-1} & \cdots & d_{n-1,n-1} & \\ c_{n1} & c_{n2} & \cdots & c_{nn} & d_{n1} & \cdots & d_{n,n-1} & d_{nn} \end{vmatrix},$$

其中未写出的元素都是零.

2. 设 $a_1, a_2, \ldots, a_n$ 是正整数. 证明, n 阶行列式

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

能被

$$1^{n-1}2^{n-2}\cdots(n-2)^2(n-1)$$

整除.

3. (Minkowski) 设 n 阶方阵 $A=(a_{ij})$ 的元素都是实的, 并且

$$a_{ii}>0, \quad a_{ij}<0, \quad i\neq j, \quad \sum_{j=1}^{n} a_{ij}>0.$$

证明方阵 A 的行列式 $\det A>0$.

4. (Lovy-Desplanques) 设 n 阶方阵 $A=(a_{ij})$ 的元素都是复数, 并且对任意 $i=1,\ldots,n$,

$$a_{ii}>\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|,$$

则方阵 A 称为**主对角占优矩阵**. 证明主对角占优矩阵的行列式不为零.

5. (Burnside) 设 n 阶方阵 $A=(a_{ij})$ 满足 $a_{ji}=-a_{ij}, 1\leqslant i, j\leqslant n$,则方阵 A 称为**斜对称方阵**. 证明,奇数阶斜对称方阵的行列式恒为零,而偶数阶斜对称方阵的行列式是一个完全平方.

6. 把 n 阶行列式

$$\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}$$

展成 λ 的多项式,并用行列式 $\det A$ 的子式表示它的关于 λ 的各次幂的系数,其中 $A=(a_{ij})$.

第三章

矩　阵

- ✍ 矩阵是线性代数研究的基本代数对象. 按照矩阵的观点，线性代数就是研究矩阵在各种意义下的分类问题及其标准形理论.
- ✍ §3.1 引进了矩阵的定义及矩阵的代数运算——矩阵的加法、数与矩阵的乘法、矩阵之间的乘法. 读者或许在这里第一次遇到不是数的一些数学对象可以象数那样进行运算，而且这些运算满足大部分数的运算所满足的运算法则.
- ✍ §3.2 考察矩阵乘积的行列式，得到 Binet-Cauchy 公式，它为行列式的计算又提供了一种有力的方法.
- ✍ 在适当的意义下，矩阵的乘法运算应当有逆运算，§3.3 处理了矩阵的求逆问题.
- ✍ 矩阵按照相抵关系进行分类，是最简单的一种矩阵分类，§3.4 利用初等变换的方法，彻底解决了矩阵在相抵下的分类问题. 读者通过这一章应当了解到，什么是矩阵的分类问题，研究矩阵分类问题时应该注意哪些问题.
- ✍ 利用初等变换，§3.4 顺便给出了求矩阵的逆的一种极其有效的方法.
- ✍ §3.5 通过例子进一步说明如何使用矩阵在相抵下的标准形来解决一些矩阵问题.
- ✍ 在 §3.6 中，利用矩阵在相抵下的标准形，解决了解线性方程组的问题，给出了线性方程组有解的判别准则以及线性方程组的解的结构.
- ✍ 矩阵的逆可以在各种意义下推广，得出各种广义逆的概念，§3.7 介绍了两种广义逆，并说明了它们的一些应用. 这一节是专门为关心计算数学、概率统计以及其它应用数学的读者而设置的. 读者完全可以跳过这一节，而不会影响对以后各章内容的理解.

§3.1　矩阵的代数运算

在介绍行列式理论和关于线性方程组的 Cramer 法则时，已经遇到 n 阶方阵的概念. 本节将给出 $m\times n$ 矩阵的定义及其代数运算规则.

定义 3.1.1　设 $\mathbb{F}$ 是数域，且 $a_{ij}\in\mathbb{F}$，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$. 把 $m\times n$ 个数 a_{ij} 排成一个 m 行 n 列的长方形表：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

它称为**数域 $\mathbb{F}$ 上的 $m\times n$ 矩阵**，简称为**矩阵**，记为 A，或者 $(a_{ij})_{m\times n}$，简记为 (a_{ij}).

矩阵 A 中位于第 i 行,第 j 列位置上的数 a_{ij} 称为**矩阵 A 的 (i,j) 元素**,或 **(i,j) 系数**. 如果 $m=n$,则 A 称为 **n 阶方阵**.

如果数域 $\mathbb{F}$ 取成实数域,则 $\mathbb{F}$ 上的 $m\times n$ 矩阵称为 $m\times n$ **实矩阵**; 如果数域 $\mathbb{F}$ 取成复数域,则 $\mathbb{F}$ 上的 $m\times n$ 矩阵称为 $m\times n$ **复矩阵**. 如无特别说明,本节所说的矩阵都指数域 $\mathbb{F}$ 上的矩阵. 数域 $\mathbb{F}$ 上的所有 $m\times n$ 矩阵的集合记为 $\mathbb{F}^{m\times n}$.

设 $A=(a_{ij}),B=(b_{ij}),A,B\in\mathbb{F}^{m\times n}$. 如果 $a_{ij}=b_{ij}$,则称矩阵 A 和 B **相等**,记为 $A=B$.

定义 3.1.2 设 $A,B\in\mathbb{F}^{m\times n}$,且

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix},\qquad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{m1} & b_{m2} & \cdots & b_{mn}\end{pmatrix}.$$

则矩阵

$$\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n}\\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn}\end{pmatrix}$$

称为**矩阵 A 与 B 之和**,记为 $A+B$.

这就定义了矩阵的加法运算. 由定义可知,如果 $A,B\in\mathbb{F}^{m\times n}$,则 $A+B\in\mathbb{F}^{m\times n}$,即 $\mathbb{F}^{m\times n}$ 对矩阵的加法是封闭的.

容易验证,矩阵的加法运算满足下列公理.

(A1) 加法结合律 对任意 $A,B,C\in\mathbb{F}^{m\times n}$,均有

$$(A+B)+C=A+(B+C);$$

(A2) 加法交换律 对任意 $A,B\in\mathbb{F}^{m\times n}$,均有

$$A+B=B+A;$$

(A3) 零矩阵 所有系数都是零的 $m\times n$ 矩阵称为**零矩阵**. 记为 $(0)_{m\times n}$,简记为 0. 对任意 $A\in\mathbb{F}^{m\times n}$,均有

$$A+0=A=0+A;$$

(A4) 负矩阵 对任意 $A=(a_{ij})\in\mathbb{F}^{m\times n}$,存在唯一的 $B=(b_{ij})\in\mathbb{F}^{m\times n}$,使得

$$A+B=0=B+A.$$

容易看出,$B=(b_{ij})$ 满足 $b_{ij}=-a_{ij}$. 矩阵 B 称为**矩阵 A 的负矩阵**,记为 $-A$.

利用负矩阵概念,可以引进矩阵的减法运算.

定义 3.1.3 设 $A=(a_{ij}),B=(b_{ij})\in\mathbb{F}^{m\times n}$. 则矩阵 $A+(-B)$ 称为矩阵 A 与 B 之差,记为 $A-B$.

显然,

$$A-B=\begin{pmatrix} a_{11}-b_{11} & a_{12}-b_{12} & \cdots & a_{1n}-b_{1n} \\ a_{21}-b_{21} & a_{22}-b_{22} & \cdots & a_{2n}-b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}-b_{m1} & a_{m2}-b_{m2} & \cdots & a_{mn}-b_{mn} \end{pmatrix}.$$

矩阵的减法运算和通常数的减法运算具有相同的性质.

定义 3.1.4 设 $\lambda \in \mathbb{F}, A=(a_{ij}) \in \mathbb{F}^{m\times n}$. 则矩阵

$$\begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix}$$

称为**纯量 $\boldsymbol{\lambda}$ 与矩阵 $\boldsymbol{A}$ 的乘积**,记为 λA.

由定义可以看出,如果 $\lambda \in \mathbb{F}, A \in \mathbb{F}^{m\times n}$,则 $\lambda A \in \mathbb{F}^{m\times n}$,即 $\mathbb{F}^{m\times n}$ 对纯量与矩阵的乘法运算是封闭的.

容易验证,纯量与矩阵的乘法满足以下的公理.

(M1) 对任意 $\lambda, \mu \in \mathbb{F}, A \in \mathbb{F}^{m\times n}$,均有

$$(\lambda\mu)A=\lambda(\mu A);$$

(M2) 对任意 $A \in \mathbb{F}^{m\times n}$,均有

$$1\cdot A=A;$$

(D1) 对任意 $\lambda, \mu \in \mathbb{F}, A \in \mathbb{F}^{m\times n}$,均有

$$(\lambda+\mu)A=\lambda A+\mu A;$$

(D2) 对任意 $\lambda \in \mathbb{F}, A, B \in \mathbb{F}^{m\times n}$,均有

$$\lambda(A+B)=\lambda A+\lambda B.$$

除此之外,还具有性质:

设 $\lambda \in \mathbb{F}, A \in \mathbb{F}^{m\times n}$,则当且仅当 $\lambda=0$ 或 $A=0$ 时,$\lambda A=0$.

和数域 $\mathbb{F}$ 上的 n 维行向量空间 $\mathbb{F}$ 相比较,可以看出,$\mathbb{F}$ 中矩阵的加法运算以及纯量与矩阵的乘法运算所满足的公理和 $\mathbb{F}$ 中向量的加法以及纯量与向量的乘法所满足的公理完全相同,因此有理由把 $\mathbb{F}$ 称为数域 $\mathbb{F}$ 上的 $\boldsymbol{m\times n}$ **维向量空间**.

其次,如果 $m=1, A \in \mathbb{F}^{1\times n}$,则 A 为 1 行 n 列的矩阵,即是数域 $\mathbb{F}$ 上的 n 维向量. 而 $1\times n$ 矩阵的加法即是 n 维向量的加法,纯量与 $1\times n$ 矩阵的乘法即是纯量与 n 维向量的乘法. 因此,$\mathbb{F}^{m\times n}$ 是 $\mathbb{F}^n$ 的自然推广.

和向量的不同点是,矩阵之间可以引进乘法.

定义 3.1.5 设 $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times p}$,并且

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},\qquad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}.$$

则 $m\times p$ 矩阵

$$\begin{pmatrix} \sum\limits_{k=1}^{n} a_{1k}b_{k1} & \sum\limits_{k=1}^{n} a_{1k}b_{k2} & \cdots & \sum\limits_{k=1}^{n} a_{1k}b_{kp} \\ \sum\limits_{k=1}^{n} a_{2k}b_{k1} & \sum\limits_{k=1}^{n} a_{2k}b_{k2} & \cdots & \sum\limits_{k=1}^{n} a_{2k}b_{kp} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{k=1}^{n} a_{mk}b_{k1} & \sum\limits_{k=1}^{n} a_{mk}b_{k2} & \cdots & \sum\limits_{k=1}^{n} a_{mk}b_{kp} \end{pmatrix}$$

称为**矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积**,记为 AB.

由矩阵乘积的定义,如果记 $AB=C=(c_{ij})$,则

$$c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}.$$

应当指出,给定矩阵 A 与 B,只有当矩阵 A 的列数等于矩阵 B 的行数,矩阵 A 与 B 的乘积 AB 才有意义. 矩阵的乘法具有以下性质.

(1) 乘法结合律 对任意的 $A\in\mathbb{F}^{m\times n}$,$B\in\mathbb{F}^{n\times p}$,$C\in\mathbb{F}^{p\times q}$,均有

$$(AB)C=A(BC).$$

证明 记 $A=(a_{ij})$,$B=(b_{ij})$,$C=(c_{ij})$,$AB=D=(d_{ij})$,$BC=E=(e_{ij})$. 容易看出,AB 和 BC 都有意义,而且 DC 和 AE 也都有意义,并且后二者都是 $m\times q$ 矩阵.

由矩阵乘法定义,

$$d_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}.$$

记 $(AB)C=H=(h_{ij})$,则

$$h_{ij}=\sum_{\ell=1}^{p} d_{i\ell}c_{\ell j}=\sum_{\ell=1}^{p}\Big(\sum_{k=1}^{n} a_{ik}b_{k\ell}\Big)c_{\ell j}=\sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{ik}b_{k\ell}c_{\ell j}.$$

另一方面,由矩阵乘法定义,

$$e_{ij}=\sum_{\ell=1}^{p} b_{i\ell}c_{\ell j}$$

因此,若记 $A(BC)=AE=G=(g_{ij})$,则

$$g_{ij}=\sum_{k=1}^{n} a_{ik}e_{kj}=\sum_{k=1}^{n} a_{ik}\Big(\sum_{\ell=1}^{p} b_{k\ell}e_{\ell j}\Big)=\sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{ik}b_{k\ell}c_{\ell j}.$$

所以

$$h_{ij}=g_{ij}.$$

由矩阵相等的定义,$H=G$,即 $(AB)C=A(BC)$. ∎

利用矩阵乘法的上述性质,可以定义 $k \geqslant 2$ 个矩阵的乘积.

设 $A_1, A_2, \ldots, A_k$ 依次是 $m \times p_1, p_1 \times p_2, \ldots, p_{k-1} \times n$ 矩阵,定义

$$A_1A_2A_3 = (A_1A_2)A_3.$$

设 $A_1A_2\cdots A_{k-1}$ 已有定义,则定义

$$A_1A_2\cdots A_k = (A_1A_2\cdots A_{k-1})A_k.$$

利用上述性质和归纳法可以证明,对于给定的 k 个矩阵 $A_1, A_2, \ldots, A_k$,它们的大小分别为 $m \times p_1, p_1 \times p_2, \ldots, p_{k-1} \times n$,则只要保持这 k 个矩阵的顺序,而不论其间如何添加括号,其乘积恒为 $A_1A_2\cdots A_k$. 例如,当 $k = 4$ 时,恒有

$$\begin{aligned} A_1A_2A_3A_4 &= (A_1A_2A_3)A_4 = (A_1A_2)(A_3A_4) \\ &= A_1(A_2A_3)A_4 = A_1(A_2A_3A_4) = A_1A_2(A_3A_4), \end{aligned}$$

等等. 特别,如果 $A_1, A_2, \ldots, A_k$ 是同一个 n 阶方阵 A,则

$$\underbrace{AAA\cdots AA}_{n\text{ 个}}$$

有意义,并记为 A^k,它称为方阵的 ***k* 次幂**.

(2) 单位方阵　主对角元素都是 1 而其它元素都是零的 n 阶方阵称为 ***n* 阶单位方阵**,记为 $I_{(n)}$,即

$$I_{(n)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}_{n \times n}.$$

在不引起混淆时可简记为 I. 容易看出,对任意 $A \in \mathbb{F}^{m \times n}$,均有

$$I_{(m)}A = A = AI_{(n)}.$$

(3) 纯量乘法　设 $\lambda \in \mathbb{F}, A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times p}$,则

$$(\lambda A)B = A(\lambda B) = \lambda(AB).$$

(4) 加乘分配律　设 $A, B \in \mathbb{F}^{m \times n}, C \in \mathbb{F}^{n \times p}, D \in \mathbb{F}^{q \times m}$,则

$$(A + B)C = AC + BC, \qquad D(A + B) = DA + DB.$$

应当指出,矩阵的乘法与通常数的乘法存在着重大区别. 对于数的乘法,交换律是成立的,即对任意 $a, b \in \mathbb{F}$,均有 $ab = ba$. 但是,对于矩阵的乘法,交换律一般并不成立. 事实上,设 $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times p}$,且 $m \neq p$. 则虽然 AB 有意义,但 BA 却无定义,因此谈不上 AB 和 BA 是否相等. 即使 $m = p$,即 AB 和 BA 都有意义,但 AB 为 m 阶方阵,BA 为 n 阶方阵,当 $m \neq n$ 时,显然,$AB \neq BA$; 而当 $m = n$ 时,仍有可能 $AB \neq BA$. 例如,

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

则
$$AB=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad BA=\begin{pmatrix}0&0\\1&0\end{pmatrix}.$$

这个例子还表明,两个非零矩阵的乘积可以是零矩阵. 这和两个非零的数的乘积一定不等于零也不相同. 由此还可以看出,如果 A,B 和 C 分别是 $m\times n,m\times n$ 和 $m\times p$ 矩阵,且 $AC=BC,C\neq 0$,也不一定有 $A=B$. 也就是说,对矩阵的乘法,消去律并不成立. 这些运算规则的不同点在进行矩阵乘法运算时是必须注意的.

设 A 是数域 $\mathbb{F}$ 上的 n 阶方阵,
$$f(\lambda)=a_0+a_1\lambda+\cdots+a_n\lambda^n$$
是数域 $\mathbb{F}$ 上的关于 λ 的多项式. 记
$$f(A)=\sum_{i=0}^{n}a_iA^i,$$
其中约定 $A^0=I_{(n)}$,则 $f(A)$ 称为**方阵 A 的多项式.**

容易验证,如果 $f(\lambda),g(\lambda),p(\lambda)$ 与 $q(\lambda)$ 是数域 $\mathbb{F}$ 上的关于 λ 的多项式,且 $f(\lambda)+g(\lambda)=p(\lambda),f(\lambda)g(\lambda)=q(\lambda)$,则
$$f(A)+g(A)=p(A);\qquad f(A)g(A)=g(A)f(A)=q(A).$$

定义 3.1.6 设 $A=(a_{ij})$ 是 $m\times n$ 矩阵,把 A 的行变成列,列变成行,得到的 $n\times m$ 矩阵称为矩阵 A 的**转置矩阵**,记为 A^{T},即
$$A^{\mathrm{T}}=\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{m1}\\a_{12}&a_{22}&\cdots&a_{m2}\\\vdots&\vdots&&\vdots\\a_{1n}&a_{2n}&\cdots&a_{mn}\end{pmatrix}.$$

容易验证,矩阵的转置具有以下性质.

(1) 对任意 $A,B\in\mathbb{F}^{m\times n}$,均有
$$(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}};$$

(2) 对任意 $\lambda\in\mathbb{F},A\in\mathbb{F}^{m\times n}$,均有
$$(\lambda A)^{\mathrm{T}}=\lambda A^{\mathrm{T}};$$

(3) 对任意 $A\in\mathbb{F}^{m\times n},B\in\mathbb{F}^{n\times p}$,均有
$$(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}},$$

一般地说,设 $A_i\in\mathbb{F}^{m_i\times m_{i+1}},i=1,2,\ldots,k$,则
$$(A_1A_2\cdots A_k)^{\mathrm{T}}=A_k^{\mathrm{T}}\cdots A_2^{\mathrm{T}}A_1^{\mathrm{T}};$$

(4) 设 $A\in\mathbb{F}^{m\times n}$,则
$$(A^{\mathrm{T}})^{\mathrm{T}}=A.$$

对于复矩阵 $A=(a_{ij})\in\mathbb{C}^{m\times n}$,可以定义矩阵 A 的**共轭矩阵** $\overline{A}$ 如下:

$$\overline{A}=\begin{pmatrix}\overline{a}_{11} & \overline{a}_{12} & \cdots & \overline{a}_{1n}\\ \overline{a}_{21} & \overline{a}_{22} & \cdots & \overline{a}_{2n}\\ \vdots & \vdots & & \vdots\\ \overline{a}_{m1} & \overline{a}_{m2} & \cdots & \overline{a}_{mn}\end{pmatrix},$$

或简记为 $\overline{A}=\left(\overline{a}_{ij}\right)$,其中 $\left(\overline{a}_{ij}\right)$ 是复数 a_{ij} 的共轭复数.

关于矩阵的共轭,以下性质成立.

(1) 设 $A,B\in\mathbb{C}^{m\times n}$,则 $\overline{(A+B)}=\overline{A}+\overline{B}$;

(2) 设 $\lambda\in\mathbb{C}$,$A\in C^{m\times n}$,则 $\overline{(\lambda A)}=\overline{\lambda}\overline{A}$;

(3) 设 $A\in\mathbb{C}^{m\times n}$,$B\in\mathbb{C}^{m\times n}$,则 $\overline{AB}=\overline{A}\overline{B}$;

(4) 设 $A\in\mathbb{C}^{m\times n}$,则 $\overline{A^{\mathrm{T}}}=(\overline{A})^{\mathrm{T}}$,并记为 $\overline{A}^{\mathrm{T}}$.

最后介绍进行矩阵代数运算的一个重要方法,即把矩阵**分块**.

给定矩阵 $A=(a_{ij})\in\mathbb{F}^{m\times n}$. 所谓对矩阵 A 进行分块是指,设想用一些水平线和竖直线把矩阵 A 的元素分隔成若干个长方形小块. 例如,设想在矩阵 A 的第 $m_1+\cdots+m_i$ 行与第 $m_1+\cdots+m_i+1$ 行之间有一条水平线,$i=1,2,\ldots,p$,$m_1+m_2+\cdots+m_p=m$,在 A 的第 $n_1+\cdots+n_j$ 列与第 $n_1+\cdots+n_j+1$ 列之间有一条竖直线,$j=1,2,\ldots,q$,$n_1+n_2+\cdots+n_q=n$. 于是,矩阵 A 便分成 $p\times q$ 个块,

$$A=\left(\begin{array}{c:c:c:c}A_{11} & A_{12} & \cdots & A_{1q}\\ \hdashline A_{21} & A_{22} & \cdots & A_{2q}\\ \hdashline \vdots & \vdots & & \vdots\\ \hdashline A_{p1} & A_{p2} & \cdots & A_{pq}\end{array}\right)$$

其中

$$A_{ij}=\begin{pmatrix}a_{s_{i-1}+1,t_{j-1}+1} & a_{s_{i-1}+1,t_{j-1}+2} & \cdots & a_{s_{i-1}+1,t_j}\\ a_{s_{i-1}+2,t_{j-1}+1} & a_{s_{i-1}+2,t_{j-1}+2} & \cdots & a_{s_{i-1}+2,t_j}\\ \vdots & \vdots & & \vdots\\ a_{s_i,t_{j-1}+1} & a_{s_i,t_{j-1}+2} & \cdots & a_{s_i,t_j}\end{pmatrix},$$

这里 $s_i=m_1+m_2+\cdots+m_i$,$t_j=n_1+n_2+\cdots+n_j$.

$m_i\times n_j$ 矩阵 A_{ij} 称为**矩阵 A 的子矩阵**,而矩阵 A 简记为 $A=\left(A_{ij}\right)$.

在进行矩阵代数运算时,可以把分块矩阵中的子矩阵当成矩阵的元素,然后进行运算. 设 $A,B\in\mathbb{F}^{m\times n}$,并且 A,B 的分块方式相同,即设

$$A=\left(A_{ij}\right),\qquad B=\left(B_{ij}\right),$$

其中 A_{ij} 和 B_{ij} 都是 $m_i\times n_j$ 矩阵,$1\leqslant i\leqslant p$,$1\leqslant j\leqslant q$. 则显然

$$A+B=\begin{pmatrix}A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1q}+B_{1q}\\ A_{21}+B_{21} & A_{22}+B_{22} & \cdots & A_{2q}+B_{2q}\\ \vdots & \vdots & & \vdots\\ A_{p1}+B_{p1} & A_{p2}+B_{p2} & \cdots & A_{pq}+B_{pq}\end{pmatrix}.$$

设 $\lambda \in \mathbb{F}, A = \left(A_{ij}\right) \in \mathbb{F}^{m\times n}$，其中 A_{ij} 是 $m_i \times n_j$ 子矩阵. 容易验证

$$\lambda A = \left(\lambda A_{ij}\right).$$

设 $A = \left(A_{ij}\right) \in \mathbb{F}^{m\times n}, B = \left(B_{ij}\right) \in \mathbb{F}^{n\times p}$，其中 A_{ij} 是 A 的 $m_i \times n_j$ 子矩阵，$1 \leqslant i \leqslant r$，$1 \leqslant j \leqslant s$，$B_{k\ell}$ 是 B 的 $n_k \times p_\ell$ 子矩阵，$1 \leqslant k \leqslant s$，$1 \leqslant \ell \leqslant t$，即 A 的子矩阵 A_{ij} 的列数等于 B 的子矩阵 B_{ij} 的行数. 记

$$C_{i\ell} = \sum_{k=1}^{s} A_{ik} B_{k\ell}.$$

则 $AB = \left(C_{i\ell}\right)$. 换句话说，在计算矩阵 A 和 B 的乘积 AB 时，可以把矩阵 A 和 B 分块，使得矩阵 A 的列的分法与矩阵 B 的行的分法相同，那么，把子矩阵 A_{ik} 和 $B_{k\ell}$ 当成通常的系数，然后按照矩阵乘法运算相乘，其结果便是矩阵 A 与 B 的乘积.

现在来证明这一事实. 记 $C = \left(C_{i\ell}\right)$，显然 C 是 $m \times p$ 矩阵，并且具有 $r \times t$ 个块（子矩阵），矩阵 C 的 (i, j) 系数 c_{ij} 一定在某个块 $C_{k\ell}$ 中. 于是，

$$m_1 + \cdots + m_{k-1} < i \leqslant m_1 + \cdots + m_{k-1} + m_k,$$
$$p_1 + \cdots + p_{\ell-1} < j \leqslant p_1 + \cdots + p_{\ell-1} + p_\ell.$$

因此，必有 i' 和 j'，$1 \leqslant i' \leqslant m_k$，$1 \leqslant j' \leqslant p_\ell$，使得

$$i = m_1 + \cdots + m_{k-1} + i', \qquad j = p_1 + \cdots + p_{\ell-1} + j'.$$

这表明，c_{ij} 是子矩阵 $C_{k\ell}$ 的 (i', j') 系数. 由于

$$C_{k\ell} = \sum_{\mu=1}^{s} A_{k\mu} B_{\mu\ell},$$

因此，它的 (i', j') 系数应是 $A_{k\mu}B_{\mu\ell}$，$\mu = 1, 2, \ldots, s$ 的 (i', j') 系数的和. 但是，

$$A_{k\mu} = \begin{pmatrix} a_{u_{k-1}+1, v_{\mu-1}+1} & a_{u_{k-1}+1, v_{\mu-1}+2} & \cdots & a_{u_{k-1}+1, v_\mu} \\ a_{u_{k-1}+2, v_{\mu-1}+1} & a_{u_{k-1}+2, v_{\mu-1}+2} & \cdots & a_{u_{k-1}+2, v_\mu} \\ \vdots & \vdots & & \vdots \\ a_{u_k, v_{\mu-1}+1} & a_{u_k, v_{\mu-1}+2} & \cdots & a_{u_k, v_\mu} \end{pmatrix},$$

$$B_{\mu\ell} = \begin{pmatrix} b_{v_{\mu-1}+1, w_{\ell-1}+1} & b_{v_{\mu-1}+1, w_{\ell-1}+2} & \cdots & b_{v_{\mu-1}+1, w_\ell} \\ b_{v_{\mu-1}+2, w_{\ell-1}+1} & b_{v_{\mu-1}+2, w_{\ell-1}+2} & \cdots & b_{v_{\mu-1}+2, w_\ell} \\ \vdots & \vdots & & \vdots \\ b_{v_\mu, w_{\ell-1}+1} & b_{v_\mu, w_{\ell-1}+2} & \cdots & b_{v_\mu, w_\ell} \end{pmatrix},$$

其中 $u_k = m_1 + \cdots + m_k$，$v_\mu = n_1 + \cdots + n_\mu$，$w_\ell = p_1 + \cdots + p_\ell$. 所以 $A_{k\mu}B_{\mu\ell}$，$\mu = 1, 2, \ldots, s$ 的 (i', j') 系数是

$$\sum_{\lambda=1}^{n_\mu} a_{u_{k-1}+i', v_{\mu-1}+\lambda} b_{v_{\mu-1}+\lambda, w_{\ell-1}+j'}.$$

于是

$$\begin{aligned} c_{ij} &= \sum_{\mu=1}^{s} \sum_{\lambda=1}^{n_\mu} a_{u_{k-1}+i', v_{\mu-1}+\lambda} b_{v_{\mu-1}+\lambda, w_{\ell-1}+j'} \\ &= \sum_{\lambda=1}^{n} a_{m_1+\cdots+m_{k-1}+i', \lambda} b_{\lambda, p_1+\cdots+p_{\ell-1}+j'} = \sum_{\lambda=1}^{n} a_{i\lambda} b_{\lambda j}. \end{aligned}$$

这就证明了 $C = AB$.

对矩阵的转置和共轭,容易验证,如果 $m \times n$ 矩阵 $A = (a_{ij})$,其中 A_{ij} 为 $m_i \times n_j$ 子矩阵,$1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q$,则

$$A^{\mathrm{T}} = \left(A_{ji}^{\mathrm{T}}\right)^{\mathrm{T}}, \qquad \overline{A} = \left(\overline{A}_{ij}\right).$$

最后给出若干特殊类型的矩阵. 如果矩阵分块后具有以下形式,

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ 0 & A_{22} & \cdots & A_{2q} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{qq} \end{pmatrix},$$

即 A 的主对角线以下的块都是零子矩阵,则 A 称为**准上三角**的. 当 $A_{11}, A_{22}, \ldots, A_{qq}$ 为一阶子方阵时,A 称为**上三角**的.

显然,当 $A_{11}, A_{22}, \ldots, A_{qq}$ 都是方阵时,A 也是方阵,并且由 Laplace 展开定理,

$$\det A = \det A_{11} \det A_{22} \cdots \det A_{qq}.$$

如果矩阵 A 具有如下分块形式:

$$A = \begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ A_{21} & A_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ A_{p1} & A_{p2} & \cdots & A_{pp} \end{pmatrix},$$

即主对角线以上的块都是零子矩阵,则 A 称为**准下三角**的. 若 $A_{11}, A_{22}, \ldots, A_{pp}$ 都是一阶子方阵,则 A 称为**下三角**的.

同样,当 $A_{11}, A_{22}, \ldots, A_{pp}$ 为子方阵时,

$$\det A = \det A_{11} \det A_{22} \cdots \det AA_{pp}.$$

如果矩阵 A 既是准上三角的,又是准下三角的,则 A 称为**准对角**的,此时记

$$A = \operatorname{diag}\left(A_{11}, A_{22}, \ldots, A_{pp}\right).$$

特别,当 $A_{11}, A_{22}, \ldots, A_{pp}$ 为一阶子方阵,即 A 的非主对角元都是零时,A 称为**对角**的,记为

$$A = \operatorname{diag}\left(a_{11}, a_{22}, \ldots, a_{nn}\right).$$

习　题　3.1

1. 求下列矩阵的乘积.

(1) $\begin{pmatrix} a & b & c \\ c & b & a \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & a & c \\ 1 & b & b \\ 1 & c & a \end{pmatrix}$;

(2) $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 3 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 & 1 \\ -1 & 1 & 0 \\ 1 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 3 & 1 & 1 \end{pmatrix}$;

(3) $\begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}^2$；

(4) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}^n$；

(5) $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}^n$；

(6) $\begin{pmatrix} 1 & \frac{\lambda}{n} \\ -\frac{\lambda}{n} & 1 \end{pmatrix}^n$；

(7) $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^n$；

(8) $\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}_{n\times n}^n$；

(9) $\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}_{n\times n}^n$；

(10) $\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2} \end{pmatrix}$，其中 $\omega^n = 1$.

2. 计算 $AB - BA$.

(1) $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$，　$B = \begin{pmatrix} 4 & 1 & 1 \\ -4 & 2 & 0 \\ 1 & 2 & 1 \end{pmatrix}$；

(2) $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix}$，　$B = \begin{pmatrix} 3 & 1 & -2 \\ 3 & -2 & 4 \\ -3 & 5 & -1 \end{pmatrix}$.

3. 求出所有和方阵 A 可交换的方阵.

(1) $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 2 \end{pmatrix}$；

(2) $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$.

4. 证明，如果方阵 A 和 B **可交换**，即 $AB = BA$，则

(1) $(A+B)^2 = A^2 + 2AB + B^2$；

(2) $A^2 - B^2 = (A+B)(A-B)$；

(3) $(A+B)^n = A^n + C_n^1 A^{n-1}B + \cdots + C_n^{n-1}AB^{n-1} + B^n$.

5. 设 n 阶对角方阵 A 的对角元两两不等. 证明，n 阶方阵 B 与 A 可变换的充分必要条件是，B 为对角方阵. 能否推广到准对角情形，即

$$A = \mathrm{diag}\left(\lambda_1 I_{(k_1)}, \lambda_2 I_{(k_2)}, \ldots, \lambda_\ell I_{(k_\ell)}\right),$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_\ell$ 两两不等，$k_1 + k_2 + \cdots + k_\ell = n$.

6. 证明下述结论：

(1) 上三角方阵的乘积仍是上三角的；

(2) 如果 $A_1, A_2, \ldots, A_n$ 是 n 个 n 阶上三角方阵，且方阵 A_i 的 (i, i) 系数为零，则

$$A_1 A_2 \cdots A_n = 0.$$

7. 设 k 是正整数. 适合 $A^k = 0$ 的方阵 A 称为**幂零**的. 使 $A^k = 0$ 成立的最小正整数 k 称为方阵 A 的**幂零指数**. 证明，上三角方阵 A 为幂零的充分必要条件是，A 的对角元素全为零，并且 A 的幂零指数不超过它的阶数.

8. 求出所有 2 阶幂零方阵.

9. 如果方阵 A 适合 $A^2 = I$，则 A 称为**对合**的. 求出所有 2 阶对合方阵.

10. 如果方阵 A 适合 $A^2 = A$，则 A 称为**幂等**的. 求出所有 2 阶幂等方阵.

11. 证明，如果 A 是 n 阶对合方阵，则 $B = \dfrac{1}{2}\left(A + I_{(n)}\right)$ 是幂等的. 反之，如果 B 是 n 阶幂等方阵，则 $A = 2B - I_{(n)}$ 是对合的.

12. 适合 $A^{\mathrm{T}} = A$ 的方阵 A 称为**对称**的. 证明，

(1) 设 $B \in \mathbb{F}^{m \times n}$，则 BB^{T} 是对称的；

(2) n 阶对称方阵 A 和 B 的乘积 AB 为对称方阵的充分必要条件是，方阵 A 和 B 可交换.

13. 适合 $A^{\mathrm{T}} = -A$ 的方阵 A 称为**斜对称**的. 证明，

(1) 斜对称方阵 A 和 B 的乘积 AB 为斜对称的充分必要条件是 $AB = -BA$；

(2) 斜对称方阵 A 和 B 的乘积 AB 为对称的充分必要条件是方阵 A 和 B 可交换.

14. 证明，任意二个方阵 A 都可以唯一地分解为对称方阵 S 和斜对称方阵 K 之和.

15. 适合 $\overline{A}^{\mathrm{T}} = A$ 的复方阵称为 **Hermite 方阵**. 证明，对任意 $A \in \mathbb{C}^{m \times n}$，$A\overline{A}^{\mathrm{T}}$ 是 Hermite 方阵.

16. 证明，对角块都是方的准上三角方阵为幂零的充分必要条件是，它的每一个对角块都是幂零的.

17. 对角元素都相等的对角方阵称为**纯量方阵**. 证明，和所有 n 阶方阵都可交换的方阵一定是纯量方阵.

18. 设 n 阶方阵 A 的每一行都恰有 2 个元素为 1，而其它元素为零，J 是元素全为 1 的 n 阶方阵. 求出所有适合 $A^2 + 2A = 2J$ 的 n 阶方阵 A.

§3.2 Binet-Cauchy 公式

本节考虑定义在 $\mathbb{F}^{n \times n}$ 上而取值在 $\mathbb{F}$ 的两个重要函数.

定义 3.2.1 n 阶方阵 $A = (a_{ij})$ 对角线上的的所有元素之和称为**方阵 A 的迹 (trace)**，记为 $\operatorname{Tr} A$，即

$$\operatorname{Tr} A = a_{11} + a_{22} + \cdots + a_{nn}.$$

容易验证，函数 $\operatorname{Tr} A$ 具有以下性质.

(1) $\operatorname{Tr}(A)$ 是方阵 A 的线性函数. 即，设 $\lambda \in \mathbb{F}$，$A, B \in \mathbb{F}^{n \times n}$，则

$$\operatorname{Tr}(A + B) = \operatorname{Tr} A + \operatorname{Tr} B, \qquad \operatorname{Tr}(\lambda A) = \lambda \operatorname{Tr} A;$$

(2) 设 $A, B \in \mathbb{F}^{n \times n}$，则

$$\operatorname{Tr}(AB) = \operatorname{Tr}(BA);$$

(3) 设 $A \in \mathbb{F}^{n\times n}$,则
$$\operatorname{Tr} A^{\mathrm{T}} = \operatorname{Tr} A;$$

(4) 设 $A \in \mathbb{C}^{n\times n}$. 如果 $\operatorname{Tr}(A\overline{A}^{\mathrm{T}}) = 0$,则 $A = 0$.

定义在 $\mathbb{F}^{n\times n}$ 上而取值在 $\mathbb{F}$ 的另一重要函数就是方阵的行列式 $\det A$. 和方阵 A 的迹 $\operatorname{Tr}(A)$ 不同,方阵 A 的行列式 $\det A$ 不是方阵 A 的线性函数. 即一般地说,对 $\lambda \in \mathbb{F}, A, B \in \mathbb{F}^{n\times n}$,等式 $\det(A+B) = \det A + \det B$ 和 $\det(\lambda A) = \lambda \det A$ 并不成立. 但是,方阵的行列式却是方阵的可乘函数.

具体地说,有以下的定理.

定理 3.2.1 设 $A, B \in \mathbb{F}^{n\times n}$,则
$$\det(AB) = \det A \det B.$$

证明 记 $A = (a_{ij}), B = (b_{ij})$. 则
$$\det(AB) = \begin{vmatrix} \sum\limits_{j_1=1}^{n} a_{1j_1}b_{j_11} & \sum\limits_{j_2=1}^{n} a_{1j_2}b_{j_22} & \cdots & \sum\limits_{j_n=1}^{n} a_{1j_n}b_{j_nn} \\ \sum\limits_{j_1=1}^{n} a_{2j_1}b_{j_11} & \sum\limits_{j_2=1}^{n} a_{2j_2}b_{j_22} & \cdots & \sum\limits_{j_n=1}^{n} a_{2j_n}b_{j_nn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{j_1=1}^{n} a_{nj_1}b_{j_11} & \sum\limits_{j_2=1}^{n} a_{nj_2}b_{j_22} & \cdots & \sum\limits_{j_n=1}^{n} a_{nj_n}b_{j_nn} \end{vmatrix}.$$
依次将 $\det(AB)$ 按第 $1, 2, \ldots, n$ 列拆开,得到
$$\det(AB) = \sum_{1\leqslant j_1, j_2, \ldots, j_n \leqslant n} b_{j_11}b_{j_22}\cdots b_{j_nn} \begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n} \end{vmatrix}.$$
当上式右端被加项的行列式的列指标 $j_1, j_2, \ldots, j_n$ 有两个相等时,该行列式应为零. 去掉右端和式中行列式为零的那些项,则和式中余下的项所相应的指标 $j_1j_2\cdots j_n$ 应是 $1, 2, \ldots, n$ 的一个排列. 因此,
$$\det(AB) = \sum_{\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}} b_{j_11}b_{j_22}\cdots b_{j_nn} \begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n} \end{vmatrix}.$$
由于对换行列式的两列,行列式的值变号. 而排列 $j_1j_2\cdots j_n$ 可以经过若干次对换变为标准排列 $12\cdots n$,其对换次数和排列 $j_1j_2\cdots j_n$ 的逆序数奇偶性相同. 按照排列 $j_1j_2\cdots j_n$ 变为标准排列 $12\cdots n$ 的对换方式对换行列式

$$\begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n} \end{vmatrix} \tag{3.2.1}$$

的列,则行列式 (3.2.1) 变为

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

因此

$$\begin{vmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n} \end{vmatrix} = \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

于是

$$\det(AB) = \sum_{\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}} \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix} b_{j_1 1} b_{j_2 2} \cdots b_{j_n n} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

上式右端中的因子

$$\sum_{\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}} \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix} b_{j_1 1} b_{j_2 2} \cdots b_{j_n n}$$

即为行列式 $\det B^{\mathrm{T}} = \det B$. ■

定理 3.2.1 可以推广到 A 和 B 为非方矩阵的情形.

定理 3.2.2(**Binet-Cauchy**) 设 $A \in \mathbb{F}^{p\times q}, B \in \mathbb{F}^{q\times p}$,则

$$\det(AB) = \begin{cases} 0, & \text{当 } q < p; \\ \det A \det B, & \text{当 } q = p; \\ \displaystyle\sum_{1\leqslant j_1<\cdots<j_p\leqslant q} A\begin{pmatrix} 1 & 2 & \cdots & p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} B\begin{pmatrix} j_1 & j_2 & \cdots & j_p \\ 1 & 2 & \cdots & p \end{pmatrix}, & \text{当 } q > p. \end{cases}$$

Binet-Cauchy 公式的证明和定理 3.2.1 基本相同,请读者自己完成.

利用 Binet-Cauchy 公式,可以证明.

定理 3.2.3 设 $A \in \mathbb{F}^{p\times q}, B \in \mathbb{F}^{q\times s}$,并且记 $C = AB$. 则矩阵 C 的 r 阶子式

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = \begin{cases} 0, & \text{当 } r > q; \\ \displaystyle\sum_{1\leqslant k_1<\cdots<k_r\leqslant q} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}, & \text{当 } r \leqslant q. \end{cases}$$

证明 设 $A = (a_{ij}), B = (b_{ij}), C = AB = (c_{ij})$,其中对于 $1 \leqslant i \leqslant p, 1 \leqslant j \leqslant s$,

$$c_{ij}=\sum_{k=1}^{q}a_{ik}b_{kj}.$$

则

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}=\det(A_1B_1)=\begin{vmatrix} \sum\limits_{k_1=1}^{q}a_{i_1k_1}b_{k_1j_1} & \sum\limits_{k_2=1}^{q}a_{i_1k_2}b_{k_2j_2} & \cdots & \sum\limits_{k_r=1}^{q}a_{i_1k_r}b_{k_1j_r} \\ \sum\limits_{k_1=1}^{q}a_{i_2k_1}b_{k_1j_1} & \sum\limits_{k_2=1}^{q}a_{i_2k_2}b_{k_2j_2} & \cdots & \sum\limits_{k_r=1}^{q}a_{i_2k_r}b_{k_1j_r} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{k_1=1}^{q}a_{i_rk_1}b_{k_1j_1} & \sum\limits_{k_2=1}^{q}a_{i_rk_2}b_{k_2j_2} & \cdots & \sum\limits_{k_r=1}^{q}a_{i_rk_r}b_{k_1j_r} \end{vmatrix},$$

对矩阵 A 的子矩阵

$$A_1=\begin{pmatrix} a_{i_11} & a_{i_12} & \cdots & a_{i_1q} \\ a_{i_21} & a_{i_22} & \cdots & a_{i_2q} \\ \vdots & \vdots & & \vdots \\ a_{i_r1} & a_{i_r2} & \cdots & a_{i_rq} \end{pmatrix}$$

和矩阵 B 的子矩阵

$$B_1=\begin{pmatrix} b_{1j_1} & b_{1j_2} & \cdots & b_{1j_r} \\ b_{2j_1} & b_{2j_2} & \cdots & b_{2j_r} \\ \vdots & \vdots & & \vdots \\ b_{qj_1} & b_{qj_2} & \cdots & b_{qj_r} \end{pmatrix}$$

应用 Binet-Cauchy 公式即得定理 3.2.3. ■

现在举例说明上述定理的应用.

例 3.2.1 n 阶方阵

$$A=\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}$$

称为**循环矩阵**,求 n 阶循环矩阵 A 的行列式.

解 设 ω 是 n 次本原单位根,即 $\omega^i\neq1, 1\leqslant i\leqslant n-1$,但 $\omega^n=1$. 令

$$f(x)=a_0+a_1x+\cdots+a_nx^n.$$

取 n 阶矩阵 B 为

$$B=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{pmatrix}.$$

容易验证,

$$AB = B\,\mathrm{diag}\big(f(1), f(\omega), f(\omega^2), \ldots, f(\omega^{n-1})\big).$$

上式两端各取行列式,则由定理 3.2.1,

$$\det A \det B = \Big(\prod_{i=0}^{n-1} f(\omega^i)\Big) \det B.$$

显然,Vandermonde 行列式

$$\det B = \prod_{0\leqslant i<j\leqslant n-1} (\omega_j - \omega_i) \neq 0.$$

因此,

$$\det A = \prod_{i=0}^{n-1} f(\omega^i).$$

■

例 3.2.2　设 $A = (a_{ij})$ 为 n 阶方阵,其中 a_{ij} 是整数 i 和 j 的最大公因数 (i, j). 证明,

$$\det A = \varphi(1)\varphi(2)\cdots\varphi(n),$$

其中 $\varphi(n)$ 是自然数 n 的 Euler 函数.

证明　取 n 阶方阵 $B = (b_{ij})$,其中

$$b_{ij} = \begin{cases} 1, & \text{当 } j \mid i; \\ 0, & \text{当 } j \nmid i. \end{cases}$$

容易看出,方阵 B 是对角元素全为 1 的下三角方阵,因此,$\det B = 1$.

再取 n 阶方阵 $C = (c_{ij})$,其中

$$c_{ij} = \varphi(i) b_{ji}.$$

显然,方阵 C 是对角元索依次为 $\varphi(1), \varphi(2), \ldots, \varphi(n)$ 的上三角方阵,因此,

$$\det C = \varphi(1)\varphi(2)\cdots\varphi(n).$$

记 $BC = D = (d_{ij})$. 则

$$d_{ij} = \sum_{k=1}^{n} b_{ik} c_{kj} = \sum_{k=1}^{n} \varphi(k) b_{ik} b_{jk}$$

因为当 $k \nmid i$ 或 $k \nmid j$ 时,$b_{ik} b_{jk} = 0$,所以,

$$d_{ij} = \sum_{k \mid i,\ k \mid j} \varphi(k) = \sum_{k \mid (i,j)} \varphi(k).$$

由 Euler 函数 $\varphi(n)$ 的性质可知,$d_{ij} = (i, j)$. 因此,$BC = D = A$. 由定理 3.2.1,

$$\det A = \det B \det C = \varphi(1)\varphi(2)\cdots\varphi(n).$$

■

例 3.2.3　利用行列式证明 **Cauchy-Буняковский 不等式**,即当 a_i 和 b_i 为实数时,

$$\big(a_1 b_1 + a_2 b_2 + \cdots + a_n b_n\big)^2 \leqslant \big(a_1 + a_2 + \cdots + a_n\big)^2 \big(b_1 + b_2 + \cdots + b_n\big)^2,$$

其中当且仅当

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} = \cdots = \frac{a_n}{b_n}$$

时等式成立.

证明 取 $2\times n$ 实矩阵

$$A=\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix},\quad 则\quad AA^{\mathrm{T}}=\begin{pmatrix} \sum\limits_{i=1}^n a_i^2 & \sum\limits_{i=1}^n a_ib_i \\ \sum\limits_{i=1}^n a_ib_i & \sum\limits_{i=1}^n b_i^2 \end{pmatrix},$$

因此

$$\det AA^{\mathrm{T}}=\Big(\sum_{i=1}^n a_i^2\Big)\Big(\sum_{i=1}^n b_i^2\Big)-\Big(\sum_{i=1}^n a_ib_i\Big)^2.$$

另一方面,由 Binet-Cauchy 公式,

$$\det AA^{\mathrm{T}}=\sum_{1\leqslant i<j\leqslant n} A\begin{pmatrix} 1 & 2 \\ i & j \end{pmatrix}A^{\mathrm{T}}\begin{pmatrix} i & j \\ 1 & 2 \end{pmatrix}=\sum_{1\leqslant i<j\leqslant n}\Big(A\begin{pmatrix} 1 & 2 \\ i & j \end{pmatrix}\Big)^2\geqslant 0,$$

所以

$$\Big(\sum_{i=1}^n a_i^2\Big)\Big(\sum_{i=1}^n b_i^2\Big)-\Big(\sum_{i=1}^n a_ib_i\Big)^2\geqslant 0.$$

因此即得 Cauchy-Буняковский 不等式. 而且当且仅当对每一对 i,j,

$$A\begin{pmatrix} 1 & 2 \\ i & j \end{pmatrix}=a_ib_j-a_jb_i=0$$

时等式成立,即当且仅当

$$\frac{a_1}{b_1}=\frac{a_2}{b_2}=\cdots=\frac{a_n}{b_n}$$

时,等式成立. ■

习 题 3.2

1. 利用矩阵乘法,直接求出下列行列式 $\det AB$,并用 Binet-Cauchy 公式加以验证.

(1) $A=\begin{pmatrix} 2 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix}$, $B=\begin{pmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{pmatrix}$; (2) $A=\begin{pmatrix} 3 & 4 \\ 2 & 1 \\ 1 & 1 \\ 2 & 3 \end{pmatrix}$, $B=\begin{pmatrix} 1 & 5 & 2 & 1 \\ 2 & 1 & 3 & 4 \end{pmatrix}$.

2. 利用等式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & 0 & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \ddots & \vdots & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

证明定理 3.2.1.

3. 计算下列行列式.

(1) 设等幂和 $s_k=x_1^k+x_2^k+\cdots+x_n^k$, $k=1,2,\ldots$, 计算

$$\begin{vmatrix} s_0 & s_1 & \cdots & s_{n-1} & 1 \\ s_1 & s_2 & \cdots & s_n & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-2} & x^{n-1} \\ s_n & s_{n+1} & \cdots & s_{2n-1} & x^n \end{vmatrix};$$

(2) 记 $\varphi_i(x) = a_{0i} + a_{1i}x + \cdots + a_{n-1,i}x^{n-1}$,且

$$D = \begin{vmatrix} a_{00} & a_{01} & \cdots & a_{0,n-1} \\ a_{10} & a_{11} & \cdots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \cdots & a_{n-1,n-1} \end{vmatrix},$$

计算

$$\begin{vmatrix} \varphi_0(x_1) & \varphi_0(x_2) & \cdots & \varphi_0(x_n) \\ \varphi_1(x_1) & \varphi_1(x_2) & \cdots & \varphi_1(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{n-1}(x_1) & \varphi_{n-1}(x_2) & \cdots & \varphi_{n-1}(x_n) \end{vmatrix};$$

(3) $\begin{vmatrix} 1^2 & 2^2 & \cdots & n^2 \\ n^2 & 1^2 & \cdots & (n-1)^2 \\ \vdots & \vdots & \ddots & \vdots \\ 2^2 & 3^2 & \cdots & 1^2 \end{vmatrix}$; (4) $\begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ -a_n & a_1 & \cdots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ -a_2 & -a_3 & \cdots & a_1 \end{vmatrix}$.

4. 求矩阵 $AB = C$ 的子式 $C\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$,其中 A 为主对角元素全为 0 而其它元素为 1 的 6 阶方阵,而

$$B = \begin{pmatrix} 1 & 3 & 5 & 7 & 9 & 1 \\ 2 & 4 & 6 & 8 & 10 & 1 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 & 1 & 2 \\ 1 & 1 & 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 0 & 1 & 2 \end{pmatrix}.$$

5. 设 $A \in \mathbb{F}^{2\times 2}$,证明

$$A^2 - (\operatorname{Tr} A)A + (\det A)I_{(2)} = 0.$$

6. 证明不存在 $A, B \in \mathbb{F}^{n\times n}$,使得 $AB - BA = I_{(n)}$.

7. 适合 $AA^{\mathrm{T}} = I_{(n)} = A^{\mathrm{T}}A$ 的 n 阶方阵 A 称为**正交**的. 证明,

(1) 正交方阵的行列式等于 ± 1;

(2) 位于正交方阵的 k 个行上的所有 k 阶子式的平方和等于 1, $k = 1, 2, \ldots, n$.

8. 适合 $A\overline{A}^{\mathrm{T}} = I_{(n)} = \overline{A}^{\mathrm{T}}A$ 的 n 阶复方阵 A 称为**酉方阵**. 证明,

(1) 酉方阵的行列式的模为 1;

(2) 位于酉方阵的 k 个行上的所有 k 阶子式的模的平方和为 1.

9. 当 $j_1 = i_1, j_2 = i_2, \ldots, j_k = i_k$ 时,矩阵 A 的子式 $A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 称为矩阵 A 的一个 k 阶**主子式**,其中 $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$. 设 $A \in \mathbb{C}^{m\times n}$. 证明,矩阵 $A\overline{A}^{\mathrm{T}}$ 的每一个主子式都是非负实数.

10. 设 $A, B \in \mathbb{F}^{n\times n}$. 证明,方阵 AB 和 BA 的所有 k 阶主子式之和相等. 由此证明,

$$\det\big(\lambda I_{(n)} - AB\big) = \det\big(\lambda I_{(n)} - BA\big).$$

11. 设 $A = (B, C) \in \mathbb{C}^{n\times n}$,其中 B 是矩阵 A 的前 k 列构成的子矩阵. 证明

$$|\det A|^2 \leqslant \left|\overline{B}^{\mathrm{T}}B\right| \cdot \left|\overline{C}^{\mathrm{T}}C\right|.$$

§3.3 可逆矩阵

在通常数的乘法中,对非零的数 a,总存在数 b,使得 $ab = 1 = ba$,数 b 即是数 a 的倒数,记为 a^{-1}. 利用倒数概念,数的除法就可以归结为数的乘法,即 $a \div b = ab^{-1}$,其中 $b \neq 0$. 在考察 n 阶方阵中是否可以引进除法时,自然应当把倒数概念推广.

定义 3.3.1 给定矩阵 A,如果存在矩阵 B,使得

$$AB = I = BA,$$

其中 I 是单位方阵. 则矩阵 A 称为**可逆的**,而 B 称为矩阵 A 的**逆矩阵**,记为 A^{-1}.

由定义容易看出,可逆矩阵一定是方阵. 长方矩阵一定不可逆. 但是,并不是所有方阵都可逆.

定理 3.3.1 n 阶方阵 $A \in \mathbb{F}^{n\times n}$ 可逆的充分必要条件是它的行列式不为零.

证明 **必要性** 设 n 阶方阵 A 可逆,则存在 n 阶方阵 B,使得 $AB = BA = I_{(n)}$,两端取行列式,得到 $\det A \det B = 1$,因此 $\det A \neq 0$.

充分性 设 $A = (a_{ij})$,且 $\det A \neq 0$. 记行列式 $\det A$ 的元素 a_{ij} 的代数余子式为 A_{ij}. 取 n 阶方阵 A^* 为

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

由定理 2.3.2,

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \delta_{ik} \det A.$$

因此,

$$AA^* = \operatorname{diag}\big(\underbrace{\det A, \det A, \dots, \det A}_{n\text{ 个}}\big) = (\det A)I_{(n)}.$$

因为 $\det A \neq 0$,所以

$$A\Big(\frac{1}{\det A}A^*\Big) = I_{(n)}.$$

同理可证

$$\Big(\frac{1}{\det A}A^*\Big)A = I_{(n)}.$$

根据可逆方阵的定义,矩阵 A 可逆,并且

$$A^{-1} = \frac{1}{\det A}A^*. \tag{3.3.1}$$

■

定理 3.3.1 的证明中出现的 n 阶方阵 A^* 称为方阵 A 的**伴随方阵 (adjoint matrix)**. 由此不难看出,如果 $A \in \mathbb{F}^{n\times n}$,则 $A^* \in \mathbb{F}^{n\times n}$,从而 $A^{-1} \in \mathbb{F}^{n\times n}$. 容易证明,对任意 n 阶方阵 A 和 B,均有

$$(AB)^* = B^* A^*.$$

另外,在一些教科书上,行列式不为零的方阵称为**非异**方阵,或**非退化**方阵,否则称为**奇异**方阵,或**退化**方阵. 因此,定理 3.3.1 也可以叙述为: 方阵 A 可逆的充分必要条件是 A 非异,或非退化. 自然应考虑的问题是,可逆方阵是否唯一?

设 B 和 C 是可逆方阵 A 的逆方阵,则 $AB = BA = I, AC = CA = I$. 因此,

$$C = CI = C(AB) = (CA)B = IB = B.$$

这表明,可逆方阵 A 的逆矩阵是唯一的. 关于逆矩阵,以下性质成立.

(1) 设 n 阶方阵 A 可逆,则它的逆矩阵 A^{-1} 也可逆,并且

$$(A^{-1})^{-1} = A;$$

(2) 穿脱原理 设 n 阶方阵 A 与 B 均可逆,则它们的乘积 AB 也可逆,并且

$$(AB)^{-1} = B^{-1}A^{-1}.$$

利用归纳法容易证明,若 n 阶方阵 $A_1, A_2, \ldots, A_k$ 可逆,则它们的乘积 $A_1A_2 \cdots A_k$ 也可逆,并且

$$\left(A_1A_2\cdots A_k\right)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1};$$

(3) 设 $\lambda \in \mathbb{F}, \lambda \neq 0$,则

$$(\lambda A)^{-1} = \lambda^{-1}A^{-1};$$

(4) 设 n 阶方阵 A 可逆,则它的转置矩阵 A^{T} 也可逆,并且

$$(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}};$$

(5) 设 n 阶方阵 A 可逆,则它的逆矩阵的行列式为 A 的行列式的倒数,即

$$\det A^{-1} = (\det A)^{-1};$$

(6) 设 m 阶方阵 A 和 n 阶方阵 B 可逆,则 $m+n$ 阶方阵 $\mathrm{diag}(A, B)$ 也可逆,并且

$$\left(\mathrm{diag}(A, B)\right)^{-1} = \mathrm{diag}\left(A^{-1}, B^{-1}\right).$$

给定可逆方阵 A,如何求它的逆矩阵 A^{-1}? 当然可以先求出方阵 A 的伴随方阵 A^*,然后按公式 (3.3.1) 求出逆矩阵 A^{-1}. 不过这种方法很麻烦,因为在求 A^{-1} 时必须计算 n^2 个 $n-1$ 阶行列式,又必须计算 n 阶行列式 $\det A$,当 n 很大时,计算量相当大. 所以通常不采用这种方法. 下面的例 3.3.1 给出利用解线性方程组来求逆矩阵的方法.

例 3.3.1 确定以下 n 阶方阵 A 的逆矩阵,其中 $a_1a_2\cdots a_n \neq 0$:

$$A = \begin{pmatrix} 0 & a_1 & 0 & \cdots & 0 \\ 0 & 0 & a_2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & a_{n-1} \\ a_n & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

解 设 $A^{-1}=X=(x_{ij})$，则由 $AX=I_{(n)}$，得到，

$$\begin{pmatrix} a_1x_{21} & a_1x_{22} & \cdots & a_1x_{2n} \\ a_2x_{31} & a_2x_{32} & \cdots & a_2x_{3n} \\ \vdots & \vdots & & \vdots \\ a_{n-1}x_{n1} & a_{n-1}x_{n2} & \cdots & a_{n-1}x_{nn} \\ a_nx_{11} & a_nx_{12} & \cdots & a_nx_{1n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

由此得到，

$$x_{21}=\frac{1}{a_1},\quad x_{32}=\frac{1}{a_2},\quad \ldots,\quad x_{n,n-1}=\frac{1}{a_{n-1}},\quad x_{1n}=\frac{1}{a_n},$$

而其它的 $x_{ij}=0$. 因此，

$$A^{-1}=\begin{pmatrix} 0 & 0 & \cdots & 0 & \frac{1}{a_n} \\ \frac{1}{a_1} & 0 & & & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{1}{a_{n-1}} & 0 \end{pmatrix}.$$ ■

求逆矩阵的方法还有很多，以后遇到时再随时介绍.

利用逆矩阵，可以将关于线性方程组的 Cramer 法则简洁清晰地表述出来. 设给定 n 个未知量 $x_1,x_2,\ldots,x_n$ 的线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n. \end{cases} \tag{3.3.2}$$

它的系数矩阵为

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

并且 $\det A\neq 0$. 记 $x=(x_1,x_2,\ldots,x_n)^{\mathrm{T}}$，$\beta=(b_1,b_2,\ldots,b_n)^{\mathrm{T}}$. 于是利用矩阵乘法，方程组 (3.3.2) 可以记为

$$Ax=\beta. \tag{3.3.3}$$

现在证明 Cramer 法则. 先证明方程组 (3.3.3) 的解存在. 事实上，因为 $\det A\neq 0$，所以 A^{-1} 存在. 取 $x=A^{-1}\beta$，代入 (3.3.3)，得到

$$A(A^{-1}\beta)=(AA^{-1})\beta=I_{(n)}\beta=\beta.$$

因此，$x=A^{-1}\beta$ 是方程组 (3.3.3) 的解.

再证明方程组 (3.3.3) 的解的唯一性. 事实上，设另有解 $\tilde{x}$，将其代入 (3.3.3)，则

$A\widetilde{x}=\beta$. 两端同乘 A^{-1},得到 $A^{-1}(A\widetilde{x})=A^{-1}\beta$,即 $(A^{-1}A)\widetilde{x}=A^{-1}\beta$,因此,

$$I_{(n)}\widetilde{x}=\widetilde{x}=A^{-1}\beta.$$

这就证明,方程组 (3.3.3) 的解是唯一的.

由于

$$x=A^{-1}\beta=\left(\frac{1}{\det A}A^*\right)\beta=\frac{1}{\det A}\begin{pmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&\ddots&\vdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{pmatrix}\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

$$=\left(\frac{1}{\det A}\sum_{j=1}^n b_jA_{j1},\ \frac{1}{\det A}\sum_{j=1}^n b_jA_{j2},\ldots\ldots,\frac{1}{\det A}\sum_{j=1}^n b_jA_{jn}\right)^{\mathrm{T}},$$

即对于 $k=1,2,\ldots,n$,

$$x_k=\frac{1}{\det A}\sum_{j=1}^n b_jA_{jk}.$$

注意到

$$\sum_{j=1}^n b_jA_{jk}=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1,k-1}&b_1&a_{1,k+1}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2,k-1}&b_2&a_{2,k+1}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots&\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{n,k-1}&b_n&a_{n,k+1}&\cdots&a_{nn}\end{vmatrix}=\Delta_k,$$

因此对于 $k=1,2,\ldots,n$,

$$x_k=\frac{\Delta_k}{\Delta}.$$

这样就重新给出了 Cramer 法则的矩阵证明.

有了逆矩阵概念,矩阵分块运算就成为矩阵论中一种重要的技巧,即**矩阵打洞**. 所谓矩阵打洞是指,把矩阵分块,然后经过适当的变换,使所得到的矩阵在某些指定的块为零子矩阵. 矩阵打洞所根据的主要是以下的 Schur 公式.

Schur 公式　设 $A\in\mathbb{F}^{r\times r}$,$B\in\mathbb{F}^{r\times(n-r)}$,$C\in\mathbb{F}^{(n-r)\times r}$,而 $D\in\mathbb{F}^{(n-r)\times(n-r)}$,并且方阵 A 可逆. 则

(1)

$$\begin{pmatrix}I_{(r)}&0\\-CA^{-1}&I_{(n-r)}\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}A&B\\0&D-CA^{-1}B\end{pmatrix};\tag{3.3.4}$$

(2)

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}\begin{pmatrix}I_{(r)}&-A^{-1}B\\0&I_{(n-r)}\end{pmatrix}=\begin{pmatrix}A&0\\C&D-CA^{-1}B\end{pmatrix};\tag{3.3.5}$$

(3)

$$\begin{pmatrix}I_{(r)}&0\\-CA^{-1}&I_{(n-r)}\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}\begin{pmatrix}I_{(r)}&-A^{-1}B\\0&I_{(n-r)}\end{pmatrix}=\begin{pmatrix}A&0\\0&D-CA^{-1}B\end{pmatrix}.\tag{3.3.6}$$

Schur 公式的正确性是不难验证的(留给读者作练习). 至于 Schur 公式的由来,则可溯源于矩阵的初等变换. 这点将在以后加以说明. 下面举例说明 Schur 公式的一些应用.

例 3.3.2 设 $A,B,C,D\in\mathbb{C}^{n\times n}$,并且 $AC=CA$. 证明

$$\det\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\det\big(AD-CB\big).$$

证明 先证明方阵 A 可逆的情形. 此时由 Schur 公式 (3.3.4),

$$\begin{pmatrix}I_{(n)} & 0\\ -CA^{-1} & I_{(n)}\end{pmatrix}\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\begin{pmatrix}A & B\\ 0 & D-CA^{-1}B\end{pmatrix}.$$

显然,

$$\det\begin{pmatrix}I_{(n)} & 0\\ -CA^{-1} & I_{(n)}\end{pmatrix}=1.$$

因此,上式两端取行列式,得到

$$\begin{aligned}\det\begin{pmatrix}A & B\\ C & D\end{pmatrix}&=\det\begin{pmatrix}A & B\\ 0 & D-CA^{-1}B\end{pmatrix}\\ &=\det A\det\big(D-CA^{-1}B\big)=\det\big(A\big(D-CA^{-1}B\big)\big).\end{aligned}$$

由于 $AC=CA$,所以

$$A\big(D-CA^{-1}B\big)=AD-\big(AC\big)A^{-1}B=AD-CAA^{-1}B=AD-CB.$$

于是,

$$\det\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\det\big(AD-CB\big).$$

再证明方阵 A 不可逆的情形. 记

$$A_\lambda=A+\lambda I_{(n)}=\begin{pmatrix}a_{11}+\lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}+\lambda & \ddots & \vdots\\ \vdots & \ddots & \ddots & a_{n-1,n}\\ a_{n1} & \cdots & a_{n,n-1} & a_{nn}+\lambda\end{pmatrix},$$

其中 $A=(a_{ij})$.

容易看出,行列式 $\det A_\lambda$ 是关于 λ 的 n 次多项式. 因此,它具有 n 个根 $\lambda_1,\lambda_2,\ldots,\lambda_n$. 如果它们不全为零,则记

$$\lambda_0=\min\big\{|\lambda_i|\;\big|\;\lambda_i\neq 0,\quad 1\leqslant i\leqslant n\big\}.$$

取正数 ε,使得 $0<\varepsilon<\lambda_0$. 于是对任意 λ, $0<\lambda<\varepsilon$, $\det A_\lambda\neq 0$,即 A_λ 可逆. 由于 $AC=CA$,故

$$A_\lambda C=CA_\lambda.$$

如果 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 全为零,则同样存在正数 ε,使得对任意 λ, $0<\lambda<\varepsilon$, A_λ 可逆,并且 $A_\lambda C=CA_\lambda$. 于是,利用上段证明的结论,

$$\det\begin{pmatrix}A_\lambda & B\\ C & D\end{pmatrix}=\det\big(A_\lambda D-CB\big).$$

上式两端都是关于 λ 的多项式,从而是 λ 的连续函数. 因此当 λ 趋于零时,上式就

化为

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det\bigl(AD - CB\bigr). \qquad \blacksquare$$

应当指出，例 3.3.2 中关于方阵 A 不可逆情形的证明所使用的方法称为**微小摄动法**. 这是矩阵论中一种值得重视的方法.

例 3.3.3　设 $A \in \mathbb{F}^{m\times n}$，$B \in \mathbb{F}^{n\times m}$. 证明，

$$\det\bigl(I_{(m)} - AB\bigr) = \det\bigl(I_{(n)} - BA\bigr).$$

证明　取 $m+n$ 阶方阵

$$\begin{pmatrix} I_{(m)} & A \\ B & I_{(n)} \end{pmatrix}.$$

由 Schur 公式 (3.3.4) 得到

$$\begin{pmatrix} I_{(m)} & 0 \\ -B & I_{(n)} \end{pmatrix}\begin{pmatrix} I_{(m)} & A \\ B & I_{(n)} \end{pmatrix} = \begin{pmatrix} I_{(m)} & A \\ 0 & I_{(n)} - BA \end{pmatrix}.$$

仿照 Schur 公式 (3.3.4)，

$$\begin{pmatrix} I_{(m)} & -A \\ 0 & I_{(n)} \end{pmatrix}\begin{pmatrix} I_{(m)} & A \\ B & I_{(n)} \end{pmatrix} = \begin{pmatrix} I_{(m)} - AB & 0 \\ B & I_{(n)} \end{pmatrix}.$$

取上述两式的行列式，得到

$$\det\begin{pmatrix} I_{(m)} & A \\ B & I_{(n)} \end{pmatrix} = \det\bigl(I_{(n)} - BA\bigr) = \det\bigl(I_{(m)} - AB\bigr). \qquad \blacksquare$$

顺便指出，例 3.3.3 给出一种计算行列式的方法. 例如

$$\begin{aligned}\begin{vmatrix} 1+x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & 1+x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & & \vdots \\ x_ny_1 & x_ny_2 & \cdots & 1+x_ny_n \end{vmatrix} &= \det\left(I_{(n)} - \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\bigl(-y_1, -y_2, \ldots, -y_n\bigr)\right) \\ &= \det\left(1 - \bigl(-y_1, -y_2, \ldots, -y_n\bigr)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) \\ &= 1 + x_1y_1 + x_2y_2 + \cdots + x_ny_n.\end{aligned}$$

例 3.3.4　n 阶方阵 A 的 n 个子式

$$A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix},$$

其中 $k = 1, 2, \ldots, n$，称为方阵 A 的**顺序主子式**.

设方阵 $A \in \mathbb{R}^{n\times n}$ 的顺序主子式都是正的，而非对角元都是负的. 证明，逆矩阵 A^{-1} 的每个元素都是正的.

证明　对阶数 n 用归纳法. 当 $n = 1$ 时，结论显然成立. 设结论对 $n-1$ 阶方阵

成立，并设 A 是 n 阶方阵. 把方阵 A 分块为

$$A = \begin{pmatrix} A_1 & \alpha \\ \beta^{\mathrm{T}} & a_{nn} \end{pmatrix},$$

其中 α 与 β 是 $(n-1)\times 1$ 矩阵，A_1 是 A 的 $n-1$ 阶子矩阵.

由假设，

$$\det A_1 = A\begin{pmatrix} 1 & 2 & \cdots & (n-1) \\ 1 & 2 & \cdots & (n-1) \end{pmatrix} > 0,$$

因此，A_1 可逆. 由 Schur 公式，

$$\begin{pmatrix} I_{(n-1)} & 0 \\ -\beta^{\mathrm{T}}A_1 & 1 \end{pmatrix}\begin{pmatrix} A_1 & \alpha \\ \beta^{\mathrm{T}} & a_{nn} \end{pmatrix}\begin{pmatrix} I_{(n-1)} & -A_1^{-1}\alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha \end{pmatrix}.$$

两端取行列式，得到

$$\det A = \left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)\det A_1.$$

因为 $\det A_1 > 0$，$\det A > 0$，故

$$a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha > 0.$$

容易验证，

$$\begin{pmatrix} I_{(n-1)} & 0 \\ -\beta^{\mathrm{T}}A_1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} I_{(n-1)} & 0 \\ \beta^{\mathrm{T}}A_1 & 1 \end{pmatrix},$$

$$\begin{pmatrix} I_{(n-1)} & -A_1^{-1}\alpha \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} I_{(n-1)} & A_1^{-1}\alpha \\ 0 & 1 \end{pmatrix}.$$

因此

$$\begin{pmatrix} A_1 & \alpha \\ \beta^{\mathrm{T}} & a_{nn} \end{pmatrix} = \begin{pmatrix} I_{(n-1)} & 0 \\ \beta^{\mathrm{T}}A_1 & 1 \end{pmatrix}\begin{pmatrix} A_1 & 0 \\ 0 & a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha \end{pmatrix}\begin{pmatrix} I_{(n-1)} & A_1^{-1}\alpha \\ 0 & 1 \end{pmatrix}.$$

于是由穿脱原理

$$A^{-1} = \begin{pmatrix} A_1 & \alpha \\ \beta^{\mathrm{T}} & a_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} I_{(n-1)} & -A_1^{-1}\alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} A_1^{-1} & 0 \\ 0 & \left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1} \end{pmatrix}\begin{pmatrix} I_{(n-1)} & 0 \\ -\beta^{\mathrm{T}}A_1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} A_1^{-1} + \left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1}A_1^{-1}\alpha\beta^{\mathrm{T}}A_1^{-1} & -\left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1}A_1^{-1}\alpha \\ -\left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1}\beta^{\mathrm{T}}A_1^{-1} & \left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1} \end{pmatrix}.$$

由于 $\left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right) > 0$，故方阵 A^{-1} 的 (n,n) 系数是正的. 由归纳假设，A_1^{-1} 的每个元素都是正的，而 α 与 β^{T} 的每个元素都是负的，从而

$$-\left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1}\beta^{\mathrm{T}}A_1^{-1} \text{ 与 } -\left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1}A_1^{-1}\alpha$$

的每个元素都是正的.

最后，由于 α 与 β^{T} 的每个元素都是负的，因此，$\alpha\beta^{\mathrm{T}}$ 的每个元素都是正的. 又 A_1^{-1} 的每个元素都是正的，所以

$$A_1^{-1} + \left(a_{nn} - \beta^{\mathrm{T}}A_1^{-1}\alpha\right)^{-1}A_1^{-1}\alpha\beta^{\mathrm{T}}A^{-1}$$

的每个元素都是正的. ■

习　题　3.3

1. 求下列方阵的逆矩阵,其中 ω 是 n 次本原单位根.

(1) $\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix}$;　　(2) $\begin{pmatrix}1&1&\cdots&1\\0&1&\cdots&1\\\vdots&\ddots&\ddots&\vdots\\0&\cdots&0&1\end{pmatrix}_{n\times n}$;

(3) $\begin{pmatrix}1&1&1&\cdots&1\\1&\omega&\omega^2&\cdots&\omega^{n-1}\\1&\omega^2&\omega^4&\cdots&\omega^{2(n-1)}\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&\omega^{n-1}&\omega^{2n-2}&\cdots&\omega^{(n-1)^2}\end{pmatrix}$;　　(4) $\begin{pmatrix}2&-1&&&&\\-1&2&-1&&&\\&-1&2&-1&&\\&&\ddots&\ddots&\ddots&\\&&&-1&2&-1\\&&&&-1&2\end{pmatrix}_{n\times n}$.

2. 设 A 是 n 阶幂零方阵. 证明方阵 $I_{(n)}-A$ 可逆. 并且如果 A 的幂零指数为 k,则

$$\left(I_{(n)}-A\right)^{-1}=I_{(n)}+A+\cdots+A^{k-1}.$$

3. 证明:设 n 阶方阵 A 不可逆,则存在 n 阶非零的方阵 B,使得 $AB=0$.

4. 证明,当且仅当对角元素全不为零时,上三角方阵可逆,并且它的逆仍是上三角的.

5. 证明:

(1) 正交方阵一定可逆,并且它的逆仍是正交方阵;

(2) 酉方阵一定可逆,并且它的逆仍是酉方阵,

(3) 可逆对称方阵的逆仍是对称方阵;　　(4) 可逆斜对称方阵的逆仍是斜对称方阵.

6. 设 $A\in\mathbb{F}^{m\times n}$,$B\in\mathbb{F}^{n\times m}$,并且 $I_{(m)}-AB$ 可逆. 证明,$I_{(n)}-BA$ 也可逆,并且

$$\left(I_{(n)}-BA\right)^{-1}=I_{(n)}+B\left(I_{(m)}-AB\right)^{-1}A.$$

7. 设 n 阶方阵 A 适合

$$a_0A^m+a_1A^{m-1}+\cdots+a_{m-1}A+a_mI_{(n)}=0,$$

其中 $m\geqslant 1$,$a_0a_m\neq 0$. 证明方阵 A 可逆,并求 A^{-1}.

8. 系数都是整数的矩阵称为**整系数矩阵**. 行列式等于 ±1 的整系数矩阵称为**幺模矩阵**. 证明,整系数矩阵 A 的逆矩阵仍是整系数矩阵的充分必要条件是 A 为幺模矩阵.

9. 设 A^* 表示 n 阶方阵 A 的伴随方阵. 证明,

(1) $(\lambda A)^*=\lambda^{n-1}A^*$,其中 λ 是纯量;　　(2) $(AB)^*=B^*A^*$,其中 B 也是 n 阶方阵.

10. 设 A_{ij} 是 n 阶方阵 $A=(a_{ij})$ 的行列式 $\det A$ 的元素 a_{ij} 的代数余子式. 证明,

$$\begin{vmatrix}A_{ik}&A_{jk}\\A_{i\ell}&A_{j\ell}\end{vmatrix}=(-1)^{i+j+k+\ell}A\begin{pmatrix}1&\cdots&(i-1)&(i+1)&\cdots&(j-1)&(j+1)&\cdots&n\\1&\cdots&(k-1)&(k+1)&\cdots&(\ell-1)&(\ell+1)&\cdots&n\end{pmatrix}\det A,$$

其中 $1\leqslant i<j\leqslant n$,$1\leqslant k<\ell\leqslant n$.

11. 设 $A\in\mathbb{F}^{m\times n}$,$B\in\mathbb{F}^{n\times m}$,$\lambda\in\mathbb{F}$. 证明

$$\lambda^n\det\left(\lambda I_{(m)}-AB\right)=\lambda^m\det\left(\lambda I_{(n)}-BA\right).$$

12. 设 $A,B\in\mathbb{C}^{n\times n}$,而复数 i 适合 $\mathrm{i}^2=-1$. 证明

$$\det\begin{pmatrix}A&-B\\B&A\end{pmatrix}=\det\left(A+\mathrm{i}B\right)\det\left(A-\mathrm{i}B\right).$$

13. 求 n 阶方阵 $A=(a_{ij})$ 的行列式,其中 $a_{ii}=x+a_i$,$i=1,2,\ldots,n$,而 $a_{ij}=a_j$,$1\leqslant i\neq j\leqslant n$.

14. 设 $A\in\mathbb{R}^{2n\times 2n}$,且

$$A\begin{pmatrix} 0 & I_{(n)} \\ -I_{(n)} & 0 \end{pmatrix}A^{\mathrm{T}} = \begin{pmatrix} 0 & I_{(n)} \\ -I_{(n)} & 0 \end{pmatrix}.$$

证明 $\det A = 1$.

§3.4 矩阵的秩与相抵

矩阵的秩是一个基本概念,它在线性空间理论和矩阵论中占有基本的重要性.它具有多种形式的等价定义. 下面的定义是 Sylvester 在 1851 年给出的.

定义 $m\times n$ 矩阵 A 的所有非零子式的最高阶数称为**矩阵 A 的秩**,记为 $\operatorname{rank} A$.

由定义可以看出:

(1) 对任意 $m\times n$ 矩阵 A,

$$\operatorname{rank} A \leqslant \min\{m,\ n\}, \qquad \operatorname{rank} A = \operatorname{rank} A^{\mathrm{T}};$$

(2) $\operatorname{rank} A = 0$ 的充分必要条件是 A 为零矩阵;

(3) 如果矩阵 A 具有非零的 r 阶子式,则 $\operatorname{rank} A \geqslant r$;

(4) 如果矩阵 A 的每一个 ℓ 阶子式全为零,则由 Laplace 展开定理,矩阵 A 的所有 t 阶子式全为零,这里 $t > \ell$,从而 $\operatorname{rank} A \leqslant \ell - 1$. 因此,如果矩阵 A 具有非零的 r 阶子式,并且所有 $r+1$ 阶子式全为零,则 $\operatorname{rank} A = r$.

这为求矩阵的秩提供一种方法. 例如,设

$$A = \begin{pmatrix} -1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -1 \\ 2 & 0 & -4 & -2 \end{pmatrix},$$

它的二阶子式

$$A\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} \neq 0,$$

而它的所有三阶子式都为零,因此 $\operatorname{rank} A = 2$.

对于 n 阶方阵,它的秩至多为 n. 秩为 n 的 n 阶方阵称为**满秩**的,否则称为**降秩**的. 显然,满秩的 n 阶方阵 A 的行列式不为零,因此,方阵 A 可逆,反之亦然. 因此,“满秩”和“可逆”是等价的概念. 对于 $m\times n$ 矩阵,秩为 m 的 $m\times n$ 矩阵称为**行满秩**的,秩为 n 的 $m\times n$ 矩阵称为**列满秩**的.

关于矩阵乘积的秩,有

定理 3.4.1 设 $A\in\mathbb{F}^{m\times n}$, $B\in\mathbb{F}^{n\times p}$, 则

$$\operatorname{rank} AB \leqslant \min\{\operatorname{rank} A,\ \operatorname{rank} B\}.$$

证明 设 $\operatorname{rank} A = r$, $\operatorname{rank} B = s$, $AB = C$. 任取矩阵 C 的 $r+1$ 阶子式

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_{r+1} \\ j_1 & j_2 & \cdots & j_{r+1} \end{pmatrix},$$

由 Binet-Cauchy 公式,

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_{r+1} \\ j_1 & j_2 & \cdots & j_{r+1} \end{pmatrix} = \sum_{1 \leqslant k_1 < \cdots < k_{r+1} \leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_{r+1} \\ k_1 & k_2 & \cdots & k_{r+1} \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_{r+1} \\ j_1 & j_2 & \cdots & j_{r+1} \end{pmatrix}.$$

因为 $\operatorname{rank} A = r$,所以 A 的每一个 $r+1$ 阶子式都为零. 因此

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_{r+1} \\ j_1 & j_2 & \cdots & j_{r+1} \end{pmatrix} = 0.$$

这表明,矩阵 $C = AB$ 的所有 $r+1$ 阶子式全为零. 因此,$\operatorname{rank} AB \leqslant r$.

同理可证,$\operatorname{rank} AB \leqslant s$. ■

由定理 3.4.1 直接得到,

定理 3.4.2　设 $A \in \mathbb{F}^{m\times n}$,$P \in \mathbb{F}^{m\times m}$,$Q \in \mathbb{F}^{n\times n}$,并且 P 和 Q 可逆,则

$$\operatorname{rank} PAQ = \operatorname{rank} PA = \operatorname{rank} AQ = \operatorname{rank} A.$$

证明　由定理 3.4.1,$\operatorname{rank} PA \leqslant \operatorname{rank} A$. 另一方面,因为 P 可逆,所以 $A = P^{-1}(PA)$. 由定理 3.4.1,$\operatorname{rank} A \leqslant \operatorname{rank} PA$. 因此,

$$\operatorname{rank} PA = \operatorname{rank} A.$$

同理可证,$\operatorname{rank} AQ = \operatorname{rank} A$. 于是,$\operatorname{rank} PAQ = \operatorname{rank} PA$. ■

给定一个 $m \times n$ 矩阵 A,尽管可以根据秩的定义确定矩阵 A 的秩,但毕竟太麻烦. 下面介绍求矩阵的秩的一般方法.

定义 3.4.1　给定 $m \times n$ 矩阵 A,对矩阵 A 实施如下的行(或列)变换:

(1) 对换矩阵 A 的某两行(或列);

(2) 矩阵 A 的某一行(或列)遍乘以某个非零的数 a,然后加到另一行(或列);

(3) 矩阵 A 的某一行(或列)遍乘以某个非零的数 b;

分别称为对矩阵 A 实施**第一、二、三种行(或列)的初等变换**,统称为**初等行(或列)变换**. 矩阵 A 的行或列的初等变换,简称为**初等变换**.

矩阵的初等变换可以通过矩阵乘法来表示. 记 n 阶方阵

$$\begin{aligned} P_{ij} &= I_{(n)} - E_{ii} - E_{jj} + E_{ij} + E_{ji}, && 1 \leqslant i \neq j \leqslant n; \\ Q_{ij}(a) &= I_{(n)} + aE_{ij}, && 1 \leqslant i, j \leqslant n, \quad a \in \mathbb{F}; \\ P_i(b) &= I_{(n)} + (b-1)E_{ii}, && 1 \leqslant i \leqslant n, \quad b \in \mathbb{F}, \end{aligned}$$

其中 E_{ij} 是 (i, j) 系数为 1 而其它系数为零的 n 阶方阵,$1 \leqslant i, j \leqslant n$.

容易看出,方阵 $P_{ij}, Q_{ij}(a), P_i(b)$ 都是可逆的,并且它们的逆矩阵分别是

$$P_{ij}^{-1} = P_{ij}, \qquad \big(Q_{ij}(a)\big)^{-1} = Q_{ij}(-a), \qquad \big(P_i(b)\big)^{-1} = P_i(b^{-1}).$$

也即与原方阵是同一类型的矩阵.

n 阶方阵 P_{ij}, $Q_{ij}(a)$ 和 $P_i(b)$ 称为**初等方阵**. 其中 P_{ij} 特别称为**初等置换方阵**,而初等置换方阵的乘积称为**置换方阵**.

可以验证,对矩阵 A 实施初等变换可以用左乘或右乘初等方阵来实现:

(1) 对换矩阵 A 的第 i 行和第 j 行,相当于用初等置换方阵 P_{ij} 左乘矩阵 A; 而

对换矩阵 A 的第 j 列和第 i 列，相当于用初等置换矩阵 P_{ij} 右乘矩阵 A；

(2) 矩阵 A 的第 j 行遍乘非零的数 a，然后加到第 i 行，相当于用 m 阶初等方阵 $Q_{ij}(a)$ 左乘矩阵 A；同样用 n 阶初等方阵 $Q_{ij}(a)$ 右乘矩阵 A，便得到矩阵 A 的相应列变换；

(3) 矩阵 A 的第 i 行遍乘以非零的数 b，相当于用 m 阶方阵 $P_i(b)$ 左乘矩阵 A；矩阵 A 的第 i 列遍乘以数 b，相当于用 n 阶方阵 $p_i(b)$ 右乘矩阵 A.

定理 3.4.3 矩阵的初等变换不改变矩阵的秩.

证明 对 $m\times n$ 矩阵 A 实施初等变换，相当于用 m 阶初等方阵 P 左乘矩阵 A，或者用 n 阶初等方阵 Q 右乘矩阵 A. 而初等方阵都是可逆的，由此由定理 3.4.2，

$$\operatorname{rank} PA = \operatorname{rank} AQ = \operatorname{rank} A.$$ ■

定理 3.4.4 设 $m\times n$ 矩阵 A 的秩为 r，则矩阵 A 可以经过有限次初等变换化为如下形式：

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

证明 如果 $A=0$，则 A 本身具有上述形式. 因此设 $A\neq 0$.

那么 A 必有某个系数 $a_{ij}\neq 0$. 对换 A 的第 1 行和第 i 行，第 1 列和第 j 列，所得到的矩阵的 $(1,1)$ 系数就是 $a_{ij}\neq 0$. 用数 a_{ij}^{-1} 乘以第 1 行，矩阵就化为

$$\begin{pmatrix} 1 & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & a'_{23} & \cdots & a'_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a'_{m1} & a'_{m2} & a'_{m3} & \cdots & a'_{mn} \end{pmatrix}.$$

再用 $-a_{i1}$ 乘以上述矩阵的第 1 行，并加到第 i 行，$i=2,3,\ldots,m$，得到矩阵

$$\begin{pmatrix} 1 & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ 0 & a''_{22} & a''_{23} & \cdots & a''_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a''_{m2} & a''_{m3} & \cdots & a''_{mn} \end{pmatrix},$$

其中 $a''_{ij}=a'_{ij}-a'_{i1}a'_{1j}$，$2\leqslant i\leqslant m$，$2\leqslant j\leqslant n$.

再用 $-a'_{1i}$ 乘以上述矩阵的第 1 列，并加到第 j 列，得到矩阵 B 为

$$B=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{m2} & b_{m3} & \cdots & b_{mn} \end{pmatrix} =: \begin{pmatrix} 1 & 0 \\ 0 & B_1 \end{pmatrix},$$

其中 $b_{ij}=a''_{ij}$，$2\leqslant i\leqslant m$，$2\leqslant j\leqslant n$，$B_1=(b_{ij})$.

如果 B_1 为零矩阵，则停止. 如果 $B_1\neq 0$，则重复上述过程，即 B_1 可以化为如下形式：

$$B_1 = \begin{pmatrix} 1 & 0 \\ 0 & C_1 \end{pmatrix}.$$

于是矩阵 B 化为

$$C = \begin{pmatrix} I_{(2)} & 0 \\ 0 & C_1 \end{pmatrix}.$$

再对矩阵 C 重复上述过程,等等. 因此矩阵 A 可以经过有限次初等变换化为

$$D = \begin{pmatrix} I_{(s)} & 0 \\ 0 & 0 \end{pmatrix}.$$

根据定理 3.4.3,

$$s = \operatorname{rank} D = \operatorname{rank} A = r.$$ ■

定理 3.4.4 有两个重要的推论.

推论 3.4.1 设 $A \in \mathbb{F}^{n\times n}$, $\operatorname{rank} A = r$. 则存在可逆的 $P \in \mathbb{F}^{m\times m}$ 和 $Q \in \mathbb{F}^{n\times n}$,使得

$$PAQ = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

证明 由于对矩阵 A 实施初等变换相当于用初等方阵左乘或右乘于 A,因此由定理 3.4.4,存在 m 阶初等方阵 $P_1, P_2, \dots, P_s$ 和 n 阶初等方阵 $Q_1, Q_2, \dots, Q_t$,使得

$$P_1P_2\cdots P_sAQ_1Q_2\cdots Q_t = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

记

$$P = P_1P_2\cdots P_s, \qquad Q = Q_1Q_2\cdots Q_t.$$

由于可逆方阵的乘积仍是可逆的,因此,P 与 Q 均是可逆的,并且

$$PAQ = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$ ■

推论 3.4.2 n 阶方阵 A 可逆当且仅当 A 可以表为有限个初等方阵的乘积.

证明 充分性是显然的,仅证必要性. 因为方阵 A 可逆,故 $\operatorname{rank} A = n$. 由定理 3.4.4,存在 n 阶初等方阵 $P_1, P_2, \dots, P_s$ 和 $Q_1, Q_2, \dots, Q_t$,使得

$$P_1P_2\cdots P_sAQ_1Q_2\cdots Q_t = I_{(n)}.$$

因此,

$$A = P_s^{-1}\cdots P_2^{-1}P_1^{-1}Q_t^{-1}\cdots Q_2^{-1}Q_1^{-1}.$$

由于初等方阵的逆方阵仍是初等方阵,因此上式表明,可逆方阵 A 可以表为有限个初等方阵的乘积. ■

推论 3.4.3 n 阶可逆方阵 A 可以经过有限次的初等变换变为 n 阶单位方阵.

证明 由推论 3.4.2,存在 n 阶初等方阵 $P_1, P_2, \dots, P_k$,使得 $A = P_1P_2\cdots P_k$. 因此

$$P_k^{-1}\cdots P_2^{-1}P_1^{-1}A = I_{(n)}.$$

由于初等方阵 P_i 的逆矩阵 P_i^{-1} 仍是初等方阵,因此上式表明,方阵 A 可以经过

有限次行的初等变换变为单位方阵. ■

推论 3.4.3 对列的初等变换也成立.

应当指出,推论 3.4.3 给出求逆矩阵的一个重要方法. 给定 n 阶可逆方阵 A,由推论 3.4.3,存在 n 阶初等方阵 $P_1, P_2, \ldots, P_k$,使得

$$P_kP_{k-1}\cdots P_1A = I_{(n)}.$$

因此,$A^{-1} = P_kP_{k-1}\cdots P_1$. 取 $n\times 2n$ 矩阵 $B = \big(A,\ I_{(n)}\big)$,则

$$\begin{aligned} P_kP_{k-1}\cdots P_1B &= \big(P_kP_{k-1}\cdots P_1A, P_kP_{k-1}\cdots P_1\big) \\ &= \big(I_{(n)}, P_kP_{k-1}\cdots P_1\big) = \big(I_{(n)}, A^{-1}\big). \end{aligned}$$

由于初等方阵左乘矩阵 B 相当于对矩阵 B 实施行的初等变换,因此上式表明,对矩阵 B 实施行的初等变换,使得它的左半部分 A 化为单位方阵,则矩阵 B 的后半部分就变成 A^{-1}. 于是就求出了方阵 A 的逆矩阵.

例 3.4.1 求 n 阶方阵 $A = (a_{ij})$ 的逆矩阵,其中 $a_{ii} = 0$,$a_{ij} = 1$,$1 \leqslant i \neq j \leqslant n$.

解 取 $n\times 2n$ 矩阵

$$B = \left(\begin{array}{ccccc:ccccc} 0 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 1 & 0 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 1 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 \end{array}\right).$$

把矩阵 B 的第 $2, 3, \ldots, n$ 行都加到第 1 行,矩阵 B 化为

$$B_1 = \left(\begin{array}{ccccc:ccccc} n-1 & n-1 & n-1 & \cdots & n-1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 & 0 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 1 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 \end{array}\right).$$

用 $\dfrac{-1}{n-1}$ 遍乘矩阵 B_1 的第 1 行,然后分别加到第 $2, 3, \ldots, n$ 行,矩阵 B_1 化为

$$B_2 = \left(\begin{array}{ccccc:ccccc} 1 & 1 & 1 & \cdots & 1 & \frac{1}{n-1} & \frac{1}{n-1} & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ 0 & -1 & 0 & \cdots & 0 & -\frac{1}{n-1} & \frac{n-2}{n-1} & -\frac{1}{n-1} & \cdots & -\frac{1}{n-1} \\ 0 & 0 & -1 & \cdots & 0 & -\frac{1}{n-1} & -\frac{1}{n-1} & \frac{n-2}{n-1} & \cdots & -\frac{1}{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & -\frac{1}{n-1} & -\frac{1}{n-1} & -\frac{1}{n-1} & \cdots & \frac{n-2}{n-1} \end{array}\right).$$

矩阵 B_2 的第 $2, 3, \ldots, n$ 行都加到第 1 行,然后用 -1 分别乘以第 $2, 3, \ldots, n$ 行,矩阵

B_2 变为

$$B_2=\left(\begin{array}{ccccc:ccccc}
1 & 0 & 0 & \cdots & 0 & -\dfrac{n-2}{n-1} & \dfrac{1}{n-1} & \dfrac{1}{n-1} & \cdots & \dfrac{1}{n-1}\\
0 & 1 & 0 & \cdots & 0 & \dfrac{1}{n-1} & -\dfrac{n-2}{n-1} & \dfrac{1}{n-1} & \cdots & \dfrac{1}{n-1}\\
0 & 0 & 1 & \cdots & 0 & \dfrac{1}{n-1} & \dfrac{1}{n-1} & -\dfrac{n-2}{n-1} & \cdots & \dfrac{1}{n-1}\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\
0 & 0 & 0 & \cdots & 1 & \dfrac{1}{n-1} & \dfrac{1}{n-1} & \dfrac{1}{n-1} & \cdots & -\dfrac{n-2}{n-1}
\end{array}\right).$$

由此求得，$A^{-1}=(c_{ij})$，其中

$$\begin{cases} c_{ij}=\dfrac{1}{n-1}, & \text{当 } 1\leqslant i\neq j\leqslant n \text{ 时；}\\[2ex] c_{ii}=-\dfrac{n-2}{n-1}, & \text{当 } i=1,2,\ldots,n \text{ 时．} \end{cases}$$ ■

现在引进矩阵相抵的概念.

定义 3.4.2　设 $A,B\in\mathbb{F}^{m\times n}$. 如果矩阵 A 可以经过有限次初等变换变为矩阵 B，则矩阵 A 和 B 是**相抵**的.

由于对矩阵 A 实施初等变换相当于用 m 阶初等方阵 P 左乘矩阵 A，或用 n 阶初等方阵 Q 右乘矩阵 A. 因此上述定义等价于说：

如果存在 m 阶初等方阵 $P_1,P_2,\ldots,P_s$ 和 n 阶初等方阵 $Q_1,Q_2,\ldots,Q_t$ 使得

$$P_1P_2\cdots P_sAQ_1Q_2\cdots Q_t=B,$$

则矩阵 A 和 B 称为相抵的.

由于 $P=P_1P_2\cdots P_s$，$Q=Q_1Q_2\cdots Q_t$ 可逆，以及推论 3.4.2，因此上述定义还等价于说：如果存在 m 阶可逆方阵 P 和 n 阶可逆方阵 Q，使得 $PAQ=B$，则矩阵 A 和 B 称为相抵的. 容易验证，矩阵的相抵满足以下性质.

(1) 自反性　对任意 $A\in\mathbb{F}^{m\times n}$，矩阵 A 和自身总是相抵的，

(2) 对称性　对任意 $A,B\in\mathbb{F}^{m\times n}$，如果矩阵 A 和 B 相抵，则矩阵 B 和 A 也相抵；

(3) 传递性　对任意 $A,B,C\in\mathbb{F}^{m\times n}$，如果矩阵 A 和 B 相抵，矩阵 B 和 C 相抵，则矩阵 A 和 C 相抵.

在所有 $m\times n$ 矩阵的集合 $\mathbb{F}^{m\times n}$ 中，所有和 $m\times n$ 矩阵 A 相抵的集合记为 R_A，它称为矩阵 A 所属的**相抵等价类**. 关于相抵等价类，以下性质成立.

性质 3.4.1　设 $B,C\in R_A$，则矩阵 B 和 C 相抵.

证明　因为 $B,C\in R_A$，所以矩阵 B 与 C 都和矩阵 A 相抵.

由相抵的对称性，矩阵 A 和 C 相抵. 由于矩阵 B 和 A 相抵，矩阵 A 和 C 相抵，因此，由相抵的传递性，矩阵 B 和 C 相抵. ■

性质 3.4.2 设 $A, B \in \mathbb{F}^{m\times n}$，则 $R_A \cap R_B$ 为空集，或者 $R_A = R_B$.

证明 设 $R_A \cap R_B \neq \varnothing$，则存在 $C \in R_A \cap R_B$，即矩阵 C 分别和矩阵 A, B 相抵. 因此，矩阵 A 和 B 相抵，即 $A \in R_B$，且 $B \in R_A$.

设 $D \in R_A$，则矩阵 D 和 A 相抵，而矩阵 A 和 B 相抵，从而矩阵 D 和 B 相抵. 因此，$D \in R_B$，即 $R_A \subseteq R_B$.

同理可证，$R_B \subseteq R_A$. 于是

$$R_A = R_B = R_A \cap R_B. \qquad \blacksquare$$

性质 3.4.2 表明，不同的相抵等价类不相交，而 $\mathbb{F}^{m\times n}$ 中每个矩阵都一定属于某个相抵等价类. 因此，$\mathbb{F}^{m\times n}$ 便分解为不交的相抵等价类的并集.

从 $\mathbb{F}^{m\times n}$ 的每一个相抵等价类中各取一个矩阵作为该相抵等价类的代表元素，所有这些元素的集合称为 $\mathbb{F}^{m\times n}$ 在相抵下的**互异代表元系**.

由性质 3.4.1，同一个相抵等价类中的矩阵都相抵，因此，相抵等价类中每一个矩阵都可以作为该等价类的代表元素. 问题是，在等价类中如何选取代表元，使它具最简单的形式？这就是矩阵在相抵下的标准形的问题.

其次，任给两个 $m \times n$ 矩阵 A 和 B，如何判定它们是否同属于一个相抵等价类，即是否相抵. 同属于一个相抵等价类的矩阵必然具有某种共同性质，这种共性应当在某种量上得到反映. 如果一种量为相抵等价类中每个矩阵所具有，那么这种量称为矩阵在相抵下的**不变量**. 如果一组相抵下的不变量足以区分不同的相抵等价类，即可以判定两个矩阵是否相抵，而且当这组不变量缺少某一个就不足以区分不同的相抵等价类，那么这组不变量就称为矩阵在相抵下的**全系不变量**. 因此，判定两个矩阵是否属于同一个相抵等价类的问题，就是寻求相抵下的全系不变量.

归结起来，寻求矩阵在相抵下的标准形和全系不变量是矩阵在相抵下的分类问题中的两个基本问题. 实际上，前面已解决了在相抵下的矩阵的分类问题，现在把结论重述一下.

定理 3.4.5 设 $A \in \mathbb{F}^{m\times n}$，$\operatorname{rank} A = r$. 则矩阵 A 相抵于如下矩阵

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.4.1}$$

而且 $A, B \in \mathbb{F}^{m\times n}$ 相抵的充分必要条件是 $\operatorname{rank} A = \operatorname{rank} B$.

证明 前一结论即是定理 3.4.4. 现在证明后一结论.

必要性 因为矩阵 A 和 B 相抵，所以矩阵 A 可以经过有限次初等变换变为矩阵 B. 由定理 3.4.3，$\operatorname{rank} A = \operatorname{rank} B$.

充分性 设 $\operatorname{rank} A = \operatorname{rank} B = r$. 定理 3.4.4 表明，矩阵 A 和 (3.4.1) 相抵，而 (3.4.1) 又和 B 相抵，由相抵的传递性，矩阵 A 和 B 相抵. ■

矩阵 (3.4.1) 称为矩阵 A 的 Hermite 标准形. 定理 3.4.5 的后一结论表明，矩阵的秩是矩阵在相抵下的不变量，而且是全系不变量.

例 3.4.2　设 $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$, $\lambda \in \mathbb{F}$. 证明

$$\lambda^n \det\bigl(\lambda I_{(m)} - AB\bigr) = \lambda^m \det\bigl(\lambda I_{(n)} - BA\bigr).$$

证明　设 $\operatorname{rank} A = r$. 则由定理 3.4.5, 矩阵 A 的 Hermite 标准形为 (3.4.1), 即存在可逆的 $P \in \mathbb{F}^{m\times m}$, $Q \in \mathbb{F}^{n\times n}$, 使得

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q = PCQ,$$

记矩阵 (3.4.1) 为 C, 于是

$$\begin{aligned}
\lambda^n \det\bigl(\lambda I_{(m)} - AB\bigr) &= \lambda^n \det\bigl(\lambda I_{(m)} - PCQB\bigr) \\
&= \lambda^n \det\bigl(P(\lambda I_{(m)} - CQBP)P^{-1}\bigr) \\
&= \lambda^n \det\bigl(\lambda I_{(m)} - CQBP\bigr).
\end{aligned}$$

记

$$QBP = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

其中 $B_{11} \in \mathbb{F}^{r\times r}$, 则上式化为

$$\begin{aligned}
\lambda^n \det\bigl(\lambda I_{(m)} - AB\bigr) &= \lambda^n \det\left(\lambda I_{(m)} - \begin{pmatrix} B_{11} & B_{12} \\ 0 & 0 \end{pmatrix}\right) \\
&= \lambda^n \det\begin{pmatrix} \lambda I_{(r)} - B_{11} & -B_{12} \\ 0 & \lambda I_{(m-r)} \end{pmatrix} \\
&= \lambda^{m+n-r} \det\bigl(\lambda I_{(r)} - B_{11}\bigr).
\end{aligned}$$

另一方面,

$$\begin{aligned}
\lambda^m \det\bigl(\lambda I_{(n)} - BA\bigr) &= \lambda^m \det\bigl(\lambda I_{(n)} - BPCQ\bigr) \\
&= \lambda^m \det\bigl(Q^{-1}(\lambda I_{(n)} - QBPC)Q\bigr) \\
&= \lambda^m \det\bigl(\lambda I_{(n)} - QBPC\bigr) \\
&= \lambda^m \det\left(\lambda I_{(n)} - \begin{pmatrix} B_{11} & 0 \\ B_{21} & 0 \end{pmatrix}\right) \\
&= \lambda^m \det\begin{pmatrix} \lambda I_{(r)} - B_{11} & 0 \\ -B_{21} & \lambda I_{(n-r)} \end{pmatrix} \\
&= \lambda^{m+n-r} \det\bigl(\lambda I_{(r)} - B_{11}\bigr).
\end{aligned}$$

于是得到

$$\lambda^n \det\bigl(\lambda I_{(m)} - AB\bigr) = \lambda^m \det\bigl(\lambda I_{(n)} - BA\bigr). \qquad \blacksquare$$

习　题　3.4

1. 求下列矩阵的秩.

(1) $\begin{pmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{pmatrix}$;　　　(2) $\begin{pmatrix} 1 & 3 & 5 & 1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \end{pmatrix}$;

(3) $\begin{pmatrix} i & j & \cdots & j \\ j & i & \ddots & \vdots \\ \vdots & \ddots & \ddots & j \\ j & \cdots & j & i \end{pmatrix}_{n\times n}$; (4) $\begin{pmatrix} 0 & i & \cdots & i \\ i & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & i \\ i & \cdots & i & 0 \end{pmatrix}_{n\times n}$.

2. 求 λ,使得矩阵 A 的秩为最小,其中

$$A = \begin{pmatrix} 3 & 1 & 1 & 4 \\ \lambda & 4 & 10 & 1 \\ 1 & 7 & 17 & 3 \\ 2 & 2 & 4 & 3 \end{pmatrix}.$$

3. 利用初等变换求下列矩阵的逆矩阵.

(1) $\begin{pmatrix} 3 & 3 & -4 & 3 \\ 0 & 6 & 1 & 1 \\ 5 & 4 & 2 & 1 \\ 2 & 3 & 3 & 2 \end{pmatrix}$; (2) $\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ n & 1 & 2 & \cdots & n-1 \\ n-1 & n & 1 & \cdots & n-2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 3 & 4 & \cdots & 1 \end{pmatrix}$;

(3) $\begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{pmatrix}$; (4) $\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}_{n\times n}$.

4. 证明,只用行的初等变换以及对换某两列,任意 $m\times n$ 矩阵 A 都可以化为

$$\begin{pmatrix} I_{(r)} & * \\ 0 & 0 \end{pmatrix},$$

其中 $r = \operatorname{rank} A$.

5. 证明,任意一个秩为 r 的矩阵都可以表为 r 个秩为 1 的矩阵之和.

6. 证明,$m\times n$ 矩阵 A 的秩为 1 的充分必要条件是 $A=\alpha\beta$,其中 α 和 β 分别是 $m\times 1$ 和 $1\times n$ 的非零矩阵.

7. 设 $A, B\in\mathbb{F}^{m\times n}$. 证明 $\operatorname{rank}(AB)\leqslant \operatorname{rank} A+\operatorname{rank} B$.

8. 设 $A\in\mathbb{R}^{m\times n}$. 证明 $\operatorname{rank} AA^{\mathrm{T}} = \operatorname{rank} A^{\mathrm{T}}A = \operatorname{rank} A$.

9. 设 n 阶方阵 A 分块为

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

其中 A_{11} 是 r 阶可逆矩阵. 证明,$\operatorname{rank} A = r$ 的充分必要条件是 $A_{22} = A_{21}A_{11}^{-1}A_{12}$.

10. 设 $A\in\mathbb{F}^{n\times n}$,$\operatorname{rank} A = r$. 从矩阵 A 中任意取出 s 个行构成 $s\times n$ 矩阵 B. 证明,$\operatorname{rank} B\geqslant r+s-n$.

11. 设 $A\in\mathbb{F}^{m\times n}$,$\operatorname{rank} A=r$,从矩阵 A 中取出 s 个行,t 个列上的交叉元素构成的 $s\times t$ 矩阵记为 B. 证明,$\operatorname{rank} B\geqslant r+s+t-m-n$.

12. 设 n 阶方阵 A 至少有 n^2-n+1 个元素为零. 证明,$\operatorname{rank} A<n$. 并确定 $\operatorname{rank} A$ 的最大值.

13. n 阶方阵 A 的伴随方阵记为 A^*. 证明,

(1) $\operatorname{rank} A^* = n$ 的充分必要条件是 $\operatorname{rank} A = n$;

(2) $\operatorname{rank} A^* = 1$ 的充分必要条件是 $\operatorname{rank} A = n-1$;

(3) $\operatorname{rank} A^* = 0$ 的充分必要条件是 $\operatorname{rank} A < n-1$;

(4) 当 $n>2$ 时，$(A^*)^*=(\det A)^{n-2}A$；当 $n=2$ 时，$(A^*)^*=A$.

14. 设 A 与 B 是行数相同的矩阵，A 和 B 并排而成的矩阵记为 (A,B). 证明，

$$\operatorname{rank}(A,B)\leqslant \operatorname{rank}A+\operatorname{rank}B.$$

15. 证明，二阶幺模矩阵 A 可以表为矩阵

$$P=\begin{pmatrix}1&1\\0&1\end{pmatrix},\qquad Q=\begin{pmatrix}0&1\\1&0\end{pmatrix}$$

的方幂的乘积.

16. 设 $A\in\mathbb{R}^{n\times n}$. 证明，如果矩阵 $A^{\mathrm{T}}A$ 的每一个 k 阶主子式都为零，则 $\operatorname{rank}A<k$.

17. 证明，n 阶方阵 A 都可以表为形如 $I_{(n)}+aE_{ij}$ 的方阵的乘积，其中 E_{ij} 是 (i,j) 系数为 1 而其它系数都为零的 n 阶方阵，$1\leqslant i,j\leqslant n$.

18. 设 A 是 $m\times n$ 整系数矩阵. 证明，有在 $\mathbb{Z}$ 可逆的 m 阶整系数矩阵 P 和 n 阶整系数矩阵 Q（逆矩阵仍是整系数的整系数矩阵称为在 $\mathbb{Z}$ 上可逆的），使得

$$A=P\begin{pmatrix}\begin{pmatrix}d_1&&&\\&d_2&&\\&&\ddots&\\&&&d_r\end{pmatrix}&0\\0&0\end{pmatrix}_{m\times n}Q,$$

其中 $d_1,d_2,\ldots,d_r$ 是正整数，并且 $d_i\mid d_{i+1}$，$i=1,2,\ldots,r-1$.

§3.5　一些例子

矩阵在相抵下的 Hermite 标准形有着广泛的应用. 下面举例说明.

例 3.5.1　$m\times n$ 矩阵 A 为列满秩当且仅当存在 m 阶可逆方阵 P，使得

$$A=P\begin{pmatrix}I_{(n)}\\0\end{pmatrix},$$

其中 0 是 $(m-n)\times n$ 零矩阵.

证明　充分性显然. 下面证明必要性. 因为 A 是列满秩的，即 $\operatorname{rank}A=n$. 因此存在 m 阶与 n 阶可逆方阵 R 与 T，使得

$$A=R\begin{pmatrix}I_{(n)}\\0\end{pmatrix}T.$$

由分块矩阵的乘法，

$$A=R\begin{pmatrix}T\\0\end{pmatrix}=R\begin{pmatrix}T&0\\0&I_{(m-n)}\end{pmatrix}\begin{pmatrix}I_{(n)}\\0\end{pmatrix}.$$

记

$$P=R\begin{pmatrix}T&0\\0&I_{(m-n)}\end{pmatrix},$$

则方阵 P 是可逆的，并且

$$A = P\begin{pmatrix} I_{(n)} \\ 0 \end{pmatrix}.$$ ■

例 3.5.2 设 $A \in \mathbb{F}^{m\times n}$. 证明 $\operatorname{rank} A = r \geqslant 1$ 的充分必要条件是,存在列满秩的 $m \times r$ 矩阵 B 与行满秩的 $r \times n$ 矩阵 C,使得 $A = BC$.

证明 **必要性** 设 $\operatorname{rank} A = r$,则存在 m 阶与 n 阶可逆方阵 P 与 Q,使得

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q.$$

记

$$B = P\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix}_{m\times r}, \qquad C = \left(I_{(r)}, 0\right)_{r\times n} Q,$$

则

$$\operatorname{rank} B = \operatorname{rank} P\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix} = r,$$

$$\operatorname{rank} C = \operatorname{rank}\left(I_{(r)}, 0\right)_{r\times n} Q = \operatorname{rank}\left(I_{(r)}, 0\right)_{r\times n} = r,$$

即 B 与 C 分别是列满秩与行满秩的,并且 $A = BC$.

充分性 设 $A = BC$,其中 B 与 C 分别是列满秩与行满秩的. 则由例 3.5.1,

$$B = P\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix}_{m\times r},$$

其中 P 为 m 阶可逆方阵. 而 C^{T} 为列满秩的,因此,

$$C^{\mathrm{T}} = Q^{\mathrm{T}}\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix}_{n\times r},$$

其中 Q^{T} 是 n 阶可逆方阵. 于是

$$A = P\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix}_{m\times r} \left(I_{(r)}, 0\right)_{r\times n} Q = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q.$$

所以 $\operatorname{rank} A = r$. ■

例 3.5.2 说明,任何一个矩阵都可以分解为一个列满秩矩阵与一个行满秩矩阵的乘积. 这一事实称为矩阵的**满秩分解定理**. 它有许多应用.

例 3.5.3 设 A 是给定 $m \times n$ 矩阵,X 是 $m \times n$ 矩阵,求矩阵方程 $A^{\mathrm{T}} X = X^{\mathrm{T}} A$ 的所有解 X.

解 设 $\operatorname{rank} A = r$,则存在 m 阶与 n 阶可逆方阵 P 与 Q,使得

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q.$$

注意,其中的 P 与 Q 不一定唯一. 取定 P_0 与 Q_0,使得

$$A = P_0\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q_0,$$

代入原矩阵方程,得到

$$Q_0^{\mathrm{T}}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}P_0^{\mathrm{T}}X = X^{\mathrm{T}}P_0\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q_0.$$

因此,

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}P_0^{\mathrm{T}}XQ_0^{-1} = (P_0^{\mathrm{T}}XQ_0^{-1})^{\mathrm{T}}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

记

$$P_0^{\mathrm{T}}XQ_0^{-1} = Y = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix},$$

其中 Y_{11} 是 r 阶方阵. 代入上式得到,

$$\begin{pmatrix} Y_{11} & Y_{12} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} Y_{11}^{\mathrm{T}} & 0 \\ Y_{12}^{\mathrm{T}} & 0 \end{pmatrix}.$$

因此,$Y_{11}^{\mathrm{T}} = Y_{11}$, $Y_{12} = 0$. 于是,

$$X = (P_0^{\mathrm{T}})^{-1}XQ_0 = (P_0^{\mathrm{T}})^{-1}\begin{pmatrix} Y_{11} & 0 \\ Y_{21} & Y_{22} \end{pmatrix}Q_0. \tag{3.5.1}$$

其中 Y_{11} 是任意的 r 阶对称方阵,Y_{21} 与 Y_{22} 分别是任意的 $(m-r)\times r$ 与 $(m-r)\times(n-r)$ 矩阵. 这说明,原矩阵方程的解 X 应具有形式 (3.5.1).

反之,容易验证,形如 (3.5.1) 的矩阵 X 一定是原矩阵方程的解. 因此,式 (3.5.1) 给出了原矩阵方程的所有解. ■

例 3.5.4　证明,秩为 r 的 n 阶实对称方阵 S (即适合 $S^{\mathrm{T}} = S$)至少有一个 r 阶主子式不为零,而且所有非零的 r 阶主子式都同号.

证明　因为 $\operatorname{rank} S = r$,所以存在 n 阶可逆方阵 P 与 Q,使得

$$S = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q.$$

由 $S^{\mathrm{T}} = S$ 得到

$$Q^{\mathrm{T}}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}P^{\mathrm{T}} = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q,$$

即

$$P^{-1}Q^{\mathrm{T}}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}(P^{-1}Q^{\mathrm{T}})^{\mathrm{T}}.$$

记

$$P^{-1}Q^{\mathrm{T}} = R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

其中 R_{11} 是 r 阶方阵,代入上式得到 $R_{11}^{\mathrm{T}} = R_{11}$, $R_{21} = 0$. 于是

$$P^{-1}Q^{\mathrm{T}} = \begin{pmatrix} R_{11} & 0 \\ R_{12}^{\mathrm{T}} & R_{22}^{\mathrm{T}} \end{pmatrix},$$

即

$$Q^{\mathrm{T}} = P\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}.$$

从而

$$Q = \begin{pmatrix} R_{11} & 0 \\ R_{12}^{\mathrm{T}} & R_{22}^{\mathrm{T}} \end{pmatrix} P^{\mathrm{T}}.$$

因此,

$$S = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} R_{11} & 0 \\ R_{12}^{\mathrm{T}} & R_{22}^{\mathrm{T}} \end{pmatrix} P^{\mathrm{T}} = P\begin{pmatrix} R_{11} & 0 \\ 0 & 0 \end{pmatrix} P^{\mathrm{T}}.$$

由于方阵 P 可逆,所以,

$$\operatorname{rank} S = \operatorname{rank}\begin{pmatrix} R_{11} & 0 \\ 0 & 0 \end{pmatrix} = \operatorname{rank} R_{11} = r.$$

由于 R_{11} 是 r 阶的,因此 R_{11} 是可逆对称方阵.

现在设 $S\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}$ 是方阵 S 的任意 r 阶主子式,$1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$.

由 Binet-Cauchy 公式,

$$S\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} = P\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ 1 & 2 & \cdots & r \end{pmatrix}^2 \det R_{11}. \tag{3.5.2}$$

由此可以看出,如果方阵 S 的所有 r 阶主子式都为零,则

$$P\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ 1 & 2 & \cdots & r \end{pmatrix} = 0.$$

于是对行列式 $\det P$ 的前 r 列作 Laplace 展开,得到

$$\det P = \sum_{1\leqslant i_1<i_2<\cdots<i_r\leqslant n} P\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ 1 & 2 & \cdots & r \end{pmatrix} P\begin{pmatrix} i_{r+1} & \cdots & i_n \\ (r+1) & \cdots & n \end{pmatrix} = 0,$$

这和方阵 P 可逆矛盾. 因此,方阵 S 至少有一个 r 阶主子式不为零.

另外,由式 (3.5.2) 可以看出,方阵 S 的任意 r 阶非零主子式都和行列式 $\det R_{11}$ 同号. 因此,对称方阵 S 的所有 r 阶非零主子式都同号. ■

例 3.5.5　证明,n 阶幂等方阵 A (即适合 $A^2 = A$)的秩等于它的迹.

证明　设 $\operatorname{rank} A = r$,则存在 n 阶可逆方阵 P 和 Q,使得

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q.$$

因为方阵 A 是幂等的,所以,

$$P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} QP \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q,$$

即

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} QP \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

记

$$QP = R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

其中 R_{11} 是 r 阶方阵. 则上式化为

$$\begin{pmatrix} R_{11} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

因此,$R_{11} = I_{(r)}$. 所以

$$Q = \begin{pmatrix} I_{(r)} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} P^{-1}.$$

于是

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_{(r)} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} P^{-1} = P\begin{pmatrix} I_{(r)} & R_{12} \\ 0 & 0 \end{pmatrix} P^{-1}.$$

由于

$$\begin{aligned} f(\lambda) &= \det\big(\lambda I_{(n)} - A\big) = \det\left(\lambda I_{(n)} - P\begin{pmatrix} I_{(r)} & R_{12} \\ 0 & 0 \end{pmatrix} P^{-1}\right) \\ &= \det P\left(\lambda I_{(n)} - \begin{pmatrix} I_{(r)} & R_{12} \\ 0 & 0 \end{pmatrix}\right) P^{-1} = \det\left(\lambda I_{(n)} - \begin{pmatrix} I_{(r)} & R_{12} \\ 0 & 0 \end{pmatrix}\right) \\ &= \lambda^{n-r}(\lambda - 1)^r. \end{aligned}$$

所以,$f(\lambda)$ 的 $n-1$ 次项系数为 $-r$. 但是,$f(\lambda) = \det\big(\lambda I_{(n)} - A\big)$ 的 $n-1$ 次项系数为

$$-\big(a_{11} + a_{22} + \cdots + a_{nn}\big),$$

其中 $a_{11}, a_{22}, \ldots, a_{nn}$ 是方阵 A 的主对角元. 而 $\operatorname{Tr} A = a_{11} + \cdots + a_{nn}$. 因此 $\operatorname{Tr} A = r$. ■

从 §3.4 知道,在讨论矩阵在相抵下的分类问题时,矩阵的初等变换起着重要作用. 矩阵的初等变换可以推广到分块矩阵. 设 $m \times n$ 矩阵 A 分块为

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

其中 A_{11} 为 $s \times t$ 子矩阵. 则

(1)
$$\begin{pmatrix} 0 & I_{(m-s)} \\ I_{(s)} & 0 \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{pmatrix},$$

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} 0 & I_{(t)} \\ I_{(n-t)} & 0 \end{pmatrix} = \begin{pmatrix} A_{12} & A_{11} \\ A_{22} & A_{21} \end{pmatrix}.$$

即将分块矩阵的每个小块当成一个元素,对换分块矩阵 A 的两行(或列)可以通过左乘(或右乘)以上述形式的分块矩阵实现;

(2)
$$\begin{pmatrix} I_{(s)} & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ BA_{21} & BA_{22} \end{pmatrix},$$

$$\begin{pmatrix} C & 0 \\ 0 & I_{(m-s)} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} CA_{11} & CA_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

即用某个矩阵左乘分块矩阵 A 的某一行,可以按上述方式左乘一个分块矩阵实现. 对列也有类似的结果;

(3) $$\begin{pmatrix} I_{(s)} & B \\ 0 & I_{(m-s)} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11}+BA_{21} & A_{12}+BA_{22} \\ A_{21} & A_{22} \end{pmatrix},$$
$$\begin{pmatrix} I_{(s)} & 0 \\ C & I_{(m-s)} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21}+CA_{11} & A_{22}+CA_{12} \end{pmatrix}.$$
即用某个矩阵左乘分块矩阵 A 的某一行,并加到另一行,可以按上述方式左乘一个分块矩阵实现. 对列也有类似的结论.

容易看出,如果在上式中令 $A_{21}+CA_{11}=0$,且设 A_{11} 可逆,则 $C=-A_{21}A_{11}^{-1}$. 于是就得到 Schur 公式 (3.3.4). 所以,Schur 公式只是对分块矩阵施行初等变换的特殊情形而已.

分块矩阵的初等变换可以用来证明有关矩阵的秩的许多命题. 在利用分块矩阵的初等变换证明秩的命题时,经常用到下述显而易见的事实. 其正确性请读者自行证明.

(1) 矩阵 A 分块后,矩阵 A 中某个块的秩不超过整个矩阵 A 的秩;

(2) 设 $A\in\mathbb{F}^{m\times n}$,$B\in\mathbb{F}^{p\times q}$,则
$$\operatorname{rank}\operatorname{diag}(A,B)=\operatorname{rank}A+\operatorname{rank}B;$$

(3) 设 $A\in\mathbb{F}^{m\times n}$,$B\in\mathbb{F}^{p\times q}$,$C\in\mathbb{F}^{m\times q}$,则
$$\operatorname{rank}\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \geqslant \operatorname{rank}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

例 3.5.6(Frobenius 秩不等式) 设 $A\in\mathbb{F}^{m\times n}$,$B\in\mathbb{F}^{n\times p}$,$C\in\mathbb{F}^{p\times q}$. 证明,
$$\operatorname{rank}AB+\operatorname{rank}BC-\operatorname{rank}B\leqslant\operatorname{rank}ABC, \tag{3.5.3}$$
其中等式成立的充分必要条件是矩阵
$$\begin{pmatrix} AB & 0 \\ 0 & BC \end{pmatrix},\qquad \begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix}$$
相抵.

证明 由 (3.5.3) 得到,
$$\operatorname{rank}AB+\operatorname{rank}BC\leqslant\operatorname{rank}ABC+\operatorname{rank}B.$$
因此式 (3.5.3) 等价于
$$\operatorname{rank}\begin{pmatrix} AB & 0 \\ 0 & BC \end{pmatrix} \leqslant \operatorname{rank}\begin{pmatrix} ABC & 0 \\ 0 & B \end{pmatrix}. \tag{3.5.4}$$
由于
$$\begin{pmatrix} I_{(m)} & A \\ 0 & I_{(n)} \end{pmatrix}\begin{pmatrix} ABC & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} I_{(q)} & 0 \\ -C & I_{(p)} \end{pmatrix}\begin{pmatrix} 0 & -I_{(q)} \\ I_{(p)} & 0 \end{pmatrix} = \begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix}, \tag{3.5.5}$$
其中方阵
$$\begin{pmatrix} I_{(m)} & A \\ 0 & I_{(n)} \end{pmatrix},\qquad \begin{pmatrix} I_{(q)} & 0 \\ -C & I_{(p)} \end{pmatrix},\qquad \begin{pmatrix} 0 & -I_{(q)} \\ I_{(p)} & 0 \end{pmatrix}$$

都是可逆的，因此由定理 3.4.2，

$$\operatorname{rank}\begin{pmatrix} ABC & 0 \\ 0 & B \end{pmatrix} = \operatorname{rank}\begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix}.$$

由事实 (3)，

$$\operatorname{rank}\begin{pmatrix} AB & 0 \\ 0 & BC \end{pmatrix} \leqslant \operatorname{rank}\begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix} = \operatorname{rank}\begin{pmatrix} ABC & 0 \\ 0 & B \end{pmatrix}. \tag{3.5.6}$$

这就证明了式 (3.5.3) 成立.

容易看出，式 (3.5.3) 中等式成立的充分必要条件是式 (3.5.6) 等式成立. ■

特别，当 B 取为 n 阶单位方阵时，Frobenius 秩不等式即为

$$\operatorname{rank} A + \operatorname{rank} C - n \leqslant \operatorname{rank} AC.$$

结合定理 3.4.1，即有

$$\operatorname{rank} A + \operatorname{rank} B - n \leqslant \operatorname{rank} AB \leqslant \min\{\operatorname{rank} A, \operatorname{rank} B\},$$

其中 $A \in \mathbb{F}^{m\times n}$，$B \in \mathbb{F}^{n\times p}$. 这就是 **Sylvester 秩不等式.**

例 3.5.7　设 n 阶方阵 $A_1, A_2, \ldots, A_k$ 满足

$$A_1 + A_2 + \cdots + A_k = I_{(n)}.$$

证明，方阵 $A_1, A_2, \ldots, A_k$ 为幂等的充分必要条件是，

$$\operatorname{rank} A_1 + \operatorname{rank} A_2 + \cdots + \operatorname{rank} A_k = n.$$

特别，当 $k = 2$ 时，记 $A = A_1$，$B = A_2 = I_{(n)} - A$. 如果方阵 A 为幂等的，则

$$B^2 = \left(I_{(n)} - A\right)^2 = I_{(n)} - 2A + A^2 = I_{(n)} - A = B.$$

因此，命题化为：n 阶方阵 A 为幂等的充分必要条件是，

$$\operatorname{rank} A + \operatorname{rank}\left(I_{(n)} - A\right) = n.$$

证明　必要性　设方阵 $A_1, A_2, \ldots, A_k$ 幂等，则由例 3.5.5，$\operatorname{rank} A_i = \operatorname{Tr} A_i$. 故

$$\operatorname{rank} A_1 + \operatorname{rank} A_2 + \cdots + \operatorname{rank} A_k = \operatorname{Tr} A_1 + \operatorname{Tr} A_2 + \cdots + \operatorname{Tr} A_k.$$

由于 $\operatorname{Tr} A$ 是方阵 A 的线性函数，所以，

$$\operatorname{rank} A_1 + \operatorname{rank} A_2 + \cdots + \operatorname{rank} A_k = \operatorname{Tr}\left(A_1 + A_2 + \cdots + A_k\right) = \operatorname{Tr} I_{(n)} = n.$$

充分性　当 $k = 2$ 时，

$$\operatorname{rank} A_1 + \operatorname{rank} A_2 = \operatorname{rank}\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A_1 & 0 \\ 0 & I_{(n)} - A_1 \end{pmatrix}.$$

由于

$$\begin{pmatrix} I_{(n)} & A_1 - I_{(n)} \\ 0 & I_{(n)} \end{pmatrix}\begin{pmatrix} 0 & I_{(n)} \\ I_{(n)} & 0 \end{pmatrix}\begin{pmatrix} I_{(n)} & I_{(n)} \\ 0 & I_{(n)} \end{pmatrix}\begin{pmatrix} A_1 & 0 \\ 0 & I_{(n)} - A_1 \end{pmatrix}$$
$$\times \begin{pmatrix} I_{(n)} & I_{(n)} \\ 0 & I_{(n)} \end{pmatrix}\begin{pmatrix} I_{(n)} & 0 \\ -A & I_{(n)} \end{pmatrix} = \begin{pmatrix} A_1^2 - A_1 & 0 \\ 0 & I_{(n)} \end{pmatrix},$$

因此

$$\operatorname{rank} A_1 + \operatorname{rank} A_2 = \operatorname{rank}\begin{pmatrix} A_1 & 0 \\ 0 & I_{(n)} - A_1 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A_1^2 - A_1 & 0 \\ 0 & I_{(n)} \end{pmatrix}$$
$$= \operatorname{rank}\left(A_1^2 - A_1\right) + \operatorname{rank} I_{(n)} = \operatorname{rank}\left(A_1^2 - A_1\right) + n.$$

所以，$\operatorname{rank}(A_1^2 - A_1) = 0$. 即 $A_1^2 = A_1$. 这就证明了方阵 A_1 是幂等的.

现在转到一般的 k. 记 $B_i = A_1 + \cdots + A_{i-1} + A_{i+1} + \cdots + A_k$，则对于 $1 \leqslant i \leqslant k$，

$$A_i + B_i = I_{(n)}.$$

因为
$$\operatorname{rank} B_i = \operatorname{rank}\left(A_1 + \cdots + A_{i-1} + A_{i+1} + \cdots + A_k\right)$$
$$\leqslant \operatorname{rank} A_1 + \cdots + \operatorname{rank} A_{i-1} + \operatorname{rank} A_{i+1} + \cdots + \operatorname{rank} A_k,$$

并且 $\operatorname{rank} A_1 + \operatorname{rank} A_2 + \cdots + \operatorname{rank} A_k = n$，所以，$\operatorname{rank} B_i \leqslant n - \operatorname{rank} A_i$，即

$$\operatorname{rank} A_i + \operatorname{rank} B_i \leqslant n.$$

另一方面，

$$\operatorname{rank} A_i + \operatorname{rank} B_i \geqslant \operatorname{rank}(A_i + B_i) = \operatorname{rank} I_{(n)} = n,$$

因此，对于 $1 \leqslant i \leqslant k$，
$$\operatorname{rank} A_i + \operatorname{rank} B_i = n.$$

这说明，方阵 A_i 与 B_i 满足 $k = 2$ 时的条件.

由上面的证明，方阵 A_i 是幂等的，$i = 1, 2, \ldots, k$. ■

例 3.5.8（Roth，1952） 设 $A \in \mathbb{F}^{m\times n}$，$B \in \mathbb{F}^{p\times q}$，$C \in \mathbb{F}^{m\times q}$，$X \in \mathbb{F}^{n\times q}$，$Y \in \mathbb{F}^{m\times p}$，则矩阵方程

$$AX - YB = C \tag{3.5.7}$$

有解的充分必要条件是矩阵

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

相抵.

证明　必要性　设 X, Y 是方程 (3.5.7) 的解. 则

$$\begin{pmatrix} I_{(m)} & -Y \\ 0 & I_{(p)} \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} I_{(n)} & X \\ 0 & I_{(q)} \end{pmatrix} = \begin{pmatrix} A & AX - YB \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}.$$

因此矩阵

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

相抵.

充分性　设矩阵

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

相抵. 则

$$\operatorname{rank}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}. \tag{3.5.8}$$

记 $\operatorname{rank} A = r, \operatorname{rank} B = s$，则存在可逆的方阵 $P \in \mathbb{F}^{m\times m}, Q \in \mathbb{F}^{n\times n}, R \in \mathbb{F}^{p\times p}$ 与 $T \in \mathbb{F}^{q\times q}$，使得

$$PAQ = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}, \qquad RBT = \begin{pmatrix} I_{(s)} & 0 \\ 0 & 0 \end{pmatrix}.$$

于是

$$\begin{pmatrix} P & 0 \\ 0 & R \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & T \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_{(s)} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.5.9}$$

$$\begin{pmatrix} P & 0 \\ 0 & R \end{pmatrix}\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & T \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 & C_{11} & C_{12} \\ 0 & 0 & C_{21} & C_{22} \\ 0 & 0 & I_{(s)} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

其中 $C_{11} \in \mathbb{F}^{r\times s}, C_{22} \in \mathbb{F}^{(m-r)\times(q-s)}$，且

$$PCT = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}.$$

记

$$K = \begin{pmatrix} C_{11} & 0 \\ C_{21} & 0 \end{pmatrix}, \qquad L = \begin{pmatrix} 0 & C_{12} \\ 0 & 0 \end{pmatrix},$$

其中 $K \in \mathbb{F}^{m\times p}, L \in \mathbb{F}^{n\times q}$，则

$$\begin{pmatrix} I_{(m)} & -K \\ 0 & I_{(p)} \end{pmatrix}\begin{pmatrix} P & 0 \\ 0 & R \end{pmatrix}\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & T \end{pmatrix}\begin{pmatrix} I_{(n)} & -L \\ 0 & I_{(q)} \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{22} \\ 0 & 0 & I_{(s)} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.5.10}$$

由 (3.5.8), (3.5.9) 与 (3.5.10), $C_{22} = 0$. 因此

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} P^{-1} & 0 \\ 0 & R^{-1} \end{pmatrix}\begin{pmatrix} I_{(m)} & K \\ 0 & I_{(p)} \end{pmatrix}\begin{pmatrix} P & 0 \\ 0 & R \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \times \begin{pmatrix} Q & 0 \\ 0 & T \end{pmatrix}\begin{pmatrix} I_{(n)} & L \\ 0 & I_{(q)} \end{pmatrix}\begin{pmatrix} Q^{-1} & 0 \\ 0 & T^{-1} \end{pmatrix}.$$

由此得到，

$$C = A(QLT^{-1}) - (-P^{-1}KR)B.$$

因此，$X = QLT^{-1}, Y = -P^{-1}KR$ 是方程 (3.5.7) 的解. ∎

习 题 3.5

1. 证明，设 A 是 n 阶方阵. 如果存在正整数 k，使得 $\operatorname{rank} A^k = \operatorname{rank} A^{k+1}$，则

$$\operatorname{rank} A^k = \operatorname{rank} A^{k+1} = \operatorname{rank} A^{k+2} = \cdots.$$

2. 设 A 和 B 为 n 阶方阵. 证明,

$$\operatorname{rank}\bigl(AB - I_{(n)}\bigr) \leqslant \operatorname{rank}\bigl(A - I_{(n)}\bigr) + \operatorname{rank}\bigl(B - I_{(n)}\bigr).$$

3. 证明,n 阶斜对称方阵 K 的秩是偶数,并且秩为 r 的斜对称方阵至少有一个 r 阶主子式不为零,同时,所有非零的 r 阶主子式都同号.

4. 设 A 和 B 为 n 阶方阵,$AB = BA = 0$,并且 $\operatorname{rank} A^2 = \operatorname{rank} A$. 证明,

$$\operatorname{rank}\bigl(A + B\bigr) = \operatorname{rank} A + \operatorname{rank} B.$$

5. 设 A 和 B 为 n 阶方阵,$AB = BA = 0$. 证明,存在正整数 k,使得

$$\operatorname{rank}\bigl(A^k + B^k\bigr) = \operatorname{rank} A^k + \operatorname{rank} B^k.$$

6. 设 $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$. 证明,$\operatorname{rank} AB = \operatorname{rank} A$ 的充分必要条件是,存在 $C \in \mathbb{F}^{m\times n}$,使得 $A = ABC$. 由此证明,如果 $\operatorname{rank} AB = \operatorname{rank} A$ 且方阵 AB 幂等,则方阵 BA 也幂等.

7. 设整数 $A_1, A_2, \dots, a_n$ 的最大公因数为 d. 证明,存在 n 阶可逆的整系数矩阵 P,使得

$$(a_1, a_2, \dots, a_n)P = \bigl(d, \underbrace{0, 0, 0, \dots, 0, 0}_{n-1\text{个}}\bigr).$$

8. 证明,存在 n 阶可逆的整系数矩阵 P,使得它的第 1 行为整数 $a_1, a_2, \dots, a_n$ 的充分必要条件是,整数 $a_1, a_2, \dots, a_n$ 互素.

§3.6 线性方程组

给定 n 个未知量 $x_1, x_2, \dots, x_n$ 的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{3.6.1}$$

其中对于 $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$, a_{ij} 和 b_i 是已知的. a_{ij} 称为方程组 (3.6.1) 的**系数**,b_i 称为方程组 (3.6.1) 的**常数项**. 记

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

则方程组 (3.6.1) 可以改写为矩阵形式

$$Ax = \beta. \tag{3.6.2}$$

$m \times n$ 矩阵 A 称为方程组 (3.6.2) 的**系数矩阵**,$m \times (n+1)$ 矩阵 (A, β) 称为方程组 (3.6.2) 的**增广矩阵**. 设

$$x^0 = \bigl(x_1^0, x_2^0, \dots, x_n^0\bigr).$$

如果把 $x_1=x_1^0, x_2=x_2^0, \dots, x_n=x_n^0$ 代入方程组 (3.6.2) 的每个方程,都使每个方程成为等式,也即 $Ax^0=\beta$,则 x^0 称为方程组 (3.6.2) 的一个**解**. 方程组 (3.6.2) 的所有解的集合称为方程组 (3.6.2) 的**通解**.

如果方程组 (3.6.2) 有解,则方程组 (3.6.2) 称为**相容的**,否则称为**不相容的**.

如果方程组 (3.6.2) 的每个常数项都为零,即 $\beta=0$,则方程组 (3.6.2) 称为**齐次的**,否则称为**非齐次的**. 容易看出,齐次线性方程组总有解

$$x=(\underbrace{0,0,0,\dots,0,0}_{n\text{ 个}})^{\mathrm{T}},$$

它称为**零解**,或者**平凡解**.

如果齐次方程组的解不是零解,则称为**非零解**,或者**非平凡解**.

对于线性方程组,值得关心的问题是:线性方程组有解的必要和充分条件是什么?在有解的情形下,线性方程组何时有唯一解,何时解不唯一?在方程组有解的条件下,它的通解是什么,如何求出它的通解?

下面将采用矩阵在相抵下的标准形理论来处理这些问题. 先讨论齐次方程组 $Ax=0$.

定理 3.6.1(**齐次方程组解的结构定理**)设 $\operatorname{rank} A=r$,则当 $r=n$ 时,那么齐次线性方程组

$$Ax=0 \tag{3.6.3}$$

只有零解. 当 $r<n$ 时,齐次方程组 (3.6.3) 具有非零解,而且它的通解依赖于 $n-r$ 个独立参数.

具体地说,设

$$A=P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q,$$

其中 P 和 Q 分别是 m 阶和 n 阶可逆方阵,并且是取定的. 则齐次方程组 (3.6.3) 的通解为

$$x=t_{r+1}Q^{-1}\varepsilon_{r+1}+\cdots+t_nQ^{-1}\varepsilon_n, \tag{3.6.4}$$

其中 $t_{r+1},\dots,t_n$ 是任意的数,ε_i 是 n 维列向量,并且对于 $i=r+1,\dots,n$,

$$\varepsilon_i=(0,\dots,0,\underset{i}{1},0,\dots,0)^{\mathrm{T}}.$$

证明　设 $r<n$. 取 $x=t_{r+1}Q^{-1}\varepsilon_{r+1}+\cdots+t_nQ^{-1}\varepsilon_n$,则

$$Ax=P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q\left(t_{r+1}Q^{-1}\varepsilon_{r+1}+\cdots+t_nQ^{-1}\varepsilon_n\right)=P\sum_{i=r+1}^{n}t_i\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}\varepsilon_i.$$

显然,当 $i=r+1,\dots,n$ 时,

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}\varepsilon_i=0,$$

因此,$Ax=0$. 这表明,形如 (3.6.4) 的 x 是齐次方程组 (3.6.3) 的解.

反之,设 x^0 是齐次方程组 (3.6.3) 的解,则

$$P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Qx^0 = 0.$$

因为方阵 P 可逆,所以

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Qx^0 = 0.$$

记 $Qx^0 = y = (y_1, y_2)^{\mathrm{T}}$,其中 y_1 是 r 维行向量. 由上式得到 $(y_1, 0)^{\mathrm{T}} = 0$. 因此 $y_1 = 0$. 所以 $Qx^0 = (0, y_2)^{\mathrm{T}}$. 记 $Qx = (0, \dots, 0, t_{r+1}, \dots, t_n)^{\mathrm{T}}$,则

$$Qx^0 = t_{r+1}\varepsilon_{r+1} + t_{r+2}\varepsilon_{r+2} + \cdots + t_n\varepsilon_n.$$

因此

$$x^0 = t_{r+1}Q^{-1}\varepsilon_{r+1} + t_{r+2}Q^{-1}\varepsilon_{r+2} + \cdots + t_nQ^{-1}\varepsilon_n,$$

即 x^0 具有形式 (3.6.4). 这就证明,形如 (3.6.4) 的 x 是齐次方程组 (3.6.3) 的通解.

当 $r = n$ 时,$m \times n$ 矩阵 A 是列满秩的. 由例 3.5.1,存在 m 阶可逆方阵 P,使得

$$A = P\big(I_{(n)},\ 0\big)^{\mathrm{T}}.$$

如果 x^0 是齐次方程组 (3.6.3) 的解,则由 $Ax^0 = 0$ 得到 $P\big(I_{(n)},\ 0\big)^{\mathrm{T}} x^0 = 0$.
因为方阵 P 可逆,所以 $\big(I_{(n)},\ 0\big)^{\mathrm{T}} x^0 = 0$,即 $\big(x^0,\ 0\big)^{\mathrm{T}} = 0$,从而 $x^0 = 0$.
因此,当 $r = n$ 时,齐次方程组 (3.6.3) 只有零解. ■

现在讨论非齐次方程组.

定理 3.6.2(**非齐次方程组的相容性定理**)给定线性方程组 (3.6.2),即 $Ax = \beta$. 则方程组 (3.6.2) 有解的充分必要条件是它的系数矩阵和增广矩阵的秩相等.

证明 设 $\operatorname{rank} A = r$,则存在 m 阶和 n 阶可逆方阵 P 和 Q,使得

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q.$$

必要性 设方程组 (3.6.2) 有解 x^0. 则

$$Ax^0 = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Qx^0 = \beta.$$

因此,

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Qx^0 = P^{-1}\beta.$$

记 $Qx^0 = (y_1, y_2)^{\mathrm{T}}$,$P^{-1}\beta = (z_1, z_2)^{\mathrm{T}}$,其中 y_1 和 z_1 都是 r 维行向量. 由上式得到,

$$(y_1, 0)^{\mathrm{T}} = (z_1, z_2)^{\mathrm{T}}.$$

因此,$y_1 = z_1$,$z_2 = 0$. 所以,

$$\begin{aligned}\operatorname{rank}(A, \beta) &= \operatorname{rank} P\left(\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q,\ P^{-1}\beta\right) = \operatorname{rank} P\left(\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q,\ \begin{pmatrix} z_1^{\mathrm{T}} \\ 0 \end{pmatrix}\right) \\ &= \operatorname{rank} P\begin{pmatrix} I_{(r)} & 0 & z_1^{\mathrm{T}} \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}.\end{aligned}$$

因为方阵 P, Q 可逆,故

$$\operatorname{rank}(A, \beta) = \operatorname{rank}\begin{pmatrix} I_{(r)} & 0 & z_1^{\mathrm{T}} \\ 0 & 0 & 0 \end{pmatrix} = r = \operatorname{rank} A.$$

充分性 设 $\operatorname{rank}(A, \beta) = \operatorname{rank} A$. 由于

$$\begin{aligned}(A, \beta) &= \left(P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q, \beta\right) = P\left(\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q, P^{-1}\beta\right) \\ &= P\left(\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}, P^{-1}\beta\right)\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} = P\begin{pmatrix} I_{(r)} & 0 & z_1^{\mathrm{T}} \\ 0 & 0 & z_2^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix},\end{aligned}$$

其中 $P^{-1}\beta = (z_1, z_2)^{\mathrm{T}}$, z_1 为 r 维行向量,所以,

$$\operatorname{rank}(A, \beta) = \operatorname{rank}\begin{pmatrix} I_{(r)} & 0 & z_1^{\mathrm{T}} \\ 0 & 0 & z_2^{\mathrm{T}} \end{pmatrix} = \operatorname{rank} A = r.$$

因此,$z_2 = 0$. 取

$$x^0 = Q^{-1}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}_{n\times m} P^{-1}\beta,$$

则

$$\begin{aligned}Ax^0 &= P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q \cdot Q^{-1}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}P^{-1}\beta = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}P^{-1}\beta \\ &= P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} z_1 \\ 0 \end{pmatrix} = P\begin{pmatrix} z_1 \\ 0 \end{pmatrix} = P(P^{-1}\beta) = \beta.\end{aligned}$$

所以 x^0 是方程组 (3.6.2) 的解. ■

定理 3.6.3(**非齐次方程组解的结构定理**)设线性方程组 (3.6.2),即 $Ax = \beta$ 有解,并且 $\operatorname{rank} A = r$. 那么,

当 $r = n$ 时,方程组 (3.6.2) 的解是唯一的; 当 $r < n$ 时,方程组 (3.6.2) 的通解依赖于 $n - r$ 个独立参数,并且它的通解由方程组 (3.6.2) 的一个特解(即一个确定的解)和相应的齐次方程组 $Ax = 0$ 的通解构成.

具体地说,设系数矩阵

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q,$$

其中 P 和 Q 分别是 m 阶和 n 阶可逆方阵,并且是取定的,则方程组 (3.6.2) 的通解为

$$x = Q^{-1}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}_{n\times m} P^{-1}\beta + t_{r+1}Q^{-1}\varepsilon_{r+1} + \cdots + t_n Q^{-1}\varepsilon_n. \tag{3.6.5}$$

其中 $t_{r+1}, \ldots, t_n$ 和 $\varepsilon_1, \ldots, \varepsilon_n$ 的意义同定理 3.6.1.

证明 设 x^0 是方程组 (3.6.2) 的一个特解. 如果 x 是方程组 (3.6.2) 的一个解,则由 $Ax^0 = \beta$ 和 $Ax = \beta$ 得到,

$$A(x - x^0) = 0.$$

记 $y = x - x^0$. 上式表明,y 是齐次方程组 $Ax = 0$ 的解. 因此,x 可以表为特解

x^0 与齐次方程组 $Ax=0$ 的解 y 的和. 反之,设 y 是齐次方程组 $Ax=0$ 的解. 则由 $Ax^0=0$, $Ay=0$ 得到,

$$A(x^0+y)=\beta,$$

从而特解 x^0 与方程组 $Ax=0$ 的解 y 之和是方程组 (3.6.2) 的解. 这就证明了方程组 (3.6.2) 的通解由它的特解和相应齐次方程组的通解所构成.

由定理 3.6.2 的证明,可取方程组 (3.6.2) 的特解为

$$x^0=Q^{-1}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}_{n\times m} P^{-1}\beta,$$

于是由定理 3.6.1 和上面的结论得到: 当 $r<n$ 时,方程组 (3.6.2) 的通解为 (3.6.5); 而当 $r=n$ 时,方程组 (3.6.2) 的解是唯一的. ■

尽管上述几个定理已经完全解决了线性方程组的理论问题,但是却未涉及线性方程组求解的具体方法. 在实际求解方程组时,往往采用**消去法**.

消去法的依据是,对方程组 (3.6.2) 的增广矩阵 (A,β) 施行一次行的初等变换,得到的矩阵记为 $(\widetilde{A},\widetilde{\beta})$. 由于对矩阵施行初等行变换,相当于左乘以一个初等方阵,因此,存在可逆方阵 R,使得

$$(\widetilde{A},\ \widetilde{\beta})=R(A,\ \beta)=(RA,\ R\beta),$$

即 $\widetilde{A}=RA$, $\widetilde{\beta}=R\beta$. 把矩阵 $(\widetilde{A},\ \widetilde{\beta})$ 看成方程组

$$\widetilde{A}x=\widetilde{\beta}$$

的增广矩阵. 由于矩阵的初等变换不改变矩阵的秩,因此,根据定理 3.6.2,当且仅当方程组 $\widetilde{A}x=\widetilde{\beta}$ 有解时,方程组 $Ax=\beta$ 有解. 而且,由于矩阵 R 可逆,因此,如果 x^0 是方程组 $\widetilde{A}x=\widetilde{\beta}$ 的解,则 x^0 也是方程组 $Ax=\beta$ 的解,反之亦然. 所以,方程组 $Ax=\beta$ 和 $\widetilde{A}x=\widetilde{\beta}$ 具有相同的通解. 于是,求方程组 $Ax=\beta$ 的解就化为求方程组 $\widetilde{A}x=\widetilde{\beta}$ 的解.

不难证明,方程组的增广矩阵 $(A,\ \beta)$ 可以经过有限次行的初等变换变为如下阶梯形式

$$(\widetilde{A},\ \widetilde{\beta})=\left(\begin{array}{ccccccccc}
\widetilde{a}_{11} & \cdots & \widetilde{a}_{1,j_2-1} & 0 & \widetilde{a}_{1,j_2+1} & \cdots & \widetilde{a}_{1,j_3-1} & 0 & \widetilde{a}_{1,j_3+1} & \cdots \\
0 & \cdots & 0 & \widetilde{a}_{2j_2} & \widetilde{a}_{2,j_2+1} & \cdots & \widetilde{a}_{2,j_3-1} & 0 & \widetilde{a}_{2,j_3+1} & \cdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \widetilde{a}_{3j_3} & \widetilde{a}_{3,j_3+1} & \cdots \\
0 & & 0 & 0 & 0 & & 0 & 0 & 0 & \cdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots
\end{array}\right.$$

$$
\left.\begin{array}{cccccccccc}
\cdots & 0 & \widetilde{a}_{1,j_{r-1}+1} & \cdots & \widetilde{a}_{1,j_r-1} & 0 & \widetilde{a}_{1,j_r+1} & \cdots & \widetilde{a}_{1n} & \widetilde{b}_1 \\
 & \vdots & \widetilde{a}_{2,j_{r-1}+1} & \cdots & \widetilde{a}_{2,j_r-1} & 0 & \widetilde{a}_{2,j_r+1} & \cdots & \widetilde{a}_{2n} & \widetilde{b}_2 \\
\cdots & 0 & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \widetilde{a}_{r-1,j_{r-1}} & \widetilde{a}_{r-1,j_{r-1}+1} & \cdots & \widetilde{a}_{r-1,j_r-1} & 0 & \widetilde{a}_{r-1,j_r+1} & \cdots & \widetilde{a}_{r-1,n} & \vdots \\
0 & 0 & 0 & \cdots & 0 & \widetilde{a}_{r,j_r} & \widetilde{a}_{r,j_r+1} & \cdots & \widetilde{a}_{rn} & \widetilde{b}_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \widetilde{b}_{r+1} \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \widetilde{b}_n
\end{array}\right),
$$

其中 $1 < j_2 < \cdots < j_r \leqslant n$，$\widetilde{a}_{11}, \widetilde{a}_{2j_2}, \ldots, \widetilde{a}_{rj_r}$ 都不为零，而 $r = \operatorname{rank} A$.

如果 $\widetilde{b}_{r+1}, \ldots, \widetilde{b}_m$ 不全为零，则由定理 3.6.2，方程组 (3.6.2) 无解；

如果 $\widetilde{b}_{r+1} = \cdots = \widetilde{b}_m = 0$，那么方程组 (3.6.2) 有解，而且由 $\widetilde{A}x = \widetilde{\beta}$ 可以求出

$$
\begin{aligned}
x_1 &= \frac{\widetilde{b}_1}{\widetilde{a}_{11}} - \frac{\widetilde{a}_{12}}{\widetilde{a}_{11}}x_2 - \cdots - \frac{\widetilde{a}_{1,j_2-1}}{\widetilde{a}_{11}}x_{j_2-1} - \cdots - \frac{\widetilde{a}_{1,j_r+1}}{\widetilde{a}_{11}}x_{j_r+1} - \cdots - \frac{\widetilde{a}_{1n}}{\widetilde{a}_{11}}x_n, \\
x_{j_2} &= \frac{\widetilde{b}_2}{\widetilde{a}_{22}} - \frac{\widetilde{a}_{2,j_2+1}}{\widetilde{a}_{22}}x_{j_2+1} - \cdots - \frac{\widetilde{a}_{2,j_3-1}}{\widetilde{a}_{22}}x_{j_3-1} - \cdots - \frac{\widetilde{a}_{2,j_r+1}}{\widetilde{a}_{22}}x_{j_r+1} - \cdots - \frac{\widetilde{a}_{2n}}{\widetilde{a}_{22}}x_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
x_{j_r} &= \frac{\widetilde{b}_r}{\widetilde{a}_{rr}} - \frac{\widetilde{a}_{r,j_r+1}}{\widetilde{a}_{rr}}x_{j_r+1} - \cdots - \frac{\widetilde{a}_{rn}}{\widetilde{a}_{rr}}x_n,
\end{aligned}
$$

其中当 $i \neq 1, j_1, \ldots, j_r$ 时，x_i 是任意的. 由此即可求得方程组 (3.6.2) 的通解.

例 3.6.1　求齐次线性方程组的通解：

$$
\begin{cases}
x_1 + x_2 + x_3 + x_4 + x_5 = 0, \\
3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = 0, \\
x_2 + 2x_3 + 2x_4 + 6x_5 = 0, \\
5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = 0.
\end{cases}
$$

解　齐次方程组的系数矩阵为

$$
A = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 \\
3 & 2 & 1 & 1 & -3 \\
0 & 1 & 2 & 2 & 6 \\
5 & 4 & 3 & 3 & -1
\end{pmatrix}.
$$

用 -3 乘矩阵 A 的第 1 行并加到第 2 行，用 -5 乘矩阵 A 的第 1 行并加到第 4 行，矩阵 A 变为

$$
B = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 \\
0 & -1 & -2 & -2 & -6 \\
0 & 1 & 2 & 2 & 6 \\
0 & -1 & -2 & -2 & -6
\end{pmatrix}.
$$

矩阵 B 的第 3 行分别加到第 2 和第 4 行，再对换第 2 和第 3 行，矩阵 B 变为

$$C=\begin{pmatrix}1&1&1&1&1\\0&1&2&2&6\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

用 -1 乘矩阵 C 的第 2 行并加到第 1 行，矩阵 C 变为

$$\widetilde{A}=\begin{pmatrix}1&0&-1&-1&-5\\0&1&2&2&6\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

因此，

$$\begin{cases}x_1-\ \ x_3-\ \ x_4-5x_5=0,\\x_2+2x_3+2x_4+6x_5=0.\end{cases}$$

于是求得通解为

$$x=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}=\begin{pmatrix}x_3+x_4+5x_5\\-2x_3-2x_4-6x_5\\x_3\\x_4\\x_5\end{pmatrix}=x_3\begin{pmatrix}1\\-2\\1\\0\\0\end{pmatrix}+x_4\begin{pmatrix}1\\-2\\0\\1\\0\end{pmatrix}+x_5\begin{pmatrix}5\\-6\\0\\0\\1\end{pmatrix},$$

其中 x_3,x_4,x_5 是独立参数. ■

例 3.6.2 求线性方程组的通解：

$$\begin{cases}2x_1+\ \ 3x_2-\ \ x_3+\ \ x_4=1,\\8x_1+12x_2-9x_3+8x_4=3,\\4x_1+\ \ 6x_2+3x_3-2x_4=3,\\2x_1+\ \ 3x_2+9x_3-7x_4=3.\end{cases}$$

解 它的增广矩阵为

$$(A,\ \beta)=\begin{pmatrix}2&3&-1&1&1\\8&12&-9&8&3\\4&6&3&-2&3\\2&3&9&-7&3\end{pmatrix}.$$

分别用 $-4,-2,-1$ 乘矩阵 (A,β) 的第 1 行，并分别加到第 2, 3, 4 行，矩阵 (A,β) 变为

$$(B,\ \gamma)=\begin{pmatrix}2&3&-1&1&1\\0&0&-5&4&-1\\0&0&5&-4&1\\0&0&10&-8&2\end{pmatrix}.$$

分别用 1, 2 乘矩阵 (B,γ) 的第 2 行，并分别加到第 3、4 行，矩阵 (B,γ) 变为

$$(C,\ \xi)=\begin{pmatrix}2&3&-1&1&1\\0&0&-5&4&-1\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

用 $-\dfrac{1}{5}$ 乘矩阵 (C,ξ) 的第 2 行. 再把第 2 行加到第 1 行,得到

$$(\widetilde{A},\ \widetilde{\beta})=\begin{pmatrix}2&3&0&\frac{1}{5}&\frac{6}{5}\\0&0&1&-\frac{4}{5}&\frac{1}{5}\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

于是得到

$$\begin{cases}2x_1+3x_2\quad+\dfrac{1}{5}x_4=\dfrac{6}{5},\\ \qquad\qquad x_3-\dfrac{4}{5}x_4=\dfrac{1}{5}.\end{cases}$$

因此

$$\begin{cases}x_1=\dfrac{3}{5}-\dfrac{3}{2}x_2-\dfrac{1}{10}x_4,\\ x_3=\dfrac{1}{5}+\dfrac{4}{5}x_4.\end{cases}$$

所以,通解为

$$x=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}\frac{3}{5}-\frac{3}{2}x_2-\frac{1}{10}x_4\\x_2\\\frac{1}{5}+\frac{4}{5}x_4\\x_4\end{pmatrix}=\begin{pmatrix}\frac{3}{5}\\0\\\frac{1}{5}\\0\end{pmatrix}+x_2\begin{pmatrix}-\frac{3}{2}\\1\\0\\0\end{pmatrix}+x_4\begin{pmatrix}-\frac{1}{10}\\0\\\frac{4}{5}\\1\end{pmatrix},$$

其中 x_2,x_4 是独立参数. ■

习　题　3.6

1. 求下列齐次线性方程组的通解:

(1) $\begin{cases}x_1+x_2-3x_4-x_5=0,\\x_1-x_2-2x_3-x_4=0,\\4x_1-2x_2+6x_3+3x_4-4x_5=0,\\2x_1+4x_2-2x_3+4x_4-7x_5=0.\end{cases}$　(2) $\begin{cases}x_1+x_2+x_3+x_4+x_5=0,\\2x_1+2x_2+x_3+x_4-2x_5=0,\\x_2+2x_3+2x_4+x_5=0,\\5x_1+4x_2-3x_3+4x_4-x_5=0.\end{cases}$

2. 求非齐次线性方程组的通解:

(1) $\begin{cases}x_1-2x_2+3x_3-4x_4=4,\\x_2-x_3+x_4=-3,\\x_1+3x_2-3x_4=1,\\-7x_2+3x_3+x_4=-3.\end{cases}$　(2) $\begin{cases}x_1+x_2-3x_3=-1,\\2x_1+x_2-2x_3=1,\\x_1+x_2+x_3=3,\\x_1+2x_2-3x_3=1.\end{cases}$

3. 选择 λ 的值,使下述线性方程组有解.

$$\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1, \\ x_1 + 2x_2 - x_3 + 4x_4 = 2, \\ x_1 + 7x_2 - 4x_3 + 11x_4 = \lambda. \end{cases}$$

4. 推广定理 3.6.2 到矩阵方程上. 即证明,设给定矩阵 $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{m\times p}$,而未知矩阵 $X \in \mathbb{F}^{n\times p}$,则矩阵方程 $AX = B$ 有解的充分必要条件是 $\operatorname{rank} A = \operatorname{rank}(A, B)$,其中 (A, B) 是矩阵 A 和 B 并排而成的矩阵.

5. 设 $A \in \mathbb{F}^{m\times m}$, $X \in \mathbb{F}^{n\times p}$. 证明,矩阵方程 $AX = 0$ 有非零解的充分必要条件是方阵 A 的行列式为零.

6. 证明,如果齐次线性方程组的系数矩阵的秩比未知量的个数小 1,则该方程组的任意两个解成比例,即相差一个数值因子.

7. 设 $A \in \mathbb{F}^{n\times(n+1)}$, $X \in \mathbb{F}^{(n+1)\times n}$. 证明,矩阵方程 $AX = I_{(n)}$ 有解的充分必要条件是,矩阵 A 为行满秩的.

8. 设齐次线性方程组

$$\sum_{i=1}^{n+1} a_{ij}x_j = 0, \qquad i = 1, 2, \ldots, n$$

的系数矩阵 $A = (a_{ij})$ 是行满秩的. 证明,它的解为

$$x_j = (-1)^{n-j} t \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n+1} \\ a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2,n+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n+1} \end{pmatrix},$$

其中 $j = 1, 2, \ldots, n+1$,而 t 是独立参数.

9. 设 n 阶方阵 A 和 B 的秩分别为 r 和 $n - r$. 求矩阵方程 $AXB = 0$ 的通解.

§3.7 矩阵的广义逆

大家知道,逆矩阵只对方阵有定义,而且即便是方阵,也不是每个方阵都可逆. 本节的目的是推广逆矩阵概念为广义逆,使得每个矩阵都有广义逆.

设给定 $A \in \mathbb{F}^{m\times n}$,未知矩阵 $X = (x_{ij}) \in \mathbb{F}^{n\times m}$,其中众 x_{ij} 是未知的. 考虑矩阵方程

$$AXA = A \tag{3.7.1}$$

的解.

定理 3.7.1 矩阵方程 (3.7.1) 恒有解. 具体地说,设

$$\operatorname{rank} A = r,$$

而且

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q, \tag{3.7.2}$$

其中 P 和 Q 分别是取定的 m 阶和 n 阶可逆矩阵. 则矩阵方程 (3.7.1) 的通解为

$$X = Q^{-1}\begin{pmatrix} I_{(r)} & B \\ C & D \end{pmatrix}_{n\times m} P^{-1}, \tag{3.7.3}$$

其中 $B \in \mathbb{F}^{m\times(m-r)}$, $C \in \mathbb{F}^{(n-r)\times r}$ 和 $D \in \mathbb{F}^{(n-r)\times(m-r)}$ 是任意的.

证明　把形如 (3.7.3) 的矩阵 X 代入矩阵方程 (3.7.1), 即可验证, 形如 (3.7.3) 的矩阵 X 的确是矩阵方程 (3.7.1) 的解. 反之, 设矩阵 X 是矩阵方程 (3.7.1) 的解, 则由 (3.7.2)

$$P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} QXP\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q.$$

因为方阵 P 和 Q 可逆, 所以,

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} QXP\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

记

$$QXP = \begin{pmatrix} E & B \\ C & D \end{pmatrix}$$

其中 E, B, C, D 分别是 $r \times r$, $r \times (m-r)$, $(n-r) \times r$ 和 $(n-r) \times (m-r)$ 矩阵. 因此,

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} E & B \\ C & D \end{pmatrix}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

所以, $E = I_{(r)}$. 于是,

$$QXB = \begin{pmatrix} I_{(r)} & B \\ C & D \end{pmatrix},$$

即

$$X = Q^{-1}\begin{pmatrix} I_{(r)} & B \\ C & D \end{pmatrix} P^{-1}. \qquad \blacksquare$$

应当指出, 当 A 为可逆方阵时, 矩阵方程 (3.7.1) 的解显然为 $X = A^{-1}$. 于是引出如下的定义.

定义 3.7.1　矩阵方程 $AXA = A$ 的解 X 称为矩阵 A 的**广义逆**, 记为 A^-.

由定理 3.7.1 可以看出, 任意矩阵 A 的广义逆 A^- 总是存在的. 而且一般地说, 矩阵 A 的广义逆 A^- 并不唯一. 事实上, 由于

$$A^- = Q^{-1}\begin{pmatrix} I_{(r)} & B \\ C & D \end{pmatrix} P^{-1},$$

其中矩阵 B, C 和 D 是任意的, 所以, $\operatorname{rank} A^- \geqslant \operatorname{rank} A$. 而且对任意正整数 k, $r \leqslant k \leqslant \min\{m, n\}$, 总可以分别取 B 和 C 为 $r \times (m-r)$ 和 $(n-r) \times r$ 零矩阵, 并取 D 为秩等于 $k - r$ 的 $(n-r) \times (m-r)$ 矩阵, 则矩阵 A 的这个广义逆 A^- 的秩 $\operatorname{rank} A^- = k$.

由定理 3.7.1 还可以得到, 矩阵 A 具有唯一的广义逆 A^- 的充分必要条件是, 矩阵 A 为可逆方阵. 现在给出矩阵的广义逆的一些应用.

例 3.7.1 (非齐次线性方程组的相容性定理) 证明方程 $Ax = \beta$ 有解的充分必要

条件是 $\beta = AA^-\beta$,其中 A 是 $m\times n$ 矩阵,β 是 $m\times 1$ 矩阵,x 是 $n\times 1$ 未知矩阵,而 A^- 是矩阵 A 的广义逆.

证明 设方程 $Ax=\beta$ 有解 x^0,则 $\beta = Ax^0$. 因此,

$$AA^-\beta = AA^-(Ax^0) = (AA^-A)x^0 = Ax^0 = \beta.$$

反之,设 $\beta = AA^-\beta$ 成立. 取 $x^0 = A^-\beta$. 于是 $Ax^0 = AA^-\beta = \beta$. ■

例 3.7.2(**非齐次线性方程组解的结构定理**)设方程 $Ax=\beta$ 有解. 则它的通解为

$$x = A^-\beta + \left(I_{(n)} - AA^-\right)z,$$

其中 A^- 是矩阵 A 的某个取定的广义逆,而 z 是任意 $n\times 1$ 矩阵.

证明 因为方程 $Ax=\beta$ 有解,因此由例 3.7.1,对于矩阵 A 的取定的广义逆 A^-,$AA^-\beta=\beta$. 取 $x = A^-\beta + \left(I_{(n)} - AA^-\right)z$,则

$$Ax = AA^-\beta + A\left(I_{(n)} - A^-A\right)z = \beta + (A - AA^-A)z = \beta.$$

即 $x = A^-\beta + \left(I_{(n)} - AA^-\right)z$ 是方程 $Ax=\beta$ 的解.

反之,设 x^0 是方程 $Ax=\beta$ 的解,即 $Ax^0=\beta$. 取 $z = x^0$,则

$$A^-\beta + \left(I_{(n)} - A^-A\right)x^0 = A^-Ax^0 + x^0 - A^-Ax^0 = x^0,$$

即 x^0 可表为所说的形式. ■

例 3.7.3(**齐次线性方程组解的结构定理**)方程 $Ax=0$ 恒有解,而且它的通解为

$$x = \left(I_{(n)} - AA^-\right)z,$$

其中 A^- 是矩阵 A 的某个取定的广义逆,而 z 是任意的 $n\times 1$ 矩阵.

证明 这是例 3.7.1 和例 3.7.2 的特殊情形. ■

对非齐次线性方程组 $Ax=\beta,\beta\neq 0$,还可以用广义逆给出另一种形式的通解.

例 3.7.4 设方程 $Ax=\beta$ 有解,其中 $\beta\neq 0$. 则它的通解为 $x=A^-\beta$,这里 A^- 是矩阵 A 的任意一个广义逆.

证明 因为方程 $Ax=\beta$ 有解,因此可设它的一个解为 x^0. 取 $x=A^-\beta$,则

$$Ax = AA^-\beta = AA^-(Ax^0) = (AA^-A)x^0 = Ax^0 = \beta.$$

因此,对矩阵 A 的任意一个广义逆 A^-,$x=A^-\beta$ 都是方程 $Ax=\beta$ 的解.

反之,设 x^0 是方程 $Ax=\beta$ 的解,下面将证明,存在矩阵 A 的一个广义逆 A^-,使得 $x^0 = A^-\beta$. 事实上,设 $\operatorname{rank} A = r$,且

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q.$$

由定理 3.7.1,

$$A^- = Q^{-1}\begin{pmatrix} I_{(r)} & B \\ C & D \end{pmatrix}_{m\times n} P^{-1}.$$

于是，由 $Ax^0=\beta$ 得到，

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Qx^0 = P^{-1}\beta.$$

记

$$Qx^0 = y = (y_1, y_2)^{\mathrm{T}}, \qquad P^{-1}\beta = z = (z_1, z_2)^{\mathrm{T}},$$

其中 y_1 和 z_1 是 $1\times r$ 矩阵．由上式得到，$y_1=z_1, z_2=0$．由于方阵 P 可逆，$\beta\neq 0$，因此 $z=P^{-1}\beta\neq 0$，所以，z_1 至少有一个元素不为零．记 $z_1=(c_1,c_2,\ldots,c_r)$，且设 $c_i\neq 0$．在

$$A^- = Q^{-1}\begin{pmatrix} I_{(r)} & B \\ C & D \end{pmatrix} P^{-1}$$

中，取矩阵 B 和 D 分别是 $r\times(m-r)$ 和 $(n-r)\times(m-r)$ 零矩阵，取 $(n-r)\times r$ 矩阵 C 的第 i 列为 $c_i^{-1}y_2^{\mathrm{T}}$，其它列为零．所得到的广义逆仍记为 A^-．于是

$$\begin{aligned} A^-\beta &= Q^{-1}\begin{pmatrix} I_{(r)} & 0 \\ C & 0 \end{pmatrix}_{m\times n} P^{-1}\beta = Q^{-1}\begin{pmatrix} I_{(r)} & 0 \\ C & 0 \end{pmatrix}\begin{pmatrix} z_1^{\mathrm{T}} \\ z_2^{\mathrm{T}} \end{pmatrix} \\ &= Q^{-1}(y_1, y_2)^{\mathrm{T}} = Q^{-1}Qx^0 = x^0, \end{aligned}$$

这就证明，方程 $Ax=\beta$ 的通解为 $x=A^-\beta$．■

这几个例子表明，利用矩阵的广义逆来讨论线性方程组的解是非常方便的，而且通解的形式特别简洁．特别，例 3.7.4 表明，当 $\beta\neq 0$ 时，方程 $Ax=\beta$ 的通解为 $x=A^-\beta$．这和方程 $Ax=\beta$ 的系数矩阵是可逆方阵的情形相类似，因为这时方程 $Ax=\beta$ 的解为 $x=A^{-1}\beta$．由此可以看出广义逆的威力．

矩阵的广义逆不只是上面所说的一种类型．还有许多其它类型的广义逆．除上面的 A^- 外，矩阵的 Moore-Penrose 广义逆也是经常遇到的．

考虑矩阵方程组

$$(\mathrm{P})\quad \begin{cases} AXA = A, & (\mathrm{P1}) \\ XAX = X, & (\mathrm{P2}) \\ \left(\overline{AX}\right)^{\mathrm{T}} = AX, & (\mathrm{P3}) \\ \left(\overline{XA}\right)^{\mathrm{T}} = XA, & (\mathrm{P4}) \end{cases}$$

其中 $m\times n$ 矩阵 A 是给定的，而 $n\times m$ 矩阵 X 是未知的．

方程组 (P) 称为 **Penrose 方程组．**

定理 3.7.2　对任意给定的 $m\times n$ 矩阵 A，Penrose 方程组 (P) 总有解，而且它的解唯一．具体地说，设矩阵 $A=BC$，其中 B 和 C 分别是列满秩和行满秩矩阵，则 Penrose 方程组 (P) 的唯一解为

$$X = \overline{C}^{\mathrm{T}}\left(C\overline{C}^{\mathrm{T}}\right)^{-1}\left(\overline{B}^{\mathrm{T}}B\right)^{-1}\overline{B}^{\mathrm{T}}. \tag{3.7.4}$$

证明　把式 (3.7.4) 代入 Penrose 方程组 (P) 的每个方程，容易看出，每个方程都成为等式，即矩阵 X 的确是 Penrose 方程组 (P) 的解．

设矩阵 X_1 和 X_2 都是 Penrose 方程组 (P) 的解．那么，

$$
\begin{aligned}
X_1 &= X_1AX_1 && \text{由方程 (P2)}\\
&= X_1AX_2AX_1 && \text{由方程 (P1)}\\
&= X_1\left(\overline{AX_2}\right)^{\mathrm{T}}\left(\overline{AX_1}\right)^{\mathrm{T}} = X_1\left(\overline{AX_1AX_2}\right)^{\mathrm{T}} && \text{由方程 (P3)}\\
&= X_1\left(\overline{AX_2}\right)^{\mathrm{T}} && \text{由方程 (P1)}\\
&= X_1AX_2 && \text{由方程 (P3)}\\
&= X_1AX_2AX_2 && \text{由方程 (P1)}\\
&= \left(\overline{X_1A}\right)^{\mathrm{T}}\left(\overline{X_2A}\right)^{\mathrm{T}}X_2 = \left(\overline{X_2AX_1A}\right)^{\mathrm{T}}X_2 && \text{由方程 (P4)}\\
&= \left(\overline{X_2A}\right)^{\mathrm{T}}X_2 && \text{由方程 (P1)}\\
&= X_2AX_2 && \text{由方程 (P4)}\\
&= X_2. && \text{由方程 (P1)}
\end{aligned}
$$

这就证明了 Penrose 方程组 (P) 的解的唯一性. ■

有必要说明的是式 (3.7.4) 的由来. 事实上,设 $\operatorname{rank} A = r$,则存在 m 阶和 n 阶可逆方阵 P 和 Q,使得

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q = P\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix}_{m\times r} \left(I_{(r)}, 0\right)_{r\times n} Q = BC,$$

其中

$$B = P\begin{pmatrix} I_{(r)} \\ 0 \end{pmatrix}_{m\times r}, \qquad C = \left(I_{(r)}, 0\right)_{r\times n} Q,$$

分别是列满秩和行满秩矩阵. 由定理 3.7.1,方程 (P1) 的解为

$$X = Q^{-1}\begin{pmatrix} I_{(r)} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}_{n\times m} P^{-1},$$

其中 X_{12}, X_{21} 和 X_{22} 分别是 $r\times(m-r), (n-r)\times r$ 和 $(n-r)\times(m-r)$ 矩阵.

把矩阵 X 代入方程 (P2),得到,

$$
\begin{aligned}
&Q^{-1}\begin{pmatrix} I_{(r)} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}P^{-1}P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}QQ^{-1}\begin{pmatrix} I_{(r)} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}P^{-1}\\
&\qquad = Q^{-1}\begin{pmatrix} I_{(r)} & X_{12} \\ X_{21} & X_{21}X_{12} \end{pmatrix}P^{-1} = Q^{-1}\begin{pmatrix} I_{(r)} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}P^{-1}.
\end{aligned}
$$

由此得到,$X_{22} = X_{21}X_{12}$. 因此,

$$X = Q^{-1}\begin{pmatrix} I_{(r)} & X_{12} \\ X_{21} & X_{21}X_{12} \end{pmatrix}P^{-1}.$$

把矩阵 A 和 X 分别代入方程 (P3) 和 (P4),得到,

$$\begin{pmatrix} I_{(r)} & 0 \\ \overline{X}_{12}^{\mathrm{T}} & 0 \end{pmatrix}\overline{P}^{\mathrm{T}}P = \overline{P}^{\mathrm{T}}P\begin{pmatrix} I_{(r)} & X_{12} \\ 0 & 0 \end{pmatrix}, \tag{3.7.5}$$

$$\begin{pmatrix} I_{(r)} & 0 \\ X_{21} & 0 \end{pmatrix}Q\overline{Q}^{\mathrm{T}} = Q\overline{Q}^{\mathrm{T}}\begin{pmatrix} I_{(r)} & \overline{X}_{12}^{\mathrm{T}} \\ 0 & 0 \end{pmatrix}. \tag{3.7.6}$$

记
$$\overline{P}^{\mathrm{T}}P=\begin{pmatrix}R_{11}&R_{12}\\ \overline{R}_{12}^{\mathrm{T}}&R_{22}\end{pmatrix}_{m\times m},\qquad Q\overline{Q}^{\mathrm{T}}=\begin{pmatrix}S_{11}&S_{12}\\ \overline{S}_{12}^{\mathrm{T}}&S_{22}\end{pmatrix}_{n\times n},$$
其中 R_{11} 和 S_{11} 都是 r 阶方阵. 由于矩阵 R_{11} 是矩阵 $\overline{P}^{\mathrm{T}}$ 的前 r 行和矩阵 P 的前 r 列的乘积,因此,
$$R_{11}=\left(I_{(r)},0\right)_{m\times m}\overline{P}^{\mathrm{T}}P\begin{pmatrix}I_{(r)}\\0\end{pmatrix}_{m\times m}=\overline{B}^{\mathrm{T}}B.$$

另一方面,由 Binet-Cauchy 公式,
$$\begin{aligned}\det R_{11}&=\overline{P}^{\mathrm{T}}P\begin{pmatrix}1&2&\cdots&r\\1&2&\cdots&r\end{pmatrix}=\sum_{1\leqslant i_1<\cdots<i_r\leqslant m}\overline{P}^{\mathrm{T}}\begin{pmatrix}1&2&\cdots&r\\i_1&i_2&\cdots&i_r\end{pmatrix}P\begin{pmatrix}i_1&i_2&\cdots&i_r\\1&2&\cdots&r\end{pmatrix}\\&=\sum_{1\leqslant i_1<\cdots<i_r\leqslant m}\left|P\begin{pmatrix}i_1&i_2&\cdots&i_r\\1&2&\cdots&r\end{pmatrix}\right|^2.\end{aligned}$$

如果 $\det R_{11}=0$,则由上式,方阵 P 的前 r 列上的每个 r 阶子式都为零. 因此,对行列式 $\det P$ 的前 r 列作 Laplace 展开,可以看出,$\det P=0$,和方阵 P 可逆矛盾. 所以,方阵 R_{11} 可逆. 同理可证,方阵 S_{11} 可逆,并且 $S_{11}=C\overline{C}^{\mathrm{T}}$.

把分块方阵 $\overline{P}^{\mathrm{T}}P$ 和 $Q\overline{Q}^{\mathrm{T}}$ 分别代入式 (3.7.5) 和式 (3.7.6),得到,
$$X_{12}=R_{11}^{-1}R_{12},\qquad X_{21}=\overline{S}_{12}^{\mathrm{T}}S_{11}^{-1}.$$
因此
$$\begin{aligned}X&=Q^{-1}\begin{pmatrix}I_{(r)}&R_{11}^{-1}R_{12}\\ \overline{S}_{12}^{\mathrm{T}}S_{11}^{-1}&\overline{S}_{12}^{\mathrm{T}}S_{11}^{-1}R_{11}^{-1}R_{12}\end{pmatrix}P^{-1}=Q^{-1}\begin{pmatrix}I_{(r)}\\ \overline{S}_{12}^{\mathrm{T}}S_{11}^{-1}\end{pmatrix}\left(I_{(r)},R_{11}^{-1}R_{12}\right)P^{-1}\\&=Q^{-1}\begin{pmatrix}S_{11}&S_{12}\\ \overline{S}_{12}^{\mathrm{T}}&S_{22}\end{pmatrix}\begin{pmatrix}S_{11}^{-1}\\0\end{pmatrix}\left(R_{11}^{-1},0\right)\begin{pmatrix}R_{11}&R_{12}\\ \overline{R}_{12}^{\mathrm{T}}&R_{22}\end{pmatrix}P^{-1}\\&=Q^{-1}\left(Q\overline{Q}^{\mathrm{T}}\right)\begin{pmatrix}I_{(r)}\\0\end{pmatrix}S_{11}^{-1}R_{11}^{-1}\left(I_{(r)},0\right)\left(\overline{P}^{\mathrm{T}}P\right)P^{-1}\\&=\left(\overline{Q}^{\mathrm{T}}\begin{pmatrix}I_{(r)}\\0\end{pmatrix}\right)S_{11}^{-1}R_{11}^{-1}\left(\left(I_{(r)},0\right)\overline{P}^{\mathrm{T}}\right)=\overline{C}^{\mathrm{T}}\left(C\overline{C}^{\mathrm{T}}\right)^{-1}\left(\overline{B}^{\mathrm{T}}B\right)^{-1}\overline{B}^{\mathrm{T}}.\end{aligned}$$
这表明,如果矩阵 X 是 Penrose 方程组 (P) 的解,则矩阵 X 应具有形式 (3.7.4).

基于定理 3.7.2,可以引进如下定义.

定义 3.7.2　对于给定的 $m\times n$ 矩阵 A,Penrose 方程组 (P) 的解 X 称为矩阵 A 的 **Moore-Penrose 广义逆**,记为 A^+.

显然当 A 为 n 阶可逆方阵时,Penrose 方程组 (P) 的解为 A^{-1},故此时 $A^+=A^{-1}$.

定理 3.7.2 表明,对任意给定的 $m\times n$ 矩阵 A,它的 Moore-Penrose 广义逆 A^+ 总是存在的. 甚至对 $m\times n$ 零矩阵 0,它的 Moore-Penrose 广义逆 0^+ 也存在,并且 $0^+=0$,这里 0 是 $n\times m$ 零矩阵. 特别,一阶零矩阵即是数 0,所以,数 0 的 Moore-Penrose 广义逆为数 0.

Moore-Penrose 广义逆和逆矩阵的有些性质是相同的,例如,
$$(A^+)^+=A,\qquad \left(\overline{A}^{\mathrm{T}}\right)^+=\overline{(A^+)}^{\mathrm{T}},\qquad (\lambda A)^+=\lambda^+A^+.$$

但是,也有一些性质不同,例如,对逆矩阵成立的穿脱原理对 Moore-Penrose 广义逆并不成立. 使用时还须留意,不能混同.

广义逆的概念早在 1920 年即已出现. 1935 年,E. H. Moore 作了系统的研究[1]. 但是,由于当时应用不广,故有湮没的危险. 直到 1955 年,R. Penrose 又重新研究了广义逆. 由于近年来广义逆的应用日趋广泛,特别是在数理统计和计算数学等的应用,它才引起普遍重视. 有兴趣的读者可以参阅有关专著(例如,I. Ben 的名著 *Generalized Inverses*),这里不拟深入介绍了.

习 题 3.7

1. 设 A 和 B 分别是 $m\times n$ 和 $m\times p$ 矩阵,X 是 $n\times p$ 未知矩阵. 证明,矩阵方程 $AX=B$ 有解的充分必要条件是,

$$B=AA^{-}B.$$

在有解时,它的通解为

$$X=A^{-}B+\left(I_{(n)}-A^{-}A\right)W,$$

其中 W 是任意的 $n\times p$ 矩阵.

2. 设 A,B 和 C 分别是 $m\times n,p\times q$ 和 $m\times q$ 矩阵,X 是 $n\times p$ 未知矩阵. 证明,矩阵方程 $AXB=C$ 有解的充分必要条件是

$$\left(I_{(m)}-AA^{-}\right)C=C\left(I_{(q)}-B^{-}B\right)=0.$$

并且在有解时,它的通解为

$$X=A^{-}CB^{-}+\left(I_{(n)}-A^{-}A\right)Y+Z\left(I_{(p)}-BB^{-}\right)+\left(I_{(n)}-A^{-}A\right)W\left(I_{(p)}-BB^{-}\right),$$

其中 Y,Z 和 W 是任意的 $n\times p$ 矩阵.

3. 设 A,B 和 C 分别是 $m\times p,q\times n$ 和 $m\times n$ 矩阵,X 和 Y 分别是 $p\times n$ 和 $m\times q$ 未知矩阵. 证明,方程 $AX-YB=C$ 有解的充分必要条件是

$$\left(I_{(m)}-AA^{-}\right)C\left(I_{(n)}-B^{-}B\right)=0.$$

而且当有解时,它的通解为

$$\begin{aligned}&X=A^{-}C+A^{-}ZB+\left(I_{(p)}-AA^{-}\right)W,\\&Y=-\left(I_{(m)}-AA^{-}\right)CB^{-}+Z-\left(I_{(m)}-AA^{-}\right)ZBB^{-},\end{aligned}$$

其中 W 和 Z 分别是任意的 $p\times n$ 和 $m\times q$ 矩阵.

4. 证明,存在 $m\times k$ 矩阵 A 和 $\ell\times n$ 矩阵的广义逆 A^{-} 和 B^{-},使得

$$\operatorname{rank}\begin{pmatrix}A&C\\0&B\end{pmatrix}=\operatorname{rank}A+\operatorname{rank}B+\operatorname{rank}\left(\left(I_{(m)}-AA^{-}\right)C\left(I_{(n)}-B^{-}B\right)\right),$$

其中 C 是 $m\times n$ 矩阵.

5. 验证 $\overline{(A^{+})}^{\mathrm{T}}=\left(\overline{A}^{\mathrm{T}}\right)^{+}$.

[1] E. H. Moore, *General Analysis*, Vol. 1, *Mem. Amer. Phil. Soc.*, Vol. 1, Philadel-Phia, 1935, P.8 and ch.3 §28.

第四章

线性空间

- ✍ 正如数学百科辞典所指出,线性代数就是线性空间的理论. 因此,线性空间是线性代数研究的最基本的几何对象,矩阵则是研究线性空间中各种几何问题的最有效的代数工具,而线性空间中各种几何问题都可以归结为相应的矩阵分类问题. 这就是线性代数的两大基本理论——矩阵理论与线性空间理论之间的实质性联系.
- ✍ §4.1 引进了数域 $\mathbb{F}$ 上的抽象线性空间概念.
- ✍ §4.2 处理了线性空间中向量间的最基本关系: 向量间的线性相关性与线性无关性,从而可以从向量的观点来看待矩阵的秩.
- ✍ §4.3 介绍了基与坐标的概念,在一组固定的基下,每一个向量都可以用坐标明确地表达出来. 利用基的概念,得到了线性空间的维数.
- ✍ §4.4 考虑同一个向量在不同基下坐标之间的关系. 这就同矩阵建立了联系. 这是将线性空间的问题与矩阵问题联系起来的关键.
- ✍ §4.5 说明了数域 $\mathbb{F}$ 上的 n 维线性空间实质上就是数域 $\mathbb{F}$ 上的 n 元数组空间.
- ✍ 最后三节介绍了与线性空间有关的一些最基本几何概念——子空间,商空间以及子空间的直和等等,为以后进一步阐述线性空间的理论奠定必要的基础.

§4.1 线性空间的定义

从第 2 章我们知道,在所有有序 n 元实数组(即 n 维实向量)集合 $\mathbb{R}^n$ 中,可以定义向量的加法,纯量与向量的乘法; 向量的加法满足结合律,交换律,有零向量,而且对每个向量 α,都存在负向量 $-\alpha$; 纯量与向量的乘法满足结合律,$1\cdot\alpha=\alpha$,以及分配律

$$\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta,\qquad (\lambda+\mu)\alpha=\lambda\alpha+\mu\alpha,$$

其中 λ 与 μ 为纯量(实数),α 为向量.

对于所有有序 n 元复数组(即 n 维复向量)集合 $\mathbb{C}^n$,同样可以定义向量的加法,纯量与向量的乘法,而且这两种运算所具有的性质和 $\mathbb{R}^n$ 的两种运算相同.

把集合的两种运算连同它们满足的公理加以概括抽象,便引出线性空间概念.

定义 4.1.1 设 V 是一个非空集合,它的元素称为**向量**. 设 $\mathbb{F}$ 是一个数域,它的元素称为**纯量**. 在 V 中定义了**向量的加法**,即对任意 $\alpha,\beta\in V$,V 中有唯一的向量 $\alpha+\beta$ 与之对应,向量 $\alpha+\beta$ 称为向量 α 与 β 的和; 在 V 中还定义了**纯量与向量的乘法**,即对任意纯量 $\lambda\in\mathbb{F}$,向量 $\alpha\in V$,V 中有唯一的向量 $\lambda\alpha$ 与之对应,向量 $\lambda\alpha$

称为纯量 λ 与向量 α 的**积**.

设 V 的向量加法,纯量与向量的乘法满足以下公理:

(A1) 加法结合律

$$(a+b)+c=a+(b+c);$$

(A2) 加法交换律

$$a+b=b+a;$$

(A3) 具有零向量 即 V 中存在向量 0,它称为**零向量**,使得对任意 $\alpha\in V$,

$$\alpha+0=0;$$

(A4) 具有负向量 即对任意 $\alpha\in V$,存在 $-\alpha\in V$,它称为向量 α 的**负向量**,使得

$$\alpha+(-\alpha)=0;$$

(M1)

$$\lambda(\mu\alpha)=(\lambda\mu)\alpha;$$

(M2)

$$1\cdot\alpha=\alpha;$$

(D1) 乘法对向量加法的**分配律**

$$\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta;$$

(D2) 乘法对纯量加法的**分配律**

$$(\lambda+\mu)\alpha=\lambda\alpha+\mu\alpha,$$

其中 $\lambda,\mu\in\mathbb{F},\alpha,\beta,\gamma\in V$,则集合 V 称为**数域 $\mathbb{F}$ 上的线性空间**.

特别,如果数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$,则 V 称为**实线性空间**; 如果数域 $\mathbb{F}$ 为复数域 $\mathbb{C}$,则 V 称为**复线性空间**.

例 4.1.1 对所有复数的集合 $\mathbb{C}$,取数域 $\mathbb{F}$ 为复数域 $\mathbb{C}$,向量加法取为通常复数的加法,纯量与向量的乘法取为通常复数的乘法.

容易验证,复数集合 $\mathbb{C}$ 的向量加法,纯量与向量的乘法满足线性空间定义中的八条公理,所以复数集合 $\mathbb{C}$ 是复线性空间.

对复数集合 $\mathbb{C}$,取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$,向量加法取为复数的加法,纯量与向量的乘法取为实数与复数的乘法.

容易验证,对集合 $\mathbb{C}$,线性空间定义中的八条公理成立. 所以集合 $\mathbb{C}$ 是实线性空间.

例 4.1.1 说明,对于同一个集合 V,只要数域 $\mathbb{F}$ 不同,作为线性空间也不同.

例 4.1.2 所有正实数的集合记为 $\mathbb{R}_+$. 取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$. 在集合 $\mathbb{R}_+$ 中规定向量加法为通常实数的加法,纯量与向量的乘法规定为通常实数的乘法. 在这两种运算下,集合 $\mathbb{R}_+$ 不构成实线性空间,因为非负数与正实数的乘积为非负数,即纯量与 $\mathbb{R}_+$ 中的向量的乘积并不封闭.

对集合 $\mathbb{R}_+$,仍取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$,在集合 $\mathbb{R}_+$ 中规定向量加法为通常实数的

乘法,即对任意 $\alpha,\beta\in\mathbb{R}_+$,规定向量 α 与 β 的和 $\alpha\oplus\beta=\alpha\beta$. 并规定纯量 $\lambda\in\mathbb{R}$ 与向量 $\alpha\in\mathbb{R}_+$ 的乘积 $\lambda\otimes\alpha=\alpha^{\lambda}$.

可以验证,集合 $\mathbb{R}_+$ 的这两种运算满足线性空间定义中的八条公理,其中零向量是实数1,向量 α 的负向量是 α^{-1}. 所以 $\mathbb{R}_+$ 是实线性空间.

例 4.1.2 说明,一个集合 V 是否成为线性空间与集合 V 的向量加法以及纯量与向量的乘法如何规定密切相关.

例 4.1.3 对数域 $\mathbb{F}$ 上的所有关于未定元 x 的多项式集合 $\mathbb{F}[x]$ 与数域 $\mathbb{F}$,规定向量加法为多项式加法,纯量与向量的乘法为数与多项式的乘法.

容易验证,集合 $\mathbb{F}[x]$ 是数域 $\mathbb{F}$ 上的线性空间,其中零向量是零多项式.

设 n 是正整数,$\mathbb{F}_n[x]$ 是数域 $\mathbb{F}$ 上的所有次数小于 n 的多项式的集合. 在多项式加法以及数与多项式的乘法下,集合 $\mathbb{F}_n[x]$ 成为数域 $\mathbb{F}$ 上的线性空间.

例 4.1.4 区间 $[0,1]$ 上的所有连续实函数的集合记为 C^2,取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$,在集合 C^2 中规定向量加法为函数的和,纯量与向量的乘法为实数与函数的乘积. 在这两种运算下,集合 C^2 成为实线性空间.

例 4.1.5 对数域 $\mathbb{F}$ 上的所有 $m\times n$ 矩阵的集合 $\mathbb{F}^{m\times n}$ 与数域 $\mathbb{F}$,规定向量加法为矩阵的加法,纯量与向量的乘法为数域 $\mathbb{F}$ 中的数与矩阵的乘法,集合 $\mathbb{F}^{m\times n}$ 使成为数域 $\mathbb{F}$ 上的线性空间.

应当指出,在线性空间定义中,并未考虑那八条公理的独立性. 事实上可以证明,如果对集合 V 中所规定的向量加法,纯量与向量的乘法,除公理 (A2) 外,其它公理都满足,那么对所规定的运算,公理 (A2) 也满足. 由于线性空间定义中的八条公理是经常使用的,所以在一般教科书中都把它们全部列出.

根据线性空间的定义可以证明,数域 $\mathbb{F}$ 上的线性空间具有以下性质:

性质 4.1.1 对任意有限多个向量作加法时,其和与向量的结合方式以及向量的先后次序无关.

在线性空间的定义中,只规定了两个向量的和. 至于多个向量的和则未加定义. 怎样规定多个向量的和?

先看四个向量的情形. 设 $\alpha,\beta,\gamma,\delta\in V$,由向量加法的定义,$\alpha+\beta$ 与 $\gamma+\delta$ 有意义,从而 $(\alpha+\beta)+(\gamma+\delta)$ 有意义; 又 $\beta+\gamma$ 有意义,因此,$\alpha+(\beta+\gamma)$ 有意义,所以,$\big(\alpha+(\beta+\gamma)\big)+\delta$ 也有意义. 于是得到向量 $(\alpha+\beta)+(\gamma+\delta)$ 与 $\big(\alpha+(\beta+\gamma)\big)+\delta$.

当然,还可以按照其它结合方式,先求出向量 $\alpha,\beta,\gamma,\delta$ 中某两个向量的和,再求出它同其它向量的和,最后求出这四个向量的和. 问题是,随着结合方式不同,这四个向量的和是否相同?

性质 4.1.1 断言,不论这四个向量的结合方式以及向量的先后次序,所得到的和总是相等的. 这样就可以把这个和规定为这四个向量的和.

性质 4.1.1 的证明 首先归纳定义向量 $\alpha_1,\alpha_2,\ldots,\alpha_k\in V$ 的标准和 $\alpha_1\oplus\alpha_2\oplus$

$\cdots\oplus\alpha_k$ 如下：当 $k=2$ 时，定义 $\alpha_1\oplus\alpha_2=\alpha_1+\alpha_2$. 假设 $\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_{k-1}$ 已经定义，则定义

$$\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_k=(\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_{k-1})+\alpha_k.$$

现在证明，

$$(\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_\ell)+(\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+k})=\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_{k+\ell}. \tag{4.1.1}$$

其中 k 与 ℓ 是正整数. 为此，对 k 用归纳法. 当 $k=1$ 时，由标准和的定义，

$$(\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_\ell)+\alpha_{\ell+1}=\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_{\ell+1}.$$

因此，当 $k=1$ 时式 (4.1.1) 成立.

假设式 (4.1.1) 对 $k-1$ 成立. 则由标准和的定义及公理 (A1)，

$$\begin{aligned}(\alpha_1\oplus\cdots\oplus\alpha_\ell)+(\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+k})&=(\alpha_1\oplus\cdots\oplus\alpha_\ell)+\big((\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+k-1})+\alpha_{\ell+k}\big)\\&=\big((\alpha_1\oplus\cdots\oplus\alpha_\ell)+(\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+k-1})\big)+\alpha_{\ell+k}.\end{aligned}$$

由归纳假设，

$$(\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_\ell)+(\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+k})=(\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_{\ell+k-1})+\alpha_{\ell+k}.$$

由标准和的定义，

$$(\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_\ell)+(\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+k})=\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_{\ell+k}.$$

这就证明了式 (4.1.1).

其次证明，对任意给定的向量 $\alpha_1,\alpha_2,\dots,\alpha_k\in V$，不论向量 $\alpha_1,\alpha_2,\dots,\alpha_k$ 的结合方式，所得到的和等于这 k 个向量的标准和.

事实上，假设 $\alpha_1,\alpha_2,\dots,\alpha_k$ 按一种结合方式求得一个和. 显然，这个和是某个向量 β 与 γ 的和 $\beta+\gamma$. 这里，β 是向量 $\alpha_1,\alpha_2,\dots,\alpha_\ell$ 按某种结合方式得到的和，而 γ 是向量 $\alpha_{\ell+1},\dots,\alpha_{\ell+r}$ 按某种结合方式得到的和，$\ell+r=k$.

对向量个数 k 用归纳法，可以假设，$\beta=\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_\ell$，$\gamma=\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+r}$. 因此由式 (4.1.1)，

$$\beta+\gamma=(\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_\ell)+(\alpha_{\ell+1}\oplus\cdots\oplus\alpha_{\ell+r})=\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_k.$$

这就证明，对于向量 $\alpha_1,\alpha_2,\dots,\alpha_k$，不论按何种结合方式，得到的和都等于标准和 $\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_k$.

最后，利用公理 (A2) 与标准和，可以证明，任意调动向量 $\alpha_1,\alpha_2,\dots,\alpha_k$ 的次序，并按任意一种结合方式，得到的和仍等于标准和 $\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_k$. ■

向量 $\alpha_1,\alpha_2,\dots,\alpha_k$ 的标准和 $\alpha_1\oplus\alpha_2\oplus\cdots\oplus\alpha_k$ 定义为 $\alpha_1,\alpha_2,\dots,\alpha_k$ 的和，记为

$$\alpha_1+\alpha_2+\cdots+\alpha_k.$$

性质 4.1.2 零向量是唯一的.

证明 设 $0',0''$ 是线性空间 V 的零向量，则由零向量的定义，

$$0'=0'+0''=0''.$$ ■

性质 4.1.3　对每个向量 $\alpha \in V$，负向量 $-\alpha$ 是唯一的.

证明　设向量 $\beta, \gamma \in V$ 是向量 α 的负向量，则由负向量的定义，

$$\beta = \beta + 0 = \beta + (\alpha + \gamma) = (\beta + \alpha) + \gamma = 0 + \gamma = \gamma. \qquad \blacksquare$$

利用负向量概念，在线性空间 V 中可以引进**减法**，即定义向量 $\alpha, \beta \in V$ 的差

$$\alpha - \beta = \alpha + (-\beta).$$

性质 4.1.4　设 $\lambda \in \mathbb{F}, \alpha \in V$. 则 $\lambda\alpha = 0$ 的必要且充分条件是，$\lambda = 0$ 或 $\alpha = 0$.

证明　**必要性**　设 $\lambda \neq 0$，则 $\lambda^{-1} \in \mathbb{F}$ 并且

$$\lambda^{-1}(\lambda\alpha) = (\lambda^{-1}\lambda)\alpha = 1 \cdot \alpha = \alpha.$$

另一方面，由于 $\lambda\alpha = 0$，故

$$\lambda^{-1} \cdot 0 = \lambda^{-1}(\lambda\alpha) = \lambda^{-1}(0 + 0) = \lambda^{-1} \cdot 0 + \lambda^{-1} \cdot 0,$$

两端同时加上 $-\lambda^{-1} \cdot 0$，则 $-\lambda^{-1} \cdot 0 = 0$. 从而 $\alpha = 0$.

充分性　设 $\lambda = 0$，则

$$0 \cdot \alpha = (0 + 0)\alpha = 0 \cdot \alpha + 0 \cdot \alpha,$$

两端同时加上 $-0 \cdot \alpha$，得到 $0 \cdot \alpha = 0$；设 $\alpha = 0$，则 $\lambda \cdot 0 = \lambda(0 + 0) = \lambda \cdot 0 + \lambda \cdot 0$，两端同时加上 $-\lambda \cdot 0$，得到 $\lambda \cdot 0 = 0$. $\blacksquare$

性质 4.1.5　设 $\lambda \in \mathbb{F}, \alpha \in V$，则 $(-\lambda)\alpha = \lambda(-\alpha) = -\lambda\alpha$.

证明　因为 $(-\lambda)\alpha + \lambda\alpha = (-\lambda + \lambda)\alpha = 0 \cdot \alpha = 0$，所以 $(-\lambda)\alpha = -\lambda\alpha$.

又因为 $\lambda(-\alpha) + \lambda\alpha = \lambda(-\alpha + \alpha) = \lambda \cdot 0 = 0$，所以 $\lambda(-\alpha) = -\lambda\alpha$. $\blacksquare$

习　题　4.1

1. 判断以下的集合 V 关于所规定的运算是否成为线性空间.

(1) 取 V 为所有实数对 (x_1, x_2) 的集合；$\mathbb{F}$ 为实数域 $\mathbb{R}$；向量的加法规定为：对 $(x_1, x_2), (y_1, y_2) \in V$，

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2 + x_1 y_2);$$

纯量与向量的乘法规定为：对任意 $\lambda \in \mathbb{F}, (x_1, x_2) \in V$，

$$\lambda(x_1, x_2) = \left(\lambda x_1,\ \lambda x_2 + \frac{1}{2}\lambda(\lambda - 1)x_1^2\right);$$

(2) 取 V 为所有实数对 (x_1, x_2) 的集合；$\mathbb{F}$ 为实数域 $\mathbb{R}$；向量的加法规定为：对任意 $(x_1, x_2), (y_1, y_2) \in V$，

$$(x_1, x_2) + (y_1, y_2) = (x_1 + x_2, y_1 + y_2);$$

纯量与向量的乘法规定为：对任意 $\lambda \in \mathbb{F}, (x_1, x_2) \in V$，

$$\lambda(x_1, x_2) = (x_1, x_2);$$

(3) 取 V 为所有满足 $f(x^2) = f^2(x)$ 的实函数集合；$\mathbb{F}$ 为实数域 $\mathbb{R}$；加法规定为：对 $f(x), g(x) \in V$，

$$(f + g)(x) = f(x) + g(x);$$

纯量与向量的乘法规定为：对 $\lambda \in \mathbb{F}, f(x) \in V$，

$$(\lambda f)(x) = \lambda f(x);$$

(4) 取 V 为所有满足 $f(-1)=0$ 的实函数集合；数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$；向量的加法规定为函数的加法；纯量与向量的乘法规定为实数与函数的乘法；

(5) 取 V 是所有满足 $a_1>0$ 的有序 n 元实数组 $(a_1,a_2,\ldots,a_n)$ 的集合；数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$；向量的加法与纯量与向量的乘法和 n 维实向量空间 $\mathbb{R}^n$ 相同；

(6) 取 V 是数域 $\mathbb{F}$ 上的所有 n 阶可逆方阵的集合；取数域为 $\mathbb{F}$；向量的加法规定为矩阵的加法，纯量与向量的乘法规定为纯量与矩阵的乘法；

(7) 给定数域 $\mathbb{F}$ 上的 n 阶方阵 A_0. 取 V 是所有满足 $A_0B=BA_0$ 的数域 $\mathbb{F}$ 上的 n 阶方阵 B 的集合；取数域为 $\mathbb{F}$；向量的加法以及纯量与向量的乘法同 (6)；

(8) 取 V 为数域 $\mathbb{F}$ 上的所有幂等方阵的集合；数域取为 $\mathbb{F}$；向量的加法，以及纯量与向量的乘法同 (6)；

(9) 取 V 是所有定义在实轴上的复值函数；数域 $\mathbb{F}$ 为复数域 $\mathbb{C}$；向量的加法规定为函数的加法，纯量与向量的乘法规定为复数与函数的乘法；

(10) 取 V 为所有定义在实轴上且满足 $f(-x)=\overline{f(x)}$ 的复函数集合，其中 $\bar{z}$ 表示复数 z 的共轭；取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$；向量的加法，以及纯量与向量的乘法同 (9).

§4.2 线性相关

从解析几何可以知道，在三维实向量空间 $\mathbb{R}^3$ 中，向量 α 和 β 共线的充分必要条件是，存在不全为零的实数 λ 和 μ，使得 $\lambda\alpha+\mu\beta=0$；向量 α,β,γ 共面的充分必要条件是，存在不全为零的实数 λ,μ 和 ν，使得 $\lambda\alpha+\mu\beta+\nu\gamma=0$. 三维实向量空间 $\mathbb{R}^3$ 中向量之间的这种共线、共面关系，推广到数域 $\mathbb{F}$ 上的线性空间 V，就是向量间的线性相关性.

定义 4.2.1 设 V 是数域 $\mathbb{F}$ 上的线性空间，$S\subseteq V$. 如果存在向量 $\alpha_1,\alpha_2,\ldots,\alpha_k\in S$，和不全为零的纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$ 使得

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k=0,$$

则**向量集合 S 称为线性相关的**.

特别，当向量集合 S 由有限个向量 $\alpha_1,\alpha_2,\ldots,\alpha_k$ 组成时，也称**向量 $\boldsymbol{\alpha_1,\alpha_2,\ldots,\alpha_k}$ 线性相关**. 不是线性相关的向量集合 S 称为**线性无关**的.

例 4.2.1 复向量空间 $\mathbb{C}^1$ 中任意两个向量都是线性相关的.

证明 设 $\alpha,\beta\in\mathbb{C}^1$. 如果 $\alpha=\beta=0$，则取 $\lambda=\mu=1\in\mathbb{C}$，于是 $\lambda\alpha+\mu\beta=0$.

如果 α 与 β 不全为零，则取 $\lambda=-\beta\in\mathbb{C}$，$\mu=-\alpha\in\mathbb{C}$. 显然，$\lambda$ 与 μ 不全为零，并且 $\lambda\alpha+\mu\beta=0$. 因此，$\mathbb{C}^1$ 中任意两个向量都线性相关. ■

例 4.2.2 设 C^2 是区间 $[0,1]$ 上的所有连续实函数的集合，它关于函数的加法，实数与函数的乘法成为实线性空间. 设 C^2 中向量集合 $S=\left\{f_i(x)=x^i \;\middle|\; i=0,1,\ldots\right\}$. 证明，向量集合 S 线性无关.

证明 反证法. 设向量集合 S 线性相关，则存在 k 个向量 $f_{i_1}(x),f_{i_2}(x),\ldots,$

$f_{i_k}(x)\in S$，和 k 个不全为零的纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{R}$，使得

$$\lambda_1 f_{i_1}(x)+\lambda_2 f_{i_2}(x)+\cdots+\lambda_k f_{i_k}(x)=0,$$

即
$$\lambda_1 x^{i_1}+\lambda_2 x^{i_2}+\cdots+\lambda_k x^{i_k}=0.$$

这表明，上式左端的多项式为零多项式．因此，它的系数 $\lambda_1,\lambda_2,\ldots,\lambda_k$ 全为零，但这是一个矛盾．■

由线性相关的定义可以得到以下的结论．

结论 4.2.1　含有零向量的向量集合 S 一定是线性相关的．

结论 4.2.2　含有线性相关向量子集的向量集合 S 一定是线性相关的．

结论 4.2.3　线性无关向量集合 S 的任何子集合 S_1 都是线性无关的．

证明　如果子集合 S_1 不是线性无关的，则 S_1 线性相关．由结论 4.2.2，向量集合 S 线性相关，矛盾．■

结论 4.2.4　向量集合 S 线性无关的充分必要条件是，S 的每一个有限子集（即由有限个向量构成的子集）都是线性无关的，即对于 S 的任意一组向量 $\alpha_1,\alpha_2,\ldots,\alpha_k$，由

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k=0$$

一定能得到
$$\lambda_1=\lambda_2=\cdots=\lambda_k=0,$$

其中纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$．

证明　必要性即是结论 4.2.2；充分性可用反证法证明．■

定义 4.2.2　设 V 是数域 $\mathbb{F}$ 上的一个线性空间，集合 $S\subseteq V$，向量 $\alpha\in V$．如果存在向量 $\alpha_1,\alpha_2,\ldots,\alpha_k\in S$，纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$，使得

$$\alpha=\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k,$$

则向量 α 称为**向量集合 S 的线性组合**，或者向量 α **可由向量集合 S 线性表出**．

特别，当 $S=\{\alpha_1,\alpha_2,\ldots,\alpha_\ell\}$ 时，则向量 α 称为**向量 $\alpha_1,\alpha_2,\ldots,\alpha_\ell$ 的线性组合**，或者称向量 α 可**由向量 $\alpha_1,\alpha_2,\ldots,\alpha_\ell$ 线性表出**．

下面的定理给出线性相关与线性组合的内在联系．

定理 4.2.1　非零向量 $\alpha_1,\alpha_2,\ldots,\alpha_k$ 线性相关的当且仅当存在某个向量 α_m，$2\leqslant m\leqslant k$，使得向量 α_m 是它前面 $m-1$ 个向量 $\alpha_1,\ldots,\alpha_{m-1}$ 的线性组合．

证明　设非零向量 $\alpha_1,\alpha_2,\ldots,\alpha_k$ 线性相关，则存在不全为零的纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$ 使得

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k=0.$$

设 λ_m 是纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k$ 中由后往前数第一个不为零的纯量，即 $\lambda_m\neq 0$，$\lambda_{m+1}=\cdots=\lambda_k=0$．由于纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k$ 不全为零，所以 λ_m 是存在的．

又由于向量 $\alpha_1,\alpha_2,\ldots,\alpha_k$ 都是非零的，因此，$2\leqslant m\leqslant k$．于是，$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+$

$\lambda_m\alpha_m = 0$. 所以,

$$\alpha_m = \left(-\frac{\lambda_1}{\lambda_m}\right)\alpha_1 + \left(-\frac{\lambda_2}{\lambda_m}\right)\alpha_2 + \cdots + \left(-\frac{\lambda_{m-1}}{\lambda_m}\right)\alpha_{m-1}.$$

即向量 α_m 是向量 $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ 的线性组合.

反之,设向量 α_m 是 $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ 的线性组合,则存在纯量 $\lambda_1, \lambda_2, \ldots, \lambda_{m-1} \in \mathbb{F}$,使得 $\alpha_m = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_{m-1}\alpha_{m-1}$. 于是,

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_{m-1}\alpha_{m-1} + (-1)\cdot\alpha_m = 0.$$

这表明,向量 $\alpha_1, \alpha_2, \ldots, \alpha_m$ 线性相关. ■

定义 4.2.3 设 $S, T \subseteq V$. 如果向量集合 T 中每个向量都可由向量集合 S 线性表出,则称**向量集合 T 可由向量集合 S 线性表出**. 如果向量集合 S 可由向量集合 T 线性表出,而向量集合 T 也可由向量集合 S 线性表出,则称**向量集合 S 与 T 等价**.

特别,当 $S = \{\alpha_1, \alpha_2, \ldots, \alpha_s\}$, $T = \{\beta_1, \beta_2, \ldots, \beta_t\}$ 时,如果向量集合 S 可由向量集合 T 线性表出,则称**向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出**. 如果向量集合 S 与 T 等价,则称**向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 与 $\beta_1, \beta_2, \ldots, \beta_t$ 等价**.

容易验证,向量集合之间的等价关系满足以下性质.

性质 4.2.1(自反性) 对任意向量集合 $S \subseteq V$,向量集合 S 和自身等价.

性质 4.2.2(对称性) 设向量集合 $S, T \subseteq V$,若 S 与 T 等价,则 T 与 S 等价.

性质 4.2.3(传递性) 设向量集合 $S, T, W \subseteq V$,如果 S 与 T 等价,T 与 W 等价,则 S 与 W 等价.

定理 4.2.2(Steinitz 替换定理) 设向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性无关,并且可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出,则 $s \leqslant t$,并且可以用向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 替换向量 $\beta_1, \beta_2, \ldots, \beta_t$ 中某 s 个向量,不妨设为 $\beta_1, \beta_2, \ldots, \beta_s$,使得向量 $\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_{s+1}, \ldots, \beta_t$ 与向量 $\beta_1, \beta_2, \ldots, \beta_t$ 等价.

证明 对 s 用归纳法. 当 $s = 1$ 时,显然,$s \leqslant t$.

由于向量 α_1 可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出,所以存在纯量 $\lambda_1, \lambda_2, \ldots, \lambda_t \in \mathbb{F}$,使得 $\alpha_1 = \lambda_1\beta_1 + \lambda_2\beta_2 + \cdots + \lambda_t\beta_t$.

如果 $\lambda_1 = \lambda_2 = \cdots = \lambda_t = 0$,则 $\alpha_1 = 0$. 显然,零向量 α_1 线性相关,和向量 α_1 线性无关的假设相矛盾. 因此,必有某个 $\lambda_k \neq 0$. 不妨设 $\lambda_1 \neq 0$. 于是,

$$\beta_1 = \frac{1}{\lambda_1}\alpha_1 + \left(-\frac{\lambda_2}{\lambda_1}\right)\beta_2 + \cdots + \left(-\frac{\lambda_t}{\lambda_1}\right)\beta_t.$$

这表明,向量 β_1 可由向量 $\alpha_1, \beta_2, \ldots, \beta_t$ 线性表出.

又显然,当 $2 \leqslant k \leqslant t$ 时,向量 β_k 可由向量 $\alpha_1, \beta_2, \ldots, \beta_t$ 线性表出. 因此,向量 $\beta_1, \beta_2, \ldots, \beta_t$ 可由向量 $a_1, \beta_2, \ldots, \beta_t$ 线性表出.

反之,向量 α_1 可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出,而当 $2 \leqslant k \leqslant t$ 时,向量 β_k 可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出. 从而向量 $\alpha_1, \beta_2, \ldots, \beta_t$ 可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表

出. 因此, 向量 $\alpha_1, \beta_2, \ldots, \beta_t$ 和向量 $\beta_1, \beta_2, \ldots, \beta_t$ 等价.

假设定理对 $s-1$ 成立. 下面证明, 定理对 s 成立.

事实上, 因为向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性无关, 因此, 向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}$ 线性无关. 又因为向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出, 因此, 向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}$ 也可由向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出.

由归纳假设, $s-1 \leqslant t$, 并且可以由向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}$ 替换向量 $\beta_1, \beta_2, \ldots, \beta_t$ 中某 $s-1$ 个, 不妨设为 $\beta_1, \beta_2, \ldots, \beta_{s-1}$, 使得 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \beta_s, \ldots, \beta_t$ 与 $\beta_1, \beta_2, \ldots, \beta_t$ 等价. 由于 α_s 可由 $\beta_1, \beta_2, \ldots, \beta_t$ 线性表出, 而 $\beta_1, \beta_2, \ldots, \beta_t$ 与 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \beta_s, \ldots, \beta_t$ 等价, 因此向量 α_s 可由 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \beta_s, \ldots, \beta_t$ 线性表出. 所以存在纯量 $\lambda_1, \lambda_2, \ldots, \lambda_t \in \mathbb{F}$, 使得

$$\alpha_s = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_{s-1}\alpha_{s-1} + \lambda_s\beta_s + \cdots + \lambda_t\beta_t.$$

如果 $s-1 = t$, 或者 $\lambda_s = \cdots = \lambda_t = 0$, 则上式化为

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_{s-1}\alpha_{s-1} + (-1)\alpha_s = 0,$$

即向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 线性相关, 矛盾. 因此, $s-1 < t$, 即 $s \leqslant t$, 并且 $\lambda_s, \ldots, \lambda_t$ 中至少有一个不为零, 不妨设 $\lambda_s \neq 0$. 于是,

$$\beta_s = \left(-\frac{\lambda_1}{\lambda_s}\right)\alpha_1 + \cdots + \left(-\frac{\lambda_{s-1}}{\lambda_s}\right)\alpha_{s-1} + \frac{1}{\lambda_s}\alpha_s + \left(-\frac{\lambda_{s+1}}{\lambda_s}\right)\beta_{s+1}\cdots + \left(-\frac{\lambda_t}{\lambda_s}\right)\beta_t.$$

因此, 向量 β_s 可由向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \alpha_s, \beta_{s+1}, \ldots, \beta_t$ 线性表出.

显然, 向量 α_k 与 β_ℓ 都可由向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \alpha_s, \beta_{s+1}, \ldots, \beta_t$ 线性表出, 其中 $1 \leqslant k \leqslant s-1, s+1 \leqslant \ell \leqslant t$. 所以向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \beta_s, \ldots, \beta_t$ 可由向量 $\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_{s+1}, \ldots, \beta_t$ 线性表出.

反之, 向量 $\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_{s+1}, \ldots, \beta_t$ 可由向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \beta_s, \ldots, \beta_t$ 线性表出. 从而向量 $\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_{s+1}, \ldots, \beta_t$ 与向量 $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \beta_s, \ldots, \beta_t$ 等价. 但是, $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}, \beta_s, \ldots, \beta_t$ 与 $\beta_1, \beta_2, \ldots, \beta_t$ 等价. 由传递性, 向量

$$\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_{s+1}, \ldots, \beta_t$$

与向量 $\beta_1, \beta_2, \ldots, \beta_t$ 等价. ■

推论 4.2.1　设 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 与 $\beta_1, \beta_2, \ldots, \beta_t$ 等价, 并且都线性无关, 则 $s = t$.

定义 4.2.4　设向量集合 $S = \{\alpha_1, \alpha_2, \ldots, \alpha_s\} \subseteq V$. 如果 S 中的向量 $\beta_1, \beta_2, \ldots, \beta_t$ 线性无关, 并且对任意 $\beta \in S$, 向量 $\beta, \beta_1, \beta_2, \ldots, \beta_t$ 线性相关, 则称向量 $\beta_1, \beta_2, \ldots, \beta_t$ 是向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的一个**极大线性无关向量组**.

定理 4.2.3　向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的任意一个极大线性无关向量组都与向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 等价, 而且向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的任意两个极大线性无关向量组所含向量的个数相同.

证明　设 $\beta_1, \beta_2, \ldots, \beta_t$ 是向量 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的极大线性无关向量组.

由定义, 向量 $\alpha_k, \beta_1, \beta_2, \ldots, \beta_t$ 线性相关, 其中 $1 \leqslant k \leqslant s$. 因此, 存在不全为零的

纯量 $\mu,\lambda_1,\ldots,\lambda_t\in\mathbb{F}$,使得

$$\mu\alpha_k+\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_t\beta_t=0.$$

如果 $\mu=0$,则 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 不全为零,并且 $\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_t\beta_t=0$,即向量组 $\beta_1,\beta_2,\ldots,\beta_t$ 线性相关. 矛盾. 因此,$\mu\neq 0$. 于是,

$$\alpha_k=\left(-\frac{\lambda_1}{\mu}\right)\beta_1+\left(-\frac{\lambda_2}{\mu}\right)\beta_2+\cdots+\left(-\frac{\lambda_t}{\mu}\right)\beta_t,$$

即向量 α_k 可由向量 $\beta_1,\beta_2,\ldots,\beta_t$ 线性表出.

由向量 α_k 的任意性,所以向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 可由 $\beta_1,\beta_2,\ldots,\beta_t$ 线性表出.

反之,向量 β_ℓ 是向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 中的某个向量,$1\leqslant\ell\leqslant t$,因此可设 $\beta_\ell=\alpha_{i_\ell}$. 所以向量 β_ℓ 可由向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 线性表出.

由向量 β_ℓ 的任意性,向量 $\beta_1,\beta_2,\ldots,\beta_t$ 可由向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 线性表出. 于是,向量 $\beta_1,\beta_2,\ldots,\beta_t$ 与向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 等价.

设 $\gamma_1,\gamma_2,\ldots,\gamma_r$ 是向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 的另一个极大线性无关向量组. 根据上述的证明可知,向量 $\gamma_1,\gamma_2,\ldots,\gamma_r$ 与向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 等价,而向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 与向量 $\beta_1,\beta_2,\ldots,\beta_t$ 等价,所以,向量 $\gamma_1,\gamma_2,\ldots,\gamma_r$ 与 $\beta_1,\beta_2,\ldots,\beta_t$ 等价. 由定理 4.2.2 的推论,$r=t$. ■

定义 4.2.5 向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 的极大线性无关向量组所含向量的个数 r 称为向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 的**秩**.

作为向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 的秩的一个应用,我们来考查矩阵的秩.

设 $A=(a_{ij})\in\mathbb{F}^{m\times n}$,这里 $\mathbb{F}^{m\times n}$ 是数域 $\mathbb{F}$ 上的所有 $m\times n$ 矩阵的集合.

矩阵 A 的第 i 行记为 $\alpha_i=(a_{i1},a_{i2},\ldots,a_{in})$,$1\leqslant i\leqslant m$. 向量 α_i 是数域 $\mathbb{F}$ 上的行向量空间 $\mathbb{F}^n$ 的向量,这里 $\mathbb{F}^n$ 是数域 $\mathbb{F}$ 上的所有有序 n 元数组的集合构成的线性空间. 向量 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 的秩称为矩阵 A 的**行秩**.

同样,记矩阵 A 的第 j 列为 $\beta_j=(a_{1j},a_{2j},\ldots,a_{mj})^{\mathrm{T}}$,$1\leqslant j\leqslant n$. 向量 β_j 是数域 $\mathbb{F}$ 上的列向量空间 $\mathbb{F}^m$ 的向量,其中 $\mathbb{F}^m$ 是数域 $\mathbb{F}$ 上的所有排成列形式的有序 m 元数组集合构成的线性空间. 向量 $\beta_1,\beta_2,\ldots,\beta_t$ 的秩称为矩阵 A 的**列秩**.

定理 4.2.4 矩阵 $A\in\mathbb{F}^{m\times n}$ 的行秩等于列秩,并等于矩阵 A 的秩.

证明 设 $\operatorname{rank}A=r$. 则矩阵 A 具有 r 阶非零子式,设

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}\neq 0, \tag{4.2.1}$$

其中 $1\leqslant i_1<i_2<\cdots<i_r\leqslant m$,$1\leqslant j_1<j_2<\cdots<j_r\leqslant n$.

考察矩阵 A 相应的行向量 $\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_r}$. 设 $\lambda_1,\lambda_2,\ldots,\lambda_r\in\mathbb{F}$,使得

$$\lambda_1\alpha_{i_1}+\lambda_2\alpha_{i_2}+\cdots+\lambda_r\alpha_{i_r}=0.$$

写成分量形式,上式即化为

$$\begin{cases} a_{i_1 1}\lambda_1 + a_{i_2 1}\lambda_2 + \cdots + a_{i_r 1}\lambda_r = 0, \\ a_{i_1 2}\lambda_1 + a_{i_2 2}\lambda_2 + \cdots + a_{i_r 2}\lambda_r = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i_1 n}\lambda_1 + a_{i_2 n}\lambda_2 + \cdots + a_{i_r n}\lambda_r = 0. \end{cases}$$

由于式 (4.2.1),上述方程组的系数矩阵的秩为 r,所以,$\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$. 这表明,行向量 $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}$ 线性无关.

其次,设 α_k 是矩阵 A 的第 k 个行向量,并且设 $\lambda_1, \lambda_2, \dots, \lambda_r, \mu \in \mathbb{F}$,使得

$$\lambda_1\alpha_{i_1} + \lambda_2\alpha_{i_2} + \cdots + \lambda_r\alpha_{i_r} + \mu\alpha_k = 0.$$

写成分量形式,上式即为

$$\begin{cases} a_{i_1 1}\lambda_1 + a_{i_2 1}\lambda_2 + \cdots + a_{i_r 1}\lambda_r + a_{k1}\mu = 0, \\ a_{i_1 2}\lambda_1 + a_{i_2 2}\lambda_2 + \cdots + a_{i_r 2}\lambda_r + a_{k2}\mu = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i_1 n}\lambda_1 + a_{i_2 n}\lambda_2 + \cdots + a_{i_r n}\lambda_r + a_{k1}\mu = 0. \end{cases}$$

由于式 (4.2.1),并且 $\operatorname{rank} A = r$,因此,当 $1 \leqslant j \leqslant n$ 时,

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_r & k \\ j_1 & j_2 & \cdots & j_r & j \end{pmatrix} = 0.$$

即上述方程组的系数矩阵的秩为 r. 因此,上述方程组具有非零解 $\lambda_1, \lambda_2, \dots, \lambda_r, \mu \in \mathbb{F}$. 所以,向量 $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}, \alpha_k$ 线性相关.

由向量 α_k 的任意性,向量 $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}$ 是向量 $\alpha_1, \alpha_2, \dots, \alpha_m$ 的极大线性无关向量组. 因此,矩阵 A 的行秩等于矩阵 A 的秩.

同理可证矩阵 A 的列秩也等于矩阵 A 的秩. ■

定理 4.2.4 给出了矩阵的秩的一种几何意义,它为应用向量的线性相关性来证明有关矩阵秩的命题开辟了一个途径.

例 4.2.3 设 $A = (a_{ij}) \in \mathbb{F}^{m\times n}$, $B = (b_{ij}) \in \mathbb{F}^{n\times p}$. 证明,

$$\operatorname{rank} AB \leqslant \min\{\operatorname{rank} A, \operatorname{rank} B\}.$$

证明 记 $C = AB$,且 $\operatorname{rank} B = t$,$\operatorname{rank} C = r$. 设矩阵 B 的 n 个行向量分别为 $\beta_1, \beta_2, \dots, \beta_n$,矩阵 C 的 m 个行向量分别为 $\gamma_1, \gamma_2, \dots, \gamma_m$,它们都是数域 $\mathbb{F}$ 上的行向量空间 $\mathbb{F}^p$ 中的向量. 由于 $C = AB$,因此,对于 $1 \leqslant i \leqslant m$,

$$\gamma_i = a_{i1}\beta_1 + a_{i2}\beta_2 + \cdots + a_{in}\beta_n.$$

这表明,向量 $\gamma_1, \gamma_2, \dots, \gamma_m$ 可由向量 $\beta_1, \beta_2, \dots, \beta_n$ 线性表出.

由于 $\operatorname{rank} C = r$,因此,由定理 4.2.3 和定理 4.2.4,向量 $\gamma_1, \gamma_2, \dots, \gamma_m$ 具有极大线性无关向量组 $\gamma_{i_1}, \gamma_{i_2}, \dots, \gamma_{i_r}$,并且向量 $\gamma_{i_1}, \gamma_{i_2}, \dots, \gamma_{i_r}$ 可由向量 $\gamma_1, \gamma_2, \dots, \gamma_m$ 线性表出,而向量 $\gamma_1, \gamma_2, \dots, \gamma_m$ 可由向量 $\beta_1, \beta_2, \dots, \beta_n$ 线性表出,从而向量 $\gamma_{i_1}, \gamma_{i_2}, \dots, \gamma_{i_r}$ 可由向量 $\beta_1, \beta_2, \dots, \beta_n$ 线性表出.

由于 $\operatorname{rank} B = t$,因此由定理 4.2.3 和定理 4.2.4,向量 $\beta_1, \beta_2, \dots, \beta_n$ 具有极大线

性无关向量组 $\beta_{j_1},\beta_{j_2},\ldots,\beta_{j_t}$，并且向量 $\beta_1,\beta_2,\ldots,\beta_n$ 可由向量 $\beta_{j_1},\beta_{j_2},\ldots,\beta_{j_t}$ 线性表出. 于是向量 $\gamma_{i_1},\gamma_{i_2},\ldots,\gamma_{i_r}$ 可由向量 $\beta_{j_1},\beta_{j_2},\ldots,\beta_{j_t}$ 线性表出. 因为向量 $\gamma_{i_1},\gamma_{i_2},\ldots,\gamma_{i_r}$ 线性无关，因此由 Steinitz 替换定理，$r\leqslant t$，即 $\operatorname{rank} C\leqslant \operatorname{rank} B$.

对矩阵 A 和 C 的列向量做同样的考虑，可以证明，$\operatorname{rank} C\leqslant \operatorname{rank} A$. ■

习 题 4.2

1. 判断下列向量是否线性无关.

(1) $\alpha_1=(2,-3,1)$，$\alpha_2=(3,-1,5)$，$\alpha_3=(1,-4,3)$；

(2) $\alpha_1=(4,-5,2,6)$，$\alpha_2=(2,-2,1,3)$，$\alpha_3=(6,-3,3,9)$，$\alpha_4=(4,-1,5,6)$；

(3) $\alpha_1=(1,0,0,2,5)$，$\alpha_2=(0,1,0,3,4)$，$\alpha_3=(0,0,1,4,7)$，$\alpha_4=(2,-3,4,11,12)$.

2. 设向量 α,β,γ 线性无关. 向量 $\alpha+\beta,\beta+\gamma,\gamma+\alpha$ 是否线性无关?

3. 设纯量 λ 满足下列条件之一，求 λ：

(1) 向量 $(1+\lambda,1-\lambda)$，$(1-\lambda,1+\lambda)\in\mathbb{C}^2$ 线性相关，

(2) 向量 $(\lambda,1,0)$，$(1,\lambda,1)$，$(0,1,\lambda)\in\mathbb{R}^3$ 线性相关.

如果在 (1) 中将 $\mathbb{C}^2$ 换成 $\mathbb{Q}^2$. 在 (2) 中将 $\mathbb{R}^3$ 换为 $\mathbb{Q}^3$，结论又怎样? 这里 $\mathbb{Q}^2$ 和 $\mathbb{Q}^3$ 分别是所有二元有理数组和三元有理数组的集合构成的有理数域上的线性空间.

4. 在什么条件下，向量 $(1,a_1,a_1^2)$，$(1,a_2,a_2^2)$，$(1,a_3,a_3^2)\in\mathbb{C}^3$ 线性相关? 将结论推广到 $\mathbb{C}^n$.

5. 设向量 $\alpha_1,\alpha_2,\ldots,\alpha_k\in\mathbb{F}^n$ 线性相关，$k\geqslant 2$. 证明对任意 $\alpha_{k+1}\in\mathbb{F}^n$，存在不全为零的纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$，使得向量

$$\alpha_1+\lambda_1\alpha_{k+1},\quad \alpha_2+\lambda_2\alpha_{k+1},\quad \ldots\ldots,\quad \alpha_k+\lambda_k\alpha_{k+1}$$

线性相关.

6. 取集合 V 为实数域 $\mathbb{R}$，数域为有理数域 $\mathbb{Q}$. 集合 V 的向量加法规定为实数的加法，纯量与向量的乘法规定为有理数与实数的乘法，则 V 成为有理数域 $\mathbb{Q}$ 上的线性空间. 证明，在线性空间 V 中，实数 1 与 α 线性无关的充分必要条件是，α 为无理数.

7. 设 V 是所有实函数构成的实数域 $\mathbb{R}$ 上的线性空间. 证明下列向量线性无关.

(1) x,x^4； (2) $x\mathrm{e}^x,\mathrm{e}^{2x}$； (3) $\sin x,\cos x$； (4) $\sin x,\mathrm{e}^x$.

8. 设 V 是所有连续实函数构成的实数域 $\mathbb{R}$ 上的线性空间. 证明，向量

$$\sin x,\sin 2x,\ldots,\sin nx,\ldots$$

线性无关.

9. 设 k 是给定的正整数，$1\leqslant k\leqslant n$，$S\subseteq\mathbb{F}^n$. 设 $\alpha\in S$，并且取向量 α 的第 $i_1,i_2,\ldots,i_k$ 个坐标 $a_{i_1},a_{i_2},\ldots,a_{i_k}$，组成向量 $\widetilde{\alpha}=\left(a_{i_1},a_{i_2},\ldots,a_{i_k}\right)\in\mathbb{F}^k$. 其中

$$1\leqslant i_1<i_2<\ldots<i_k\leqslant n. \tag{$*$}$$

当向量 α 遍历集合 S 中的向量，并且 $i_1i_2\cdots i_k$ 遍历自然数 $1,2,\ldots,n$ 中所有满足条件 $(*)$ 的数组 $i_1i_2\cdots i_k$ 时，便得到向量集合 $\widetilde{S}\subseteq\mathbb{F}^k$. 证明，如果向量集合 $\widetilde{S}$ 线性相关，则向量集合 S 也线性相关；如果向量集合 $\widetilde{S}$ 线性无关，则向量集合 S 也线性无关. (本题似有误)

10. 设 $t\leqslant n$，而 t 个 n 维行向量 $\alpha_i=(a_{i1},a_{i2},\ldots,a_{in})$，$i=1,2,\ldots,t$，满足

$$2|a_{ii}|>\sum_{k=1}^{n}|a_{ik}|.$$

证明向量 $\alpha_1,\alpha_2,\ldots,\alpha_t$ 线性无关.

11. 设数域 $\mathbb{F}$ 上的线性空间 V 中向量 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 线性无关. 添加向量 $\beta \in V$ 到向量 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 中. 证明, 在向量序列 $\beta, \alpha_1, \alpha_2, \ldots, \alpha_k$ 中, 能够由前面的向量线性表出的向量不多于 1 个.

12. 求向量 $\alpha_1 = (4, -1, 3, -2)$, $\alpha_2 = (8, -2, 6, -4)$, $\alpha_3 = (3, -1, 4, -2)$, $\alpha_4 = (6, -2, 8, -4)$ 的所有极大线性无关向量组.

13. 设 A 是 n 阶方阵. 证明 $\operatorname{rank} A^n = \operatorname{rank} A^{n+1} = \operatorname{rank} A^{n+2} = \cdots$.

14. 设 A 和 B 都是 $m \times p$ 矩阵. 证明 $\operatorname{rank}(A+B) \leqslant \operatorname{rank} A + \operatorname{rank} B$.

§4.3　基与坐标

大家知道, 在解析几何里, 如果在三维实向量空间 $\mathbb{R}^3$ 中建立直角坐标系, 则三个坐标轴上的单位向量分别为 $\varepsilon_1 = (1,0,0)$, $\varepsilon_2 = (0,1,0)$, $\varepsilon_3 = (0,0,1)$. 如果存在实数 $\lambda_1, \lambda_2, \lambda_3$ 使得 $\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2 + \lambda_3\varepsilon_3 = 0$, 则可得到 $\lambda_1 = \lambda_2 = \lambda_3 = 0$. 这表明, 向量 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 线性无关. 另一方面, 由于对任意 $\alpha \in \mathbb{R}^3$, 在这个坐标系下向量 α 的坐标为 (a_1, a_2, a_3), 故

$$\alpha = (a_1, a_2, a_3) = a_1\varepsilon_1 + a_2\varepsilon_2 + a_3\varepsilon_3.$$

因此, 向量空间 $\mathbb{R}^3$ 中任意一个向量 α 都可由向量 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 线性表出. 于是, 用线性空间语言讲, 所谓在空间 $\mathbb{R}^3$ 中设立坐标系, 相当于在 $\mathbb{R}^3$ 中选取一组线性无关的向量, 使得 $\mathbb{R}^3$ 中任意向量都可由它们线性表出.

把空间 $\mathbb{R}^3$ 中坐标系的概念推广到线性空间, 便引出线性空间的基概念.

定义 4.3.1　设 S 是数域 $\mathbb{F}$ 上的线性空间 V 的向量集合. 如果向量集合 S 线性无关, 而且 V 中每个向量都可由向量集合 S 线性表出, 则向量集合 S 称为**线性空间 V 的一组基**, S 中的向量称为**基向量**.

如果线性空间 V 的基 S 由有限多个基向量组成, 则线性空间 V 称为**有限维**的. 不是有限维的线性空间称为**无限维**的.

注　如果数域 $\mathbb{F}$ 上的线性空间 V 只含一个向量, 则由线性空间的定义, V 由零向量组成, 即 $V = \{0\}$. 此时, 称 $V = \{0\}$ 为**零维线性空间**. 上述定义中数域 $\mathbb{F}$ 上的线性空间是指非零维的.

定理 4.3.1　设数域 $\mathbb{F}$ 上的有限维线性空间 V 具有一组基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, $n \geqslant 1$. 则 V 中任意一个线性无关向量集合 S 都是有限的, 并且 S 所含向量的数目不超过 n.

证明　设线性无关向量集合 S 至少含有 $n+1$ 个向量 $\beta_1, \beta_2, \ldots, \beta_{n+1}$.

由结论 4.2.3, 向量 $\beta_1, \beta_2, \ldots, \beta_{n+1}$ 线性无关. 由基的定义, 向量 $\beta_1, \beta_2, \ldots, \beta_{n+1}$ 可由向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性表出. 由 Steinitz 替换定理, $n+1 \leqslant n$, 不可能. 因此, 向量集合 S 至多含有 n 个向量. ■

由定理 4.3.1 立即得到,

推论 4.3.1 有限维线性空间 V 任意两组基所含向量的个数相同.

根据推论 4.3.1,可以引进如下定义.

定义 4.3.2 数域 $\mathbb{F}$ 上的有限维线性空间 V 的一组基中所含向量的个数称为 **V 的维数**,记为 $\dim_{\mathbb{F}} V$,或简记为 $\dim V$.

利用维数概念,定理 4.3.1 可以叙述为

推论 4.3.2 设 V 是数域 $\mathbb{F}$ 上的 n 维线性空间,则 V 中任意 $n+1$ 个向量都线性相关.

定理 4.3.2 数域 $\mathbb{F}$ 上的 n 维线性空间 V 中任意一个线性无关向量集合 S 都可以扩充为 V 的一组基.

换言之,设 $S=\{\alpha_1,\alpha_2,\ldots,\alpha_r\}\subseteq V$ 线性无关,则存在 $\alpha_{r+1},\alpha_{r+2},\ldots,\alpha_n\in V$,使得 $\{\alpha_1,\alpha_2,\ldots,\alpha_r,\alpha_{r+1}\ldots,\alpha_n\}$ 是 V 的基.

证明 因为向量集合 S 线性无关,因此,由定理 4.3.1,S 所含向量的个数 $r\leqslant n$. 记 $S=\{\alpha_1,\alpha_2,\ldots,\alpha_r\}$. 设 $\beta_1,\beta_2,\ldots,\beta_n$ 是 V 的一组基. 于是向量 $\alpha_1,\alpha_2,\ldots,\alpha_r$ 可由向量 $\beta_1,\beta_2,\ldots,\beta_n$ 线性表出. 由 Steinitz 替换定理,向量 $\beta_1,\beta_2,\ldots,\beta_n$ 中存在向量 $\beta_{i_{r+1}},\beta_{i_{r+2}},\ldots,\beta_{i_n}$,使得向量

$$\{\alpha_1,\alpha_2,\ldots,\alpha_r,\beta_{i_{r+1}},\beta_{i_{r+2}},\ldots,\beta_{i_n}\} \tag{$*$}$$

与向量 $\beta_1,\beta_2,\ldots,\beta_n$ 等价. 从而向量 $\beta_1,\beta_2,\ldots,\beta_n$ 可由向量 $(*)$ 线性表出. 如果向量 $(*)$ 线性相关,则向量 $(*)$ 的极大线性无关向量组 S_1 所含向量的个数 $s<n$.

另一方面,由于向量 $\beta_1,\beta_2,\ldots,\beta_n$ 可由向量 $(*)$ 线性表出,而向量 $(*)$ 可由向量集合 S_1 线性表出. 因此,由 Steinitz 替换定理,$n\leqslant s$,矛盾. 这就证明,向量集合 $(*)$ 线性无关.

另外,由于向量集合 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 是 V 的一组基,所以 V 中每个向量都可由向量 $\beta_1,\beta_2,\ldots,\beta_n$ 线性表出. 由于向量集合 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 与 $(*)$ 等价,所以 V 中每个向量也都可由向量 $(*)$ 线性表出. 因此,向量集合 $(*)$ 是 V 的基. ■

现在设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的基. 由基的定义,对任意向量 $\alpha\in V$,存在纯量 $a_1,a_2,\ldots,a_n\in\mathbb{F}$,使得

$$\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n.$$

于是,向量 α 便确定数域 $\mathbb{F}$ 上的有序 n 元数组 $(a_1,a_2,\ldots,a_n)$.

注意,n 元数组 $(a_1,a_2,\ldots,a_n)$ 是由向量 α 所唯一确定的. 事实上,设另有一组纯量 $b_1,b_2,\ldots,b_n$,使得 $\alpha=b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n$. 则

$$(a_1-b_1)\alpha_1+(a_2-b_2)\alpha_2+\ldots+(a_n-b_n)\alpha_n=0.$$

因为向量 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 线性无关,所以 $a_i-b_i=0$,即 $a_i=b_i$,$1\leqslant i\leqslant n$.

由向量 α 所唯一确定的 n 元数组 $(a_1,a_2,\ldots,a_n)$ 称为向量 α 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的**坐标**,而 a_i 称为向量 α 的**第 i 个坐标分量**,$i=1,2,\ldots,n$.

设向量 $\alpha,\beta\in V$ 在基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的坐标分别为 $(a_1,a_2,\dots,a_n)$ 与 $(b_1,b_2,\dots,b_n)$，则 $\alpha=a_1\alpha_1+a_2\alpha_2+\dots+a_n\alpha_n$，$\beta=b_1\alpha_1+b_2\alpha_2+\dots+b_n\alpha_n$. 因此，

$$\alpha+\beta=(a_1+b_1)\alpha_1+(a_2+b_2)\alpha_2+\cdots+(a_n+b_n)\alpha_n.$$

所以向量 $\alpha+\beta$ 在基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的坐标为 $(a_1+b_1,a_2+b_2,\dots,a_n+b_n)$.

设纯量 $\lambda\in\mathbb{F}$，则

$$\lambda\alpha=\lambda a_1\alpha_1+\lambda a_2\alpha_2+\cdots+\lambda a_n\alpha_n.$$

因此，向量 $\lambda\alpha$ 在基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的坐标为 $(\lambda a_1,\lambda a_2,\dots,\lambda a_n)$.

记数域 $\mathbb{F}$ 上的所有有序 n 元数组构成的线性空间为 $\mathbb{F}^n$. 在数域 $\mathbb{F}$ 上的 n 维线性空间 V 中取定基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$. 定义线性空间 V 到 $\mathbb{F}^n$ 的映射 η 如下：对任意 $\alpha\in V$，

$$\eta(\alpha)=(a_1,a_2,\dots,a_n),$$

其中 $(a_1,a_2,\dots,a_n)$ 是向量 α 在基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的坐标. 从上面的讨论可以知道，线性空间 V 到 $\mathbb{F}^n$ 的映射 η 是一个双射，并且对任意 $\alpha,\beta\in V$，

$$\eta(\alpha+\beta)=\eta(\alpha)+\eta(\beta).$$

对任意 $\lambda\in\mathbb{F},\alpha\in V$，

$$\eta(\lambda\alpha)=\lambda\eta(\alpha).$$

例 4.3.1　证明，所有复数的集合 $\mathbb{C}$ 作为复线性空间是一维的，而作为实线性空间是二维的.

证明　把复数集合 $\mathbb{C}$ 当成复线性空间，取向量 ε 为复数 1. 显然，向量 ε 线性无关，而且，对于任意 $\beta\in\mathbb{C}$，$\beta=\beta\cdot\varepsilon$. 这表明，向量 β 可由向量 ε 线性表出. 由基的定义，$\{\varepsilon\}$ 是复线性空间 $\mathbb{C}$ 的一组基. 因此，$\dim_{\mathbb{C}}\mathbb{C}=1$.

把复数集合 $\mathbb{C}$ 当成实线性空间，取向量 ε_1 为复数 1，向量 ε_2 为复数 i，$\mathrm{i}^2=-1$. 设 $\lambda_1,\lambda_2\in\mathbb{R}$，使得

$$\lambda_1\varepsilon_1+\lambda_2\varepsilon_2=\lambda_1+\mathrm{i}\lambda_2=0,$$

则 $\lambda_1=\lambda_2=0$. 因此向量 $\varepsilon_1,\varepsilon_2$ 线性无关. 设 $\alpha\in\mathbb{C}$，则 $\alpha=a_1+a_2\mathrm{i}=a_1\varepsilon_1+a_2\varepsilon_2$，其中 $a_1,a_2\in\mathbb{R}$. 因此，实线性空间 $\mathbb{C}$ 中任意向量都可由向量 $\varepsilon_1,\varepsilon_2$ 线性表出. 所以 $\{\varepsilon_1,\varepsilon_2\}$ 是实线性空间 $\mathbb{C}$ 的基，即 $\dim_{\mathbb{R}}\mathbb{C}=2$. ■

例 4.3.2　证明：数域 $\mathbb{F}$ 上的所有关于未定元 x 的多项式构成的数域 $\mathbb{F}$ 上的线性空间 $\mathbb{F}[x]$ 是无限维的.

证明　取 $\mathbb{F}[x]$ 中的向量 $\alpha_i=x^i$，$i=0,1,\dots$. 记向量集合

$$S=\{\alpha_0,\alpha_1,\alpha_2,\dots\}.$$

在 S 中任取有限个向量 $\alpha_{i_1},\alpha_{i_2},\dots,\alpha_{i_k}$，$0\leqslant i_1<i_2<\cdots<i_k$. 设 $\lambda_1,\lambda_2,\dots,\lambda_k\in\mathbb{F}$，使得

$$\lambda_1\alpha_{i_1}+\lambda_2\alpha_{i_2}+\cdots+\lambda_k\alpha_{i_k}=0,$$

即

$$f(x)=\lambda_1x^{i_1}+\lambda_2x^{i_2}+\cdots+\lambda_kx^{i_k}=0.$$

这表明，多项式 $f(x)$ 是零多项式，从而 $\lambda_1=\lambda_2=\cdots=\lambda_k=0$. 因此，向量集合 $S_1=\{\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_k}\}$ 线性无关. 由于 S_1 的任意性，向量集合 S 线性无关.

对任意 $\alpha\in\mathbb{F}[x]$，显然 α 是关于未定元 x 的多项式 $f(x)$. 设 $\deg f(x)=n$，则

$$f(x)=a_0+a_1x^1+\cdots+a_nx^n=a_0\alpha_0+a_1\alpha_1+\cdots+a_n\alpha_n.$$

这表明，$\mathbb{F}[x]$ 中每个向量都可由向量集合 S 线性表出. 因此，向量集合 S 是 $\mathbb{F}[x]$ 的基. 所以，$\mathbb{F}[x]$ 是无限维的. ■

例 4.3.3 记数域 $\mathbb{F}$ 上的所有次数都小于 n 的多项式构成的数域 $\mathbb{F}$ 上的线性空间为 $\mathbb{F}_n[x]$. 证明

$$\dim\mathbb{F}_n[x]=n.$$

证明 取 $\mathbb{F}_n[x]$ 中向量集合

$$S_2=\{1,x,x^2,\ldots,x^{n-1}\},\qquad S=\{1,x,\ldots,x^n,\ldots\}.$$

由例 4.3.2，向量集合 S 线性无关，因此向量集合 S_2 线性无关. 另外，设 $f(x)\in\mathbb{F}_n[x]$，则 $\deg f(x)\leqslant n-1$，因此，

$$f(x)=a_0\cdot 1+a_1x+\cdots+a_{n-1}x^{n-1},\qquad a_0,a_1,\ldots,a_{n-1}\in\mathbb{F}.$$

这表明，$\mathbb{F}_n[x]$ 中每个向量都可由向量集合 S_2 线性表出. 因此，向量集合 S_2 是 $\mathbb{F}_n[x]$ 的基. 所以 $\dim\mathbb{F}_n[x]=n$. ■

例 4.3.4 数域 $\mathbb{F}$ 上的所有 $m\times n$ 矩阵的集合 $\mathbb{F}^{m\times n}$ 是数域 $\mathbb{F}$ 上的线性空间. 证明：

$$\dim\mathbb{F}^{m\times n}=m\times n.$$

证明 第 i 行和第 j 列交叉元素为 1，其它元素都为零的 $m\times n$ 矩阵记为 E_{ij}. 记

$$S=\{E_{ij}\mid 1\leqslant i\leqslant m,\ 1\leqslant j\leqslant n\}.$$

如果存在 $a_{ij}\in\mathbb{F}$，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$，使得

$$\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}E_{ij}=0,$$

则这一等式左端是一个 $m\times n$ 矩阵 $A=(a_{ij})$，并且 A 为零矩阵. 因此，$a_{ij}=0$，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$. 这表明，向量集合 S 线性无关.

其次，设 $A=(a_{ij})\in\mathbb{F}^{m\times n}$，则

$$A=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}E_{ij},$$

因此，$\mathbb{F}^{m\times n}$ 中每个向量都可由向量集合 S 线性表出. 所以向量集合 S 是 $\mathbb{F}^{m\times n}$ 的基. 于是，$\dim\mathbb{F}^{m\times n}=mn$. ■

例 4.3.5 数域 $\mathbb{F}$ 上的所有有序 n 元数组的集合 $\mathbb{F}^n$ 是数域 $\mathbb{F}$ 上的线性空间. 证明：

$$\dim\mathbb{F}^n=n.$$

证明　第 i 个数为 1，其它的数都为零的有序 n 元数组记为 ε_i．设 $a_i \in \mathbb{F}, 1 \leqslant i \leqslant n$，使得

$$a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n = 0.$$

则 $(a_1, a_2, \ldots, a_n) = 0$．从而 $a_1 = a_2 = \cdots = a_n = 0$．因此，向量集合

$$S = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$$

线性无关．其次，设 $\alpha \in \mathbb{F}^n$，则

$$\alpha = (a_1, a_2, \ldots, a_n) = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n,$$

其中 $a_i \in \mathbb{F}, 1 \leqslant i \leqslant n$．因此，$\mathbb{F}^n$ 中每个向量都可由向量集合 S 线性表出．所以，S 是 $\mathbb{F}^n$ 的基．因此，$\dim \mathbb{F}^n = n$．■

习　题　4.3

1. 证明，在四维实行向量空间 $\mathbb{R}^4$ 中，向量 $\alpha_1 = (1,1,0,0)$，$\alpha_2 = (0,0,1,1)$，$\alpha_3 = (1,0,0,4)$，$\alpha_4 = (0,0,0,2)$ 构成一组基．并求标准基向量

$$\varepsilon_1 = (1,0,0,0), \qquad \varepsilon_2 = (0,1,0,0), \qquad \varepsilon_3 = (0,0,1,0), \qquad \varepsilon_4 = (0,0,0,1)$$

在基 $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ 下的坐标．

2. 证明，在三维复向量空间 $\mathbb{C}$ 中，向量 $\alpha_1 = (2\mathrm{i}, 1, 0)$，$\alpha_2 = (2, -1, 1)$，$\alpha_3 = (0, 1+\mathrm{i}, 1-\mathrm{i})$ 构成一组基．并求标准基向量

$$\varepsilon_1 = (1,0,0), \qquad \varepsilon_2 = (0,1,0), \qquad \varepsilon_3 = (0,0,1)$$

在基 $(\alpha_1, \alpha_2, \alpha_3)$ 下的坐标．

3. 在数域 $\mathbb{F}$ 上的 n 维向量空间 $\mathbb{F}^n$ 中，求向量 $\alpha = (a_1, a_2, \ldots, a_n)$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标，其中对于 $j = 1, 2, \ldots, n$，

$$\alpha_j = (\underbrace{1,1,1,\ldots,1,1}_{\text{共 } j \text{ 个}}, 0, \ldots, 0).$$

4. 在数域 $\mathbb{F}$ 上的所有 2 阶方阵构成的线性空间 $\mathbb{F}^{2\times 2}$ 中，求一组基 $\{A_1, A_2, A_3, A_4\}$，使得对每个 j，

$$A_j^2 = A_j.$$

5. 证明，在所有次数不超过 n 的多项式构成的线性空间 $\mathbb{F}_{n+1}[x]$ 中，向量 $1, (x+a)^1, \ldots, (x+a)^n$ 构成一组基，其中 $a \in \mathbb{F}$．并求向量 $f(x) = a_0 + a_1x + \cdots + a_nx^n$ 在这组基下的坐标．

6. 证明，所有实数的集合作为有理数域 $\mathbb{Q}$ 上的线性空间是无限维的；所有复数的集合作为有理数域 $\mathbb{Q}$ 上的线性空间也是无限维的．

§4.4 基变换与坐标变换

§4.3 讨论了数域 $\mathbb{F}$ 上的 n 维线性空间 V 的向量在 V 的一组基下的坐标. 一般地说,同一个向量在不同基下的坐标是不同的. 问题是,同一个向量在不同基下的坐标之间有什么关系? 这就是本节所要讨论的问题.

为讨论方便,今后把向量在一组基下的坐标写成列向量形式.

设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 分别是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的基,设向量 $\alpha\in V$ 在这两组基下的坐标分别为

$$x=(x_1,x_2,\dots,x_n)^{\mathrm{T}}\quad 与\quad y=(y_1,y_2,\dots,y_n)^{\mathrm{T}}.$$

由于 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 V 的基,因此向量 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 可由向量 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 线性表出,所以可设

$$\begin{aligned}\beta_1&=b_{11}\alpha_1+b_{21}\alpha_2+\cdots+b_{n1}\alpha_n,\\ \beta_2&=b_{12}\alpha_1+b_{22}\alpha_2+\cdots+b_{n2}\alpha_n,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \beta_n&=b_{1n}\alpha_1+b_{2n}\alpha_2+\cdots+b_{nn}\alpha_n.\end{aligned}$$

上式可以写成矩阵形式如下:

$$(\beta_1,\beta_2,\dots,\beta_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\ b_{21}&b_{22}&\cdots&b_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ b_{n1}&b_{n2}&\cdots&b_{nn}\end{pmatrix}.$$

记 $B=(b_{ij})$,于是上式即为

$$(\beta_1,\beta_2,\dots,\beta_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)B. \tag{4.4.1}$$

式 (4.4.1) 称为由基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 到基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 的**基变换公式**,矩阵 B 称为由基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 到基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 的**过渡矩阵**.

定理 4.4.1 设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 分别是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的基. 设方阵 A 是由基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 到基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 的过渡矩阵. 则方阵 A 可逆,并且由基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 到基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 的过渡矩阵为 A^{-1}.

证明 由于方阵 $A=(a_{ij})$ 是由基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 到基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 的过渡矩阵,因此,

$$(\alpha_1,\alpha_2,\dots,\alpha_n)=(\beta_1,\beta_2,\dots,\beta_n)A. \tag{4.4.2}$$

设由基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 到基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 的过渡矩阵为 $B=(b_{ij})$,则

$$(\beta_1,\beta_2,\dots,\beta_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)B. \tag{4.4.3}$$

由式 (4.4.2) 与式 (4.4.3),对于 $i,j=1,2,\dots,n$,

$$\alpha_k=\sum_{i=1}^n a_{ik}\beta_i,\quad \beta_j=\sum_{k=1}^n b_{kj}\alpha_k.$$

因此,

$$\beta_j=\sum_{k=1}^n b_{kj}\Big(\sum_{i=1}^n a_{ik}\beta_i\Big)=\sum_{i=1}^n\Big(\sum_{k=1}^n a_{ik}b_{kj}\Big)\beta_i.$$

由于向量 $\beta_1,\beta_2,\dots,\beta_n$ 线性无关,因此,

$$\sum_{k=1}^n a_{ik}b_{kj}=\delta_{ij},$$

其中 $1\leqslant i,j\leqslant n$, δ_{ij} 是 Kronecker 符号.

所以 $AB=I_{(n)}$. 从而方阵 A 可逆,并且 $B=A^{-1}$. ■

记数域 $\mathbb{F}$ 上的 n 维线性空间 V 的所有基的集合为 $\mathcal{B}$,数域 $\mathbb{F}$ 上的所有 n 阶可逆方阵的集合为 $\mathrm{GL}_n(\mathbb{F})$.

取定 $\{\beta_1,\beta_2,\dots,\beta_n\}\in\mathcal{B}$,则对于任意 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}\in\mathcal{B}$,据定理 4.4.1,由

$$(\alpha_1,\alpha_2,\dots,\alpha_n)=(\beta_1,\beta_2,\dots,\beta_n)A.$$

所确定的矩阵 A 可逆. 通过上式,可以定义集合 $\mathcal{B}$ 到 $\mathrm{GL}_n(\mathbb{F})$ 的映射 η 如下: 对于任意 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}\in\mathcal{B}$,令

$$\eta(\alpha_1,\alpha_2,\dots,\alpha_n)=A.$$

容易看出,如果 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 与 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n\}$ 是 V 的两组不同的基,则由基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 分别到基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 和基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n\}$ 的过渡矩阵 A 与 $\widetilde{A}$ 也不同,即

$$\eta(\alpha_1,\alpha_2,\dots,\alpha_n)\neq\eta(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n).$$

因此,映射 η 是单射. 其次设 $A\in\mathrm{GL}_n(\mathbb{F})$,则由 $(\beta_1,\beta_2,\dots,\beta_n)A$ 便确定 V 的 n 个向量,记为 $\alpha_1,\alpha_2,\dots,\alpha_n$,即

$$(\alpha_1,\alpha_2,\dots,\alpha_n)=(\beta_1,\beta_2,\dots,\beta_n)A.$$

如果存在纯量 $\lambda_1,\lambda_2,\dots,\lambda_n\in\mathbb{F}$,使得

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_n=0,$$

则记

$$x=(\lambda_1,\lambda_2,\dots,\lambda_n)^{\mathrm{T}},$$

便得到

$$(\alpha_1,\alpha_2,\dots,\alpha_n)x=0\implies(\beta_1,\beta_2,\dots,\beta_n)Ax=0.$$

由于 $\beta_1,\beta_2,\dots,\beta_n$ 线性无关,因此 $Ax=0$. 由于 $A\in\mathrm{GL}_n(\mathbb{F})$,故 $x=0$,即

$$\lambda_1=\lambda_2=\cdots=\lambda_n=0.$$

这表明,向量 $\alpha_1,\alpha_2,\dots,\alpha_n$ 线性无关. 由于 $\dim V=n$,所以 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}\in\mathcal{B}$. 这就证明了对任意 $A\in\mathrm{GL}_n(\mathbb{F})$,存在 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}\in\mathcal{B}$,使得

$$\eta(\alpha_1,\alpha_2,\dots,\alpha_n)=A.$$

因此映射 η 是满射.

于是,在 V 中取定一组基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$, V 中所有基的集合 $\mathcal{B}$ 与数域 $\mathbb{F}$ 上的所有 n 阶可逆方阵的集合 $\mathrm{GL}_n(\mathbb{F})$ 之间便存在一个一一对应.

下一定理给出同一个向量在不同基下的坐标间的关系.

定理 4.4.2 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 分别是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的基,由基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 到基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 的过渡矩阵为 A.

设 $\alpha\in V$ 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 下的坐标分别为 x 与 y,则

$$y=Ax. \tag{4.4.4}$$

证明 由假设,

$$\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)x=(\beta_1,\beta_2,\ldots,\beta_n)y,$$

并且

$$(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\beta_1,\beta_2,\ldots,\beta_n)A.$$

因此,

$$\alpha=(\beta_1,\beta_2,\ldots,\beta_n)Ax=(\beta_1,\beta_2,\ldots,\beta_n)y.$$

由于向量 α 在同一组基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 下的坐标是唯一的,所以 $y=Ax$. ■

定理 4.4.2 中的式 (4.4.4) 称为同一个向量在不同基下的**坐标变换公式**.

习 题 4.4

1. 求四维实向量空间 $\mathbb{R}^4$ 中由基 $\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$ 到基 $\{\beta_1,\beta_2,\beta_3,\beta_4\}$ 的过渡矩阵,并求向量 $\alpha=(1,-1,1,-1)$ 在这两组基下的坐标.

(1) $\begin{cases}\alpha_1=(1,0,0,0), & \beta_1=(1,1,0,0),\\ \alpha_2=(0,1,0,0), & \beta_2=(1,0,1,0),\\ \alpha_3=(0,0,1,0), & \beta_3=(1,0,0,1),\\ \alpha_4=(0,0,0,1), & \beta_4=(1,1,1,1);\end{cases}$ (2) $\begin{cases}\alpha_1=(1,2,-1,0), & \beta_1=(2,1,0,1),\\ \alpha_2=(1,-1,1,1), & \beta_2=(0,1,2,2),\\ \alpha_3=(-1,2,1,1), & \beta_3=(-2,1,1,2),\\ \alpha_4=(-1,-1,0,1), & \beta_4=(1,3,1,2).\end{cases}$

2. 在数域 $\mathbb{F}$ 上的所有关于 $\cos x$ 的次数不超过 n 的多项式构成的线性空间中,试写出由基

$$\{1,\cos x,\cos^2 x,\ldots,\cos^n x\}\quad \text{到基}\quad \{1,\cos x,\cos 2x,\ldots,\cos nx\}$$

的过渡矩阵.

3. 设 V 是所有定义在当轴上由复值函数构成的复线性空间,在 V 中取向量

$$f_1(x)=1,\quad f_2(x)=\mathrm{e}^{\mathrm{i}x},\quad f_3(x)=\mathrm{e}^{-\mathrm{i}x},\quad g_1(x)=1,\quad g_2(x)=\cos x,\quad g_3(x)=\sin x,$$

其中 $\mathrm{i}^2=-1$. 证明,向量 f_1,f_2,f_3 和 g_1,g_2,g_3 分别是线性无关的,并求三阶可逆方阵 A,使得

$$(g_1,g_2,g_3)=(f_1,f_2,f_3)A.$$

4. 在四维实向量空间 $\mathbb{R}^4$ 的标准基 $\{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\}$ 下,超球面的方程为 $x_1^2+x_2^2+x_3^2+x_4^2=1$. 设

$$\alpha_1=(1,1,1,1),\quad \alpha_2=(1,1,-1,-1),\quad \alpha_3=(1,-1,1,-1),\quad \alpha_4=(1,-1,-1,1).$$

试求该超球面在基 $\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$ 下的方程.

5. 在数域 $\mathbb{F}$ 上的 n 维行向量空间 $\mathbb{F}^n$ 中,给定 n 个向量 $\alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{F}^n$,便可确定数域 $\mathbb{F}$ 上的 n 阶方阵 A: $A=(\alpha_1,\alpha_2,\ldots,\alpha_n)^{\mathrm{T}}$; 反之亦然. 证明,$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 $\mathbb{F}^n$ 的基的充分必要条件为方阵 A 可逆.

§4.5　同　构

给定数域 $\mathbb{F}$,数域 $\mathbb{F}$ 上的有限维线性空间当然有很多. 因此,自然希望能对数域 $\mathbb{F}$ 上的所有有限维线性空间的集合进行分类,使得同属于一个类的线性空间具有相同的结构.

什么是线性空间的结构? 从线性空间的定义可知,数域 $\mathbb{F}$ 上的线性空间 V 首先是一个集合 V,其次在集合 V 中定义了满足八条公理的两种代数运算. 所以在比较同一个数域 $\mathbb{F}$ 上的两个线性空间 V_1 与 V_2 的结构时,自然首先要考察,作为集合,V_1 与 V_2 是否能够建立一一对应; 其次再考察 V_1 与 V_2 的两种代数运算. 这就引出线性空间同构的概念.

定义 4.5.1　设 V_1 与 V_2 是同一个数域 $\mathbb{F}$ 上的两个线性空间. 如果存在 V_1 到 V_2 上的一一对应 η,它把 V_1 中的向量 α 映为 V_2 的向量 $\eta(\alpha)$,使得对任意 $\alpha,\beta\in V_1$, $\lambda\in\mathbb{F}$,都有

$$\eta(\alpha+\beta)=\eta(\alpha)+\eta(\beta), \tag{4.5.1}$$

$$\eta(\lambda\alpha)=\lambda\eta(\alpha), \tag{4.5.2}$$

则线性空间 V_1 与 V_2 称为**同构**的,而映射 η 称为 V_1 到 V_2 上的一个**同构映射**.

满足上述定义中条件 (4.5.1) 的映射 η 称为**保加法**的; 满足条件 (4.5.2) 的映射称为**保乘法**的. 因此同构映射 η 是线性空间 V_1 到 V_2 上的保加法与保乘法的一一映射(即双射).

由同构映射的定义,易知线性空间 V_1 到 V_2 上的同构映射 η 具有下列性质.

性质 4.5.1　当且仅当 $\alpha=0$ 时,$\eta(\alpha)=0$.

证明　设 $\beta\in V_1$,因为同构映射 η 是保加法的,因此 $\eta(\beta)=\eta(\beta+0)=\eta(\beta)+\eta(0)$. 所以 $\eta(0)=0$. 由于同构映射 η 是双射,所以,如果 $\eta(\alpha)=0$,则 $\alpha=0$. ∎

性质 4.5.2　设 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$, $\alpha_1,\alpha_2,\ldots,\alpha_k\in V$,则

$$\eta\big(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k\big)=\lambda_1\eta(\alpha_1)+\lambda_2\eta(\alpha_2)+\cdots+\lambda_k\eta(\alpha_k).$$

证明　对 k 用归纳法. 当 $k=1$ 时,因为同构映射 η 保乘法,所以 $\eta(\lambda_1\alpha_1)=\lambda_1\eta(\alpha_1)$,因此结论对 $k=1$ 成立. 假设结论对 $k-1$ 成立. 由于

$$\eta(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k)=\eta\big((\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_{k-1}\alpha_{k-1})+\lambda_k\alpha_k\big),$$

并且同构映射 η 保加法,所以

$$\eta(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k)=\eta(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_{k-1}\alpha_{k-1})+\eta(\lambda_k\alpha_k).$$

由归纳假设,以及同构映射 η 保乘法,故

$$\eta(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k)=\lambda_1\eta(\alpha_1)+\lambda_2\eta(\alpha_2)+\cdots+\lambda_{k-1}\eta(\alpha_{k-1})+\lambda_k\eta(\alpha_k).\quad\blacksquare$$

注 如果线性空间 V_1 到 V_2 的映射 η 满足；对任意 $\alpha,\beta\in V_1,\lambda,\mu\in\mathbb{F}$,

$$\eta(\lambda\alpha+\mu\beta)=\lambda\eta(\alpha)+\mu\eta(\beta),$$

则映射 η 称为**保线性关系的**. 性质 4.5.2 表明,同时保加法与保乘法的映射一定保线性关系. 反之可以证明,保线性关系的映射一定同时保加法与保乘法,即保线性等价于同时保加法与保乘法.

性质 4.5.3 向量 $\alpha_1,\alpha_2,\dots,\alpha_k\in V_1$ 线性相关的充分必要条件是,向量

$$\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_k)\in V_2$$

线性相关.

证明 设向量 $\alpha_1,\alpha_2,\dots,\alpha_k$ 线性相关,则存在不全为零的纯量 $\lambda_1,\lambda_2,\dots,\lambda_k\in\mathbb{F}$ 使得

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k=0.$$

由性质 4.5.2,

$$\eta(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k)=\lambda_1\eta(\alpha_1)+\lambda_2\eta(\alpha_2)+\cdots+\lambda_k\eta(\alpha_k)=0.$$

因此,向量 $\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_k)$ 线性相关.

反之,设 $\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_k)$ 线性相关,则存在不全为零的纯量 $\lambda_1,\lambda_2,\dots,\lambda_k\in\mathbb{F}$,使得

$$\lambda_1\eta(\alpha_1)+\lambda_2\eta(\alpha_2)+\cdots+\lambda_k\eta(\alpha_k)=0.$$

由性质 4.5.2,

$$\lambda_1\eta(\alpha_1)+\lambda_2\eta(\alpha_2)+\cdots+\lambda_k\eta(\alpha_k)=\eta(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k)=0.$$

由性质 4.5.1,$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_k\alpha_k=0$. 因此,$\alpha_1,\alpha_2,\dots,\alpha_k$ 线性相关. ■

性质 4.5.4 设 η 是数域 $\mathbb{F}$ 上的有限维线性空间 V_1 到 V_2 上的同构映射,则 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 V_1 的基的必要且充分条件为 $\{\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_n)\}$ 是 V_2 的基,从而 $\dim V_1=\dim V_2$.

证明 设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 V_1 的基,因此向量 $\alpha_1,\alpha_2,\dots,\alpha_n$ 线性无关. 由性质 4.5.3,向量 $\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_n)$ 线性无关.

另外,设 $\beta\in V_2$,由于 η 是 V_1 到 V_2 上的双射,所以存在 $\alpha\in V_1$,使得 $\eta(\alpha)=\beta$.

由于 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 V_1 的基,因此 $\alpha=a_1\alpha_1+a_2\alpha_2+\dots+a_n\alpha_n$,其中 $\alpha_1,\alpha_2,\dots,\alpha_n\in\mathbb{F}$. 由性质 4.5.2,

$$\beta=\eta(\alpha)=a_1\eta(\alpha_1)+a_2\eta(\alpha_2)+\cdots+a_n\eta(\alpha_n).$$

即 β 可由向量 $\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_n)$ 线性表出. 这就证明了 $\{\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_n)\}$ 是线性空间 V_2 的基.

反之,设 $\{\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_n)\}$ 是 V_2 的基,则向量 $\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_n)$ 线性无关. 由性质 4.5.3,向量 $\alpha_1,\alpha_2,\dots,\alpha_n$ 线性无关.

另外,设 $\alpha\in V_1$,因为 $\{\eta(\alpha_1),\eta(\alpha_2),\dots,\eta(\alpha_n)\}$ 为 V_2 的基,因此,

$$\eta(\alpha) = a_1\eta(\alpha_1) + a_2\eta(\alpha_2) + \cdots + a_n\eta(\alpha_n),$$

其中 $a_1, a_2, \ldots, a_n \in \mathbb{F}$.

由性质 4.5.2,$\eta(\alpha) = \eta(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n)$. 由于 η 是双射,所以

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n.$$

即向量 α 可由向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性表出. 因此,$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 V_1 的基. ■

线性空间之间的同构关系是数域 $\mathbb{F}$ 上的所有线性空间集合中的一种关系,它满足以下性质.

性质 4.5.5(自反性) 数域 $\mathbb{F}$ 上的线性空间 V 与 V 自身同构.

证明 定义线性空间 V 到自身的映射 ε 如下: 对任意 $\alpha \in V$,令

$$\varepsilon(\alpha) = \alpha.$$

映射 ε 称为 V 到自身的恒等映射. 显然,恒等映射 ε 是 V 到自身的一一映射. 并且对任意 $\alpha, \beta \in V$,$\lambda \in \mathbb{F}$,均有

$$\varepsilon(\alpha + \beta) = \alpha + \beta = \varepsilon(\alpha) + \varepsilon(\beta), \qquad \varepsilon(\lambda\alpha) = \lambda\alpha = \lambda\varepsilon(\alpha).$$

因此,ε 是线性空间 V 到自身上的同构映射,从而线性空间 V 与自身同构. ■

性质 4.5.6(对称性) 若 $\mathbb{F}$ 上的线性空间 V_1 与 V_2 同构,则 V_2 与 V_1 也同构.

证明 因为线性空间 V_1 与 V_2 同构,所以存在 V 到 V_2 上的同构映射 η.

定义 V_2 到 V_1 的映射 σ 如下: 设 $\beta \in V_2$,因为 η 是 V_1 到 V_2 上的满射,因此存在 $\alpha \in V_1$,使得 $\eta(\alpha) = \beta$,于是令

$$\sigma(\beta) = \alpha.$$

由于 η 是一一的,所以适合 $\eta(\alpha) = \beta$ 的 α 是唯一的. 因此映射 σ 有确切定义. 映射 σ 即是映射 η 的逆映射.

映射 σ 是单射. 事实上,设 $\beta_1, \beta_2 \in V_2$,且 $\sigma(\beta_1) = \sigma(\beta_2)$. 记 $\alpha_1 = \sigma(\beta_1)$,$\alpha_2 = \sigma(\beta_2)$. 由映射 σ 的定义,$\beta_1 = \eta(\alpha_1)$,$\beta_2 = \eta(\alpha_2)$. 因为 $\alpha_1 = \alpha_2$,故 $\beta_1 = \beta_2$.

映射 σ 是满射. 事实上,设 $\alpha \in V_1$,则 $\beta = \eta(\alpha) \in V_2$. 由映射 σ 的定义,$\sigma(\beta) = \alpha$.

映射 σ 是保加法的. 事实上,设 $\beta_1, \beta_2 \in V_2$,则存在 $\alpha_1, \alpha_2 \in V_1$,使得 $\eta(\alpha_1) = \beta_1$,$\eta(\alpha_2) = \beta_2$. 因为映射 η 保加法,所以 $\beta_1 + \beta_2 = \eta(\alpha_1) + \eta(\alpha_2) = \eta(\alpha_1 + \alpha_2)$. 由映射 σ 的定义,$\sigma(\beta_1 + \beta_2) = \alpha_1 + \alpha_2$. 由于 $\eta(\alpha_1) = \beta_1$,$\eta(\alpha_2) = \beta_2$,故 $\sigma(\beta_1) = \alpha_1$,$\sigma(\beta_2) = \alpha_2$. 所以,

$$\sigma(\beta_1 + \beta_2) = \alpha_1 + \alpha_2 = \sigma(\beta_1) + \sigma(\beta_2).$$

映射 σ 是保乘法的. 事实上,设 $\lambda \in \mathbb{F}$,$\beta \in V_2$,则存在 $\alpha \in V_1$,使得 $\eta(\alpha) = \beta$,从而 $\sigma(\beta) = \alpha$. 因为映射 η 保乘法,所以 $\eta(\lambda\alpha) = \lambda\eta(\alpha) = \lambda\beta$. 由映射 σ 的定义,

$$\sigma(\lambda\beta) = \lambda\alpha = \lambda\sigma(\beta).$$

这就证明,映射 σ 是线性空间 V_2 到 V_1 上的同构映射. 所以 V_2 与 V_1 同构. ■

性质 4.5.7（传递性）设 V_1, V_2 和 V_3 是数域 $\mathbb{F}$ 上的线性空间. 如果 V_1 与 V_2 同构，V_2 与 V_3 同构，则 V_1 与 V_3 同构.

证明 设 η 和 ζ 分别是 V_1 到 V_2 和 V_2 到 V_3 上的同构映射.

定义 V_1 到 V_3 的映射 ξ 如下：对任意 $\alpha \in V_1$，令

$$\xi(\alpha) = \zeta\big(\eta(\alpha)\big).$$

对任意 $\alpha, \beta \in V_1$，$\alpha \neq \beta$，由于 η 是 V_1 到 V_2 上的同构映射，因此 η 是 V_1 到 V_2 上的单射，所以 $\eta(\alpha) \neq \eta(\beta)$. 因为 ζ 是 V_2 到 V_3 上的单射，所以 $\zeta\big(\eta(\alpha)\big) \neq \zeta\big(\eta(\beta)\big)$，即 $\xi(\alpha) \neq \xi(\beta)$. 因此映射 ξ 是 V_1 到 V_3 上的单射.

对任意 $\gamma \in V_3$，由于 ζ 是 V_2 到 V_3 上的满射，因此存在 $\beta \in V_2$，使得 $\gamma = \zeta(\beta)$. 又 η 是 V_1 到 V_2 上的满射，所以存在 $\alpha \in V_1$，使得 $\beta = \eta(\alpha)$. 因此 $\gamma = \zeta\big(\eta(\alpha)\big) = \xi(\alpha)$. 这就证明，映射 ξ 是 V_1 到 V_3 上的满射.

对任意 $\alpha, \beta \in V_1$，$\lambda, \mu \in \mathbb{F}$，由于

$$\begin{aligned}\xi(\lambda\alpha + \mu\beta) &= \zeta\big(\eta(\lambda\alpha + \mu\beta)\big) = \zeta\big(\lambda\eta(\alpha) + \mu\eta(\beta)\big)\\ &= \lambda\zeta\big(\eta(\alpha)\big) + \mu\zeta\big(\eta(\beta)\big) = \lambda\xi(\alpha) + \mu\xi(\beta),\end{aligned}$$

因此映射 ξ 是保线性关系的，自然是保加法与乘法的.

这就证明，映射 ξ 是线性空间 V_1 到 V_3 上的同构映射，从而 V_1 与 V_3 同构. ■

由于数域 $\mathbb{F}$ 上的线性空间之间的同构关系满足自反性、对称性和传递性，所以同构关系是数域 $\mathbb{F}$ 上的线性空间之间的一种等价关系. 按照同构关系可以对线性空间进行分类：

彼此同构的线性空间归在同一个类，彼此不同构的线性空间归在不同的类.

和矩阵在相抵关系下分类相类似，基本的问题是：

如何判定数域 $\mathbb{F}$ 上的两个线性空间 V_1 和 V_2 是否属于同一个类，即如何判定线性空间 V_1 和 V_2 是否同构? 在线性空间的同构类中怎样选取代表元?

对此，有

定理 4.5.1 数域 $\mathbb{F}$ 上的任意一个 n 维线性空间 V 都同构于数域 $\mathbb{F}$ 上的 n 维向量空间 $\mathbb{F}^n$.

证明 在线性空间 V 中取定一组基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. 于是，向量 $\alpha \in V$ 便具有唯一的坐标 $(a_1, a_2, \ldots, a_n)$.

定义线性空间 V 到 $\mathbb{F}^n$ 的映射 η 如下：对任意向量 $\alpha \in V$，令

$$\eta(\alpha) = (a_1, a_2, \ldots, a_n),$$

其中 $(a_1, a_2, \ldots, a_n)$ 是向量 α 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标.

由于 V 中不同的向量 α 和 β 的坐标 $(a_1, a_2, \ldots, a_n)$ 和 $(b_1, b_2, \ldots, b_n)$ 不相等，因此当 $\alpha \neq \beta$ 时，$\eta(\alpha) \neq \eta(\beta)$. 所以映射 η 是 V 到 $\mathbb{F}^n$ 的单射.

其次，对任意 $(a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$，向量 $\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n \in V$，因此

$\eta(\alpha)=(a_1,a_2,\ldots,a_n)$,所以映射 η 是 V 到 $\mathbb{F}^n$ 上的满射.

最后,由于 V 中向量 α 与 β 的和 $\alpha+\beta$ 的坐标等于向量 α 与 β 的坐标之和,所以 $\eta(\alpha+\beta)=\eta(\alpha)+\eta(\beta)$; 而纯量 λ 与向量 α 的乘积 $\lambda\alpha$ 的坐标等于纯量 λ 与向量 α 的坐标的乘积,因此 $\eta(\lambda\alpha)=\lambda\eta(\alpha)$. 所以映射 η 是保加法和乘法的.

这就证明,映射 η 是线性空间 V 到 $\mathbb{F}^n$ 上的同构映射,所以 V 与 $\mathbb{F}^n$ 同构. ■

定理 4.5.2 数域 $\mathbb{F}$ 上的有限维线性空间 V_1 与 V_2 同构当且仅当

$$\dim V_1=\dim V_2.$$

证明 必要性即是性质 4.5.4. 下证充分性.

设 $\dim V_1=\dim V_2=n$,则由定理 4.5.1,线性空间 V_1 和 V_2 都同构于 $\mathbb{F}^n$. 由对称性,$\mathbb{F}^n$ 同构于 V_2. 由传递性,线性空间 V_1 与 V_2 同构. ■

定理 4.5.2 给出了数域 $\mathbb{F}$ 上的有限维线性空间 V_1 和 V_2 属于同一个同构类的判别准则. 定理 4.5.1 表明,在数域 $\mathbb{F}$ 上的所有 n 维线性空间构成的同构类中,可以取 n 维向量空间 $\mathbb{F}^n$ 作为它的代表元. 尽管数域 $\mathbb{F}$ 上的 n 维线性空间比较抽象,但从结构上看,可以用数域 $\mathbb{F}$ 上的 n 维向量空向来理解.

习 题 4.5

1. 证明,所有实数的集合 $\mathbb{R}$ 作为实线性空间与 §4.1 例 4.1.2 中的实线性空间 $\mathbb{R}_+$ 同构.

2. 如果有理数域 $\mathbb{Q}$ 上的线性空间 V_1 和 V_2 之间存在一一对应,那么线性空间 V_1 和 V_2 一定同构吗?

3. 设 V 是 n 维复线性空间. 取集合 V,数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$. 定义 V 中向量的加法为复线性空间 V 的向量加法,而纯量 λ 与向量的乘法定义为实数与 V 中的向量的乘法. 如此得到的实线性空间记为 V^*. 试确定 $\dim_{\mathbb{R}} V^*$.

4. 设 V 是 n 维实线性空间. 如果保留 V 的向量加法,但在纯量 λ 乘以向量时,限定纯量 λ 只取有理数,如此得到的有理域 $\mathbb{Q}$ 上的线性空间记为 $\widetilde{V}$. 线性空间 $\widetilde{V}$ 是否是有限维的?

§4.6 子 空 间

大家知道,在三维实向量空间 $\mathbb{R}^3$ 中,取定非零向量,则对于任意纯量 $\lambda\in\mathbb{R}$,向量 $\lambda\alpha$ 与 α 共线. 反之,如果向量 β 与 α 共线,则存在纯量 $\lambda\in\mathbb{R}$,使得 $\beta=\lambda\alpha$. 因此 $\mathbb{R}^3$ 中所有与向量 α 共线的向量集合为 $S=\left\{\lambda\alpha \,\middle|\, \lambda\in\mathbb{R}\right\}$. 向量集合 S 具有以下两个特点:

(1) 与向量 α 共线的向量之和仍与向量 α 共线,即对任意 $\alpha_1,\alpha_2\in S$,均有 $\alpha_1+\alpha_2\in S$. 换言之,向量集合 S 对 $\mathbb{R}^3$ 的向量加法是封闭的;

(2) 与向量 α 共线的向量 β 的任意纯量倍 $\lambda\beta$ 仍与向量 α 共线,其中 $\lambda\in\mathbb{R}$,即对任意 $\lambda\in\mathbb{R}$,$\beta\in S$,均有 $\lambda\beta\in S$. 换言之,向量集合对 $\mathbb{R}^3$ 的纯量与向量的乘法是封闭的.

同样,如果向量 α 与 β 不共线,则所有与向量 α 及 β 共面的向量集合为 $\widetilde{S} = \left\{\lambda\alpha + \mu\beta \mid \lambda, \mu \in \mathbb{R}\right\}$. 向量集合 $\widetilde{S}$ 具有如下两个特点.

(1) 与向量 α 及 β 共面的向量之和仍与向量 α 及 β 共面,即对任意 $\alpha_1, \alpha_2 \in \widetilde{S}$,均有 $\alpha_1 + \alpha_2 \in \widetilde{S}$. 换言之,向量集合 $\widetilde{S}$ 对 $\mathbb{R}^3$ 的向量加法是封闭的;

(2) 与向量 α 及 β 共面的向量 γ 的任意纯量倍 $\lambda\gamma$ 仍与向量 α 及 β 共面,其中 $\lambda \in \mathbb{R}$,即对任意 $\lambda \in \mathbb{R}, \gamma \in \widetilde{S}$,均有 $\lambda\gamma \in \widetilde{S}$. 换言之,向量集合 $\widetilde{S}$ 对 $\mathbb{R}^3$ 的纯量与向量的乘法是封闭的.

把三维实向量空间 $\mathbb{R}^3$ 中上述向量集合 S 和 $\widetilde{S}$ 所具有的共性加以抽象,就成为线性空间的子空间概念.

定义 4.6.1 设 U 是数域 $\mathbb{F}$ 上的线性空间 V 的非空向量集合. 如果 U 对于线性空间 V 的向量加法和纯量与向量的乘法是封闭的,即对任意 $\alpha, \beta \in U$,有 $\alpha + \beta \in U$,并且对任意 $\lambda \in \mathbb{F}, \alpha \in U$,有 $\lambda\alpha \in U$,则 U 称为**线性空间 V 的子空间**.

应当指出,线性空间 V 的子空间 U 一定含有零向量. 事实上,由子空间的定义,向量集合 U 非空,因此存在向量 $\alpha \in U$. 由于 U 对纯量与向量的乘法是封闭的,所以 $(-1)\alpha = -\alpha \in U$. 由于 U 对向量的加法封闭,因此

$$\alpha + (-\alpha) = 0 \in U.$$

其次,线性空间 V 的子空间 U 本身是数域 $\mathbb{F}$ 上的线性空间.

事实上,由于子空间 U 对线性空间 V 的向量加法和纯量与向量的乘法封闭,因此可以规定 U 的向量加法为 V 的向量加法,而 U 的纯量与向量的乘法取为 V 的纯量与向量的乘法. 由于子空间 U 是线性空间 V 的向量集合,因此向量集合 U 对上面取定的两种代数运算自然满足线性空间定义中的八条公理. 所以子空间 U 本身也是线性空间.

另外,如果线性空间 V 是 n 维的,则 V 中任意 $n+1$ 个向量一定线性相关. 因此,子空间 U 中任意 $n+1$ 个向量也一定线性相关,所以 U 中线性无关的向量集合所含向量的个数不超过 n,从而 $\dim U \leqslant \dim V$.

容易验证,线性空间 V 本身是 V 的一个子空间,只由零向量构成的向量集合 $\{0\}$ 是 V 的子空间. 后者称为 V 的**零子空间**,记为 O.

V 和零子空间 O 称为 V 的**平凡子空间**. 不是平凡的子空间称为**真子空间**.

例 4.6.1 设 $\mathbb{F}_n[x]$ 是数域 $\mathbb{F}$ 上的所有关于未定元 x 的次数小于 n 的多项式构成的线性空间. 设正整数 $m \leqslant n-1$,取定 $x_1, x_2, \ldots, x_m \in \mathbb{F}$. $\mathbb{F}_n[x]$ 中所有满足 $f(x_1) = f(x_2) = \cdots = f(x_m) = 0$ 的多项式 $f(x)$ 的集合记为 U. 证明,U 是 $\mathbb{F}_n[x]$ 的子空间.

证明 首先,显然零多项式 $0 \in U$,因此 $U \neq \varnothing$.

其次,设 $f, g \in U$,则 $f(x_1) = \cdots = f(x_m) = 0, g(x_1) = \cdots = g(x_m) = 0$. 因此,

$$(f+g)(x_i) = f(x_i) + g(x_i) = 0.$$

所以，$f+g\in U$，即 U 对 $\mathbb{F}_n[x]$ 的向量加法封闭.

最后，设 $\lambda\in\mathbb{F}$，$f\in U$，则 $f(x_1)=\cdots=f(x_m)=0$. 因此对于 $i=1,2,\ldots,m$，

$$(\lambda f)(x_i)=\lambda f(x_i)=0.$$

所以 $\lambda f\in U$，即 U 对 $\mathbb{F}_n[x]$ 纯量与向量的乘法封闭. 从而 U 是 $\mathbb{F}_n[x]$ 的子空间. ■

例 4.6.2　证明，在数域 $\mathbb{F}$ 上的所有关于未定元 x 的多项式构成的线性空间 $\mathbb{F}[x]$ 中，所有满足 $f(-x)=f(x)$ 的多项式集合 U 是 $\mathbb{F}[x]$ 的子空间.

证明　首先，显然零多项式 $0\in U$.

其次，设 $f,g\in U$，则 $f(-x)=f(x)$，$g(-x)=g(x)$. 因此，

$$(f+g)(-x)=f(-x)+g(-x)=f(x)+g(x)=(f+g)(x),$$

所以 $f+g\in U$，即 U 对 $\mathbb{F}[x]$ 的向量加法是封闭的.

最后，设 $\lambda\in\mathbb{F}$，$f\in U$，则由 $f(-x)=f(x)$ 得到，

$$(\lambda f)(-x)=\lambda f(-x)=\lambda f(x)=(\lambda f)(x).$$

因此 $\lambda f\in U$，即 U 对 $\mathbb{F}[x]$ 的纯量与向量的乘法封闭.

所以 U 是 $\mathbb{F}[x]$ 的子空间. ■

例 4.6.3　设 A 是数域 $\mathbb{F}$ 上的 $m\times n$ 矩阵，$x=(x_1,x_2,\ldots,x_n)^{\mathrm{T}}$，其中 $x_1,x_2,\ldots,x_n$ 是未知量. 齐次方程组 $Ax=0$ 的所有解的集合记为 V_A.

证明，V_A 是数域 $\mathbb{F}$ 上的 n 维列向量空间 $\mathbb{F}^n$ 的子空间，并求 $\dim V_A$.

解　显然，齐次方程组 $Ax=0$ 具有零解 $0\in\mathbb{F}^n$，即 $0\in U$，因此 $U\neq\varnothing$.

设 $x,y\in V_A$，则 $Ax=0$，$Ay=0$. 所以 $A(x+y)=Ax+Ay=0$，因此 $x+y\in V_A$.

设 $\lambda\in\mathbb{F}$，$x\in V_A$，则由 $A(\lambda x)=\lambda Ax=0$ 得到 $\lambda x\in V_A$. 所以 V_A 是 $\mathbb{F}^n$ 的子空间. 它称为齐次方程组 $Ax=0$ 的**解空间**.

由 §3.6 齐次线性方程组解的结构定理可知，如果 $\operatorname{rank}A=r$，并且

$$A=P\begin{pmatrix}I_{(r)} & 0\\ 0 & 0\end{pmatrix}Q,$$

其中 P 和 Q 分别是 m 阶和 n 阶可逆方阵，则方程组 $Ax=0$ 的通解为

$$x=t_{r+1}Q^{-1}\varepsilon_{r+1}+t_{r+2}Q^{-1}\varepsilon_{r+2}+\cdots+t_nQ^{-1}\varepsilon_n, \tag{4.6.1}$$

其中 $t_{r+1},t_{r+2},\ldots,t_n\in\mathbb{F}$，而 $\varepsilon_i=(0,\ldots,0,\underset{i}{1},0,\ldots,0)^{\mathrm{T}}$.

由于

$$AQ^{-1}\varepsilon_i=P\begin{pmatrix}I_{(r)} & 0\\ 0 & 0\end{pmatrix}\varepsilon_i=0,$$

因此，$Q^{-1}\varepsilon_i\in V_A$，$i=r+1,r+2,\ldots,n$. 式 (4.6.1) 表明，$V_A$ 中任意向量 x 都可由向量 $Q^{-1}\varepsilon_{r+1},Q^{-1}\varepsilon_{r+2},\ldots,Q^{-1}\varepsilon_n$ 线性表出. 另外，如果存在 $\lambda_{r+1},\lambda_{r+2},\ldots,\lambda_n\in\mathbb{F}$ 使得

$$\lambda_{r+1}Q^{-1}\varepsilon_{r+1}+\lambda_{r+2}Q^{-1}\varepsilon_{r+2}+\cdots+\lambda_nQ^{-1}\varepsilon_n=0,$$

则上式两端同时左乘方阵 Q，便得到

$$\lambda_{r+1}\varepsilon_{r+1}+\lambda_{r+2}\varepsilon_{r+2}+\cdots+\lambda_n\varepsilon_n=0,$$

从而 $\lambda_{r+1}=\lambda_{r+2}=\cdots=\lambda_n=0$. 所以向量 $Q^{-1}\varepsilon_{r+1},Q^{-1}\varepsilon_{r+2},\ldots,Q^{-1}\varepsilon_n$ 线性无关.

于是,$\{Q^{-1}\varepsilon_{r+1},Q^{-1}\varepsilon_{r+2},\ldots,Q^{-1}\varepsilon_n\}$ 是 V 的基. 从而 $\dim V_A=n-\operatorname{rank}A$. ■

设齐次方程组 $Ax=0$ 的解空间 V 的一组基为 $\{\alpha_1,\alpha_2,\ldots,\alpha_k\}$,则向量 $\alpha_1,\alpha_2,\ldots,\alpha_k$ 称为方程组 $Ax=0$ 的**基础解系**.

例如,例 4.6.3 中向量 $Q^{-1}\varepsilon_{r+1},Q^{-1}\varepsilon_{r+2},\ldots,Q^{-1}\varepsilon_n$ 即是 $Ax=0$ 的一个基础解系.

现在介绍子空间的运算.

定义 4.6.2 设 I 是下标集合,$\{V_\nu \mid \nu\in I\}$ 是数域 $\mathbb{F}$ 上的线性空间 V 的子空间集合. V 中所有属于每个子空间 V_ν 的向量集合称为所有子空间 V_ν 的**交**,记为

$$\bigcap_{\nu\in I}V_\nu.$$

定理 4.6.1 任意多个子空间的交是子空间.

证明 设 $\{V_\nu \mid \nu\in I\}$ 是线性空间 V 的子空间集合. 由于 V_ν 是子空间,因此,零向量 $0\in V_\nu$,所以

$$0\in\bigcap_{\nu\in I}V_\nu\neq\varnothing.$$

设 α 与 β 是所有子空间 V_ν 的交的两个向量,则 $\alpha,\beta\in V_\nu$. 由于 V_ν 是子空间,因此,$\alpha+\beta\in V_\nu$. 所以

$$\alpha+\beta\in\bigcap_{\nu\in I}V_\nu,$$

即所有子空间 V_ν 的交对线性空间 V 的向量加法封闭.

又设 $\lambda\in\mathbb{F}$,α 是所有子空间 V_ν 的交的某个向量,因此 $\alpha\in V_\nu$. 由于 V_ν 是子空间,所以 $\lambda\alpha\in V_\nu$. 因此

$$\lambda\alpha\in\bigcap_{\nu\in I}V_\nu,$$

即所有子空间 V_ν 的交对线性空间 V 的纯量与向量的乘法封闭.

这就证明,子空间 V_ν 的交是子空间. ■

在考虑线性空间 V 的子空间运算时,自然要考虑子空间的并. 但子空间的并一般不是子空间. 例如,在二维实向量空间 $\mathbb{R}^2$ 中,子空间 $V_1=\{\lambda\varepsilon_1 \mid \lambda\in\mathbb{R}\}$ 和 $V_2=\{\mu\varepsilon_2 \mid \mu\in\mathbb{R}\}$ 的并 $V_1\cup V_2$ 不再是子空间,因为 $\varepsilon_1\in V_1$,$\varepsilon_2\in V_2$,但 $\varepsilon_1+\varepsilon_2\notin V_1\cup V_2$,即集合 $V_1\cup V_2$ 对 $\mathbb{R}^2$ 的向量加法并不封闭.

所以,在子空间的运算中不考虑子空间的并,而是用子空间的和来代替.

定义 4.6.3 设 V_1 和 V_2 是数域 $\mathbb{F}$ 上的线性空间 V 的子空间,则集合

$$\{\alpha+\beta \mid \alpha\in V_1,\quad \beta\in V_2\}$$

称为**子空间 V_1 与 V_2 的和**,记为 V_1+V_2.

定理 4.6.2 线性空间 V 的子空间 V_1 与 V_2 的和 V_1+V_2 是 V 的子空间.

证明 因为子空间 V_1 与 V_2 都含有零向量,因此零向量

$$0=0+0\in V_1+V_2\neq\varnothing.$$

设向量 $\gamma_1, \gamma_2 \in V_1 + V_2$，则存在向量 $\alpha_1, \alpha_2 \in V_1$，$\beta_1, \beta_2 \in V_2$，使得 $\gamma_1 = \alpha_1 + \beta_1$，$\gamma_2 = \alpha_2 + \beta_2$. 因为 V_1 与 V_2 是子空间，所以，$\alpha_1 + \alpha_2 \in V_1$，$\beta_1 + \beta_2 \in V_2$，因此，

$$\gamma_1 + \gamma_2 = (\alpha_1 + \beta_1) + (\alpha_2 + \beta_2) = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) \in V_1 + V_2.$$

即集合 $V_1 + V_2$ 对线性空间 V 的向量加法封闭.

设 $\lambda \in \mathbb{F}$，$\gamma \in V_1 + V_2$，则存在 $\alpha \in V_1$ 和 $\beta \in V_2$，使得 $\gamma = \alpha + \beta$. 由于 V_1 与 V_2 是子空间，所以 $\lambda\alpha \in V_1$，$\lambda\beta \in V_2$. 于是

$$\lambda\gamma = \lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta \in V_1 + V_2,$$

即集合 $V_1 + V_2$ 对线性空间 V 的纯量与向量的乘法封闭.

因此，子空间 V_1 与 V_2 的和 $V_1 + V_2$ 是线性空间 V 的子空间. ■

两个子空间的和的概念可以推广到有限多个子空间. 其定义如下.

设 $V_1, V_2, \ldots, V_k$ 是数域 $\mathbb{F}$ 上的线性空间 V 的子空间，则集合

$$\left\{\alpha_1 + \alpha_2 + \cdots + \alpha_k \,\middle|\, \alpha_i \in V_i,\quad i = 1, 2, \ldots, k\right\}$$

称为**子空间 $\boldsymbol{V_1, V_2, \ldots, V_k}$ 的和**，记为 $V_1 + V_2 + \cdots + V_k$.

可以证明，V 的子空间 $V_1, V_2, \ldots, V_k$ 的和 $V_1 + V_2 + \cdots + V_k$ 是 V 的子空间.

关于子空间的交与和的维数，有以下重要定理.

定理 4.6.3（维数公式） 设 V_1 与 V_2 是线性空间 V 的子空间，则

$$\dim V_1 + \dim V_2 = \dim(V_1 \cap V_2) + \dim(V_1 + V_2).$$

证明　设 $\dim V_1 = r$，$\dim V_2 = s$，$\dim(V_1 \cap V_2) = t$，并且 $\{\alpha_1, \alpha_2, \ldots, \alpha_t\}$ 是子空间 $V_1 \cap V_2$ 的一组基. 由于 $V_1 \cap V_2 \subseteq V_1$，$V_1 \cap V_2 \subseteq V_2$，因此向量 $\alpha_1, \alpha_2, \ldots, \alpha_t$ 分别是子空间 V_1 与 V_2 的线性无关向量，所以向量 $\alpha_1, \alpha_2, \ldots, \alpha_t$ 可以分别扩充为子空间 V_1 与 V_2 的基.

设 $\{\alpha_1, \alpha_2, \ldots, \alpha_t, \beta_1, \ldots, \beta_{r-t}\}$ 与 $\{\alpha_1, \alpha_2, \ldots, \alpha_t, \gamma_1, \ldots, \gamma_{s-t}\}$ 分别是子空间 V_1 与 V_2 的基. 首先，向量 $\alpha_1, \alpha_2, \ldots, \alpha_t, \beta_1, \ldots, \beta_{r-t}, \gamma_1, \ldots, \gamma_{s-t}$ 线性无关. 事实上，设

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_t\alpha_t + b_1\beta_1 + \cdots + b_{r-t}\beta_{r-t} + c_1\gamma_1 + \cdots + c_{s-t}\gamma_{s-t} = 0,$$

其中 $a_i, b_j, c_k \in \mathbb{F}$，$1 \leqslant i \leqslant t$，$1 \leqslant j \leqslant r - t$，$1 \leqslant k \leqslant s - t$，则记

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_t\alpha_t + b_1\beta_1 + \cdots + b_{r-t}\beta_{r-t} = -(c_1\gamma_1 + \cdots + c_{s-t}\gamma_{s-t}).$$

显然，$\alpha \in V_1 \cap V_2$. 由于 $\{\alpha_1, \alpha_2, \ldots, \alpha_t\}$ 是子空间 $V_1 \cap V_2$ 的基，因此，

$$\alpha = -(c_1\gamma_1 + \cdots + c_{s-t}\gamma_{s-t}) = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_t\alpha_t,$$

其中 $d_1, d_2, \ldots, d_t \in \mathbb{F}$. 所以，

$$d_1\alpha_1 + d_2\alpha_2 + \cdots + d_t\alpha_t + c_1\gamma_1 + \cdots + c_{s-t}\gamma_{s-t} = 0.$$

由于 $\{\alpha_1, \alpha_2, \ldots, \alpha_t, \gamma_1, \ldots, \gamma_{s-t}\}$ 是子空间 V_2 的基，所以，

$$d_1 = d_2 = \cdots = d_t = c_1 = c_2 = \cdots = c_{s-t} = 0.$$

因此，$\alpha=0$，即

$$a_1\alpha_1+a_2\alpha_2+\cdots+a_t\alpha_t+b_1\beta_1+\cdots+b_{r-t}\beta_{r-t}=0.$$

由于 $\{\alpha_1,\alpha_2,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t}\}$ 是子空间 V_1 的基，所以，

$$a_1=a_2=\cdots=a_t=b_1=b_2=\cdots=b_{r-t}=0.$$

前面已证 $c_1=c_2=\cdots=c_{s-t}=0$，因此 $\alpha_1,\alpha_2,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}$ 线性无关.

由于 $V_1,V_2,V_1\cap V_2\subseteq V_1+V_2$，因此向量

$$\alpha_1,\alpha_2,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}\in V_1+V_2.$$

其次，对任意 $\gamma\in V_1+V_2$，向量 γ 可由向量 $\alpha_1,\alpha_2,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}$ 线性表出. 事实上，因为 $\gamma\in V_1+V_2$，所以存在 $\alpha\in V_1$ 及 $\beta\in V_2$，使得 $\gamma=\alpha+\beta$.

由于 $\{\alpha_1,\alpha_2,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t}\}$ 与 $\{\alpha_1,\alpha_2,\ldots,\alpha_t,\gamma_1,\ldots,\gamma_{s-t}\}$ 分别是子空间 V_1 与 V_2 的基，所以，

$$\begin{aligned}\alpha&=a_1\alpha_1+a_2\alpha_2+\cdots+a_t\alpha_t+b_1\beta_1+\cdots+b_{r-t}\beta_{r-t},\\ \beta&=c_1\alpha_1+c_2\alpha_2+\cdots+c_t\alpha_t+d_1\gamma_1+\cdots+d_{s-t}\gamma_{s-t},\end{aligned}$$

其中 $a_i,b_j,c_k,d_\ell\in\mathbb{F}$，$1\leqslant i\leqslant t$，$1\leqslant j\leqslant r-t$，$1\leqslant k\leqslant t$，$1\leqslant \ell\leqslant s-t$. 因此，

$$\begin{aligned}\gamma=(a_1+c_1)\alpha_1+(a_2+c_2)\alpha_2+\cdots+(a_t+c_t)\alpha_t&\\ +b_1\beta_1+\cdots+b_{r-t}\beta_{r-t}+d_1\gamma_1+\cdots+d_{s-t}\gamma_{s-t}.&\end{aligned}$$

这就证明，子空间 V_1+V_2 中任一向量 γ 都可由 $\alpha_1,\alpha_2,\ldots,\alpha_t,\beta_1,\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}$ 线性表出. 因此，$\{\alpha_1,\alpha_2,\ldots,\alpha_t,\beta_1\ldots,\beta_{r-t},\gamma_1,\ldots,\gamma_{s-t}\}$ 是子空间 V_1+V_2 的基. 所以，

$$\dim(V_1+V_2)=r+s-t=\dim V_1+\dim V_2-\dim(V_1\cap V_2).$$ ■

推论 4.6.1 设 V_1 和 V_2 是线性空间 V 的子空间，则

$$\dim(V_1+V_2)\leqslant\dim V_1+\dim V_2,$$

其中当且仅当它们的交 $V_1\cap V_2$ 为零子空间时等式成立.

推论 4.6.2 设 V_1 和 V_2 是线性空间 V 的子空间，则

$$\dim(V_1\cap V_2)\geqslant\dim V_1+\dim V_2-\dim V.$$

推论 4.6.3 设 V_1 和 V_2 是线性空间 V 的子空间，并且 $\dim V_1+\dim V_2>\dim V$，则它们的交 $V_1\cap V_2$ 含有非零向量.

在子空间中，由向量集合生成的子空间具有重要的作用.

定义 4.6.4 设 S 是数域 $\mathbb{F}$ 上的线性空间 V 的非空向量集合. V 中所有包含向量集合 S 的子空间的交称为**由向量集合 S 生成的子空间**，记为 $V(S)$.

特别，如果 $S=\{\alpha_1,\alpha_2,\ldots,\alpha_k\}$，则由 S 生成的子空间记为 $V(\alpha_1,\alpha_2,\ldots,\alpha_k)$，并且称为**由向量 $\boldsymbol{\alpha_1,\alpha_2,\ldots,\alpha_k}$ 生成的子空间**.

由于 $S \subseteq V$,因此所有包含 S 的子空间集合非空,所以 $V(S)$ 是有意义的. 其次,由于 $V(S)$ 是所有包含 S 的子空间的交,所以 $S \subseteq V(S)$,并且任意一个包含 S 的子空间一定包含子空间 $V(S)$. 这表明,由 S 生成的子空间是所有包含 S 的子空间中最小的一个. 由向量集合 S 生成的子空间 $V(S)$ 究竟由哪些向量组成? 下一定理给出了答案.

定理 4.6.4 设 S 是数域 $\mathbb{F}$ 上的线性空间 V 的向量集合. 则由 S 生成的子空间 $V(S)$ 等于所有 S 中有限个向量的线性组合的集合.

证明 记所有 S 中有限个向量的线性组合的集合记为 V_0.

设 $\alpha \in S$,则由集合 V_0 的定义,$0\alpha = 0 \in V_0$,其中 $0 \in \mathbb{F}$. 因此,$V_0 \neq \varnothing$.

显然,S 中两个有限个向量的线性组合之和仍是 S 中有限个向量的线性组合,所以,如果 $\alpha, \beta \in V_0$,则 $\alpha + \beta \in V_0$. 又 S 中有限个向量的线性组合的纯量倍仍是 S 中有限个向量的线性组合,所以,如果 $\lambda \in \mathbb{F}$,$\alpha \in V_0$,则 $\lambda\alpha \in V_0$.

这表明,集合 V_0 是 V 的子空间. 显然,$S \subseteq V_0$,所以 V_0 是 V 中包含 S 的子空间. 由 $V(S)$ 的最小性,$V(S) \subseteq V_0$.

另一方面,由于 $V(S)$ 是包含 S 的子空间,因此,S 中有限个向量的线性组合一定属于 $V(S)$,所以,$V_0 \subseteq V(S)$. 从而,$V(S) = V_0$. ■

特别地,若 $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$,由定理 4.6.4,由 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 生成的子空间为

$$V(\alpha_1, \alpha_2, \ldots, \alpha_n) = \{a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n \mid a_1, a_2, \ldots, a_n \in \mathbb{F}\}.$$

如果向量集合 $S = V_1 \cup V_2 \cup \cdots \cup V_m$,其中 $V_1, V_2, \ldots, V_m$ 是线性空间 V 的子空间,则由定理 4.6.4 可以推出,

$$\begin{aligned} V(V_1 \cup V_2 \cup \cdots \cup V_n) &= \{a_1\alpha_1 + a_2\alpha_2 + \cdots + a_m\alpha_m \mid a_i \in \mathbb{F},\ \ \alpha_i \in V_i,\ \ 1 \leqslant i \leqslant m\} \\ &= V_1 + V_2 + \cdots + V_n, \end{aligned}$$

即由子空间 $V_1, V_2, \ldots, V_m$ 生成的子空间(也即由向量集合 $V_1 \cup V_2 \cup \cdots \cup V_m$ 生成的子空间)等于它们的和.

由于由 S 生成的子空间是所有包含 S 的子空间,因此,当 S 本身是子空间时,由 S 生成的子空间即是子空间 S 自身. 所以,由两个子空间的交生成的子空间即是它们的交.

定理 4.6.5 设 S 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的向量集合. 则 S 的极大线性无关向量组是由 S 生成的子空间 $V(S)$ 的基.

证明 由于 $\dim V = n$,所以线性空间 V 中任意 $n+1$ 个向量线性相关. 因此,S 的极大线性无关向量组所含向量的个数是有限的. 于是可设 $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ 是 S 的极大线性无关向量组.

由于 $S \subseteq V(S)$,所以向量 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 是 $V(S)$ 的线性无关向量. 由定理 4.6.4,$V(S)$ 中每个向量都是 S 中有限个向量的线性组合,而 S 中每个向量又是向量

$\alpha_1,\alpha_2,\dots,\alpha_r$ 的线性组合,因此 $V(S)$ 中每个向量都是向量 $\alpha_1,\alpha_2,\dots,\alpha_r$ 的线性组合. 所以 $\{\alpha_1,\alpha_2,\dots,\alpha_r\}$ 是 $V(S)$ 的基. ■

由定理 4.6.5 的证明可以知道,如果 $\{\alpha_1,\alpha_2,\dots,\alpha_r\}$ 是 S 的极大线性无关向量组,则

$$V(S)=V(\alpha_1,\alpha_2,\dots,\alpha_r).$$

即由 S 生成的子空间等于由 S 的极大线性无关向量组生成的子空间.

例 4.6.4 证明,无限数域 $\mathbb{F}$ 上的线性空间 V 不能被它的有限个真子空间所覆盖,即如果 $V_1,V_2,\dots,V_k$ 是线性空间 V 的真子空间,则存在向量 $\alpha\in V$,使得 $\alpha\notin V_1\cup V_2\cup\cdots\cup V_k$.

证明 对子空间个数 k 用归纳法. 当 $k=1$ 时结论显然成立. 假设结论对 k 成立,下面证明,结论对 $k+1$ 成立. 如果 $V_{k+1}\subseteq V_1\cup V_2\cup\cdots\cup V_k$,则

$$V_1\cup V_2\cup\cdots\cup V_k\cup V_{k+1}=V_1\cup V_2\cup\cdots\cup V_k.$$

由归纳假设,存在向量 $\alpha\in V$,使得

$$\alpha\notin V_1\cup V_2\cup\cdots\cup V_k=V_1\cup V_2\cup\cdots\cup V_k\cup V_{k+1}.$$

因此,当 $V_{k+1}\subseteq V_1\cup V_2\cup\cdots\cup V_k$ 时结论成立.

如果存在向量 $\beta\in V_{k+1}$,但 $\beta\notin V_1\cup V_2\cup\cdots\cup V_k$,由于 V_{k+1} 是 V 的真子空间,因此存在向量 $\gamma\in V$,但 $\gamma\notin V_{k+1}$.

考虑所有形如 $\beta+\lambda\gamma$ 的向量,这里 $\lambda\neq 0,\lambda\in\mathbb{F}$,其中至多有一个属于 V_{k+1}.

因为否则将有 $\lambda_1,\lambda_2\in\mathbb{F},\lambda_1\neq\lambda_2$,使得 $\beta+\lambda_1\gamma,\beta+\lambda_2\gamma\in V_{k+1}$,从而

$$(\beta+\lambda_1\gamma)-(\beta+\lambda_2\gamma)=(\lambda_1-\lambda_2)\gamma\in V_{k+1}.$$

即 $\gamma\in V_{k+1}$,不可能.

因此,必有无数多个形如 $\beta+\lambda\gamma$ 的向量不在 V_{k+1} 中. 如果不在 V_{k+1} 中的形如 $\beta+\lambda\gamma$ 的向量都属于 $V_1\cup V_2\cup\cdots\cup V_k$,则由抽屉原理,其中至少有两个向量,例如 $\beta+\mu_1\gamma,\beta+\mu_2\gamma$ 同在某个子空间 V_i 中,$1\leqslant i\leqslant k$. 这里 $\mu_1\neq\mu_2$. 由于 $\beta\notin V_1\cup V_2\cup\cdots\cup V_k$,所以,$\mu_1\neq 0,\mu_2\neq 0$. 因此,$\mu_1^{-1}\beta+\gamma,\mu_2^{-1}\beta+\gamma$ 同属于 V_i. 于是,

$$\left(\frac{1}{\mu_1}\beta+\gamma\right)-\left(\frac{1}{\mu_2}\beta+\gamma\right)=\left(\frac{1}{\mu_1}-\frac{1}{\mu_2}\right)\beta\in V_i\implies\beta\in V_i,$$

即 $\beta\in V_1\cup V_2\cup\cdots\cup V_k$,不可能.

这就证明,必有一个向量 $\beta+\lambda_0\gamma\notin V_{k+1}$ 且 $\beta+\lambda_0\gamma\notin V_1\cup V_2\cup\cdots\cup V_k$. 从而

$$\beta+\lambda_0\gamma\notin V_1\cup V_2\cup\cdots\cup V_k\cup V_{k+1}$$ ■

最后用一个利用解空间的维数证明矩阵秩不等式的例子结束本节.

例 4.6.5 设 A 和 B 分别是数域 $\mathbb{F}$ 上的 $m\times n$ 和 $n\times p$ 矩阵. 证明

$$\operatorname{rank}A+\operatorname{rank}B-n\leqslant\operatorname{rank}AB.$$

证明 由例 4.6.3,解空间

$$V_A = \left\{x \in \mathbb{F}^n \,\middle|\, Ax = 0\right\},\quad V_B = \left\{y \in \mathbb{F}^p \,\middle|\, By = 0\right\},\quad V_{AB} = \left\{z \in \mathbb{F}^p \,\middle|\, ABz = 0\right\}.$$

的维数分别是 $\dim V_A = n - \operatorname{rank} A$, $\dim V_B = p - \operatorname{rank} B$, $\dim V_{AB} = p - \operatorname{rank} AB$. 记

$$V_0 = \left\{x \in \mathbb{F}^n \,\middle|\, x = By,\quad y \in V_{AB}\right\}.$$

容易验证, V_0 是 $\mathbb{F}^n$ 的子空间, 并且 $V_0 \subseteq V_A$. 如果能够证明,

$$\dim V_0 = \dim V_{AB} - \dim V_B, \tag{4.6.2}$$

则结论已然成立. 下面证明 (4.6.2) 成立.

记 $\dim V_B = s$, $\dim V_{AB} = t$. 由于 $V_B \subseteq V_{AB}$, 所以, 子空间 V_B 的基 $\{y_1, y_2, \ldots, y_s\}$ 可以扩充成 V_{AB} 的基 $\{y_1, y_2, \ldots, y_s, y_{s+1}, \ldots, y_t\}$.

显然, $By_{s+1}, By_{s+2}, \ldots, By_t \in V_0$. 设

$$\lambda_{s+1}By_{s+1} + \lambda_{s+2}By_{s+2} + \cdots + \lambda_t By_t = B(\lambda_{s+1}y_{s+1} + \cdots + \lambda_t y_t) = 0,$$

其中 $\lambda_{s+1}, \lambda_{s+2}, \ldots, \lambda_t \in \mathbb{F}$. 则 $\lambda_{s+1}y_{s+1} + \cdots + \lambda_t y_t \in V_B$.

由于 $\{y_1, y_2, \ldots, y_s\}$ 是 V_B 的基, 所以,

$$\lambda_{s+1}y_{s+1} + \lambda_{s+2}y_{s+2} + \cdots + \lambda_t y_t = \lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_s y_s,$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{F}$. 因此,

$$\lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_s y_s + (-\lambda_{s+1})y_{s+1} + (-\lambda_{s+2})y_{s+2} + \cdots + (-\lambda_t)y_t = 0.$$

由于 $\{y_1, y_2, \ldots, y_s, y_{s+1}, \ldots, y_t\}$ 是 V_{AB} 的基, 所以,

$$\lambda_{s+1} = \lambda_{s+2} = \cdots = \lambda_t = 0.$$

这表明, V_0 中向量 $By_{s+1}, By_{s+2}, \ldots, By_t$ 线性无关.

其次, 设 $x \in V_0$, 则存在向量 $y \in V_{AB}$, 使得 $x = By$.

由于 $y \in V_{AB}$, 而 $\{y_1, y_2, \ldots, y_s, y_{s+1}, \ldots, y_t\}$ 是 V_{AB} 的基, 因此,

$$y = a_1 y_1 + a_2 y_2 + \cdots + a_s y_s + a_{s+1} y_{s+1} + \cdots + a_t y_t,$$

其中 $a_1, a_2, \ldots, a_t \in \mathbb{F}$. 所以,

$$x = By = a_1 By_1 + a_2 By_2 + \cdots + a_s By_s + a_{s+1} By_{s+1} + \cdots + a_t By_t.$$

由于向量 $y_1, y_2, \ldots, y_s \in V_B$, 所以 $By_1 = By_2 = \cdots = By_s = 0$. 因此,

$$x = a_{s+1} By_{s+1} + a_{s+2} By_{s+2} + \cdots + a_t By_t.$$

这就证明了 $\{By_{s+1}, By_{s+2}, \ldots, By_t\}$ 是子空间 V_0 的基. 于是

$$\dim V_0 = \dim V_{AB} - \dim V_B. \qquad \blacksquare$$

习　题　4.6

1. 在数域 $\mathbb{F}$ 上的所有 n 阶方阵构成的线性空间 $\mathbb{F}^{n\times n}$ 中, 所有满足 $\operatorname{Tr} A = 0$ 的方阵的集合记为 W. 证明, W 是 $\mathbb{F}^{n\times n}$ 的子空间. 并求 $\dim W$.

2. 在数域 $\mathbb{F}$ 上的所有 n 阶方阵构成的线性空间 $\mathbb{F}^{n\times n}$ 中,所有对称方阵的集合记为 S,所有斜对称方阵的集合记为 K. 证明,S 和 K 都是 $\mathbb{F}^{n\times n}$ 的子空间; $S+K=\mathbb{F}^{n\times n}$; $S\cap K=0$; 并求 $\dim S$, $\dim K$.

3. 在 $\mathbb{F}^{2\times 2}$ 中,所有形如

$$\begin{pmatrix} x & -x \\ y & z \end{pmatrix}$$

的矩阵的集合记为 V_1,所有形如

$$\begin{pmatrix} a & b \\ -a & c \end{pmatrix}$$

的矩阵的集合记为 V_2.

证明,V_1 和 V_2 都是 $\mathbb{F}^{2\times 2}$ 的子空间; 并求 $\dim V_1$, $\dim V_2$, $\dim(V_1+V_2)$, $\dim V_1\cap V_2$.

4. 在数域 $\mathbb{F}$ 上的所有关于 x 的多项式构成的线性空间 $\mathbb{F}[x]$ 中,有满足 $f(-x)=f(x)$ 的多项式 $f(x)$ 的集合记为 W,所有满足 $f(-x)=-f(x)$ 的多项式 $f(x)$ 的集合记为 U. 证明. W 与 U 都是 $\mathbb{F}[x]$ 的子空间,并且 $W\cap U=0$, $W+U=\mathbb{F}[x]$.

5. $\mathbb{F}^n$ 中下列子集合是否是子空间? 如果是子空间,则确定它的维数,并给出一组基; 如果不是子空间,则写出它所生成的子空间,并给出一组基.

(1) $W=\left\{(a_1,a_2,\dots,a_n)\in\mathbb{F}^n \;\middle|\; a_1+a_2+\cdots+a_n=0\right\}$;

(2) $W=\left\{(a_1,a_2,\dots,a_n)\in\mathbb{F}^n \;\middle|\; a_1,a_2,\dots,a_n \text{ 不同时大于零,或不同时小于零}\right\}$;

(3) $W=\left\{(a_1,a_2,\dots,a_n)\in\mathbb{F}^n \;\middle|\; \text{有某个 } i(1\leqslant i\leqslant n), \text{使 } a_i>0\right\}$.

6. 设 U,V,W 是线性空间 L 的子空间. 证明,

(1) 等式 $U\cap(V+W)=(U\cap V)+(U\cap W)$ 不一定成立;

(2) 等式 $U\cap\big(V+(U\cap W)\big)=(U\cap V)+(U\cap W)$ 恒成立.

7. 设 U 和 W 是线性空间 V 的子空间. 证明,等式 $U\cup W=U+W$ 成立的充分必要条件是 $U\subseteq W$,或者 $W\subseteq U$.

8. 设 U,V 和 W 是线性空间 L 的子空间. 证明,

$$(U+V)\cap(U+W)=U+(U+V)\cap W.$$

9. 证明,数域 $\mathbb{F}$ 上的无限维线性空间 V 一定含有无限维真子空间.

10. 求下列向量 $\alpha_1,\alpha_2,\alpha_3$ 与向量 β_1,β_2,β_3 生成的子空间 W_1 与 W_2 的维数,并给出子空间 $W_1\cap W_2$ 与 W_1+W_2 的一组基.

(1) $\alpha_1=(1,2,1,-2)$, $\alpha_2=(2,3,1,0)$, $\alpha_3=(1,2,2,-3)$;
$\beta_1=(1,1,1,1)$, $\beta_2=(1,0,1,-1)$, $\beta_3=(1,3,0,-4)$;

(2) $\alpha_1=(1,1,0,0)$, $\alpha_2=(0,1,1,0)$, $\alpha_3=(0,0,1,1)$;
$\beta_1=(1,0,1,0)$, $\beta_2=(0,2,1,1)$, $\beta_3=(1,2,1,2)$.

11. 设 $V_0,V_1,\dots,V_k,\dots$ 是数域 $\mathbb{F}$ 上的有限维线性空间 V 的子空间. 如果 $V_{i+1}\subseteq V_i$, $i=0,1,\dots$,则子空间序列 $\left\{V_i \;\middle|\; i=0,1,\dots\right\}$ 称为**子空间降序列**. 证明,对任意子空间降序列 $\left\{V_i \;\middle|\; i=0,1,\dots\right\}$,恒存在正整数 k, $0\leqslant k\leqslant\dim V$,使得

$$V_k=V_{k+1}=V_{k+2}=\cdots. \qquad \text{(本题似有误)}$$

12. 设线性空间 V 中的向量 α,β 和 γ 满足 $\alpha+\beta+\gamma=0$. 证明,$V(\alpha,\beta)=V(\beta,\gamma)$.

13. 设 α,β 是线性空间 V 中的向量,W 是 V 的子空间. 向量 α 与子空间 W 生成的子空间记为 U,向量 β 与子空间 W 生成的子空间记为 K. 证明,如果 $\beta\in U$,但 $\beta\notin W$,则 $\alpha\in K$.

14. 设 A 和 B 分别是 $m\times n$ 和 $n\times p$ 矩阵. 证明,等式 $\operatorname{rank}B=\operatorname{rank}AB$ 的充分必要条件是,方程组 $ABx=0$ 的解一定是方程组 $Bx=0$ 的解.

15. 设 A,B 和 C 分别是 $m\times n, n\times p$ 和 $p\times q$ 矩阵. 证明,

$$\operatorname{rank} AB + \operatorname{rank} BC \leqslant \operatorname{rank} ABC + \operatorname{rank} B.$$

16. B 及 C 的意义同上题. 证明,如果 $\operatorname{rank} B = \operatorname{rank} AB$,则 $\operatorname{rank} BC = \operatorname{rank} ABC$.

17. 设 A 是 n 阶方阵,k 是正整数,并且 $\operatorname{rank} A^k = \operatorname{rank} A^{k+1}$. 证明,

$$\operatorname{rank} A^k = \operatorname{rank} A^{k+1} = \operatorname{rank} A^{k+2} = \cdots.$$

18. 设 $P_1,P_2,\ldots,P_k,Q_1,Q_2,\ldots,Q_k$ 都是 n 阶方阵,且 $P_iQ_j = Q_jP_i$,$\operatorname{rank} P_i = \operatorname{rank} P_iQ_i$,其中 $1\leqslant i,j\leqslant k$. 证明,

$$\operatorname{rank} P_1P_2\cdots P_k = \operatorname{rank} P_1P_2\cdots P_kQ_1Q_2\cdots Q_k.$$

19. 设 A 是 n 阶复方阵,方阵 $G = \overline{A}^{\mathrm{T}}A$ 称为方阵 A 的 Gram 方阵. 证明,$\operatorname{rank} G = \operatorname{rank} A$.

§4.7　直　和

在子空间 V_1 与 V_2 的和 V_1+V_2 中,子空间 V_1 与 V_2 的交为零子空间的情形特别重要.

定义 4.7.1　设 V_1 与 V_2 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的子空间. 如果交 $V_1\cap V_2 = 0$,则子空间 V_1 与 V_2 的和 V_1+V_2 称为 **V_1 与 V_2 的直和**,记为 $V_1\oplus V_2$.

关于子空间 V_1 与 V_2 的直和 $V_1\oplus V_2$,有

定理 4.7.1　下列命题等价.

(1) 和 V_1+V_2 是直和;

(2) 和 V_1+V_2 中每个向量 α 都可以唯一地表为 $\alpha=\alpha_1+\alpha_2$,其中 $\alpha_1\in V_1$,$\alpha_2\in V_2$;

(3) 和 V_1+V_2 中零向量可以唯一地表为 $0 = 0_1+0_2$,其中 0_1 和 0_2 分别是 V_1 和 V_2 的零向量(当然它们是线性空间 V 的零向量);

(4) $$\dim(V_1+V_2) = \dim V_1 + \dim V_2.$$

证明　命题 (1) 与 (4) 的等价性是维数定理的推论 4.6.1.

(1) $\Longrightarrow$ (2)　由子空间和的定义,和 V_1+V_2 中每个向量 α 都可以表为 $\alpha = \alpha_1+\alpha_2$,其中 $\alpha_1\in V_1$,$\alpha_2\in V_2$.

如果向量 α 还可以表为 $\alpha=\beta_1+\beta_2$,其中 $\beta_1\in V_1$,$\beta_2\in V_2$,则

$$\alpha_1-\beta_1 = \beta_2-\alpha_2 \in V_1\cap V_2.$$

由于和 V_1+V_2 是直和,所以 $V_1\cap V_2 = 0$,因此 $\alpha_1=\beta_1$,$\alpha_2=\beta_2$. 这就证明,向量 α 的分解式 $\alpha=\alpha_1+\alpha_2$ 是唯一的.

(2) $\Longrightarrow$ (3)　显然 (3) 是 (2) 的特殊情形.

(3) $\Longrightarrow$ (1)　设和 V_1+V_2 不是直和,则由定义,$V_1\cap V_2\neq 0$,于是存在非零向量 $\alpha\in V_1\cap V_2$. 因此

$$0 = 0_1+0_2 = \alpha+(-\alpha).$$

其中 $\alpha\in V_1$,$-\alpha\in V_2$. 即零向量 0 具有两个不同的分解式,矛盾. ■

定理 4.7.2 设 V_1 是线性空间 V 的子空间. 则存在子空间 $V_2 \subseteq V_1$, 使得 $V = V_1 \oplus V_2$.

证明 如果 V_1 是零子空间, 则取 $V_2 = V$. 显然, $V_1 \cap V_2 = 0$, 并且 $V = V_1 + V_2$, 因此 $V = V_1 \oplus V_2$.

如果 V_1 不是零子空间, 则设 $\dim V_1 = r, 1 \leqslant r \leqslant n$, 其中 $n = \dim V$. 于是, 子空间 V_1 具有基 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$. 把 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ 扩充成线性空间 V 的基

$$\{\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n\}.$$

V 中由向量 $\alpha_{r+1}, \ldots, \alpha_n$ 生成的子空间 $V(\alpha_{r+1}, \ldots, \alpha_n)$ 为

$$V(\alpha_{r+1}, \ldots, \alpha_n) = \{a_{r+1}\alpha_{r+1} + \cdots + a_n\alpha_n \mid a_i \in \mathbb{F}, r+1 \leqslant i \leqslant n\}.$$

取 $V_2 = V(\alpha_{r+1}, \ldots, \alpha_n)$. 显然, $V_1 \cap V_2 = 0$.

其次, 设 $a \in V$, 则因 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 是 V 的基, 故

$$\alpha = (a_1\alpha_1 + a_2\alpha_2 + \cdots + a_r\alpha_r) + (a_{r+1}\alpha_{r+1} + \cdots + a_n\alpha_n),$$

其中 $a_1\alpha_1 + a_2\alpha_2 + \cdots + a_r\alpha_r \in V_1, a_{r+1}\alpha_{r+1} + \cdots + a_n\alpha_n \in V_2$.

于是, $V = V_1 + V_2$. 又 $V_1 \cap V_2 = 0$, 所以 $V = V_1 \oplus V_2$. ■

如果 V_1 和 V_2 是线性空间 V 的子空间, 并且 $V = V_1 \oplus V_2$, 则子空间 V_2 称为**子空间 V_1 的补**. 当然, V_2 也是 V_1 的补. 定理 4.7.2 说明, 对线性空间 V 的每一个子空间 V_1, 它的补是存在的, 并且补的维数为 $\dim V - \dim V_1$.

应当指出, 子空间 V_1 的补并不一定唯一. 例如, 在二维实行向量空间 $\mathbb{R}^2$ 中, 取 $V_1 = \{a\varepsilon_1 \mid a \in \mathbb{R}\}, V_2 = \{a\varepsilon_2 \mid a \in \mathbb{R}\}$, 其中 $\varepsilon_1 = (1,0), \varepsilon_2 = (0,1)$. 容易验证, V_2 是子空间 V_1 的补. 取 $\widetilde{V}_2 = \{a\beta \mid a \in \mathbb{R}\}$, 其中 $\beta = \varepsilon_1 + \varepsilon_2$, 则对任意 $\alpha \in \mathbb{R}^2$, 均有

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 = (a_1 - a_2)\varepsilon_1 + a_2(\varepsilon_1 + \varepsilon_2) = (a_1 - a_2)\varepsilon_1 + a_2\beta,$$

因此, $\mathbb{R}^2 = V_1 + \widetilde{V}_2$. 又设 $\alpha \in V_1 \cap \widetilde{V}_2$, 则

$$\alpha = a_1\varepsilon_1 = a_2\beta = a_2(\varepsilon_1 + \varepsilon_2),$$

其中 $a_1, a_2 \in \mathbb{R}$. 因此, $(a_1 - a_2)\varepsilon_1 - a_2\varepsilon_2 = 0$, 所以 $a_1 = a_2 = 0$, 即 $\alpha = 0$. 从而 $V_1 \cap \widetilde{V}_2 = 0$. 因此, $V = V_1 \oplus \widetilde{V}_2$, 即 $\widetilde{V}_2$ 也是子空间 V_1 的补.

两个子空间的直和概念可以推广到有限个子空间的情形.

定义 4.7.2 设 $V_1, V_2, \ldots, V_k$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的子空间. 如果子空间 $V_1, V_2, \ldots, V_k$ 的和 $V_1 + V_2 + \cdots + V_k$ 中每个向量 α 都可以唯一地表为

$$\alpha = a_1 + a_2 + \cdots + a_k,$$

其中 $\alpha_i \in V_i, 1 \leqslant i \leqslant k$. 则和 $V_1 + V_2 + \cdots + V_k$ 称为**子空间 $V_1, V_2, \ldots, V_k$ 的直和**, 记为

$$V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

定理 4.7.3 子空间 $V_1, V_2, \ldots, V_k$ 的和 $V_1 + V_2 + \cdots + V_k$ 为直和的当且仅当

$$\dim(V_1 + V_2 + \cdots + V_k) = \dim V_1 + \dim V_2 + \cdots + \dim V_k.$$

证明 **必要性** 设 $\dim V_i = n_i$, 且 $\{\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in_i}\}$ 是 V_i 的基. 设

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} a_{ij}\alpha_{ij} = 0,$$

其中 $a_{ij} \in \mathbb{F}, 1 \leqslant j \leqslant n_i, 1 \leqslant i \leqslant k$, 则因和 $V_1 + V_2 + \cdots + V_k$ 是直和, 所以直和

$$V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

中零向量的分解式是唯一的, 因此,

$$\sum_{j=1}^{n_i} a_{ij}\alpha_{ij} = 0.$$

由于 $\{\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in_i}\}$ 是 V_i 的基, 所以, $a_{i1} = a_{i2} = \cdots = a_{in_i} = 0$. 这表明, 向量集合 $\{\alpha_{ij} \mid 1 \leqslant j \leqslant n_i, 1 \leqslant i \leqslant k\}$ 线性无关.

其次设 $\alpha \in V_1 + V_2 + \cdots + V_k$, 则存在 $\beta_i \in V_i, i = 1, 2, \ldots, k$, 使得

$$\alpha = \beta_1 + \beta_2 + \cdots + \beta_k.$$

由于 $\beta_i \in V_i$, $\{\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in_i}\}$ 是 V_i 的基, 所以

$$\beta_i = \sum_{j=1}^{n_i} a_{ij}\alpha_{ij} \implies \alpha = \sum_{i=1}^{k}\sum_{j=1}^{n_i} a_{ij}\alpha_{ij}.$$

即和 $V_1 + \cdots + V_k$ 中每个向量都可由向量集合 $\{\alpha_{ij} \mid 1 \leqslant j \leqslant n_i, 1 \leqslant i \leqslant k\}$ 线性表出.

这就证明, $\{\alpha_{ij} \mid 1 \leqslant j \leqslant n_i, 1 \leqslant i \leqslant k\}$ 是和 $V_1 + V_2 + \cdots + V_k$ 的基, 于是得到

$$\dim(V_1 + V_2 + \cdots + V_k) = \dim V_1 + \dim V_2 + \cdots + \dim V_k.$$

充分性 仍设 $\dim V_i = n_i$, $\{\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in_i}\}$ 是 V_i 的基, $i = 1, 2, \ldots, k$. 由于

$$\dim(V_1 + V_2 + \cdots + V_k) = \dim V_1 + \dim V_2 + \cdots + \dim V_k,$$

所以 $\{\alpha_{ij} \mid 1 \leqslant j \leqslant n_i, 1 \leqslant i \leqslant k\}$ 是和 $V_1 + V_2 + \cdots + V_k$ 的基.

设和 $V_1 + V_2 + \cdots + V_k$ 中向量 α 具有两个分解式

$$\alpha = \beta_1 + \beta_2 + \cdots + \beta_k = \gamma_1 + \gamma_2 + \cdots + \gamma_k,$$

其中 $\beta_i, \gamma_i \in V_i$, 则

$$(\beta_1 - \gamma_1) + (\beta_2 - \gamma_2) + \cdots + (\beta_k - \gamma_k) = 0.$$

因为 $\{\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in_i}\}$ 是 V_i 的基, 所以,

$$\alpha_i - \beta_i = \sum_{j=1}^{n_i} a_{ij}\alpha_{ij},$$

其中 $a_{ij} \in \mathbb{F}$. 因此,

$$\sum_{i=1}^{k}(\alpha_i - \beta_i) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} a_{ij}\alpha_{ij} = 0.$$

由于 $\{\alpha_{ij} \mid 1 \leqslant j \leqslant n_i, 1 \leqslant i \leqslant k\}$ 是和 $V_1 + V_2 + \cdots + V_k$ 的基, 所以 $a_{ij} = 0$. 因此, $\alpha_i - \beta_i = 0$, 即 $\alpha_i = \beta_i$.

这表明, 和 $V_1 + V_2 + \cdots + V_k$ 中每个向量表为子空间 $V_1, V_2, \ldots, V_k$ 中向量的和

的表法唯一. 因此,和 $V_1 + V_2 + \cdots + V_k$ 是直和. ■

从定理 4.7.3 的证明可以看出,关于子空间 $V_1 + V_2 + \cdots + V_k$ 的直和的定义条件可以减弱,即有

定理 4.7.4 子空间 $V_1, V_2, \ldots, V_k$ 的和为直和当且仅当和 $V_1 + V_2 + \cdots + V_k$ 中零向量可以唯一地表为子空间 $V_1, V_2, \ldots, V_k$ 中向量(当然是零向量)的和.

证明 **必要性** 由直和的定义,和 $V_1 + V_2 + \cdots + V_k$ 中任意向量可以唯一地表为子空间 $V_1, V_2, \ldots, V_k$ 中向量的和,因此,和 $V_1 + V_2 + \cdots + V_k$ 中零向量当然也可以唯一地表为子空间 $V_1, V_2, \ldots, V_k$ 中向量的和.

充分性 设 $\alpha \in V_1 + V_2 + \cdots + V_k$ 具有两种表示方式,即设

$$\alpha = \beta_1 + \beta_2 + \cdots + \beta_k = \gamma_1 + \gamma_2 + \cdots + \gamma_k,$$

则

$$(\beta_1 - \gamma_1) + (\beta_2 - \gamma_2) + \cdots + (\beta_k - \gamma_k) = 0.$$

因为和 $V_1 + V_2 + \cdots + V_k$ 中零向量表为子空间 $V_1, V_2, \ldots, V_k$ 中向量的和的表法唯一,而零向量显然具有如下的表示方式:

$$0 = 0_1 + 0_2 + \cdots + 0_k,$$

其中 0_i 是子空间 V_i 的零向量(也是 V 的零向量),所以 $\alpha_i - \beta_i = 0_i = 0$,即 $\alpha_i = \beta_i$.

这就证明,和 $V_1 + V_2 + \cdots + V_k$ 中每个向量表为子空间 $V_1, V_2, \ldots, V_k$ 中向量的和的表法唯一,即和 $V_1 + V_2 + \cdots + V_k$ 是直和. ■

上面介绍的是同一个线性空间中子空间的直和. 下面讨论数域 $\mathbb{F}$ 上的不同线性空间的直和.

设 U 和 W 是同一数域 $\mathbb{F}$ 上的两个线性空间. 取空间 U 的向量 α 作为第一个分量,空间 W 的向量 β 作为第二个分量,组成有序向量偶 (α, β). 所有这样的有序向量偶集合称为**线性空间 U 与 W 的 Descartes 积**,记为 $U \times W$,即

$$U \times W = \left\{(\alpha, \beta) \,\middle|\, \alpha \in U, \beta \in W\right\}.$$

在集合 $U \times W$ 中规定向量加法如下: 设 $(\alpha_1, \beta_2), (\alpha_2, \beta_2) \in U \times W$,定义

$$(\alpha_1, \beta_2) + (\alpha_2, \beta_2) = (\alpha_1 + \alpha_2, \beta_1 + \beta_2);$$

纯量与向量的乘法规定为: 设 $\lambda \in \mathbb{F}, (\alpha, \beta) \in U \times W$,则

$$\lambda(\alpha, \beta) = (\lambda\alpha, \lambda\beta).$$

容易验证,集合 $U \times W$ 在如此规定的向量加法,纯量与向量的乘法下成为数域 $\mathbb{F}$ 上的线性空间. 在线性空间 $U \times W$ 中,记

$$\widetilde{U} = \left\{(\alpha, \beta) \in U \times W \,\middle|\, \beta = 0\right\}, \qquad \widetilde{W} = \left\{(\alpha, \beta) \in U \times W \,\middle|\, \alpha = 0\right\}.$$

容易验证,$\widetilde{U}$ 和 $\widetilde{W}$ 都是 $U \times W$ 的子空间,并且 $\widetilde{U} \cap \widetilde{W} = 0$. 对于任意 $(\alpha, \beta) \in U \times W$,都有

$$(\alpha, \beta) = (\alpha, 0) + (0, \beta),$$

其中 $(\alpha,0)\in\widetilde{U},(0,\beta)\in\widetilde{W}$,因此 $U\times W=\widetilde{U}\oplus\widetilde{W}$.

定义线性空间 U 到 $\widetilde{U}$ 的映射为: 对任意 $\alpha\in U$,令

$$\eta(\alpha)=(\alpha,0)\in\widetilde{U}.$$

显然,η 是 U 到 $\widetilde{U}$ 上的双射. 设 $\alpha_1,\alpha_2\in U$,则

$$\eta(\alpha_1+\alpha_2)=(\alpha_1+\alpha_2,0)=(\alpha_1,0)+(\alpha_2,0)=\eta(\alpha_1)+\eta(\alpha_2),$$

因此,映射 η 保加法. 又设 $\lambda\in\mathbb{F},\alpha\in U$,则

$$\eta(\lambda\alpha)=(\lambda\alpha,0)=\lambda(\alpha,0)=\lambda\eta(\alpha),$$

所以映射 η 保乘法.

这就证明,映射 η 是线性空间 U 到 $\widetilde{U}$ 上的同构映射,即 U 和 $\widetilde{U}$ 同构.

同样可以定义线性空间 W 到 $\widetilde{W}$ 的映射 ζ,对任意 $\beta\in W$,令

$$\zeta(\beta)=(0,\beta)\in W.$$

映射 ζ 是 W 到 $\widetilde{W}$ 上的同构映射. 因此,W 和 $\widetilde{W}$ 同构.

在同构映射 η 和 ζ 下,可以把 U 中向量 α 和它在 $\widetilde{U}$ 中的象 $\zeta(\alpha)=(\alpha,0)$,以及 W 中向量 β 和它在 $\widetilde{W}$ 中的象 $\zeta(\beta)=(0,\beta)$ 分别等同起来. 也就是说,可以把 U 和 $\widetilde{U}$,W 和 $\widetilde{W}$ 分别视为同一个空间. 所以在同构意义下,线性空间 Descartes 积 $U\times W$ 是线性空间 U 和 W 的直和,并记为 $U\times W=U\oplus W$.

习　题　4.7

1. 在 n 维实向量空间 $\mathbb{R}^n$ 中,记

$$V=\left\{(a_1,a_2,\ldots,a_n)\in\mathbb{R}^n \,\middle|\, a_1+a_2+\cdots+a_n=0\right\},$$
$$W=\left\{(a_1,a_2,\ldots,a_n)\in\mathbb{R}^n \,\middle|\, a_1=a_2=\cdots=a_n\right\}.$$

V 和 W 显然是 $\mathbb{R}^n$ 的子空间. 证明,$\mathbb{R}^n=V\oplus W$.

2. 在数域 $\mathbb{F}$ 上的 $2n$ 维向量空间 $\mathbb{F}^{2n}$ 中,记

$$V=\left\{(a_1,a_2,\ldots,a_{2n})\in\mathbb{F}^{2n} \,\middle|\, a_1=a_2=\cdots=a_n\right\},$$
$$W=\left\{(a_1,a_2,\ldots,a_{2n})\in\mathbb{F}^{2n} \,\middle|\, a_i=-a_{n+i},\ \ 1\leqslant i\leqslant n\right\}.$$

V 和 W 显然是 $\mathbb{F}^{2n}$ 的子空间. 证明,$\mathbb{F}^{2n}=V\oplus W$.

3. 设 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 都是四维复向量空间 $\mathbb{C}^4$ 中的向量,$\mathbb{C}^4$ 中分别由向量集合 $\{\alpha_1,\alpha_2\}$ 和 $\{\alpha_3,\alpha_4\}$ 生成的子空间记为 V 和 W. 试判断 $\mathbb{C}^4=V\oplus W$ 是否成立?

(1) $\alpha_1=(0,1,0,1),\alpha_2=(0,0,1,0),\alpha_3=(1,0,1,0),\alpha_4=(1,1,0,0)$;

(2) $\alpha_1=(-1,1,1,0),\alpha_2=(0,1,-1,1),\alpha_3=(1,0,0,0),\alpha_4=(0,0,0,1)$;

(3) $\alpha_1=(1,0,0,1),\alpha_2=(0,1,1,0),\alpha_3=(1,0,1,0),\alpha_4=(0,1,0,1)$.

4. 设 U,V 和 W 是线性空间 L 的子空间. 如果其中每一个都与另外两个之和的交是零子空间,则这三个子空间称为**无关的**.

证明,$L=U\oplus(V\oplus W)$ 的当且仅当 U,V,W 是无关的,并且 $L=U+V+W$.

5. 举例说明,线性空间的子空间 U,V,W 两两之交为零子空间,但 U,V,W 并不一定无关.

6. 证明,三个子空间无关的充分必要条件是它们的和的维数等于它们的维数的和.

7. 设 $V_1, V_2, \ldots, V_k$ 是线性空间 V 的子空间. 证明下列命题等价.

(1) 和 $V_1 + V_2 + \cdots + V_k$ 是直和;

(2) $V_j \cap (V_1 + V_2 + \cdots + V_{j-1} + V_{j+1} + \cdots + V_k) = 0$, $j = 1, 2, \ldots, k$;

(3) $V_1 \cap V_2 = 0, (V_1 + V_2) \cap V_3 = 0, \ldots, (V_1 + \cdots + V_{k-1}) \cap V_k = 0$.

§4.8 商 空 间

设 V 是数域 $\mathbb{F}$ 上的有限维线性空间 V 的子空间, 向量 $\alpha, \beta \in V$. 如果向量 $\alpha - \beta \in W$, 则向量 α 与 β 称为**模 W 同余**, 记为 $\alpha \equiv \beta \pmod W$.

易知 V 中向量之间的模 W 同余关系具有以下性质.

(1) 自反性 对任意向量 $\alpha \in V$, $\alpha \equiv \alpha \pmod W$;

(2) 对称性 设向量 $\alpha, \beta \in V$, 且 $\alpha \equiv \beta \pmod W$, 则 $\beta \equiv \alpha \pmod W$;

(3) 传递性 设向量 $\alpha, \beta, \gamma \in V$, 并且 $\alpha \equiv \beta \pmod W$, $\beta \equiv \gamma \pmod W$, 则

$$\alpha \equiv \gamma \pmod W.$$

上述性质的证明留给读者作练习.

由于线性空间 V 中任意两个向量要么模 W 同余, 要么模 W 不同余, 这样便在线性空间 V 的向量之间引进了一种关系, 即同余关系. 同余关系满足自反性、对称性和传递性, 因此 V 中向量便按照同余关系划分为同余类: 即在同一个同余类的向量彼此模 W 同余, 而在不同的同余类中的向量一定模 W 不同余.

向量 α 所在的同余类记为 $\overline{\alpha}$. 同余类 $\overline{\alpha}$ 由哪些向量构成? 记

$$\alpha + W = \left\{\alpha + \gamma \,\middle|\, \gamma \in W\right\}.$$

设 $\beta \in \alpha + W$, 则 $\beta = \alpha + \gamma, \gamma \in W$. 因此 $\beta - \alpha = \gamma \in W$, 即 $\beta \equiv \alpha \pmod W$, 所以, $\beta \in \overline{\alpha}$. 其次设 $\beta \in \overline{\alpha}$, 则 $\beta \equiv \alpha \pmod W$, 因此存在 $\gamma \in W$, 使得 $\beta - \alpha = \gamma$, 即 $\beta = \alpha + \gamma$, 于是 $\beta \in \alpha + W$. 所以, $\overline{\alpha} \subseteq a + W$, 从而 $\overline{\alpha} = \alpha + W$.

所有模 W 的同余类集合记为 V/W. 在集合 V/W 中规定同余类的加法如下: 对任意 $\overline{\alpha}, \overline{\beta} \in V/W$, 令

$$\overline{\alpha} + \overline{\beta} = \overline{\alpha + \beta}.$$

纯量与同余类的乘法规定为: 对任意纯量 $\lambda \in \mathbb{F}, \overline{\alpha} \in V/W$, 令

$$\lambda\overline{\alpha} = \overline{\lambda\alpha}.$$

应当指出, 上面规定的同余类加法是在同余类 $\overline{\alpha}$ 与 $\overline{\beta}$ 中各取出一个向量 α 与 β (它们分别称为同余类 $\overline{\alpha}$ 与 $\overline{\beta}$ 的**代表元**), 然后用和 $\alpha + \beta$ 所在的同余类 $\overline{\alpha + \beta}$ 作为同余类 $\overline{\alpha}$ 与 $\overline{\beta}$ 的和 $\overline{\alpha} + \overline{\beta}$. 自然产生的问题是, 如果在同余类 $\overline{\alpha}$ 与 $\overline{\beta}$ 中各另取一个向量 α' 与 β', 和 $a' + b'$ 所在的同余类 $\overline{\alpha' + \beta'}$ 是否和同余类 $\overline{\alpha + \beta}$ 相同?

如果同余类 $\overline{\alpha' + \beta'}$ 和同余类 $\overline{\alpha + \beta}$ 不相同, 那么这样规定的同余类的和是不确

定的. 如果不论同余类 $\overline{a}$ 和 $\overline{\beta}$ 的代表元 α' 与 β' 如何选取,同余类 $\overline{\alpha'+\beta'}$ 都和 $\overline{\alpha+\beta}$ 相同,那么就说这样规定的同余类的和与同余类代表元的选取无关,于是同余类的加法便有确切意义.

现在证明,上面规定的同余类加法与同余类代表元的选取无关. 事实上,设 $\alpha'\in\overline{\alpha}$,$\beta'\in\overline{\beta}$. $\alpha'\equiv\alpha \pmod W$,$\beta'\equiv\beta \pmod W$,即 $\alpha'-\alpha$,$\beta'-\beta\in W$,因此,

$$(\alpha'+\beta')-(\alpha+\beta)=(\alpha'-\alpha)+(\beta'-\beta)\in W.$$

所以 $\alpha'+\beta'\equiv\alpha+\beta \pmod W$. 于是 $\overline{\alpha'+\beta'}=\overline{\alpha+\beta}$.

同样可以证明,纯量与同余类的乘法与同余类的代表元选取无关. 因此在集合 V/W 中同余类的加法、纯量与同余类的乘法都有确切意义.

容易验证,同余类集合 V/W 在上述规定的加法与乘法下是数域 $\mathbb{F}$ 上的线性空间,其中 W 是线性空间 V/W 的零向量.

数域 $\mathbb{F}$ 上的线性空间 V/W 称为线性空间 V 关于子空间 W 的**商空间**.

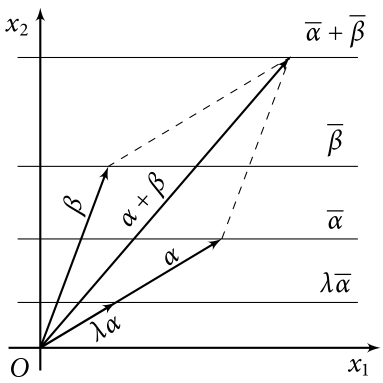

图 4.8.1

例如,在二维实向量空间 $\mathbb{R}^2$ 中,记

$$W=\{(x_1,0)\mid x_1\in\mathbb{R}\}.$$

W 显然是 $\mathbb{R}^2$ 的子空间. 从几何上看,$\mathbb{R}^2$ 就是通常的 Euclid 平面,W 是平面上的坐标轴: x_1 轴. 设向量 $\alpha=(x_1,x_2)\in\mathbb{R}^2$. 已经知道,向量 α 所在的模 W 同余类 $\overline{\alpha}$ 为 $\overline{\alpha}=\alpha+W=\{\alpha+\beta\mid\beta\in W\}$,所以,

$$\overline{\alpha}=\{(y_1,x_2)\mid y_1\in\mathbb{R}\}.$$

这表明,向量 $\alpha=(x_1,x_2)$ 所在的模 W 同余类是 Euclid 平面上过点 (x_1,x_2) 并平行于 x_1 轴的直线. 因此在此时,商空间 $\mathbb{R}^2/W$ 即是 Euclid 平面上所有平行于 x_1 轴的直线的集合(见图 4.8.1).

在商空间 $\mathbb{R}^2/W$ 中,同余类 $\overline{\alpha}$ 与 $\overline{\beta}$ 的和为 $\overline{\alpha+\beta}$,即直线 $\overline{\alpha}$ 与 $\overline{\beta}$ 之和是过向量 $\alpha+\beta$ 的终点并平行于 x_1 轴的直线. 纯量 λ 与同余类 $\overline{\alpha}$ 的乘积 $\lambda\overline{\alpha}$ 为 $\overline{\lambda\alpha}$,即纯量 λ 与直线 $\overline{\alpha}$ 的乘积 $\lambda\overline{\alpha}$ 是过向量 $\lambda\alpha$ 的终点并平行于 x_1 轴的直线.

设 W 是线性空间 V 的子空间,W' 为其补,即 $V=W\oplus W'$,V/W 是 V 模 W 的商空间. 在 W' 与 V/W 之间可以自然地规定映射 η 如下: 对任意向量 $\alpha\in W'$,令

$$\eta(\alpha)=\overline{\alpha}.$$

映射 η 称为**自然映射**.

设 $\overline{\alpha}\in V/W$,则 $\alpha\in V=W\oplus W'$,因此 $\alpha=\beta+\gamma$,其中 $\beta\in W$,$\gamma\in W'$. 由于 $\alpha-\gamma=\beta\in W$,所以 $\alpha\equiv\gamma \pmod W$. 因此 $\overline{\gamma}=\overline{\alpha}$. 即对任意 $\overline{\alpha}\in V/W$,总存在向量 $\gamma\in W'$,使得 $\overline{\gamma}=\overline{\alpha}$. 由于 $\eta(\gamma)=\overline{\gamma}=\overline{\alpha}$,所以,映射 η 是满射.

其次，设 $\gamma_1,\gamma_2\in W'$，并且 $\eta(\gamma_1)=\eta(\gamma_2)$，则 $\overline{\gamma}_1=\overline{\gamma}_2$. 因此，$\gamma_1\equiv\gamma_2 \pmod W$. 即 $\gamma_1-\gamma_2\in W$. 又 $\gamma_1,\gamma_2\in W'$，故 $\gamma_1-\gamma_2\in W'$. 因此，$\gamma_1-\gamma_2\in W\cap W'$. 由于 W' 是 W 的补. 所以 $W\cap W'=0$，由此 $\gamma_1-\gamma_2=0$，即 $\gamma_1=\gamma_2$. 这表明，映射 η 是单射.

最后，由于对任意 $\gamma_1,\gamma_2\in W'$，

$$\eta(\gamma_1+\gamma_2)=\overline{\gamma_1+\gamma_2}=\overline{\gamma}_1+\overline{\gamma}_2=\eta(\gamma_1)+\eta(\gamma_2),$$

所以，映射 η 是保加法的.

由于对任意 $\lambda\in\mathbb{F},\gamma\in W'$，

$$\eta(\lambda\gamma)=\overline{\lambda\gamma}=\lambda\overline{\gamma}=\lambda\eta(\gamma),$$

所以，映射 η 是保乘法的.

因此，W' 到 V/W 上的映射 η 是同构映射，从而 W' 和 V/W 同构.

总结上面的讨论，我们得到，

定理 4.8.1 设 W 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的子空间，W' 是 W 的补. 则子空间 W' 和 V 关于 W 的商空间 V/W 同构，并且 W' 到 V/W 上的自然映射 η 是同构映射.

由定理 4.8.1 立即得到，

推论 4.8.1 设 W 是线性空间 V 的子空间，则 V 关于 W 的商空间 V/W 的维数为

$$\dim(V/W)=\dim V-\dim W.$$

习 题 4.8

1. 在数域 $\mathbb{F}$ 上的所有关于未定元 x 的多项式构成的线性空间 $\mathbb{F}[x]$ 中，记

$$\mathbb{F}_n[x]=\{f(x)\in\mathbb{F}[x]\mid \deg f(x)<n\},\qquad W=\{f(x)\in\mathbb{F}[x]\mid f(-x)=f(x)\}.$$

$\mathbb{F}_n[x]$ 和 W 都是 $\mathbb{F}[x]$ 的子空间. 商空间 $\mathbb{F}[x]/\mathbb{F}_n[x]$ 和 $\mathbb{F}[x]/W$ 是否是有限维的?

2. 设 A 是数域 $\mathbb{F}$ 上的 $m\times n$ 矩阵，β 是数域 $\mathbb{F}$ 上的 n 维列向量空间 $\mathbb{F}^n$ 中的向量，并且

$$\operatorname{rank}(A,\ \beta)=\operatorname{rank}A.$$

记方程组 $Ax=0$ 的解空间为 V_A. 设向量 α 是方程组 $Ax=\beta$ 的一个特解. 证明，方程组 $Ax=\beta$ 的所有解构成 $\mathbb{F}^n$ 中向量 α 所在的模 V_A 同余类.

3. 设 W 是数域 $\mathbb{F}$ 上 n 维列向量空间 $\mathbb{F}^n$ 的子空间. 证明，存在数域 $\mathbb{F}$ 上的 $m\times n$ 矩阵 A，使得齐次方程组 $Ax=0$ 的解空间 V_A 为 W.

4. 设 W 是数域 $\mathbb{F}$ 上的 n 维列向量空间 $\mathbb{F}^n$ 中的子空间. 证明，对于 $\mathbb{F}^n$ 中每个向量 α，总存在数域 $\mathbb{F}$ 上的 $m\times n$ 矩阵 A 和向量 $\beta\in\mathbb{F}^n$，使得线性方程组 $Ax=\beta$ 的所有解的集合就是向量 α 所在的模 W 同余类.

第五章

线性变换

- 变换历来是数学中最主要的基本概念之一. 线性代数只讨论线性空间之间的变换.
- §5.1—§5.4 研究两个线性空间之间的线性变换——这里称为线性映射,而将变换一词留给线性空间到它自身上的线性映射.
- 在取定的基下,每个线性变换可以用一个矩阵表示,而同一个线性映射在不同的基下的矩阵表示恰好是彼此相抵的矩阵. 于是,寻求线性映射的最简单的矩阵表示,就归结为矩阵相抵下的分类问题. 在研究矩阵在相抵下的分类时遇到的不变量——矩阵的秩,从几何的观点看,就是线性映射的象空间的维数.
- 于是,这几节的内容以明确的方式展示了线性代数中几何理论与矩阵理论之间的联系. 真正弄懂这种联系,将会极大地有助于读者理解全部线性代数的内容.
- §5.5 开始讨论线性变换,这部分内容是线性代数中的重点,也是线性代数中的难点.
- 本书在讨论线性变换时是从几何入手的,也就是从线性空间关于线性变换的分解开始的. 为此需要一些准备知识.
- §5.6 引入了线性变换的不变子空间的概念.
- 在涉及线性变换的一维不变子空间时,自然而然地产生了线性变换的特征值、特征向量、特征多项式以及特征子空间等一系列与线性变换紧密联系的重要概念. 这些是 §5.7 与 §5.8 的内容.
- 由于线性变换的特征值在各种问题上的重要性,§5.9 对特征值的界作了一些估计.

§5.1 映 射

为本章需要,先复习一下映射概念,并介绍一些有关映射的术语.

给定集合 S 和 T. 所谓**集合 S 到 T 的映射 η** 是指集合 S 到集合 T 的一个对应规律,使得对任意给定的元素 $\alpha \in S$,按照对应规律 η,可以确定集合 T 的唯一一个元素 β 与 α 相对应. 集合 S 到 T 的映射 η 记为

$$\eta: S \to T.$$

元素 β 称为**元素 α 在映射 η 下的象**,元素 α 称为**元素 β 的原象**. 当 α 遍历集合 S 的元素时,所有 α 的象的集合称为**集合 S 在映射 η 下的象**,记为 $\operatorname{Im}\eta$,或者 $\eta(S)$. 即

$$\operatorname{Im}\eta = \{\eta(\alpha) \mid \alpha \in S\}.$$

元素 β 的所有原象的集合记为 $\eta^{-1}(\beta)$,即

$$\eta^{-1}(\beta) = \{\alpha \in S \mid \eta(\alpha) = \beta\}.$$

给定映射 $\eta: S \to T$. 如果对任意 $\alpha_1, \alpha_2 \in S$, $\alpha_1 \neq \alpha_2$, 均有 $\eta(\alpha_1) \neq \eta(\alpha_2)$, 则 η 称为**单射**. 可以看出,当且仅当由 $\eta(\alpha_1) = \eta(\alpha_2)$ 可以推出 $\alpha_1 = \alpha_2$ 时, η 是单射.

给定映射 $\eta: S \to T$. 如果对任意 $\beta \in T$, 存在 $\alpha \in S$, 使得 $\eta(\alpha) = \beta$, 则 η 称为**满射**. 容易看出,当且仅当 $\eta(S) = T$ 时,映射 η 是满射. 如果 η 是满射,则说映射 η 是 S 到 T 上的; 如果 η 不是满射,则说映射 η 是 S 到 T 内的.

如果映射 $\eta: S \to T$ 既是单射,又是满射,则 η 称为**双射**,或者 S 到 T 上的一一对应.

给定集合 U, V 和 W, 以及映射 $\eta: U \to V$, $\xi: V \to W$, 定义 U 到 W 的合成映射 $\xi\eta$ 如下: 设 $u \in U$, 则令

$$(\xi\eta)(u) = \xi(\eta(u)).$$

映射 $\xi\eta: U \to W$ 称为映射 η 和 ξ 的**乘积**.

例如映射 $f: \mathbb{R} \to \mathbb{R}$ 和 $g: \mathbb{R} \to \mathbb{R}$ 是函数,则映射 f 与 g 的乘积 gf 就是复合函数.

应当指出,映射的乘积一般是不可交换的,即等式 $\xi\eta = \eta\xi$ 一般并不成立. 其原因是,一般地说,尽管映射 η 与 ξ 的乘积 $\xi\eta$ 有意义,但乘积 $\eta\xi$ 不一定有意义. 即使乘积 $\xi\eta$ 与 $\eta\xi$ 都有意义,也不能保证 $\xi\eta = \eta\xi$. 这只要注意复合函数一般与函数复合的次序有关就够了.

对映射的乘积,有

定理 5.1.1 设 S, T, U 和 V 是集合, 且 $\xi: S \to T$, $\eta: T \to U$ 与 $\zeta: U \to V$ 是映射, 则 $\zeta(\eta\xi) = (\zeta\eta)\xi$. 换句话说,映射的乘法满足结合律.

证明 由映射乘积的定义,对任意 $s \in S$,

$$\big(\zeta(\eta\xi)\big)(s) = \zeta\big(\eta\xi(s)\big) = \zeta\Big(\eta\big(\xi(s)\big)\Big),$$

并且

$$\big((\zeta\eta)\xi\big)(s) = (\zeta\eta)\big(\xi(x)\big) = \zeta\big(\eta(\xi(s))\big).$$

因此 $\big(\zeta(\eta\xi)\big)(s) = \big((\zeta\eta)\xi\big)(s)$, 即 $\zeta(\eta\xi) = (\zeta\eta)\xi$. ■

集合 S 到自身的双射称为集合 S 的**变换**.

定义集合 S 到自身的映射 ε_S 如下: 设 $s \in S$, 则令

$$\varepsilon_S(s) = s.$$

显然, ε_S 是集合 S 的变换,它称为集合 S 的**恒等变换**,或者**单位变换**.

关于单位变换 ε_S, 有

定理 5.1.2 对任意映射 $\eta: S \to T$, 均有 $\eta\varepsilon_S = \varepsilon_T\eta = \eta$.

证明 显然,对任意 $s \in S$,

$$(\eta\varepsilon_S)(s) = \eta\big(\varepsilon_S(s)\big) = \eta(s).$$

同样,对任意 $s \in S$,

$$(\varepsilon_T\eta)(s) = \varepsilon_T\big(\eta(s)\big) = \eta(s).$$

因此 $\eta\varepsilon_S = \eta = \varepsilon_T\eta$. ■

最后讨论逆映射. 设 $\eta: S \to T$ 是映射,如果存在 $\xi: T \to S$,使得 $\xi\eta = \varepsilon_S$, $\eta\xi = \varepsilon_T$,则映射 $\eta: S \to T$ 称为**可逆映射**,映射 $\xi: T \to S$ 称为 η 的逆映射,并记为 η^{-1}.

关于可逆映射,有

定理 5.1.3　映射 $\eta: S \to T$ 可逆的充分必要条件是映射 η 为双射.

证明　设映射 $\eta: S \to T$ 可逆,则存在 η 的逆映射 $\xi: T \to S$,使得 $\xi\eta = \varepsilon_S$, $\eta\xi = \varepsilon_T$. 设 $s_1, s_2 \in S$,如果 $\eta(s_1) = \eta(s_2)$,则

$$\begin{aligned} s_1 &= \varepsilon_S(s_1) = (\xi\eta)(s_1) = \xi\big(\eta(s_1)\big) = \xi\big(\eta(s_2)\big) \\ &= (\xi\eta)(s_2) = \varepsilon_S(s_2) = s_2. \end{aligned}$$

这表明 η 是单射. 其次设 $t \in T$,则

$$t = \varepsilon_T(t) = (\eta\xi)(t) = \eta\big(\xi(t)\big).$$

记 $\xi(t) = s$,则 $s \in S$,且 $\eta(s) = t$. 因此映射 η 是满射. 所以 η 是双射.

反之,设 $\eta: S \to T$ 是双射. 由于 η 是满射,因此对任意 $t \in T$,存在 $s \in S$,使得 $\eta(s) = t$. 由于 η 是单射,因此 s 是唯一的. 定义映射 $\xi: T \to S$ 如下: 设 $t \in T$,如果 $s \in S$,且 $\eta(s) = t$,则令 $\xi(t) = s$. 由映射 ξ 的定义,对任意 $s \in S$,

$$s = \xi\big(\eta(s)\big) = (\xi\eta)(s),$$

因此 $\xi\eta = \varepsilon_S$; 对任意 $t \in T$,存在 $s \in S$,使得 $\eta(s) = t$. 因此 $\xi(t) = s$. 所以

$$t = \eta(s) = \eta\big(\xi(t)\big) = (\eta\xi)(t),$$

即 $\eta\xi = \varepsilon_T$. 所以 ξ 是 η 的逆映射,η 是可逆的. ■

定理 5.1.4　设映射 $\eta: S \to T$ 可逆,则 η 的逆映射是唯一的.

证明　设 $\xi_1: T \to S$ 和 $\xi_2: T \to S$ 是 η 的逆映射,则 $\xi_1\eta = \varepsilon_S$, $\eta\xi_1 = \varepsilon_T$,且 $\xi_2\eta = \varepsilon_S$, $\eta\xi_2 = \varepsilon_T$. 因此

$$\xi_1 = \xi_1\varepsilon_T = \xi_1(\eta\xi_2) = (\xi_1\eta)\xi_2 = \varepsilon_S\xi_2 = \xi_2.$$ ■

定理 5.1.5　设映射 $\eta: S \to T$ 和 $\xi: T \to U$ 可逆,则它们的乘积 $\xi\eta: S \to U$ 也可逆.

证明　设 η 和 ξ 的逆映射分别是 η^{-1} 和 ξ^{-1}. 则 $\eta\eta^{-1} = \varepsilon_T$, $\eta^{-1}\eta = \varepsilon_S$, $\xi\xi^{-1} = \varepsilon_U$, $\xi^{-1}\xi = \varepsilon_T$. 因此

$$(\xi\eta)(\eta^{-1}\xi^{-1}) = \xi\big((\eta\eta^{-1})\xi^{-1}\big) = \xi(\varepsilon_T\xi^{-1}) = \xi\xi^{-1} = \varepsilon_U,$$

而

$$(\eta^{-1}\xi^{-1})(\xi\eta) = \eta^{-1}\big((\xi^{-1}\xi)\eta\big) = \eta^{-1}(\varepsilon_T\eta) = \eta^{-1}\eta = \varepsilon_S.$$

所以映射 $\xi\eta$ 可逆,并且 $\xi\eta$ 的逆映射为 $\eta^{-1}\xi^{-1}$,即 $(\xi\eta)^{-1} = \eta^{-1}\xi^{-1}$. ■

§5.2 线性映射

设 U 和 V 是数域 $\mathbb{F}$ 上的线性空间,且 $\mathscr{A}: U \to V$ 是映射.

定义 5.2.1 如果映射 $\mathscr{A}: U \to V$ 满足:

LM (1) 对任意 $\alpha_1, \alpha_2 \in U$,

$$\mathscr{A}(\alpha_1 + \alpha_2) = \mathscr{A}(\alpha_1) + \mathscr{A}(\alpha_2);$$

LM (2) 对任意 $\lambda \in \mathbb{F}, \alpha \in U$,

$$\mathscr{A}(\lambda\alpha) = \lambda\mathscr{A}(\alpha),$$

则映射 $\mathscr{A}$ 称为线性空间 U 到 V 的**线性映射**.

例 5.2.1 设 V 是属于 $\mathbb{F}$ 上的 n 维线性空间,$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 U 的基,$\alpha \in U$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 $(a_1, a_2, \ldots, a_n)^{\mathrm{T}}$. 定义 U 到数域 $\mathbb{F}$ 上 n 维列向量空间 $\mathbb{F}^n$ 的映射 $\mathscr{A}$ 如下:

$$\mathscr{A}(\alpha) = (a_1, a_2, \ldots, a_n)^{\mathrm{T}}.$$

则映射 $\mathscr{A}: U \to \mathbb{F}^n$ 是线性映射.

证明 设 $\alpha, \beta \in U$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标分别为 $(a_1, a_2, \ldots, a_n)^{\mathrm{T}}, (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$,则

$$\mathscr{A}(\alpha) = (a_1, a_2, \ldots, a_n)^{\mathrm{T}}, \qquad \mathscr{A}(\beta) = (b_1, b_2, \ldots, b_n)^{\mathrm{T}}.$$

显然,$\alpha + \beta$ 与 $\lambda\alpha$ 的坐标分别为 $(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)^{\mathrm{T}}$ 与 $(\lambda a_1, \lambda a_2, \ldots, \lambda a_n)^{\mathrm{T}}$,其中 $\lambda \in \mathbb{F}$. 因此

$$\begin{aligned}
\mathscr{A}(\alpha + \beta) &= (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)^{\mathrm{T}} \\
&= (a_1, a_2, \ldots, a_n)^{\mathrm{T}} + (b_1, b_2, \ldots, b_n)^{\mathrm{T}} \\
&= \mathscr{A}(\alpha) + \mathscr{A}(\beta), \\
\mathscr{A}(\lambda\alpha) &= (\lambda a_1, \lambda a_2, \ldots, \lambda a_n)^{\mathrm{T}} = \lambda(a_1, a_2, \ldots, a_n)^{\mathrm{T}} \\
&= \lambda\mathscr{A}(\alpha).
\end{aligned}$$

因此映射 $\mathscr{A}: U \to \mathbb{F}^n$ 满足 LM (1) 与 LM (2). 所以映射 $\mathscr{A}: U \to \mathbb{F}^n$ 是线性的. ■

例 5.2.2 设 $\mathbb{F}$ 是数域 $\mathbb{F}^n$ 上 n 维行向量空间,且正整数 $m < n$. 定义映射 $\mathscr{A}: \mathbb{F}^n \to \mathbb{F}^m$ 如下:设 $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$,则令

$$\mathscr{A}(x) = (x_1, x_2, \ldots, x_m).$$

容易验证,映射 $\mathscr{A}: \mathbb{F}^m \to \mathbb{F}^m$ 是线性的,它称为 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 上的**投影变换**.

例 5.2.3 设 V 是数域 $\mathbb{F}$ 上线性空间,U 与 W 是 V 的子空间,且 $V = U \oplus W$. 定义映射 $\mathscr{A}: V \to U$ 如下:设 $\alpha \in V$,则存在唯一一对向量 β 与 γ,$\beta \in U$,$\gamma \in W$,使得 $\alpha = \beta + \gamma$,令 $\mathscr{A}(\alpha) = \beta$. 容易验证,映射 $\mathscr{A}: V \to U$ 是线性的,它称为线性空间 V 到

子空间 U 上的**投影变换**. 注意投影变换 $\mathscr{A}: V \to U$ 是满射.

例 5.2.4　设 V 和 U 是数域 $\mathbb{F}$ 上线性空间. 定义映射 $\mathscr{A}: V \to U$ 如下,设 $\alpha \in V$,则令 $\mathscr{A}(\alpha) = 0$. 显然映射 $\mathscr{A}$ 是线性的. 它称为**零映射**.

例 5.2.5　设 $\mathbb{F}^{m\times n}$ 是数域 $\mathbb{F}$ 上所有 $m \times n$ 阶矩阵构成的线性空间,P 和 Q 分别是给定的 m 阶和 n 阶方阵. 定义映射 $\mathscr{A}: \mathbb{F}^{m\times n} \to \mathbb{F}^{m\times n}$ 如下: 对任意 $X \in \mathbb{F}^{m\times n}$,令 $\mathscr{A}(X) = PXQ$. 容易验证,映射 $\mathscr{A}$ 是线性的.

例 5.2.6　设 $\mathbb{F}[x]$ 是数域 $\mathbb{F}$ 上所有关于 x 的多项式构成的线性空间. 定义映射 $\mathscr{D}: \mathbb{F}[x] \to \mathbb{F}[x]$ 如下: 设 $f(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{F}[x]$,则令

$$\mathscr{D}\big(f(x)\big) = a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

映射 $\mathscr{D}$ 称为**微商变换**. 容易验证,映射 $\mathscr{D}$ 是线性的.

性质 5.2.1　设 0 是线性空间 U 的零向量,则 $\mathscr{A}(0) = 0$.

证明　由于 $0 = 0 + 0$,故由 LM (1),$\mathscr{A}(0) = \mathscr{A}(0) + \mathscr{A}(0)$. 因此 $\mathscr{A}(0) = 0$. ■

性质 5.2.2　设 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$, $\alpha_1, \alpha_2, \ldots, \alpha_k \in U$,则

$$\mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k) = \lambda_1\mathscr{A}(\alpha_1) + \lambda_2\mathscr{A}(\alpha_2) + \cdots + \lambda_k\mathscr{A}(\alpha_k).$$

证明　对 k 用归纳法. 当 $k = 1$ 时上式即为 LM (2),故结论对 $k = 1$ 成立.

设结论对 $k - 1$ 成立. 现在证明结论对 k 成立. 由于

$$\mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k) = \mathscr{A}\big((\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_{k-1}\alpha_{k-1}) + \lambda_k\alpha_k\big),$$

故由 LM (1),

$$\mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k) = \mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_{k-1}\alpha_{k-1}) + \mathscr{A}(\lambda_k\alpha_k).$$

由归纳假设和 LM (2),

$$\mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k) = \lambda_1\mathscr{A}(\alpha_1) + \lambda_2\mathscr{A}(\alpha_2) + \cdots + \lambda_{k-1}\mathscr{A}(\alpha_{k-1}) + \lambda_k\mathscr{A}(\alpha_k).$$

即结论对 k 也成立. ■

性质 5.2.3　设向量组 $\alpha_1, \alpha_2, \ldots, \alpha_k \in U$ 线性相关,则 $\mathscr{A}(\alpha_1), \ldots, \mathscr{A}(\alpha_k) \in V$ 也线性相关.

证明　因为 $\alpha_1, \alpha_2, \ldots, \alpha_k \in U$ 线性相关,故存在不全为零的 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$,使得 $\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k = 0$. 由性质 5.2.1 和性质 5.2.2,

$$\begin{aligned}\mathscr{A}(0) &= \mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k)\\ &= \lambda_1\mathscr{A}(\alpha_1) + \lambda_2\mathscr{A}(\alpha_2) + \cdots + \lambda_k\mathscr{A}(\alpha_k) = 0.\end{aligned}$$

因此 $\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \ldots, \mathscr{A}(\alpha_k)$ 线性相关. ■

定理 5.2.1　设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维和 m 维线性空间,$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 U 的基. 设 $\beta_1, \beta_2, \ldots, \beta_n$ 是 V 中任意给定的 n 个向量. 则存在唯一线性映射 $\mathscr{A}: U \to V$,使得 $\mathscr{A}(\alpha_j) = \beta_j$, $j = 1, 2, \ldots, n$.

证明 设 $\alpha \in U$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 $(a_1, a_2, \ldots, a_n)^{\mathrm{T}}$. 定义映射 $\mathscr{A}: U \to V$ 如下：设 $\alpha \in U$，则令

$$\mathscr{A}(\alpha) = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n.$$

设 $\alpha, \beta \in U$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标分别为 $(a_1, a_2, \ldots, a_n)^{\mathrm{T}}$ 和 $(b_1, b_2, \ldots, b_n)^{\mathrm{T}}$，则 $\alpha + \beta$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 $(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)^{\mathrm{T}}$. 因此

$$\mathscr{A}(\alpha) = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n,$$
$$\mathscr{A}(\beta) = b_1\beta_1 + b_2\beta_2 + \cdots + b_n\beta_n.$$

且
$$\begin{aligned}\mathscr{A}(\alpha + \beta) &= (a_1 + b_1)\beta_1 + (a_2 + b_2)\beta_2 + \cdots + (a_n + b_n)\beta_n \\ &= (a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n) + (b_1\beta_1 + b_2\beta_2 + \cdots + b_n\beta_n) \\ &= \mathscr{A}(\alpha) + \mathscr{A}(\beta).\end{aligned}$$

因此映射 $\mathscr{A}: U \to V$ 满足 LM (1).

设 $\lambda \in \mathbb{F}$，且 $\alpha \in U$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 $(a_1, a_2, \ldots, a_n)^{\mathrm{T}}$. 则 $\lambda\alpha$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 $(\lambda a_1, \lambda a_2, \ldots, \lambda a_n)^{\mathrm{T}}$. 因此

$$\begin{aligned}\mathscr{A}(\lambda\alpha) &= \lambda a_1\beta_1 + \lambda a_2\beta_2 + \cdots + \lambda a_n\beta_n = \lambda(a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n) \\ &= \lambda\mathscr{A}(\alpha).\end{aligned}$$

因此映射 $\mathscr{A}: U \to V$ 满足 LM (2). 这就证明，映射 $\mathscr{A}: U \to V$ 是线性的.

显然基向量 α_j 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为

$$(0, \ldots, 0, \underset{j}{1}, 0, \ldots, 0)^{\mathrm{T}}.$$

因此 $\mathscr{A}(\alpha_j) = \beta_j$. 所以映射 $\mathscr{A}$ 是满足 $\mathscr{A}(\alpha_j) = \beta_j$，$j = 1, 2, \ldots, n$ 的线性映射.

现在设线性映射 $\mathscr{B}: U \to V$ 也满足 $\mathscr{B}(\alpha_j) = \beta_j$，$j = 1, 2, \ldots, n$，且 $\alpha \in U$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 $(a_1, a_2, \ldots, a_n)^{\mathrm{T}}$. 因此

$$\begin{aligned}\mathscr{B}(\alpha) &= \mathscr{B}(a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n) \\ &= a_1\mathscr{B}(\alpha_1) + a_2\mathscr{B}(\alpha_2) + \cdots + a_n\mathscr{B}(\alpha_n) \\ &= a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n = \mathscr{A}(\alpha).\end{aligned}$$

由 α 的任意性，$\mathscr{B} = \mathscr{A}$. ■

定理 5.2.1 表明，如果在线性空间 U 中取定一组基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$，则线性映射 $\mathscr{A}: U \to V$ 由基向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 在 $\mathscr{A}$ 下的象所唯一决定.

现在给出线性映射 $\mathscr{A}: U \to V$ 的矩阵表示. 设 U 和 V 分别是数域 $\mathbb{F}$ 上 n 维和 m 维线性空间，$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 和 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 分别是 U 和 V 的基. 显然，$\mathscr{A}(\alpha_j) \in V$. 由于 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 是 V 的基，所以

$$\mathscr{A}(\alpha_1) = a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1m}\beta_m,$$
$$\mathscr{A}(\alpha_2) = a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2m}\beta_m,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\mathscr{A}(\alpha_n) = a_{n1}\beta_1 + a_{n2}\beta_2 + \cdots + a_{nm}\beta_m.$$

上式可以写出矩阵形式如下:

$$\begin{aligned}\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) &= \big(\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \ldots, \mathscr{A}(\alpha_n)\big) \\ &= (\beta_1, \beta_2, \ldots, \beta_n)\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} \\ &=: (\beta_1, \beta_2, \ldots, \beta_n)A.\end{aligned}$$

即

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_n)A. \tag{5.2.1}$$

矩阵 A 称为线性变换 $\mathscr{A}$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵.

设 $\alpha \in U$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 $x^{\mathrm{T}} = (x_1, x_2, \ldots, x_n)$,则

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = (\alpha_1, \alpha_2, \ldots, \alpha_n)x.$$

因此

$$\begin{aligned}\mathscr{A}(\alpha) &= \mathscr{A}(x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n) \\ &= x_1\mathscr{A}(\alpha_1) + x_2\mathscr{A}(\alpha_2) + \cdots + x_n\mathscr{A}(\alpha_n) \\ &= \big(\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \ldots, \mathscr{A}(\alpha_n)\big)x \\ &= \mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n)x.\end{aligned}$$

由式 (5.2.1),

$$\mathscr{A}(\alpha) = (\beta_1, \beta_2, \ldots, \beta_m)Ax.$$

这表明,向量 α 在 $\mathscr{A}$ 下的像 $\mathscr{A}(\alpha)$ 在 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的坐标 y 为

$$y = Ax. \tag{5.2.2}$$

所有 U 到 V 的线性映射集合记为 $\mathscr{L}_{m\times n}$,数域 $\mathbb{F}$ 上所有 $m \times n$ 矩阵的集合记为 $\mathbb{F}^{m\times n}$. 设 $\mathscr{A} \in \mathscr{L}_{m\times n}$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵为 A.

定义 $\mathscr{L}_{m\times n}$ 到 $\mathbb{F}^{m\times n}$ 的映射 η 如下:对任意 $\mathscr{A} \in \mathscr{L}_{m\times n}$,令

$$\eta(\mathscr{A}) = A.$$

显然,对任意 $\mathscr{A}, \mathscr{B} \in \mathscr{L}_{m\times n}$,$\mathscr{A} \neq \mathscr{B}$,$\mathscr{A}$ 与 $\mathscr{B}$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵 A 与 B 满足 $A \neq B$. 因此 $\eta(\mathscr{A}) \neq \eta(\mathscr{B})$. 即映射 η 是单射. 另一方面,设 $A \in \mathbb{F}^{m\times n}$. 记

$$(\gamma_1, \gamma_2, \ldots, \gamma_n) = (\beta_1, \beta_2, \ldots, \beta_m)A.$$

显然,$\gamma_1, \gamma_2, \ldots, \gamma_n \in V$. 由定理 5.2.1,存在 $\mathscr{A} \in \mathscr{L}_{m\times n}$,使得 $\mathscr{A}(\alpha_j) = \gamma_j$. 因此

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\gamma_1, \gamma_2, \ldots, \gamma_n) = (\beta_1, \beta_2, \ldots, \beta_m)A.$$

这表明 $\mathscr{A}$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵为 A.

所以 $\eta(\mathscr{A})=A$,即映射 η 是满射. 于是 η 是双射.

这表明,数域 $\mathbb{F}$ 上一个 $m\times n$ 矩阵表示数域 $\mathbb{F}$ 上 n 维线性空间 U 到数域 $\mathbb{F}$ 上 m 维线性空间 V 的一个线性映射. 这就是给出了矩阵的一种几何意义.

设线性映射 $\mathscr{A}:U\to V$ 在 U 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 V 的基 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 下和 U 的基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n\}$ 与 V 的基 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m\}$ 下的矩阵分别为 A 与 B,即

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\beta_1,\beta_2,\ldots,\beta_m)A, \tag{5.2.3}$$
$$\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m)B. \tag{5.2.4}$$

矩阵 A 与 B 有什么联系? 下面回答这一问题.

设由基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 到基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n\}$ 的过渡矩阵为 Q,由基 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m\}$ 到 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 的过渡矩阵为 P,即

$$(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)Q, \tag{5.2.5}$$
$$(\beta_1,\beta_2,\ldots,\beta_m)=(\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m)P. \tag{5.2.6}$$

显然 Q 与 P 分别是 n 阶与 m 阶可逆方阵. 记 $Q=(q_{ij})$ 的第 j 列为 $q_j=(q_{1j},q_{2j},\ldots,q_{nj})^{\mathrm{T}}$. 由式 (5.2.5),$\widetilde{\alpha}_j=(\alpha_1,\alpha_2,\ldots,\alpha_n)q_j$.

由于映射 $\mathscr{A}$ 是线性的,所以 $\mathscr{A}(\widetilde{\alpha}_j)=\big(\mathscr{A}(\alpha_1),\ldots,\mathscr{A}(\alpha_n)\big)q_j$. 因此,

$$\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=\big(\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\ldots,\mathscr{A}(\alpha_n)\big)Q.$$

把式 (5.2.3) 代入上式得到,

$$\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\beta_1,\beta_2,\ldots,\beta_m)AQ.$$

再把式 (5.2.6) 代入上式得到,

$$\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m)PAQ.$$

由于线性映射 $\mathscr{A}$ 由基向量 $\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n$ 在 $\mathscr{A}$ 下的象所唯一确定,因此 $B=PAQ$. 即线性映射 $\mathscr{A}$ 的矩阵 A 与 B 是相抵的.

反之,设矩阵 A 与 B 相抵,即设 $B=PAQ$,其中 P 和 Q 分别是 m 阶与 n 阶可逆方阵. 在 U 与 V 中分别取基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与基 $\{\beta_1,\beta_2,\ldots,\beta_m\}$. 由式 (5.2.5) 和式 (5.2.6) 分别确定 U 的基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n\}$ 和 V 的基 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m\}$. 由

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\beta_1,\beta_2,\ldots,\beta_m)A,$$
$$\mathscr{B}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m)B$$

可以定义 U 到 V 的线性映射 $\mathscr{A}$ 与 $\mathscr{B}$. 由于 $B=PAQ$,故

$$\mathscr{B}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m)PAQ.$$

由式 (5.2.6),

$$\mathscr{B}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\beta_1,\beta_2,\ldots,\beta_m)AQ.$$

即

$$\mathscr{B}\big((\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)Q^{-1}\big)=(\beta_1,\beta_2,\ldots,\beta_m)A.$$

由式 (5.2.5),

$$\mathscr{B}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\beta_1,\beta_2,\dots,\beta_m)A.$$

由定理 5.2.1，$\mathscr{B}=\mathscr{A}$. 这表明，相抵的矩阵是同一个线性映射在不同基下的矩阵. 这就给出了矩阵相抵的几何意义.

利用矩阵在相抵下的标准形，可以证明

定理 5.2.2　设线性映射 $\mathscr{A}:U\to V$ 在 U 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 与 V 的基 $\{\beta_1,\beta_2,\dots,\beta_m\}$ 下的矩阵为 A，且 $\operatorname{rank}A=r$. 则 U 与 V 中分别存在基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_m\}$ 与 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m\}$，使得

$$\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n)=(\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m)\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}.$$

证明　由假设，

$$\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\beta_1,\beta_2,\dots,\beta_m)A.$$

因为 $\operatorname{rank}A=r$，因此存在 m 阶与 n 阶可逆方阵 P 与 Q，使得

$$PAQ=\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}.$$

令

$$(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)Q,$$
$$(\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m)=(\beta_1,\beta_2,\dots,\beta_m)P^{-1}.$$

由于 P,Q 可逆，所以 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n\}$ 与 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m\}$ 分别是 U 与 V 的基，而且 $\mathscr{A}$ 在 U 的基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n\}$ 与 V 的基 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m\}$ 下的矩阵即为

$$PAQ=\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}.$$ ■

习　题　5.2

1. 设 $\mathbb{F}[x]$ 是数域 $\mathbb{F}$ 上所有一元多项式 $f(x)$ 的集合. 定义映射 $\mathscr{A}:\mathbb{F}[x]\to\mathbb{F}[x]$ 如下：设 $f(x)\in\mathbb{F}[x]$，则令

$$\mathscr{A}(f(x))=f(x+1)-f(x).$$

证明 $\mathscr{A}$ 是线性映射.

2. 数域 $\mathbb{F}$ 上所有 n 阶方阵构成的线性空间记为 $\mathbb{F}^{n\times n}$. 定义映射 $\mathscr{A}:\mathbb{F}^{n\times n}\to\mathbb{F}^{n\times n}$ 如下：设 $X\in\mathbb{F}^{n\times n}$，则令

$$\mathscr{A}(X)=AX-XA,$$

其中 A 是 $\mathbb{F}^{n\times n}$ 中给定的方阵. 证明 $\mathscr{A}$ 是线性映射.

3. 数域 $\mathbb{F}$ 上 k 维列向量空间记为 $\mathbb{F}^k$. 定义映射 $\mathscr{A}:\mathbb{F}^n\to\mathbb{F}^m$ 如下：设 $\alpha\in\mathbb{F}^n$，则令 $\mathscr{A}(\alpha)=A\alpha$，其中 A 是 $\mathbb{F}^{m\times n}$ 中固定的矩阵. 证明 $\mathscr{A}$ 是线性映射，并且 $\mathscr{A}$ 为零映射当且仅当 A 为零矩阵.

4. 设 $\mathbb{C}$ 是复数域. 求映射 $\mathscr{A}:\mathbb{C}\to\mathbb{C}$，使得当 $\mathbb{C}$ 视为实数域上线性空间时 $\mathscr{A}$ 是线性的，而当 $\mathbb{C}$ 视为复数域上线性空间时 $\mathscr{A}$ 不是线性的.

5. 设 $\mathbb{F}^n$ 是数域 $\mathbb{F}$ 上 n 维行向量空间，π 是自然数集合 $\mathbb{N}=\{1,2,\dots,n\}$ 到自身上的映射. 定义映射 $\mathscr{A}_\pi:\mathbb{F}^n\to\mathbb{F}^n$ 如下：设 $x=(x_1,x_2,\dots,x_n)\in\mathbb{F}^n$，则令

$$\mathscr{A}_\pi(x) = \left(x_{\pi(1)}, x_{\pi(2)}, \dots, x_{\pi(n)}\right).$$

证明 $\mathscr{A}_\pi$ 是线性映射,并求 $\mathscr{A}_\pi$ 在基 $\left\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\right\}$ 下的方阵,其中 $\varepsilon_j = (0, \dots, 0, \underset{j}{1}, 0, \dots, 0)$.

记 N 到自身上所有双射的集合为 S_n. 设 $\pi \in S_n$,则 $\pi(1)\pi(2)\cdots\pi(n)$ 是 $1, 2, \dots, n$ 的一个排列. 排列 $\pi(1)\pi(2)\cdots\pi(n)$ 的奇偶性符号记为 $\operatorname{sgn}(\pi)$. 证明

$$\mathscr{A}(x) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\mathscr{A}_\pi(x)$$

是 $\mathbb{F}^n$ 到 $\mathbb{F}^n$ 的线性映射,并求出 $\mathscr{A}$ 在基 $\left\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\right\}$ 下的方阵.

6. 是否存在线性映射 $\mathscr{A}: \mathbb{R}^3 \to \mathbb{R}^2$,使得 $\mathscr{A}\big((1,-1,1)\big) = (1,0)$,$\mathscr{A}\big((1,1,1)\big) = (0,1)$?

7. 在实数域 $\mathbb{R}$ 上 2 维行向量空间 $\mathbb{R}^2$ 中,记

$$\alpha_1 = (1,-1), \qquad \alpha_2 = (2,-1), \qquad \alpha_3 = (-3,1);$$
$$\beta_1 = (1,0), \qquad \beta_2 = (0,1), \qquad \beta_3 = (1,1).$$

是否存在线性映射 $\mathscr{A}: \mathbb{R}^2 \to \mathbb{R}^2$,使得 $\mathscr{A}(\alpha_j) = \beta_j$, $j = 1, 2, 3$?

8. 证明,映射 $\mathscr{A}: U \to V$ 为线性映射的充分必要条件是,对任意 $\lambda, \mu \in \mathbb{F}$, $\alpha, \beta \in U$,均有

$$\mathscr{A}(\lambda\alpha + \mu\beta) = \lambda\mathscr{A}(\alpha) + \mu\mathscr{A}(\beta).$$

9. 在所有 2 阶实方阵构成的线性空间 $\mathbb{R}^{2\times2}$ 中,取

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

定义映射 $\mathscr{A}: \mathbb{R}^{2\times2} \to \mathbb{R}^{2\times2}$ 与 $\mathscr{B}: \mathbb{R}^{2\times2} \to \mathbb{R}^{2\times2}$ 如下: 设 $X \in \mathbb{R}^{2\times2}$,则令

$$\mathscr{A}(X) = AX, \qquad \mathscr{B}(X) = XA.$$

证明映射 $\mathscr{A}$ 与 $\mathscr{B}$ 都是线性的,并求出它们在基

$$\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

下的方阵.

10. $\mathbb{R}^{2\times2}$ 中所有形如

$$\begin{pmatrix} a & b \\ -b & d \end{pmatrix}$$

的方阵集合记为 V,其中 $a, b \in \mathbb{R}$. 显然 V 是实数域 $\mathbb{R}$ 上线性空间. 视复数域 $\mathbb{C}$ 为实数域 $\mathbb{R}$ 上线性空间. 定义映射 $\mathscr{A}: \mathbb{C} \to V$ 如下: 设 $\alpha = a + \mathrm{i}b \in \mathbb{C}$, $a, b \in \mathbb{R}$,则令

$$\mathscr{A}(\alpha) = \begin{pmatrix} a & b \\ -b & d \end{pmatrix}.$$

证明 $\mathscr{A}$ 是实线性空间 $\mathbb{C}$ 到 V 上的可逆线性映射,并且对任意 $\alpha, \beta \in \mathbb{C}$,

$$\mathscr{A}(\alpha\beta) = \mathscr{A}(\alpha)\mathscr{A}(\beta).$$

11. 定义映射 $\operatorname{Tr}: \mathbb{F}^{n\times n} \to \mathbb{F}$ 如下: 设 $A = (a_{ij}) \in \mathbb{F}^{n\times n}$,则令

$$\operatorname{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

证明映射 Tr 是线性映射,并且对任意 $A, B \in \mathbb{F}^{n\times n}$,

$$\operatorname{Tr}(AB) = \operatorname{Tr}(BA).$$

12. 设 $\mathscr{A}: \mathbb{F}^{n\times n} \to \mathbb{F}$ 是线性映射,并且对任意 $A, B \in \mathbb{F}^{n\times n}$, $\mathscr{A}(AB) = \mathscr{A}(BA)$. 证明 $\mathscr{A} = \lambda\operatorname{Tr}$,其中 $\lambda \in \mathbb{F}$.

§5.3　线性映射的代数运算

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间,$\mathscr{A}: U \to V$ 与 $\mathscr{B}: U \to V$ 是线性映射. 定义**映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的和** $\mathscr{A}+\mathscr{B}$ 如下,设 $\alpha \in U$,则令

$$(\mathscr{A}+\mathscr{B})(\alpha)=\mathscr{A}(\alpha)+\mathscr{B}(\alpha).$$

映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的和 $\mathscr{A}+\mathscr{B}$ 是 U 到 V 的线性映射. 事实上,对任意 $\lambda \in \mathbb{F}, \alpha, \beta \in U$,由和 $\mathscr{A}+\mathscr{B}$ 的定义,

$$\begin{aligned}(\mathscr{A}+\mathscr{B})(\alpha+\beta)&=\mathscr{A}(\alpha+\beta)+\mathscr{B}(\alpha+\beta),\\(\mathscr{A}+\mathscr{B})(\lambda\alpha)&=\mathscr{A}(\lambda\alpha)+\mathscr{B}(\lambda\alpha).\end{aligned}$$

因为 $\mathscr{A}$ 与 $\mathscr{B}$ 是线性映射,因此

$$\begin{aligned}(\mathscr{A}+\mathscr{B})(\alpha+\beta)&=\mathscr{A}(\alpha)+\mathscr{A}(\beta)+\mathscr{B}(\alpha)+\mathscr{B}(\beta)\\&=\big(\mathscr{A}(\alpha)+\mathscr{B}(\alpha)\big)+\big(\mathscr{A}(\beta)+\mathscr{B}(\beta)\big),\\(\mathscr{A}+\mathscr{B})(\lambda\alpha)&=\lambda\mathscr{A}(\alpha)+\lambda\mathscr{B}(\alpha)\\&=\lambda\big(\mathscr{A}(\alpha)+\mathscr{B}(\alpha)\big).\end{aligned}$$

由和 $\mathscr{A}+\mathscr{B}$ 的定义,

$$\begin{aligned}(\mathscr{A}+\mathscr{B})(\alpha+\beta)&=(\mathscr{A}+\mathscr{B})(\alpha)+(\mathscr{A}+\mathscr{B})(\beta),\\(\mathscr{A}+\mathscr{B})(\lambda\alpha)&=\lambda(\mathscr{A}+\mathscr{B})(\alpha).\end{aligned}$$

即映射 $\mathscr{A}+\mathscr{B}$ 满足 LM (1) 与 LM (2). 因此 $\mathscr{A}+\mathscr{B}: U \to V$ 是线性映射.

设 $\lambda \in \mathbb{F}, \mathscr{A}: U \to V$ 是线性映射. 定义**纯量 $\boldsymbol{\lambda}$ 与映射 $\mathscr{A}$ 的乘积 $\boldsymbol{\lambda}\mathscr{A}$** 如下:设 $\alpha \in U$,则令

$$(\lambda\mathscr{A})(\alpha)=\lambda\mathscr{A}(\alpha).$$

纯量 λ 与映射 $\mathscr{A}$ 的乘积 $\lambda\mathscr{A}$ 是 U 到 V 的线性映射. 事实上,对任意 $\mu \in \mathbb{F}$, $\alpha, \beta \in U$,由乘积 $\lambda\mathscr{A}$ 的定义,

$$\begin{aligned}(\lambda\mathscr{A})(\alpha+\beta)&=\lambda\mathscr{A}(\alpha+\beta),\\(\lambda\mathscr{A})(\mu\alpha)&=\lambda\mathscr{A}(\mu\alpha).\end{aligned}$$

因为 $\mathscr{A}: U \to V$ 是线性的,因此

$$\begin{aligned}(\lambda\mathscr{A})(\alpha+\beta)&=\lambda\big(\mathscr{A}(\alpha)+\mathscr{B}(\beta)\big)=\lambda\mathscr{A}(\alpha)+\lambda\mathscr{A}(\beta),\\(\lambda\mathscr{A})(\mu\alpha)&=\mu\lambda\mathscr{A}(\alpha).\end{aligned}$$

由 $\lambda\mathscr{A}$ 的定义,

$$\begin{aligned}(\lambda\mathscr{A})(\alpha+\beta)&=(\lambda\mathscr{A})(\alpha)+(\lambda\mathscr{A})(\beta),\\(\lambda\mathscr{A})(\mu\alpha)&=\mu(\lambda\mathscr{A})(\alpha).\end{aligned}$$

即乘积 $\lambda\mathscr{A}$ 满足 LM (1) 与 LM (2). 所以 $\lambda\mathscr{A}: U \to V$ 是线性映射.

所有 U 到 V 的线性映射集合记为 $\mathscr{L}_{m\times n}$. 容易验证,集合 $\mathscr{L}_{m\times n}$ 在上述的加法以及纯量与映射的乘法下构成数域 $\mathbb{F}$ 上线性空间.

定理 5.3.1 $$\dim \mathscr{L}_{m\times n} = mn.$$

证明 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 分别是 U 与 V 的基.

对给定的 $i,j,1\leqslant i\leqslant n,1\leqslant j\leqslant m$,定义映射 $\mathscr{E}_{ij}\colon U\to V$ 如下: 设 $\alpha\in U$,且 α 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为 $(a_1,a_2,\ldots,a_n)^{\mathrm{T}}$,则令

$$\mathscr{E}_{ij}(\alpha)=a_i\beta_j.$$

容易验证,映射 $\mathscr{E}_{ij}\in\mathscr{L}_{m\times n}$. 下面证明,$\{\mathscr{E}_{ij}\}$ 是 $\mathscr{L}_{m\times n}$ 的基.

首先证明,$\{\mathscr{E}_{ij}\}$ 线性无关. 事实上,设 $\lambda_{ij}\in\mathbb{F}$,使得

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_{ij}\mathscr{E}_{ij}=\mathscr{O},$$

其中 $\mathscr{O}\in\mathscr{L}_{m\times n}$ 是 U 到 V 的零映射. 则对任意 $\alpha\in U$,且 α 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为 $(a_1,a_2,\ldots,a_n)^{\mathrm{T}}$,有

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_{ij}\mathscr{E}_{ij}(\alpha)=\sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_{ij}a_i\beta_j=\sum_{j=1}^{m}\Big(\sum_{i=1}^{n}a_i\lambda_{ij}\Big)\beta_j=0.$$

由于 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 是 V 的基,因此对任意 $j=1,2,\ldots,m$,

$$\sum_{i=1}^{n}a_i\lambda_{ij}=0.$$

因为 $\alpha\in U$ 是任意的,所以可取 $\alpha=\alpha_k$,则 α 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为 $(0,\ldots,0,\underset{k}{1},0,\ldots,0)^{\mathrm{T}}$. 因此上式化为 $\lambda_{kj}=0$. 这就证得 $\{\mathscr{E}_{ij}\}$ 线性无关.

其次证明,任意 $\mathscr{A}\in\mathscr{L}_{m\times n}$ 均可由 $\{\mathscr{E}_{ij}\}$ 线性表出.

事实上,设 $\mathscr{A}\colon U\to V$ 在 U 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 V 的基 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 下的矩阵为 $A=(a_{ij})$,即

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\beta_1,\beta_2,\ldots,\beta_m)A.$$

考虑 U 到 V 的映射

$$\mathscr{A}-\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ji}\mathscr{E}_{ij}.$$

设 $\alpha\in U$ 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为 $a=(a_1,a_2,\ldots,a_n)^{\mathrm{T}}$,则

$$\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)a.$$

因此

$$\mathscr{A}(\alpha)=(\beta_1,\beta_2,\ldots,\beta_m)Aa=\sum_{i=1}^{n}\sum_{j=1}^{m}a_ia_{ji}\beta_j.$$

所以

$$\begin{aligned}\Big(\mathscr{A}-\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ji}\mathscr{E}_{ij}\Big)(\alpha)&=\mathscr{A}(\alpha)-\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ji}\mathscr{E}_{ij}(\alpha)\\&=\sum_{i=1}^{n}\sum_{j=1}^{m}a_ia_{ji}\beta_j-\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ji}a_i\beta_j=0.\end{aligned}$$

因此
$$\mathscr{A}-\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ji}\mathscr{E}_{ij}=\mathscr{O}\in\mathscr{L}_{m\times n},\quad 即\quad \mathscr{A}=\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ji}\mathscr{E}_{ij}.$$
这就证明,任意 $\mathscr{A}\in\mathscr{L}_{m\times n}$ 均可由 $\left\{\mathscr{E}_{ij}\right\}$ 线性表出.

综上所述,$\left\{\mathscr{E}_{ij}\,\middle|\,1\leqslant i\leqslant n,1\leqslant j\leqslant m\right\}$ 是 $\mathscr{L}_{m\times n}$ 的基,从而 $\dim\mathscr{L}_{m\times n}=mn$. ■

设 $\mathscr{A}\colon U\to V$ 与 $\mathscr{B}\colon V\to W$ 是线性映射. 作为映射,可以定义**映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的乘积** $\mathscr{B}\mathscr{A}\colon U\to W$,即对任意 $\alpha\in U$,
$$(\mathscr{B}\mathscr{A})(\alpha)=\mathscr{B}\big(\mathscr{A}(\alpha)\big).$$
设 $\lambda\in\mathbb{F},\alpha,\beta\in U$,则
$$(\mathscr{B}\mathscr{A})(\alpha+\beta)=\mathscr{B}\big(\mathscr{A}(\alpha+\beta)\big),$$
$$(\mathscr{B}\mathscr{A})(\lambda\alpha)=\mathscr{B}\big(\mathscr{A}(\lambda\alpha)\big).$$
由于映射 $\mathscr{A}$ 与 $\mathscr{B}$ 是线性的,所以
$$(\mathscr{B}\mathscr{A})(\alpha+\beta)=\mathscr{B}\big(\mathscr{A}(\alpha)\big)+\mathscr{B}\big(\mathscr{A}(\beta)\big)=(\mathscr{B}\mathscr{A})(\alpha)+(\mathscr{B}\mathscr{A})(\beta),$$
$$(\mathscr{B}\mathscr{A})(\lambda\alpha)=\lambda\mathscr{B}\big(\mathscr{A}(\alpha)\big)=\lambda(\mathscr{B}\mathscr{A})(\alpha).$$
所以 $\mathscr{B}\mathscr{A}\colon U\to W$ 是线性映射.

与映射的乘积相同,线性映射的乘积满足结合律,但交换律并不成立. 即一般地说,对线性映射 $\mathscr{A}$ 与 $\mathscr{B}$,等式 $\mathscr{B}\mathscr{A}=\mathscr{A}\mathscr{B}$ 不成立.

对零映射 $\mathscr{O}\colon U\to V$ 与线性映射 $\mathscr{A}\colon V\to W$,恒有 $\mathscr{A}\mathscr{O}=\mathscr{O}$,其中右端的 $\mathscr{O}$ 是 U 到 W 的零映射. 而对线性映射 $\mathscr{A}\colon W\to U$,恒有 $\mathscr{O}\mathscr{A}=\mathscr{O}$,其中右端的 $\mathscr{O}$ 是 W 到 V 的零映射.

作为集合,线性空间 U 到自身的单位映射记为 $\mathscr{E}_U$. 容易验证,$\mathscr{E}_U\colon U\to U$ 是线性映射. 并且对任意线性映射 $\mathscr{A}\colon U\to V$,恒有 $\mathscr{E}_V\mathscr{A}=\mathscr{A}=\mathscr{A}\mathscr{E}_U$.

利用线性映射的乘法所满足的结合律,可以定义 k 个线性映射的乘积,$k\geqslant 3$. 给定的 k 个线性映射 $\mathscr{A}_i\colon V_{i-1}\to V_i,1\leqslant i\leqslant k$,它们的乘积 $\mathscr{A}_k\cdots\mathscr{A}_2\mathscr{A}_1$ 是线性空间 V_0 到 V_k 的线性映射. 特别地,当 k 个线性映射 $\mathscr{A}_i$ 都是线性空间 V 到自身的同一个线性映射 $\mathscr{A}$ 时,则记
$$\mathscr{A}^k\coloneqq\underbrace{\mathscr{A}\mathscr{A}\mathscr{A}\cdots\mathscr{A}\mathscr{A}}_{k\text{ 个}}.$$
$\mathscr{A}^k$ 是 V 到自身的线性映射,并称为 **$\mathscr{A}$ 的 k 次幂**. 约定 $\mathscr{A}$ 的零次幂 $\mathscr{A}^0=\mathscr{E}_V$. 设多项式 $f(\lambda)=\sum\limits_{i=1}^{m}a_i\lambda^i$,则记
$$f(\mathscr{A})=\sum_{i=0}^{m}a_i\mathscr{A}^i,$$
它称为**线性映射 $\mathscr{A}$ 的多项式**.

设 $f(\lambda),g(\lambda),p(\lambda),q(\lambda)$ 是数域 $\mathbb{F}$ 上关于 λ 的多项式,并且 $f(\lambda)+g(\lambda)=p(\lambda)$, $f(\lambda)g(\lambda)=q(\lambda)$. 容易验证,
$$f(\mathscr{A})+g(\mathscr{A})=p(\mathscr{A}),\qquad f(\mathscr{A})g(\mathscr{A})=g(\mathscr{A})f(\mathscr{A})=q(\mathscr{A}).$$

现在给出线性映射的和,纯量与线性映射的乘积以及线性映射乘积的矩阵表示. 设线性映射 $\mathscr{A}: U \to V$ 与 $\mathscr{B}: U \to V$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵分别为 A 与 B,即

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)A,$$
$$\mathscr{B}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)B,$$

其中 $A = (a_{ij}), B = (b_{ij}) \in \mathbb{F}^{m\times n}$. 于是对于任意 $j = 1, 2, \ldots, n$,

$$\mathscr{A}(\alpha_j) = \sum_{k=1}^{m} a_{jk}\beta_k, \quad \mathscr{B}(\alpha_j) = \sum_{k=1}^{m} b_{jk}\beta_k.$$

由 $\mathscr{A} + \mathscr{B}$ 的定义,

$$(\mathscr{A} + \mathscr{B})(\alpha_j) = \mathscr{A}(\alpha_i) + \mathscr{B}(\alpha_j) = \sum_{k=1}^{m} a_{jk}\beta_k + \sum_{k=1}^{m} b_{jk}\beta_k = \sum_{k=1}^{m}(a_{jk} + b_{jk})\beta_k.$$

因此

$$(\mathscr{A} + \mathscr{B})(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)(A + B).$$

所以线性映射 $\mathscr{A} + \mathscr{B}$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵分别为 $A + B$.

对于 $\lambda \in \mathbb{F}$,

$$(\lambda\mathscr{A})(\alpha_j) = \lambda\mathscr{A}(\alpha_j) = \lambda\sum_{k=1}^{m} a_{jk}\beta_k = \sum_{k=1}^{m}(\lambda a_{jk})\beta_k.$$

因此

$$(\lambda\mathscr{A})(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)(\lambda A).$$

所以线性映射 $\lambda\mathscr{A}$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\beta_1, \beta_2, \ldots, \beta_m$ 下的矩阵为 λA.

设线性映射 $\mathscr{A}: U \to V$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵为 A,设线性映射 $\mathscr{B}: V \to W$ 在 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 与 W 的基 $\{\gamma_1, \gamma_2, \ldots, \gamma_\ell\}$ 下的矩阵为 B. 即

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)A,$$
$$\mathscr{B}(\beta_1, \beta_2, \ldots, \beta_m) = (\gamma_1, \gamma_2, \ldots, \gamma_\ell)B,$$

其中 $A = (a_{ij}) \in \mathbb{F}^{m\times n}, B = (b_{jk}) \in \mathbb{F}^{\ell\times m}$. 于是,

$$\mathscr{A}(\alpha_j) = \sum_{k=1}^{m} a_{jk}\beta_k, \quad \mathscr{B}(\beta_k) = \sum_{i=1}^{\ell} b_{ki}\gamma_i.$$

因此

$$\begin{aligned}(\mathscr{B}\mathscr{A})(\alpha_j) &= \mathscr{B}\big(\mathscr{A}(\alpha_j)\big) = \mathscr{B}\Big(\sum_{k=1}^{m} a_{jk}\beta_k\Big) = \sum_{k=1}^{m} a_{jk}\mathscr{B}(\beta_k) \\ &= \sum_{k=1}^{m}\sum_{i=1}^{\ell} a_{jk}b_{ki}\gamma_i = \sum_{i=1}^{\ell}\Big(\sum_{k=1}^{m} a_{jk}b_{ki}\Big)\gamma_i.\end{aligned}$$

所以

$$(\mathscr{B}\mathscr{A})(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_\ell)(BA).$$

这表明,线性映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的乘积 $\mathscr{B}\mathscr{A}$ 的矩阵表示等于线性映射 $\mathscr{B}$ 的矩阵表示 B 与线性映射 $\mathscr{A}$ 的矩阵表示 A 的乘积.

最后讨论可逆的线性映射. 设 $\mathscr{A}: U \to V$ 是线性映射. 如果作为映射,$\mathscr{A}$ 可逆,则 $\mathscr{A}$ 称为**可逆的线性映射**. 关于可逆的线性映射,由

定理 5.3.2　设线性 $\mathscr{A}: U \to V$ 可逆,则 $\dim U = \dim V$.

证明　由于线性映射 $\mathscr{A}$ 可逆,故由定理 5.1.3,$\mathscr{A}$ 是双射. 又映射 $\mathscr{A}$ 是线性的,因此 $\mathscr{A}$ 是 U 到 V 上的同构映射,即线性空间 U 与 V 同构. 所以 $\dim U = \dim V$. ■

定理 5.3.3　设线性映射 $\mathscr{A}: U \to V$ 可逆,则 $\mathscr{A}$ 的逆映射是可逆线性映射.

证明　设 $\mathscr{B}: V \to U$ 是映射 $\mathscr{A}: U \to V$ 的逆映射. 则 $\mathscr{B}$ 显然可逆,并且 $\mathscr{B}\mathscr{A} = \mathscr{E}_U$,$\mathscr{A}\mathscr{B} = \mathscr{E}_V$.

设 $\widetilde{\alpha}, \widetilde{\beta} \in V$,则存在唯一的 $\alpha, \beta \in U$,使得 $\mathscr{A}(\alpha) = \widetilde{\alpha}$,$\mathscr{A}(\beta) = \widetilde{\beta}$. 因此,

$$
\begin{aligned}
\mathscr{B}(\widetilde{\alpha} + \widetilde{\beta}) &= \mathscr{B}\big(\mathscr{A}(\alpha) + \mathscr{B}(\beta)\big) = \mathscr{B}\big(\mathscr{A}(\alpha + \beta)\big) \\
&= (\mathscr{B}\mathscr{A})(\alpha + \beta) = \mathscr{E}_U(\alpha + \beta) = \alpha + \beta \\
&= \mathscr{E}_U(\alpha) + \mathscr{E}_U(\beta) = (\mathscr{B}\mathscr{A})(\alpha) + (\mathscr{B}\mathscr{A})(\beta) \\
&= \mathscr{B}\big(\mathscr{A}(\alpha)\big) + \mathscr{B}\big(\mathscr{A}(\beta)\big) = \mathscr{B}(\widetilde{\alpha}) + \mathscr{B}(\widetilde{\beta}), \\
\mathscr{B}(\lambda\widetilde{\alpha}) &= \mathscr{B}\big(\lambda\mathscr{A}(\alpha)\big) = \mathscr{B}\big(\mathscr{A}(\lambda\alpha)\big) = (\mathscr{B}\mathscr{A})(\lambda\alpha) \\
&= \mathscr{E}_U(\lambda\alpha) = \lambda\alpha = \lambda\mathscr{E}_U(\alpha) \\
&= \lambda(\mathscr{B}\mathscr{A})(\alpha) = \lambda\mathscr{B}\big(\mathscr{A}(\alpha)\big) = \lambda\mathscr{B}(\widetilde{\alpha}),
\end{aligned}
$$

其中 $\lambda \in \mathbb{F}$. 因此 $\mathscr{B}: V \to U$ 是线性映射. ■

定理 5.3.4　设 $\mathscr{A}: U \to V$ 与 $\mathscr{B}: V \to W$ 是可逆线性映射. 则乘积 $\mathscr{B}\mathscr{A}: U \to W$ 是可逆线性映射.

证明　作为映射,$\mathscr{A}$ 与 $\mathscr{B}$ 可逆,因此由定理 5.1.5,$\mathscr{B}\mathscr{A}$ 可逆. 又映射,$\mathscr{A}$ 与 $\mathscr{B}$ 是线性的,所以 $\mathscr{B}\mathscr{A}$ 是线性的. 因此 $\mathscr{B}\mathscr{A}: U \to W$ 是可逆线性映射. ■

现在给出可逆线性映射 $\mathscr{A}: U \to V$ 的逆映射 $\mathscr{B}: V \to U$ 的矩阵表示.

由于映射 $\mathscr{A}$ 可逆,所以 $\dim U = \dim V$. 设线性映射 $\mathscr{A}$ 在 U 的基 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \dots, \beta_n\}$ 下的矩阵为 A,即

$$\mathscr{A}(\alpha_1, \alpha_2, \dots, \alpha_n) = (\beta_1, \beta_2, \dots, \beta_n)A.$$

设 $\mathscr{B}: V \to U$ 在 V 的基 $\{\beta_1, \beta_2, \dots, \beta_n\}$ 与 U 的基 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 下的矩阵为 B,即

$$\mathscr{B}(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)B.$$

因为 $\mathscr{B}\mathscr{A} = \mathscr{E}_U$,$\mathscr{A}\mathscr{B} = \mathscr{E}_V$,因此,

$$
\begin{aligned}
(\mathscr{B}\mathscr{A})(\alpha_1, \alpha_2, \dots, \alpha_n) &= (\alpha_1, \alpha_2, \dots, \alpha_n)BA = (\alpha_1, \alpha_2, \dots, \alpha_n)I_{(n)}, \\
(\mathscr{A}\mathscr{B})(\beta_1, \beta_2, \dots, \beta_n) &= (\beta_1, \beta_2, \dots, \beta_n)AB = (\beta_1, \beta_2, \dots, \beta_n)I_{(n)}.
\end{aligned}
$$

所以 $AB = I_{(n)} = BA$. 即可逆线性映射 $\mathscr{A}$ 的矩阵表示 A 是可逆方阵,并且 $\mathscr{A}$ 的逆映射 $\mathscr{B}$ 的矩阵表示是方阵 A 的逆方阵. 以后记线性映射 $\mathscr{A}$ 的逆映射为 $\mathscr{A}^{-1}$.

习 题 5.3

1. 设 $\{\varepsilon_1, \varepsilon_2\}$ 是数域 $\mathbb{F}$ 上 2 维行向量空间 $\mathbb{F}^2$ 的基,线性映射 $\mathscr{A}: \mathbb{F}^2 \to \mathbb{F}^2$ 与 $\mathscr{B}: \mathbb{F}^2 \to \mathbb{F}^2$ 分别把 $\{\varepsilon_1, \varepsilon_2\}$ 映为 $\{\varepsilon_2, 0\}$ 与 $\{0, \varepsilon_2\}$. 证明 $\mathscr{A}\mathscr{B} = \mathscr{O}$,但 $\mathscr{B}\mathscr{A} = \mathscr{A}$.

2. 定义微商映射 $\mathscr{D}: \mathbb{F}[x] \to \mathbb{F}[x]$ 如下: 设 $f(x) \in \mathbb{F}[x]$,则令 $\mathscr{D}\big(f(x)\big) = f'(x)$,其中 $f'(x)$ 是 $f(x)$ 的微商. 定义**积分映射** $\mathscr{S}: \mathbb{F}[x] \to \mathbb{F}[x]$ 如下: 设 $f(x) \in \mathbb{F}[x]$,则令

$$\mathscr{S}\big(f(x)\big) = \int_0^x f(t)\,\mathrm{d}t.$$

证明映射 $\mathscr{D}$ 与 $\mathscr{S}$ 是线性的,且 $\mathscr{D}$ 是满射,但不是单射; 而 $\mathscr{S}$ 是单射,但不是满射. 求 $\mathscr{S}\mathscr{D} - \mathscr{D}\mathscr{S}$.

3. 定义映射 $\mathscr{A}: \mathbb{F}[x] \to \mathbb{F}[x]$ 如下,设 $f(x) \in \mathbb{F}[x]$,则令 $\mathscr{A}\big(f(x)\big) = xf(x)$. 证明 $\mathscr{A}$ 是线性的. 设 $\mathscr{D}$ 是微商映射. 证明 $\mathscr{D}\mathscr{A} - \mathscr{A}\mathscr{D}$ 是单位映射.

4. 设 U 与 V 分别是数域 $\mathbb{F}$ 上 m 维与 n 维线性空间. 取定 $\alpha \in U$. 所有满足 $\mathscr{A}(\alpha) = 0$ 的线性映射 $\mathscr{A}: U \to V$ 的集合记为 K. 证明 K 在线性映射的加法以及纯量与线性映射的乘法下成为数域 $\mathbb{F}$ 上的线性空间. 求 $\dim K$.

5. 设 $\mathscr{A}: V \to V$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性映射. 所有满足 $\mathscr{A}\mathscr{B} = \mathscr{O}$ 的线性映射 $\mathscr{B}: V \to V$ 的集合记为 R. 证明集合 R 在线性映射的加法以及纯量与线性映射的乘法下成为数域 $\mathbb{F}$ 上的线性空间. 选择适当的线性映射 $\mathscr{A}$,使得 $\dim R = 0$,或 n,或 n^2.

6. 设 $\mathscr{D}: \mathbb{F}[x] \to \mathbb{F}[x]$ 是微商映射,$\mathscr{S}: \mathbb{F}[x] \to \mathbb{F}[x]$ 是积分映射. 确定映射 $\mathscr{D}^n\mathscr{S}^n$ 与 $\mathscr{S}^n\mathscr{D}^n$, $n = 1, 2, \ldots$.

7. 定义映射 $\mathscr{A}: \mathbb{F}_n[x] \to \mathbb{F}_n[x]$ 如下,设 $f(x) \in \mathbb{F}_n[x]$,则令 $\mathscr{A}\big(f(x)\big) = f(x+1)$. 证明,

$$\mathscr{A} = \mathscr{E} + \frac{\mathscr{D}}{1!} + \frac{\mathscr{D}^2}{2!} + \cdots + \frac{\mathscr{D}^{n-1}}{(n-1)!},$$

其中 $\mathscr{E}: \mathbb{F}_n[x] \to \mathbb{F}_n[x]$ 是单位映射,$\mathscr{D}: \mathbb{F}_n[x] \to \mathbb{F}_n[x]$ 是微商映射.

8. 设 V 是数域 $\mathbb{F}$ 上 n 维线性空间. 所有线性映射 $\mathscr{A}: V \to V$ 构成数域 $\mathbb{F}$ 上 n^2 维线性空间记为 U. 取定 $\mathscr{A} \in U$. 定义映射 $P_{\mathscr{A}}: U \to U$ 如下: 设 $X \in U$,则令 $P_{\mathscr{A}}(\mathscr{X}) = \mathscr{A}\mathscr{X}$. 证明 $P_{\mathscr{A}}$ 是线性的. 对映射 $Q: U \to U$,是否存在 $\mathscr{A} \in U$,使得 $Q = P_{\mathscr{A}}$?

§5.4 象 与 核

在本节中,恒假设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间.

定义 5.4.1 设 $\mathscr{A}: U \to V$ 是线性映射. 集合

$$\operatorname{Im}\mathscr{A} = \left\{\mathscr{A}(\alpha) \in V \;\middle|\; \alpha \in U\right\}$$

称为 ***U* 在 $\mathscr{A}$ 下的象**,或者 **$\mathscr{A}$ 的值域**,也记为 $\mathscr{A}(U)$. 集合

$$\operatorname{Ker}\mathscr{A} = \left\{\alpha \in U \;\middle|\; \mathscr{A}(\alpha) = 0 \in V\right\}$$

称为 **$\mathscr{A}$ 的核**,也记为 $\mathscr{A}^{-1}(0)$.

定理 5.4.1 设 $\mathscr{A}: U \to V$ 是线性映射,则 U 在 $\mathscr{A}$ 下的象 $\operatorname{Im}\mathscr{A}$ 是 V 的子空间,而 $\mathscr{A}$ 的核 $\operatorname{Ker}\mathscr{A}$ 是 U 的子空间.

证明 因为 $\mathscr{A}$ 是线性映射,因此 $\mathscr{A}(0) = 0$,即 $0 \in \operatorname{Im}\mathscr{A}$,即 $\operatorname{Im}\mathscr{A}$ 非空.

设 $\widetilde{\alpha},\widetilde{\beta}\in\operatorname{Im}\mathscr{A}$，则存在 $\alpha,\beta\in U$，使得 $\mathscr{A}(\alpha)=\widetilde{\alpha},\mathscr{A}(\beta)=\widetilde{\beta}$. 因此

$$\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta)=\widetilde{\alpha}+\widetilde{\beta},$$
$$\mathscr{A}(\lambda\alpha)=\lambda\mathscr{A}(\alpha)=\lambda\widetilde{\alpha},$$

其中 $\lambda\in\mathbb{F}$. 所以 $\widetilde{\alpha}+\widetilde{\beta}\in\operatorname{Im}\mathscr{A},\lambda\widetilde{\alpha}\in\operatorname{Im}\mathscr{A}$. 这就证明，$\operatorname{Im}\mathscr{A}$ 是 V 的子空间.

其次由 $\mathscr{A}(0)=0$ 可知，$0\in\operatorname{Ker}\mathscr{A}$，故 $\operatorname{Ker}\mathscr{A}$ 非空.

设 $\alpha,\beta\in\operatorname{Ker}\mathscr{A}$，则 $\mathscr{A}(\alpha)=0,\mathscr{A}(\beta)=0$. 由于 A 是线性映射，因此

$$\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta)=0,$$
$$\mathscr{A}(\lambda\alpha)=\lambda\mathscr{A}(\alpha)=\lambda\cdot 0=0,$$

其中 $\lambda\in\mathbb{F}$. 所以 $\alpha+\beta\in\operatorname{Ker}\mathscr{A},\lambda\alpha\in\operatorname{Ker}\mathscr{A}$. 这就证明，$\operatorname{Ker}\mathscr{A}$ 是 U 的子空间. ■

定理 5.4.2　设 $\mathscr{A}:U\to V$ 是线性映射，则 U 模 $\operatorname{Ker}\mathscr{A}$ 的商空间 $U/\operatorname{Ker}\mathscr{A}$ 同构于 $\operatorname{Im}\mathscr{A}$.

证明　设 $\alpha\in U$ 所在的模 $\operatorname{Ker}\mathscr{A}$ 的同余类为 $\overline{\alpha}$，显然 $\overline{\alpha}=\alpha+\operatorname{Ker}\mathscr{A}$.

定义映射 $\eta:U/\operatorname{Ker}\mathscr{A}\to\operatorname{Im}\mathscr{A}$ 如下：设 $\overline{\alpha}\in U/\operatorname{Ker}\mathscr{A}$，则令 $\eta(\overline{\alpha})=\mathscr{A}(\alpha)$.

设 $\beta\in\overline{\alpha}$，则 $\beta=\alpha+\gamma,\gamma\in\operatorname{Ker}\mathscr{A}$. 因此

$$\mathscr{A}(\beta)=\mathscr{A}(\alpha+\gamma)=\mathscr{A}(\alpha).$$

这表明，映射 η 的定义与同余类 $\overline{\alpha}$ 的代表元选取无关，因此映射 η 有确切定义.

设 $\overline{\alpha},\overline{\beta}\in U/\operatorname{Ker}\mathscr{A}$，且 $\eta(\overline{\beta})=\eta(\overline{\alpha})$. 则 $\mathscr{A}(\alpha)=\mathscr{A}(\beta)$，即 $\mathscr{A}(\alpha)-\mathscr{A}(\beta)=0$. 因为 $\mathscr{A}$ 是线性的，因此 $\mathscr{A}(\alpha-\beta)=0$. 所以 $\alpha-\beta\in\operatorname{Ker}\mathscr{A}$. 这表明，$\overline{\alpha}=\overline{\beta}$，即映射 η 是单射. 设 $\alpha^*\in\operatorname{Im}\mathscr{A}$，则存在 $\alpha\in U$，使得 $\mathscr{A}(\alpha)=\alpha^*$. 因此 $\eta(\overline{\alpha})=\mathscr{A}(\alpha)=\alpha^*$. 这表明映射 η 是满射. 从而 η 是双射.

现在设 $\overline{\alpha},\overline{\beta}\in U/\operatorname{Ker}\mathscr{A}$，则

$$\eta(\overline{\alpha}+\overline{\beta})=\eta(\overline{\alpha+\beta})=\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta)=\eta(\overline{\alpha})+\eta(\overline{\beta}),$$
$$\eta(\lambda\overline{\alpha})=\eta(\overline{\lambda\alpha})=\mathscr{A}(\lambda\alpha)=\lambda\mathscr{A}(\alpha)=\lambda\eta(\overline{\alpha}),$$

其中 $\lambda\in\mathbb{F}$. 这表明 η 保持加法与数乘.

综上所述 η 是 $U/\operatorname{Ker}\mathscr{A}$ 到 $\operatorname{Im}\mathscr{A}$ 上的同构映射. ■

由定理 5.4.2 以及定理 4.5.2 与推论 4.8.1 立即得到

定理 5.4.3　设 $\mathscr{A}:U\to V$ 是线性映射. 则

$$\dim U=\dim\operatorname{Im}\mathscr{A}+\dim\operatorname{Ker}\mathscr{A}.$$

证明　由推论 4.8.1，

$$\dim\left(U/\operatorname{Ker}\mathscr{A}\right)=\dim U-\dim\operatorname{Ker}\mathscr{A}.$$

由定理 5.4.2 和定理 4.5.2，

$$\dim\left(U/\operatorname{Ker}\mathscr{A}\right)=\dim\operatorname{Im}\mathscr{A}.$$

综合以上两式即得结论. ■

定义 5.4.2 设 $\mathscr{A}: U \to V$ 是线性映射,则 $\dim \operatorname{Im} \mathscr{A}$ 称为 $\mathscr{A}$ **的秩**,记为 $\rho(\mathscr{A})$;而 $\dim \operatorname{Ker} \mathscr{A}$ 称为 $\mathscr{A}$ **的零度**(**nullity**),记为 $\nu(\mathscr{A})$.

定理 5.4.4 设 $\mathscr{A}: U \to V$ 是线性映射,并且 $\rho(\mathscr{A}) = r$. 则存在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$,使得

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

证明 因为 $\rho(\mathscr{A}) = r$,故由定理 5.4.3,

$$\nu(\mathscr{A}) = \dim \operatorname{Ker} \mathscr{A} = n - r.$$

取 $\operatorname{Ker} \mathscr{A}$ 的基 $\{\alpha_{r+1}, \alpha_{r+2}, \ldots, \alpha_n\}$. 因为 $\operatorname{Ker} \mathscr{A}$ 是 U 的子空间,并且 $\dim U = n$,因此 $\{\alpha_{r+1}, \alpha_{r+2}, \ldots, \alpha_n\}$ 可扩充为 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n\}$. 记

$$\mathscr{A}(\alpha_j) = \beta_j, \tag{5.4.1}$$

其中 $j = 1, 2, \ldots, r$. 显然 $r \leqslant \dim V = m$,并且 $\{\beta_1, \beta_2, \ldots, \beta_r\} \subseteq \operatorname{Im} \mathscr{A}$. 设

$$\lambda_1\beta_1 + \lambda_2\beta_2 + \cdots + \lambda_r\beta_r = 0,$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$. 则

$$\begin{aligned} 0 &= \lambda_1\beta_1 + \lambda_2\beta_2 + \cdots + \lambda_r\beta_r = \lambda_1\mathscr{A}(\alpha_1) + \lambda_2\mathscr{A}(\alpha_2) + \cdots + \lambda_r\mathscr{A}(\alpha_r) \\ &= \mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_r\alpha_r). \end{aligned}$$

因此 $\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_r\alpha_r \in \operatorname{Ker} \mathscr{A}$. 由于 $\{\alpha_{r+1}, \alpha_{r+2}, \ldots, \alpha_n\}$ 是 $\operatorname{Ker} \mathscr{A}$ 的基,所以

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_r = \lambda_{r+1}\alpha_{r+1} + \lambda_{r+2}\alpha_{r+2} + \cdots + \lambda_n\alpha_n.$$

即

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_r\alpha_r + (-\lambda_{r+1})\alpha_{r+1} + (-\lambda_{r+2})\alpha_{r+2} - \cdots + (-\lambda_n)\alpha_n = 0.$$

因为 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 U 的基,所以

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0.$$

这就证明 $\{\beta_1, \beta_2, \ldots, \beta_r\}$ 线性无关. 由于 $\rho(\mathscr{A}) = \dim \operatorname{Im} \mathscr{A} = r$,因此 $\{\beta_1, \beta_2, \ldots, \beta_r\}$ 是 $\operatorname{Im} \mathscr{A}$ 的基. 由于 $\operatorname{Im} \mathscr{A} \subseteq V$,所以 $\{\beta_1, \beta_2, \ldots, \beta_r\}$ 可扩充为 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$. 由式 (5.4.1) 以及 $\alpha_{r+1}, \alpha_{r+2}, \ldots, \alpha_n \in \operatorname{Ker} \mathscr{A}$ 得到

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}. \qquad \blacksquare$$

定理 5.4.5 设 A 是线性映射 $\mathscr{A}: U \to V$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 与 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_m\}$ 下的矩阵. 则

$$\operatorname{rank} A = \dim \operatorname{Im} \mathscr{A}.$$

证明 由假设,

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_m)A.$$

考虑齐次线性方程组 $Ax=0$ 的解空间 $V_A=\{x\in\mathbb{F}^n \mid Ax=0\}$，其中 $\mathbb{F}^n$ 是数域 $\mathbb{F}$ 上 n 维列向量空间.

设 $\alpha\in\operatorname{Ker}\mathscr{A}$ 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为 $x\in\mathbb{F}^n$. 则 $\mathscr{A}(\alpha)$ 在基 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 下的坐标为 Ax. 由于 $\alpha\in\operatorname{Ker}(\mathscr{A})$，故 $\mathscr{A}(\alpha)=0$. 因此 $Ax=0$，即 $x\in V_A$.

于是定义映射 $\eta:\operatorname{Ker}\mathscr{A}\to V_A$ 如下：设 $\alpha\in U$，且 α 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为 x，则令 $\eta(\alpha)=x$.

容易验证，η 是 $\operatorname{Ker}\mathscr{A}$ 到 V_A 上的同构映射. 因此 $\nu(\mathscr{A})=\dim V_A$. 由例 4.6.3，

$$\nu(\mathscr{A})=\dim V_A=n-\operatorname{rank}A=\dim U-\operatorname{rank}A.$$

由定理 5.4.3，$\operatorname{rank}A=\dim\operatorname{Im}\mathscr{A}$. ■

容易看出，线性映射 $\mathscr{A}:U\to V$ 为满射当且仅当 $\operatorname{Im}\mathscr{A}=V$. 可以证明，

定理 5.4.6　线性映射 $\mathscr{A}:U\to V$ 为单射的充分必要条件是 $\operatorname{Ker}\mathscr{A}=0$.

证明　设 $\mathscr{A}:U\to V$ 是单射，且 $\alpha\in\operatorname{Ker}\mathscr{A}$. 则 $\mathscr{A}(\alpha)=0$，从而 $\alpha=0$，即 $\operatorname{Ker}\mathscr{A}=0$. 反之设 $\operatorname{Ker}\mathscr{A}=0$. 如果 $\alpha,\beta\in U$，且 $\mathscr{A}(\alpha)=\mathscr{A}(\beta)$，则 $\mathscr{A}(\alpha-\beta)=0$，故 $\alpha-\beta\in\operatorname{Ker}\mathscr{A}$. 因此 $\alpha-\beta=0$，即 $\alpha=\beta$. 所以 $\mathscr{A}$ 为单射. ■

定理 5.4.7　线性映射 $\mathscr{A}:U\to V$ 可逆的充分必要条件是

$$\dim V=\dim U\quad \text{且}\quad \operatorname{Ker}\mathscr{A}=0.$$

证明　设线性映射 $\mathscr{A}:U\to V$ 可逆，则由定理 5.1.3，$\mathscr{A}$ 是双射. 因此 $\operatorname{Im}\mathscr{A}=V$. 且由定理 5.4.6，$\operatorname{Ker}\mathscr{A}=0$. 由定理 5.4.3 即得

$$\dim V=\dim\operatorname{Im}\mathscr{A}=\dim U.$$

反之设 $\operatorname{Ker}\mathscr{A}=0$，且 $\dim V=\dim U$. 由定理 5.4.6，$\mathscr{A}:U\to V$ 为单射. 由定理 5.4.3，

$$\dim\operatorname{Im}\mathscr{A}=\dim U=\dim V.$$

由于 $\operatorname{Im}\mathscr{A}\subseteq V$，故由 $\dim V=\dim\operatorname{Im}\mathscr{A}$ 得到 $\operatorname{Im}\mathscr{A}=V$. ■

利用线性映射的象与核的结论，可以处理矩阵在相抵下的标准形理论. 为此先证明

定理 5.4.8　设 $A,B\in\mathbb{F}^{m\times n}$ 相抵，则 $\operatorname{rank}A=\operatorname{rank}B$.

证明　设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 分别是 U 与 V 的基. 由定理 5.2.1，存在线性映射 $\mathscr{A}:U\to V$，使得

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\beta_1,\beta_2,\ldots,\beta_m)A.$$

由定理 5.4.5，$\operatorname{rank}A=\dim\operatorname{Im}\mathscr{A}$.

因为矩阵 A 与 B 相抵，因此存在 m 阶与 n 阶可逆方阵 P 与 Q，使得 $B=PAQ$. 设

$$(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)Q,$$
$$(\beta_1,\beta_2,\ldots,\beta_m)=(\widetilde{\beta}_1,\widetilde{\beta}_2,\ldots,\widetilde{\beta}_m)P.$$

显然 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n\}$ 与 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m\}$ 分别是 U 与 V 的基,并且

$$\begin{aligned}\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n) &= \big(\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n)\big)Q\\ &= (\beta_1,\beta_2,\dots,\beta_m)AQ = (\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m)PAQ\\ &= (\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m)B.\end{aligned}$$

这表明 B 是线性 $\mathscr{A}:U\to V$ 在 U 的基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n\}$ 与 V 的基 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m\}$ 下的矩阵. 因此由定理 5.4.5,

$$\operatorname{rank} B = \dim \operatorname{Im}\mathscr{A} = \operatorname{rank} A. \qquad \blacksquare$$

定理 5.4.9 设 $A\in\mathbb{F}^{m\times n}$,且 $\operatorname{rank} A = r$. 则矩阵 A 相抵于矩阵

$$\begin{pmatrix} I_{(r)} & 0\\ 0 & 0\end{pmatrix}.$$

证明 设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\dots,\beta_m\}$ 分别是 U 与 V 的基. 由定理 5.2.1,存在线性映射 $\mathscr{A}:U\to V$,使得

$$\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n) = (\beta_1,\beta_2,\dots,\beta_m)A.$$

由定理 5.4.5, $\operatorname{rank} A = \dim\operatorname{Im}\mathscr{A} = r$.

定理 5.4.4,存在 U 的基 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n\}$ 与 V 的基 $\{\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m\}$,使得

$$\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n) = (\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m)\begin{pmatrix} I_{(r)} & 0\\ 0 & 0\end{pmatrix}.$$

设

$$\begin{aligned}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n) &= (\alpha_1,\alpha_2,\dots,\alpha_n)Q,\\ (\beta_1,\beta_2,\dots,\beta_m) &= (\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m)P,\end{aligned}$$

其中 P 与 Q 分别是 m 阶与 n 阶可逆方阵. 则

$$\begin{aligned}\mathscr{A}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\dots,\widetilde{\alpha}_n) &= \big(\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n)\big)Q = (\beta_1,\beta_2,\dots,\beta_m)AQ\\ &= (\widetilde{\beta}_1,\widetilde{\beta}_2,\dots,\widetilde{\beta}_m)PAQ.\end{aligned}$$

因此

$$PAQ = \begin{pmatrix} I_{(r)} & 0\\ 0 & 0\end{pmatrix}. \qquad \blacksquare$$

由定理 5.4.8 与定理 5.4.9 立即得到,矩阵的秩是矩阵在相抵下的全系不变量.

应当指出,利用矩阵在相抵下的标准形理论,也可以证明前面提到的关于线性映射的象与核的定理 5.4.3 与定理 5.4.4. 请读者自行证之,这里不拟赘述.

我们已经知道,给定 $A\in\mathbb{F}^{m\times n}$,取定数域 $\mathbb{F}$ 上 n 维线性空间 U 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 以及数域 $\mathbb{F}$ 上 m 维线性空间 V 的基 $\{\beta_1,\beta_2,\dots,\beta_m\}$,则存在线性映射 $\mathscr{A}:U\to V$,使得

$$\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n) = (\beta_1,\beta_2,\dots,\beta_m)A.$$

而且 $\rho(\mathscr{A}) = \operatorname{rank} A$. 因此我们可以利用有关线性映射 $\mathscr{A}$ 的象 $\operatorname{Im}\mathscr{A}$ 与核 $\operatorname{Ker}\mathscr{A}$ 的结论来论证有关矩阵的秩的命题.

例 5.4.1　设 $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times p}$. 证明

$$\operatorname{rank}(AB) \leqslant \min\{\operatorname{rank} A, \operatorname{rank} B\}.$$

证明　设 $\{\alpha_1, \alpha_2, \ldots, \alpha_p\}, \{\beta_1, \beta_2, \ldots, \beta_n\}$ 与 $\{\gamma_1, \gamma_2, \ldots, \gamma_m\}$ 依次是数域 $\mathbb{F}$ 上 p 维, n 维与 m 维线性空间 U, V 与 W 的基. 由

$$\mathscr{B}(\alpha_1, \alpha_2, \ldots, \alpha_p) = (\beta_1, \beta_2, \ldots, \beta_n)B$$

与

$$\mathscr{A}(\beta_1, \beta_2, \ldots, \beta_n) = (\gamma_1, \gamma_2, \ldots, \gamma_m)A$$

分别确定线性映射 $\mathscr{B}: U \to V$ 与 $\mathscr{A}: V \to W$. 于是线性映射 $\mathscr{A}\mathscr{B}: U \to W$ 的矩阵记为 AB, 即

$$\mathscr{A}\mathscr{B}(\alpha_1, \alpha_2, \ldots, \alpha_p) = (\gamma_1, \gamma_2, \ldots, \gamma_m)AB.$$

考虑象 $\operatorname{Im}\mathscr{A}$ 与 $\operatorname{Im}\mathscr{A}\mathscr{B}$. 显然 $\operatorname{Im}\mathscr{A}\mathscr{B} \subseteq \operatorname{Im}\mathscr{A}$. 因此 $\dim\operatorname{Im}\mathscr{A}\mathscr{B} \leqslant \dim\operatorname{Im}\mathscr{A}$. 根据定理 5.4.5,

$$\operatorname{rank} AB \leqslant \operatorname{rank} A.$$

再考虑核 $\operatorname{Ker}\mathscr{B}$ 与 $\operatorname{Ker}\mathscr{A}\mathscr{B}$. 显然, $\operatorname{Ker}\mathscr{B} \subseteq \operatorname{Ker}\mathscr{A}\mathscr{B}$. 因此

$$\dim\operatorname{Ker}\mathscr{B} \leqslant \dim\operatorname{Ker}\mathscr{A}\mathscr{B}.$$

由定理 5.4.3 与定理 5.4.5,

$$\begin{aligned}\dim\operatorname{Ker}\mathscr{B} &= p - \dim\operatorname{Im}\mathscr{B} = p - \operatorname{rank} B \\ &\leqslant \dim\operatorname{Ker}\mathscr{A}\mathscr{B} = p - \operatorname{rank} AB.\end{aligned}$$

所以,

$$\operatorname{rank} AB \leqslant \operatorname{rank} B.$$ ■

例 5.4.2　设 A 是数域 $\mathbb{F}$ 上 n 阶方阵. 证明,

$$\operatorname{rank} A^n = \operatorname{rank} A^{n+1} = \operatorname{rank} A^{n+2} = \cdots.$$

证明　取数域 $\mathbb{F}$ 上 n 维线性空间 V 的一组基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, 由

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$$

便确定 V 到自身的一个线性映射 $\mathscr{A}: V \to V$.

考虑 V 到自身的线性映射序列 $\mathscr{A}, \mathscr{A}^2, \ldots, \mathscr{A}^k, \ldots$. 它们在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的矩阵依次为 $A, A^2, \ldots, A^k, \ldots$. 由定理 5.4.3 与定理 5.4.5,

$$\operatorname{rank} A^k = n - \dim\operatorname{Ker}\mathscr{A}^k. \tag{5.4.2}$$

于是, 为证明结论, 只须证明,

$$\dim\operatorname{Ker}\mathscr{A}^n = \dim\operatorname{Ker}\mathscr{A}^{n+1} = \dim\operatorname{Ker}\mathscr{A}^{n+2} = \cdots,$$

即

$$\operatorname{Ker}\mathscr{A}^n = \operatorname{Ker}\mathscr{A}^{n+1} = \operatorname{Ker}\mathscr{A}^{n+2} = \cdots.$$

这启发我们考察核序列

$$\operatorname{Ker}\mathscr{A},\operatorname{Ker}\mathscr{A}^2,\ldots,\operatorname{Ker}\mathscr{A}^k,\ldots.$$

首先证明

$$\operatorname{Ker}\mathscr{A}\subseteq\operatorname{Ker}\mathscr{A}^2\subseteq\cdots\subseteq\operatorname{Ker}\mathscr{A}^k\subseteq\cdots\subseteq V. \tag{5.4.3}$$

事实上,设 $\alpha\in\operatorname{Ker}\mathscr{A}^k$,则 $\mathscr{A}^k(\alpha)=0$. 因此 $\mathscr{A}^{k+1}(\alpha)=\mathscr{A}\big(\mathscr{A}^k(\alpha)\big)=\mathscr{A}(0)=0$. 所以 $\alpha\in\operatorname{Ker}\mathscr{A}^{k+1}$. 于是 $\operatorname{Ker}\mathscr{A}^k\subseteq\operatorname{Ker}\mathscr{A}^{k+1}$, $k=1,2,\ldots$. 由此即得 (5.4.3).

由 (5.4.3) 得到

$$\dim\operatorname{Ker}\mathscr{A}\leqslant\dim\operatorname{Ker}\mathscr{A}^2\leqslant\cdots\leqslant\dim\operatorname{Ker}\mathscr{A}^k\leqslant\cdots\leqslant\dim V=n.$$

因此必有某个正整数 k,使得

$$\dim\operatorname{Ker}\mathscr{A}^k=\dim\operatorname{Ker}\mathscr{A}^{k+1}.$$

由于 $\operatorname{Ker}\mathscr{A}^k\subseteq\operatorname{Ker}\mathscr{A}^{k+1}$,因此,

$$\operatorname{Ker}\mathscr{A}^k=\operatorname{Ker}\mathscr{A}^{k+1}.$$

现在证明 $\operatorname{Ker}\mathscr{A}^{k+1}=\operatorname{Ker}\mathscr{A}^{k+2}$.

设 $\alpha\in\operatorname{Ker}\mathscr{A}^{k+2}$,则 $\mathscr{A}^{k+2}(\alpha)=\mathscr{A}^{k+1}\big(\mathscr{A}(\alpha)\big)=0$. 因此 $\mathscr{A}(\alpha)\in\operatorname{Ker}\mathscr{A}^{k+1}=\operatorname{Ker}\mathscr{A}^k$. 所以 $\mathscr{A}^k\big(\mathscr{A}(\alpha)\big)=\mathscr{A}^{k+1}(\alpha)=0$. 即 $\alpha\in\operatorname{Ker}\mathscr{A}^{k+1}$. 因此 $\operatorname{Ker}\mathscr{A}^{k+2}\subseteq\operatorname{Ker}\mathscr{A}^{k+1}$. 但是已经证得 $\operatorname{Ker}\mathscr{A}^{k+1}\subseteq\operatorname{Ker}\mathscr{A}^{k+2}$,所以 $\operatorname{Ker}\mathscr{A}^{k+1}=\operatorname{Ker}\mathscr{A}^{k+2}$.

同理可证,$\operatorname{Ker}\mathscr{A}^{k+2}=\operatorname{Ker}\mathscr{A}^{k+3}$,等等. 于是得到,

$$\operatorname{Ker}\mathscr{A}^k=\operatorname{Ker}\mathscr{A}^{k+1}=\operatorname{Ker}\mathscr{A}^{k+2}=\cdots.$$

式 (5.4.3) 中使上式成立的最小正整数仍记为 k,即

$$\operatorname{Ker}\mathscr{A}\subsetneqq\operatorname{Ker}\mathscr{A}^2\subsetneqq\cdots\subsetneqq\operatorname{Ker}\mathscr{A}^{k-1}\subsetneqq\operatorname{Ker}\mathscr{A}^k=\operatorname{Ker}\mathscr{A}^{k+1}=\cdots\subseteq V.$$

则

$$\dim\operatorname{Ker}\mathscr{A}<\dim\operatorname{Ker}\mathscr{A}^2<\cdots<\dim\operatorname{Ker}\mathscr{A}^{k-1}<\dim\operatorname{Ker}\mathscr{A}^k=\cdots\leqslant n.$$

因此 $k\leqslant n$. 这表明,

$$\operatorname{Ker}\mathscr{A}^n=\operatorname{Ker}\mathscr{A}^{n+1}=\operatorname{Ker}\mathscr{A}^{n+2}=\cdots.$$

由式 (5.4.2) 即得结论. ■

对于 V 到自身的线性映射序列 $\mathscr{A},\mathscr{A}^2,\ldots,\mathscr{A}^k,\ldots$,如果考虑相应的象序列

$$\operatorname{Im}\mathscr{A},\operatorname{Im}\mathscr{A}^2,\ldots,\operatorname{Im}\mathscr{A}^k,\ldots,$$

可以给出例 5.4.2 的另一证明. 请读者自行证之.

例 5.4.3 设 A 是数域 $\mathbb{F}$ 上 n 阶方阵. 证明,存在数域 $\mathbb{F}$ 上 n 阶可逆方阵 P,使得 PA 是幂等方阵(即满足 $(PA)^2=PA$).

证法 1 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的基. 由

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A$$

可以确定 V 的一个线性映射 $\mathscr{A}$.

设 $\operatorname{rank} A = r$,则由定理 5.4.3 与定理 5.4.5,$\dim \operatorname{Ker} \mathscr{A} = n - r$.

设 $\{\beta_{r+1}, \dots, \beta_n\}$ 是核 $\operatorname{Ker} \mathscr{A}$ 的基,它可以扩充为 V 的基 $\{\beta_1, \dots, \beta_r, \beta_{r+1}, \dots, \beta_n\}$.

和定理 5.4.4 的证明一样,可以证明,$\{\mathscr{A}(\beta_1), \dots, \mathscr{A}(\beta_r)\}$ 是象 $\operatorname{Im} \mathscr{A}$ 的基. 记 $\mathscr{A}(\beta_j) = \gamma_j$. 把 $\{\gamma_1, \dots, \gamma_r\}$ 扩充成 V 的基 $\{\gamma_1, \dots, \gamma_r, \gamma_{r+1}, \dots, \gamma_n\}$. 由定理 5.2.1,由

$$\mathscr{P}(\gamma_1, \gamma_2, \dots, \gamma_n) = (\beta_1, \beta_2, \dots, \beta_n)$$

可以确定一个可逆线性映射 $\mathscr{P}: V \to V$. 于是,当 $1 \leqslant j \leqslant r$ 时,

$$\mathscr{P}\mathscr{A}(\beta_j) = \mathscr{P}\big(\mathscr{A}(\beta_j)\big) = \mathscr{P}(\gamma_j) = \beta_j,$$
$$(\mathscr{P}\mathscr{A})^2(\beta_j) = \beta_j = \mathscr{P}\mathscr{A}(\beta_j);$$

而当 $r + 1 \leqslant j \leqslant n$ 时,

$$\mathscr{P}\mathscr{A}(\beta_j) = \mathscr{P}\big(\mathscr{A}(\beta_j)\big) = \mathscr{P}(0) = 0,$$
$$(\mathscr{P}\mathscr{A})^2(\beta_j) = 0 = \mathscr{P}\mathscr{A}(\beta_j).$$

因此,

$$(\mathscr{P}\mathscr{A})^2 = \mathscr{P}\mathscr{A}.$$

记

$$\mathscr{P}(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P.$$

因为 $\mathscr{P}$ 是可逆线性映射,所以方阵 P 是可逆的. 因此,

$$\mathscr{P}\mathscr{A}(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)PA.$$

由于 $(\mathscr{P}\mathscr{A})^2 = \mathscr{P}\mathscr{A}$,所以 $(PA)^2 = PA$. ■

证法 2 设 $\operatorname{rank} A = r$,则存在 n 阶可逆方阵 Q 与 R,使得

$$QAR = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

因此

$$RQA = R(QAR)R^{-1} = R\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}R^{-1}.$$

记 $P = RQ$. 显然 n 阶方阵 P 可逆,并且 $(PA)^2 = PA$. ■

例 5.4.4 设 U_1 是数域 $\mathbb{F}$ 上 n 维线性空间 U 的 k 维子空间,$\mathscr{A}: U \to V$ 是线性映射. 记 $\mathscr{A}(U_1) = \{\mathscr{A}(\alpha) \mid \alpha \in U_1\}$. 显然 $\mathscr{A}(U_1)$ 是 V 的子空间. 证明,

$$\dim \mathscr{A}(U_1) \geqslant \dim U_1 - n + \rho(\mathscr{A}).$$

证明 考虑映射 $\mathscr{A}|_{U_1}: U_1 \to \mathscr{A}(U_1)$ 如下:设 $\alpha \in U_1$,则令 $\mathscr{A}|_{U_1}(\alpha) = \mathscr{A}(\alpha)$.

显然映射 $\mathscr{A}|_{U_1}$ 是线性映射,它称为线性映射 $\mathscr{A}$ 在子空间 U_1 上的限制. 容易验证,$\operatorname{Im} \mathscr{A}|_{U_1} = \mathscr{A}(U_1)$,且 $\operatorname{Ker} \mathscr{A}|_{U_1} = \{\alpha \in U_1 \mid \mathscr{A}(\alpha) = 0\}$. 因此 $\operatorname{Ker} \mathscr{A}|_{U_1} \subseteq \operatorname{Ker} \mathscr{A}$,于是

$$\dim \operatorname{Ker} \mathscr{A}|_{U_1} \leqslant \dim \operatorname{Ker} \mathscr{A}.$$

由定理 5.4.3 和定理 5.4.5,

$$\dim U_1 - \dim \operatorname{Im} \mathscr{A}|_{U_1} \leqslant \dim U - \dim \operatorname{Im} \mathscr{A}.$$

因此,

$$\dim \mathscr{A}(U_1) = \dim \operatorname{Im} \mathscr{A}\Big|_{U_1} \geqslant \dim U_1 - \dim U + \dim \operatorname{Im} \mathscr{A}$$
$$\geqslant \dim U_1 - n + \rho(\mathscr{A}).$$ ■

习 题 5.4

1. 设 $\mathscr{D}:\mathbb{F}_n[x] \to \mathbb{F}_n[x]$ 是微商映射. 求 $\rho(\mathscr{A})$ 与 $\nu(\mathscr{A})$. 等式 $\mathbb{F}_n[x] = \operatorname{Im}\mathscr{D} \oplus \operatorname{Ker}\mathscr{D}$ 是否成立?

2. 取

$$A = \begin{pmatrix} 1 & -1 \\ -4 & 4 \end{pmatrix}.$$

定义线性映射 $\mathscr{A}:\mathbb{C}^{2\times 2} \to \mathbb{C}^{2\times 2}$ 如下: 设 $X \in \mathbb{C}^{2\times 2}$,则令 $\mathscr{A}(X) = AX$. 求 $\rho(X)$.

3. 设 $\mathscr{A}: V \to V$ 是数域上 n 维线性空间 V 到自身的线性映射,且 $\rho(\mathscr{A}^2) = \rho(\mathscr{A})$. 证明

$$\operatorname{Im}\mathscr{A} \cap \operatorname{Ker}\mathscr{A} = 0.$$

4. 设 W 是数域 $\mathbb{F}$ 上 n 维线性空间 V 到自身的所有线性映射构成的线性空间,$\mathscr{A} \in W$,且 $\rho(\mathscr{A}) = k$. 定义线性映射 $\mathscr{T}_{\mathscr{A}}: W \to W$ 如下: 设 $\mathscr{X} \in W$,则令 $\mathscr{T}_{\mathscr{A}}(\mathscr{X}) = \mathscr{A}\mathscr{X}$. 求 $\rho\big(\mathscr{T}_{\mathscr{A}}\big)$ 与 $\nu\big(\mathscr{T}_{\mathscr{A}}\big)$.

5. 设 $\mathscr{A}: U \to U$ 是线性映射. 证明存在线性映射 $\mathscr{B}: U \to U$,使得 $\mathscr{A}\mathscr{B} = 0$,且

$$\rho(\mathscr{A}) + \rho(\mathscr{B}) = \dim U.$$

6. 设 $\mathscr{A}, \mathscr{B}, \mathscr{C}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 U 到自身的线性映射. 证明

$$\rho(\mathscr{A}\mathscr{B}) + \rho(\mathscr{B}\mathscr{C}) \leqslant \rho(\mathscr{B}) + \rho(\mathscr{A}\mathscr{B}\mathscr{C}).$$

7. 设 $\mathscr{A}: U \to U$ 是线性映射,$\rho(\mathscr{A}) = 1$. 证明,存在唯一 $\lambda \in \mathbb{F}$,使得 $\mathscr{A}^2 = \lambda\mathscr{A}$; 而且当 $\lambda \neq 1$ 时,$\mathscr{E}_U - \mathscr{A}$ 是可逆线性映射.

8. 设 $V_0, V_1, \ldots, V_{n+1}$ 是数域 $\mathbb{F}$ 上有限维线性空间,$V_0 = V_{n+1} = 0$. 设 $\mathscr{A}_i: V_i \to V_{i+1}$ 是线性映射,$i = 0, 1, 2, \ldots, n$. 并且 $\operatorname{Ker}\mathscr{A}_{i+1} = \operatorname{Im}\mathscr{A}_i$, $i = 0, 1, 2, \ldots, n-1$. 证明,

$$\sum_{i=1}^{n} (-1)^i \dim V_i = 0.$$

9. 设 $\mathscr{A}:\mathbb{F}^3 \to \mathbb{F}^3$ 定义如下: 对任意 $(x, y, z) \in \mathbb{F}^3$,令 $\mathscr{A}\big((x, y, z)\big) = (0, x+y, 0)$. $\mathbb{F}^3$ 中由向量 $\varepsilon_1 = (1, 0, 0)$ 与 $\varepsilon_2 = (0, 1, 0)$ 生产的子空间分别记为 U 与 V. 等式 $\mathscr{A}(U \cap V) = \mathscr{A}(U) \cap \mathscr{A}(V)$ 是否成立?

10. 设 $A, B \in \mathbb{F}^{m\times n}$. 证明 $\operatorname{rank}(A+B) = \operatorname{rank} A + \operatorname{rank} B$ 的充分必要条件是,存在数域 $\mathbb{F}$ 上 m 阶与 n 阶可逆方阵 P 与 Q 使得

$$PAQ = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}, \quad PBQ = \begin{pmatrix} 0 & 0 \\ 0 & I_{(s)} \end{pmatrix},$$

其中 $r = \operatorname{rank} A$, $s = \operatorname{rank} B$, 且 $r + s \leqslant \min\{m, n\}$.

§5.5　线性变换

本节恒假设 V 是数域 $\mathbb{F}$ 上 n 维线性空间.

我们知道,集合 S 到自身的映射通常也称为集合 S 的变换. 因此线性映射 $\mathscr{A}: V \to V$ 也称为线性空间 V 的**线性变换**. 由于线性变换是线性映射的特例,所以前面叙述的关于线性映射的结论对线性变换也成立. 当然,由于线性变换是特殊类型的线性映射,所以关于线性变换的结论也带有某种特殊性. 这里只总结一下那些带有特殊性的结论. 首先考虑线性变换 $\mathscr{A}: V \to V$ 的矩阵表示.

设 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 V 的基. 则 $\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \ldots, \mathscr{A}(\alpha_n) \in V$. 因此

$$\begin{aligned}
\mathscr{A}(\alpha_1) &= a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n, \\
\mathscr{A}(\alpha_2) &= a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2n}\alpha_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\mathscr{A}(\alpha_n) &= a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + a_{nn}\alpha_n.
\end{aligned}$$

即

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}.$$

记 $A = (a_{ij}) \in \mathbb{F}^{n\times n}$. 则上式为

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A. \tag{5.5.1}$$

方阵 A 称为线性变换 $\mathscr{A}: V \to V$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的矩阵.

数域 $\mathbb{F}$ 上 n 维线性空间 V 的所有线性变换集合记为 $\mathscr{L}_n(V)$. 容易验证, $\mathscr{L}_n(V)$ 在线性变换(即线性映射)的加法与纯量和线性变换的乘法下成为数域 $\mathbb{F}$ 上 n^2 维线性空间. 数域 $\mathbb{F}$ 上所有 n 阶方阵集合记为 $\mathbb{F}^{n\times n}$. 定义映射 $\eta: \mathscr{L}_n(V) \to \mathbb{F}^{n\times n}$ 如下: 设 $\mathscr{A} \in \mathscr{L}_n(V)$,且 $\mathscr{A}$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A,即式 (5.5.1) 成立,则令 $\eta(\mathscr{A}) = A$.

容易验证,映射 η 是单射. 并由式 (5.5.1) 可以确定一个线性变换 $\mathscr{A} \in \mathscr{L}_n(V)$,使得 $\eta(\mathscr{A}) = A$,即映射 η 是一个满射. 不但如此,映射 η 是 $\mathscr{L}_n(V)$ 到 $\mathbb{F}^{n\times n}$ 上的同构映射.

设 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 是 V 的另一组基,线性变换 $\mathscr{A}: V \to V$ 在基 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 下的方阵为 B,即

$$\mathscr{A}(\beta_1, \beta_2, \ldots, \beta_n) = (\beta_1, \beta_2, \ldots, \beta_n)B. \tag{5.5.2}$$

设基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 到基 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 的过渡矩阵为 P,即

$$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P, \tag{5.5.3}$$

其中 P 是数域 $\mathbb{F}$ 上 n 阶可逆方阵. 于是

$$\mathscr{A}(\beta_1, \beta_2, \dots, \beta_n) = \big(\mathscr{A}(\alpha_1, \alpha_2, \dots, \alpha_n)\big)P.$$

由 (5.5.1),得

$$= (\alpha_1, \alpha_2, \dots, \alpha_n)AP.$$

由 (5.5.3),得

$$= (\beta_1, \beta_2, \dots, \beta_n)P^{-1}AP. \tag{5.5.4}$$

比较式 (5.5.2) 与式 (5.5.4),得到 $B = P^{-1}AP$.

定义 5.5.1 设 $A, B \in \mathbb{F}^{n\times n}$,如果存在 n 阶可逆方阵 $P \in \mathbb{F}^{n\times n}$,使得

$$B = P^{-1}AP,$$

则方阵 A, B 称为**相似的**.

于是我们得到:

定理 5.5.1 设 $\mathscr{A}$ 是 V 上线性变换,则 $\mathscr{A}$ 在 V 的不同基下的方阵是相似的.

反之,设 $A, B \in \mathbb{F}^{n\times n}$ 相似,基存在可逆方阵 $P \in \mathbb{F}^{n\times n}$,使得 $B = P^{-1}AP$. 在数域 $\mathbb{F}$ 上 n 维线性空间 V 中取定基 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$,由式 (5.5.1),可以确定线性变换 $\mathscr{A}: V \to V$. 由式 (5.5.3) 可以确定 V 的基 $\{\beta_1, \beta_2, \dots, \beta_n\}$. 于是由式 (5.5.4),线性变换 $\mathscr{A}: V \to V$ 在基 $\{\beta_1, \beta_2, \dots, \beta_n\}$ 下的方阵记为 $B = P^{-1}AP$. 这就证明,

定理 5.5.2 设 $A \in \mathbb{F}^{n\times n}$,则由式 (5.5.1) 可以确定一个线性变换 $\mathscr{A}: V \to V$. 而且相似的方阵是同一个线性变换在不同基下的方阵.

方阵之间的相似关系是集合 $\mathbb{F}^{n\times n}$ 中元素间的一种关系. 容易验证,方阵之间的相似关系具有如下的性质:

(1) 自反性 对任意 $A \in \mathbb{F}^{n\times n}$,方阵 A 与自身相似;

(2) 对称性 设 $A, B \in \mathbb{F}^{n\times n}$ 相似,则 B, A 相似;

(3) 传递性 设 $A, B, C \in \mathbb{F}^{n\times n}$,且 A, B 相似,B, C 相似,则 A, C 相似.

因此利用方阵之间的相似关系,可以把 $\mathbb{F}^{n\times n}$ 分类: 彼此相似的方阵归在同一个方阵类,而彼此不相似的方阵归在不同的方阵类. 自然产生的问题是:

(1) 如何判定两个方阵是否同属于一个方阵类? 也即方阵 A 与 B 相似的充分必要条件是什么?

(2) 在每个方阵相似类中每个方阵类里如何选择一个方阵作为它的代表元? 当然,每个方阵相似类中每个方阵都可以作为这个方阵相似类的代表元,不过我们总希望作为代表元的方阵具有某种特性.

这就是方阵在相似下的标准形问题. 由定理 5.5.2 可以看出,方阵在相似下的标准形问题之一是,判断两个方阵所表示的线性变换是否相同; 由定理 5.5.1 可以看出,方阵在相似下的标准形问题之二是,选择 V 的一组基,使得线性变换在这组基下的方阵具有简单的形式. 方阵在相似下的标准形理论是线性代数与矩阵论的一个重要组成部分. 在线性代数与矩阵论的教材中,处理方阵在相似下的

标准形理论的方法多种多样,有的侧重用几何语言讲述,有的侧重用矩阵语言讲述,各有千秋. 在第 6 章中将首先用几何方法导出方阵在相似下的标准形,然后再用矩阵方法给出求矩阵在相似下的标准形的方法. 当然在讲述方阵在相似下的标准形理论之前,还需用几节的篇幅介绍一些准备知识.

现在转到可逆线性变换. 所谓**可逆线性变换** $\mathscr{A}\colon V\to V$ 是指线性空间 V 到自身的可逆线性映射. 设可逆线性变换 $\mathscr{A}\colon V\to V$ 在 V 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的方阵为 A,即

$$\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)A,$$

则 A 可逆,并且 $\mathscr{A}$ 的逆变换 $\mathscr{A}^{-1}$ 在同一组基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的方阵为 A^{-1},即

$$\mathscr{A}^{-1}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)A^{-1}.$$

由定理 5.1.3,线性映射 $\mathscr{A}\colon U\to V$ 可逆的充分必要条件是 $\mathscr{A}$ 为双射. 对于线性变换,条件可以减弱,即有

定理 5.5.3　设 $\mathscr{A}\colon V\to V$ 是线性变换,则下述三条件等价:

(1) 线性变换 $\mathscr{A}$ 可逆;　**(2)** $\operatorname{Ker}\mathscr{A}=0$;　**(3)** $\operatorname{Im}\mathscr{A}=V$.

证明　**(1) $\Longrightarrow$ (2)**　因为线性变换 $\mathscr{A}$ 可逆,因此由定理 5.4.7,$\operatorname{Ker}\mathscr{A}=0$.

(2) $\Longrightarrow$ (3)　设 $\operatorname{Ker}\mathscr{A}=0$,则由定理 5.4.3,$\dim\operatorname{Im}\mathscr{A}=\dim V=n$. 由于 $\operatorname{Im}\mathscr{A}\subseteq V$,因此 $\operatorname{Im}\mathscr{A}=V$.

(3) $\Longrightarrow$ (1)　因为 $\operatorname{Im}\mathscr{A}=V$,故由定理 5.4.3,$\operatorname{Ker}\mathscr{A}=0$,因此由定理 5.4.6,映射 $\mathscr{A}$ 是单射. 而 $\operatorname{Im}\mathscr{A}=V$ 表明,映射 $\mathscr{A}$ 是满射. 所以 $\mathscr{A}$ 是双射,从而线性变换 $\mathscr{A}$ 可逆. ■

定理 5.5.4　线性变换 $\mathscr{A}\colon V\to V$ 可逆当且仅当线性变换 $\mathscr{A}$ 把 V 的基变为 V 的基,即如果 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 V 的基,则 $\{\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\dots,\mathscr{A}(\alpha_n)\}$ 是 V 的基.

证明　设线性变换 $\mathscr{A}$ 可逆,且存在 $\lambda_1,\lambda_2,\dots,\lambda_n\in\mathbb{F}$,使得

$$\lambda_1\mathscr{A}(\alpha_1)+\lambda_2\mathscr{A}(\alpha_2)+\cdots+\lambda_n\mathscr{A}(\alpha_n)=0.$$

因为 $\mathscr{A}$ 是线性的,所以

$$\mathscr{A}(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_n)=0.$$

由于 $\mathscr{A}$ 可逆,故由定理 5.5.3,$\operatorname{Ker}\mathscr{A}=0$. 因此 $\lambda_1\alpha_1+\cdots+\lambda_n\alpha_n=0$,但 $\alpha_1,\dots,\alpha_n$ 线性无关,所以

$$\lambda_1=\lambda_2=\cdots=\lambda_n=0.$$

这就证明,$\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\dots,\mathscr{A}(\alpha_n)$ 线性无关.

又因为 $\dim V=n$,所以 $\{\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\dots,\mathscr{A}(\alpha_n)\}$ 是 V 的基.

反之,设 $\{\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\dots,\mathscr{A}(\alpha_n)\}$ 是 V 的基,$\alpha\in\operatorname{Ker}\mathscr{A}$,且 $\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n$,则

$$\mathscr{A}(\alpha)=a_1\mathscr{A}(\alpha_1)+a_2\mathscr{A}(\alpha_2)+\cdots+a_n\mathscr{A}(\alpha_n)=0.$$

因为 $\mathscr{A}(\alpha_1),\mathscr{A}(\alpha_2),\dots,\mathscr{A}(\alpha_n)$ 线性无关,所以 $a_1=a_2=\cdots=a_n=0$,即 $\alpha=0$. 因此

$\operatorname{Ker}\mathscr{A}=0$. 由定理 5.5.3,线性变换 $\mathscr{A}$ 可逆. ■

习 题 5.5

1. 设线性变换 $\mathscr{A}:\mathbb{C}^2\to\mathbb{C}^2$ 在基 $\{(1,0),(0,1)\}$ 下的方阵为

$$\begin{pmatrix}1&1\\1&1\end{pmatrix}.$$

求 $\mathscr{A}$ 在基 $\{(1,1),(1,-1)\}$ 与 $\{(1,0),(1,1)\}$ 下的方阵.

2. 设线性变换 $\mathscr{A}:\mathbb{C}^3\to\mathbb{C}^3$ 在基 $\{(1,0,0),(0,1,0),(0,0,1)\}$ 下的方阵为

$$A=\begin{pmatrix}0&1&1\\1&0&-1\\-1&-1&0\end{pmatrix}.$$

求 $\mathscr{A}$ 在基 $\{(0,1,-1),(1,-1,1),(-1,1,0)\}$ 下的方阵.

3. 设数域 $\mathbb{F}$ 上 n 维线性空间 V 是子空间 U 与 W 的直和,则对任意 $\alpha\in V$,存在唯一一对向量 β 与 γ,$\beta\in U,\gamma\in W$,使得 $\alpha=\beta+\gamma$. 定义映射 $\mathscr{A}:V\to V$ 如下:设 $\alpha\in V$,则令 $\mathscr{A}(\alpha)=\beta$. 映射 $\mathscr{A}$ 称为 V 沿子空间 W 在 U 上的投影变换. 证明

(1) 投影变换 $\mathscr{A}$ 是线性变换;

(2) 线性变换 $\mathscr{B}:V\to V$ 为投影变换当且仅当 $\mathscr{B}$ 为幂等变换,即 $\mathscr{B}$ 满足 $\mathscr{B}^2=\mathscr{B}$;

(3) 线性变换 $\mathscr{B}:V\to V$ 为投影变换当且仅当 $\mathscr{I}-\mathscr{B}$ 为投影变换,其中 $\mathscr{I}$ 是单位映射.

4. 设 $U=\{(x_1,x_2)\in\mathbb{C}^2\mid x_1=x_2\}$,$V_1=\{(x_1,x_2)\in\mathbb{C}^2\mid x_1=0\}$,$V_2=\{(x_1,x_2)\in\mathbb{C}^2\mid x_2=0\}$. 显然,$U,V_1$ 与 V_2 是 $\mathbb{C}^2$ 的子空间,并且 $\mathbb{C}^2=U\oplus V_1=U\oplus V_2$. 设 $\mathscr{A}_1:V\to V$ 与 $\mathscr{A}_2:V\to V$ 分别是 $\mathbb{C}^2$ 沿 V_1 与 V_2 在 U 上的投影变换. 证明,$\mathscr{A}_1\mathscr{A}_2=\mathscr{A}_2$,$\mathscr{A}_2\mathscr{A}_1=\mathscr{A}_1$.

5. 证明线性变换 $\mathscr{A}:V\to V$ 在 V 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 下的方阵相等的充分必要条件是,$\mathscr{A}$ 在基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的方阵与基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 到 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 的过渡矩阵可交换.

6. 设 $\mathscr{A}:V\to V$ 是线性变换. 证明存在 $f(x)\in\mathbb{F}[x]$,使得 $f(\mathscr{A})=0$.

7. 设 A,B,C 与 $D\in\mathbb{F}^{n\times n}$. 定义映射 $\mathscr{P}:\mathbb{F}^{n\times n}\to\mathbb{F}^{n\times n}$ 如下:设 $X\in\mathbb{F}^{n\times n}$,则令

$$\mathscr{P}(X)=AXB+CX+XD.$$

证明 $\mathscr{P}$ 是线性变换,并且当 $C=D=0$ 时,$\mathscr{P}$ 可逆的充分必要条件是,方阵 A 与 B 可逆.

8. 设 $\mathscr{A}:V\to V$ 与 $\mathscr{B}:V\to V$ 是幂等变换,即 $\mathscr{A}^2=\mathscr{A}$,$\mathscr{B}^2=\mathscr{B}$. 证明:

(1) $\operatorname{Im}\mathscr{A}=\operatorname{Im}\mathscr{B}$ 的充分必要条件是 $\mathscr{A}\mathscr{B}=\mathscr{B}$,$\mathscr{B}\mathscr{A}=\mathscr{A}$;

(2) $\operatorname{Ker}\mathscr{A}=\operatorname{Ker}\mathscr{B}$ 的充分必要条件是 $\mathscr{A}\mathscr{B}=\mathscr{A}$,$\mathscr{B}\mathscr{A}=\mathscr{B}$.

9. 设 $\mathscr{A}:V\to V$ 是线性变换,且 k 是正整数. 证明 $\operatorname{Im}\mathscr{A}^k=\operatorname{Im}\mathscr{A}^{2k}$ 的充分必要条件是 $V=\operatorname{Im}\mathscr{A}^k\oplus\operatorname{Ker}\mathscr{A}^k$.

10. 设 $A,B\in\mathbb{F}^{n\times n}$ 相似. 证明 A^{T} 与 B^{T} 相似,A^2 与 B^2 相似,并且当 A,B 可逆时,A^{-1} 与 B^{-1} 也相似.

11. 设 $A\in\mathbb{F}^{n\times n}$ 满足 $A^2=0$. 证明,方阵 A 相似于方阵

$$\begin{pmatrix}0&A_{r\times(n-r)}\\0&0\end{pmatrix}.$$

12. 证明秩为 r 的幂等方阵 A(即 A 满足 $A^2=A$)相似于

$$\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}.$$

13. 设 $A \in \mathbb{C}^{n\times n}$, $f(x), g(x) \in \mathbb{C}[x]$, 且 $\big(f(\lambda), g(\lambda)\big) = 1$. 证明,

$$\operatorname{rank} f(A) + \operatorname{rank} g(A) = n + \operatorname{rank} f(A)g(A).$$

§5.6　不变子空间

给定线性变换 $\mathscr{A}: V \to V$. 容易验证,如果 U 是 V 的子空间,则

$$\mathscr{A}(U) = \big\{\mathscr{A}(\alpha) \,\big|\, \alpha \in U\big\}$$

是 V 的子空间. 记 V 的所有子空间集合记为 S. 定义映射 $\eta_{\mathscr{A}}: S \to S$ 如下: 设 $U \in S$, 则令

$$\eta_{\mathscr{A}}(U) = \mathscr{A}(U).$$

映射 $\eta_{\mathscr{A}}$ 称为**由线性变换 $\mathscr{A}$ 诱导的映射**.

值得关心的是这样的子空间,U 它在映射 $\eta_{\mathscr{A}}$ 下仍映回到 U,即

$$\eta_{\mathscr{A}}(U) = \mathscr{A}(U) \subseteq U.$$

定义 5.6.1　设 $\mathscr{A}: V \to V$ 是线性变换,U 是 V 的子空间. 如果对任意 $\alpha \in U$,均有 $\mathscr{A}(\alpha) \in U$,即 $\mathscr{A}(U) \subseteq U$,则 U 称为线性变换 $\mathscr{A}$ 的**不变子空间**.

显然,对线性空间 V 自身,$\mathscr{A}(V) \subseteq V$. 因此 V 是线性变换 $\mathscr{A}$ 的不变子空间. 同样,对 V 的零子空间 0,$\mathscr{A}(0) = 0$,因此零子空间 0 是线性变换 $\mathscr{A}$ 的子空间. 它们称为 $\mathscr{A}$ 的平凡不变子空间. 除 V 本身与零子空间外,其它不变子空间称为 $\mathscr{A}$ 的非平凡不变子空间.

由于 $\mathscr{A}(V) \subseteq V$,故 $\mathscr{A}\big(\mathscr{A}(V)\big) \subseteq \mathscr{A}(V)$,因此线性变换 $\mathscr{A}$ 的象 $\operatorname{Im}\mathscr{A}$ 是 $\mathscr{A}$ 的不变子空间. 设 $\alpha \in \operatorname{Ker}\mathscr{A}$,则 $\mathscr{A}(\alpha) = 0$,因此 $\mathscr{A}\big(\mathscr{A}(\alpha)\big) = 0$,所以 $\mathscr{A}(\alpha) \in \operatorname{Ker}\mathscr{A}$. 因此线性变换 $\mathscr{A}$ 的核 $\operatorname{Ker}\mathscr{A}$ 也是 A 的不变子空间.

定理 5.6.1　线性变换 $\mathscr{A}$ 的有限多个不变子空间之和是 $\mathscr{A}$ 的不变子空间; $\mathscr{A}$ 的任意多个不变子空间之交是 $\mathscr{A}$ 的不变子空间.

证明　设 $U_1, U_2, \ldots, U_k$ 是 $\mathscr{A}$ 的 k 个不变子空间,且 $\alpha \in U_1 + U_2 + \cdots + U_k$. 则存在 $\alpha_j \in U_j$,使得 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_k$.

由于 U_j 是 $\mathscr{A}$ 的不变子空间,因此 $\mathscr{A}(\alpha_j) \in U_j$,所以

$$\mathscr{A}(\alpha) = \mathscr{A}(\alpha_1) + \mathscr{A}(\alpha_2) + \cdots + \mathscr{A}(\alpha_k) \in U_1 + U_2 + \cdots + U_k.$$

因此 $U_1 + U_2 + \cdots + U_k$ 是 $\mathscr{A}$ 的不变子空间.

设 I 是下标集合,$\big\{U_\nu \,\big|\, \nu \in I\big\}$ 是 V 的子空间集合,其中每个子空间 U_ν 是 $\mathscr{A}$ 的不变子空间. 设 $\alpha \in \bigcap_{\nu\in I} U_\nu$. 则 $\alpha \in U_\nu$, $\nu \in I$. 由于 U_ν 是 $\mathscr{A}$ 的不变子空间,因此 $\mathscr{A}(\alpha) \in U_\nu$,因此 $\mathscr{A}(\alpha) \in U_\nu$, $\nu \in I$. 所以 $\mathscr{A}(\alpha) \in \bigcap_{\nu\in I} U_\nu$. 因此 $\bigcap_{\nu\in I} U_\nu$ 是 $\mathscr{A}$ 的不变子空间. ■

定理 5.6.2 设 $\mathscr{A}$ 是 V 上的线性变换，U 是 $\mathscr{A}$ 的不变子空间. 设 $\{\alpha_1,\alpha_2,\dots,\alpha_k\}$ 是 U 的基. 则 $\mathscr{A}$ 在 V 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_k,\alpha_{k+1},\dots,\alpha_n\}$ 下的方阵是准上三角方阵.

证明 因为 U 是 $\mathscr{A}$ 的不变子空间，所以对 $1\leqslant j\leqslant k$，$\mathscr{A}(\alpha_j)\in U$，由于 $\{\alpha_1,\alpha_2,\dots,\alpha_k\}$ 是 U 的基，因此

$$\begin{aligned}\mathscr{A}(\alpha_1)&=a_{11}\alpha_1+a_{12}\alpha_2+\cdots+a_{1k}\alpha_k,\\ \mathscr{A}(\alpha_2)&=a_{21}\alpha_1+a_{22}\alpha_2+\cdots+a_{2k}\alpha_k,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \mathscr{A}(\alpha_k)&=a_{k1}\alpha_1+a_{k2}\alpha_2+\cdots+a_{kk}\alpha_k.\end{aligned}$$

而当 $k+1\leqslant j\leqslant n$ 时，由于 $\{\alpha_1,\alpha_2,\dots,\alpha_k,\alpha_{k+1},\dots,\alpha_n\}$ 为 V 的基，故

$$\begin{aligned}\mathscr{A}(\alpha_{k+1})&=a_{k+1,1}\alpha_1+a_{k+1,2}\alpha_2+\cdots+a_{k+1,k}\alpha_k+a_{k+1,k+1}\alpha_{k+1}+\cdots+a_{k+1,n}\alpha_n,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \mathscr{A}(\alpha_n)&=a_{n,1}\alpha_1+a_{n,2}\alpha_2+\cdots+a_{n,k}\alpha_k+a_{n,k+1}\alpha_{k+1}+\cdots+a_{n,n}\alpha_n.\end{aligned}$$

因此

$$\mathscr{A}(\alpha_1,\dots,\alpha_n)=(\alpha_1,\dots,\alpha_n)\begin{pmatrix}a_{11}&\cdots&a_{k1}&a_{k+1,1}&\cdots&a_{n1}\\ \vdots&\ddots&\vdots&\vdots&&\vdots\\ a_{1k}&\cdots&a_{kk}&a_{k+1,k}&\cdots&a_{nk}\\ 0&\cdots&0&a_{k+1,k+1}&\cdots&a_{n,k+1}\\ \vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\ 0&\cdots&0&a_{k+1,n}&\cdots&a_{nn}\end{pmatrix}. \tag{5.6.1}$$

■

定理 5.6.3 设线性映射 $\mathscr{A}:V\to V$ 在 V 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的方阵是准上三角的，即式 (5.6.1) 成立. 则由向量 $\alpha_1,\alpha_2,\dots,\alpha_k$ 生成的子空间 U 是 $\mathscr{A}$ 的不变子空间.

证明 由式 (5.6.1)，对于 $1\leqslant j\leqslant k$，

$$\mathscr{A}(\alpha_j)=\sum_{\ell=1}^{k}a_{j\ell}\alpha_\ell.$$

因此，对任意 $\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_k\alpha_k\in U$，均有

$$\mathscr{A}(\alpha)=\mathscr{A}\Big(\sum_{j=1}^{k}a_j\alpha_j\Big)=\sum_{j=1}^{k}a_j\mathscr{A}(\alpha_j)=\sum_{j=1}^{k}\sum_{\ell=1}^{k}a_ja_{j\ell}\alpha_\ell=\sum_{\ell=1}^{k}\Big(\sum_{j=1}^{k}a_ja_{j\ell}\Big)\alpha_j.$$

所以 $\mathscr{A}(\alpha)\in U$，即 U 是 $\mathscr{A}$ 的不变子空间. ■

定义 5.6.2 设 U 是线性变换 $\mathscr{A}$ 的不变子空间. 定义映射 $\mathscr{A}|_U:U\to U$ 如下：设 $\alpha\in U$，则令 $\mathscr{A}|_U(\alpha)=\mathscr{A}(\alpha)$. 映射 $\mathscr{A}|_U$ 称为线性变换 $\mathscr{A}$ 在不变子空间 U 上的**限制**.

由于 U 是线性变换 $\mathscr{A}$ 的不变子空间，因此对 $\alpha\in U$，$\mathscr{A}(\alpha)\in U$，所以映射 $\mathscr{A}|_U$ 是有意义的. 由于 $\mathscr{A}$ 是线性变换，因此 $\mathscr{A}|_U$ 也是线性变换. 定理 5.6.2 说明，如

果 $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ 是 $\mathscr{A}$ 的不变子空间 U 的基，且 $\mathscr{A}$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$ 下的方阵是准上三角的，即

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \tag{5.6.2}$$

其中 A_{11} 是 k 阶方阵，则线性变换 $\mathscr{A}|_U$ 在 U 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ 下的方阵即为 A_{11}.

现在考虑式 (5.6.2) 中 $n-k$ 阶方阵 A_{22} 的意义.

设 U 是线性变换 $\mathscr{A}: V \to V$ 的不变子空间，V/U 是 V 模 U 的商空间，V 中向量 α 所属模 U 的同余类记为 $\overline{\alpha}$. 定义映射 $\mathscr{A}|_{V/U}: V/U \to V/U$ 如下：设 $\overline{\alpha} \in V/U$，则令

$$\mathscr{A}|_{V/U}(\overline{\alpha}) = \overline{\mathscr{A}(\alpha)}.$$

注意，如果向量 $\beta \in \overline{\alpha}$，则存在向量 $\gamma \in U$，使得 $\beta = \alpha + \gamma$. 因此 $\mathscr{A}(\beta) = \mathscr{A}(\alpha) + \mathscr{A}(\gamma)$. 由于 U 是 $\mathscr{A}$ 的不变子空间，所以 $\mathscr{A}(\gamma) \in U$. 这表明，$\mathscr{A}(\beta) \equiv \mathscr{A}(\alpha) \pmod U$，也即 $\overline{\mathscr{A}(\beta)} = \overline{\mathscr{A}(\alpha)}$. 因此映射 $\mathscr{A}|_{V/U}$ 的定义并不依赖于同余类 $\overline{\alpha}$ 的代表元的选取. 也就是说，映射 $\mathscr{A}|_{V/U}$ 是有确切定义的.

映射 $\mathscr{A}|_{V/U}$ 称为由线性变换 $\mathscr{A}$ 诱导的**商变换**.

定理 5.6.4　记号同定理 5.6.2，且式 (5.6.2) 成立. 商变换 $\mathscr{A}|_{V/U}$ 是线性变换，并且在基 $\{\overline{\alpha}_{k+1}, \overline{\alpha}_{k+2}, \ldots, \overline{\alpha}_n\}$ 下的方阵即为 A_{22}.

证明　设 $\overline{\alpha}, \overline{\beta} \in V/U$，则 $\overline{\alpha} + \overline{\beta} = \overline{\alpha + \beta}$. 因此，

$$\begin{aligned}\mathscr{A}|_{V/U}(\overline{\alpha} + \overline{\beta}) &= \mathscr{A}|_{V/U}(\overline{\alpha + \beta}) = \overline{\mathscr{A}(\alpha + \beta)} = \overline{\mathscr{A}(\alpha)} + \overline{\mathscr{A}(\beta)} \\ &= \mathscr{A}|_{V/U}(\alpha) + \mathscr{A}|_{V/U}(\beta).\end{aligned}$$

设 $\lambda \in \mathbb{F}$，$\overline{\alpha} \in V/U$，则 $\lambda\overline{\alpha} = \overline{\lambda\alpha}$. 所以

$$\mathscr{A}|_{V/U}(\lambda\overline{\alpha}) = \mathscr{A}|_{V/U}(\overline{\lambda\alpha}) = \overline{\mathscr{A}(\lambda\alpha)} = \overline{\lambda\mathscr{A}(\alpha)} = \lambda\overline{\mathscr{A}(\alpha)} = \lambda\mathscr{A}|_{V/U}(\overline{\alpha}).$$

因此 $\mathscr{A}|_{V/U}$ 确实是线性变换.

设 $\lambda_{k+1}, \lambda_{k+2}, \ldots, \lambda_n \in \mathbb{F}$，使得 $\lambda_{k+1}\overline{\alpha}_{k+1} + \lambda_{k+2}\overline{\alpha}_{k+2} + \cdots + \lambda_n\overline{\alpha}_n = 0$. 则

$$\overline{\lambda_{k+1}\alpha_{k+1} + \lambda_{k+2}\alpha_{k+2} + \cdots + \lambda_n\alpha_n} = 0.$$

因此，

$$\lambda_{k+1}\alpha_{k+1} + \lambda_{k+2}\alpha_{k+2} + \cdots + \lambda_n\alpha_n \in U.$$

由于 $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ 是 U 的基，所以

$$\lambda_{k+1}\alpha_{k+1} + \lambda_{k+2}\alpha_{k+2} + \cdots + \lambda_n\alpha_n = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k,$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$. 因此

$$(-\lambda_1)\alpha_1 + (-\lambda_2)\alpha_2 + \cdots + (-\lambda_k)\alpha_k + \lambda_{k+1}\alpha_{k+1} + \lambda_{k+2}\alpha_{k+2} + \cdots + \lambda_n\alpha_n = 0.$$

由于 $\{\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$ 是 V 的基，所以 $\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_n = 0$. 因此 $\overline{\alpha}_{k+1}, \overline{\alpha}_{k+2}, \ldots, \overline{\alpha}_n \in V/U$ 线性无关. 由推论推论 4.8.1，$\dim V/U = n-k$，故 $\{\overline{\alpha}_{k+1}, \overline{\alpha}_{k+2}, \ldots, \overline{\alpha}_n\}$ 是 V/U 的基. 现在记式 (5.6.2) 中的方阵为 $A = (a_{ij})$，则对任意 $k+1 \leqslant j \leqslant n$，

$$\mathscr{A}(\alpha_j) = \sum_{\ell=1}^{n} a_{j\ell}\alpha_\ell.$$

因此

$$\mathscr{A}\big|_{V/U}(\overline{\alpha}_j) = \overline{\mathscr{A}(\alpha_j)} = \sum_{\ell=1}^{n} a_{j\ell}\overline{\alpha}_\ell = \sum_{\ell=k+1}^{n} a_{j\ell}\overline{\alpha}_\ell,$$

最后一个等号成立是因为，当 $1 \leqslant \ell \leqslant k$ 时，$\alpha_\ell \in U$，从而 $\overline{\alpha}_\ell = 0$. 即

$$\mathscr{A}\big|_{V/U}(\overline{\alpha}_{k+1}, \overline{\alpha}_{k+2}, \ldots, \overline{\alpha}_n) = (\overline{\alpha}_{k+1}, \overline{\alpha}_{k+2}, \ldots, \overline{\alpha}_n)A_{22}.$$ ■

应当指出，线性变换 $\mathscr{A}: V \to V$ 的不变子空间 U 的补 W（即 W 满足 $V = U \oplus W$）不一定是 $\mathscr{A}$ 的不变子空间.

定理 5.6.5 设 U 与 W 是线性变换 $\mathscr{A}: V \to V$ 的不变子空间，$U \oplus W = V$，且 $\{\alpha_1, \ldots, \alpha_k\}$ 与 $\{\alpha_{k+1}, \ldots, \alpha_n\}$ 分别是 U 与 W 的基，则 $\mathscr{A}$ 在 V 的基 $\{\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$ 下的方阵是准对角的，即有

$$\mathscr{A}(\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n) = (\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n)\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}, \tag{5.6.3}$$

其中 A_{11} 与 A_{22} 分别是 k 阶与 $n-k$ 阶方阵.

反之，设 $\{\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_n\}$ 是 V 的基，且 $\mathscr{A}$ 在此基下的方阵是准对角的，即式 (5.6.3) 成立，则 V 中分别由 $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ 与 $\{\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_n\}$ 生成的子空间 U 与 W 是 $\mathscr{A}$ 的不变子空间，并且

$$\mathscr{A}\big|_U(\alpha_1, \alpha_2, \ldots, \alpha_k) = (\alpha_1, \alpha_2, \ldots, \alpha_k)A_{11}, \tag{5.6.4}$$

$$\mathscr{A}\big|_U(\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_n) = (\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_n)A_{22}. \tag{5.6.5}$$

证明 证明方法与定理 5.6.2、定理 5.6.3 相同. 请读者自己完成. ■

在定理 5.6.5 的条件下，线性变换 $\mathscr{A}$ 记为

$$\mathscr{A} = \mathscr{A}\big|_U \oplus \mathscr{A}\big|_W,$$

并称线性变换 $\mathscr{A}$ 分解为 $\mathscr{A}\big|_U$ 与 $\mathscr{A}\big|_W$ 的直和.

例 5.6.1 设线性变换 $\mathscr{A}: V \to V$ 与 $\mathscr{B}: V \to V$ 可交换，即 $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$. 证明，线性变换 $\mathscr{B}$ 的象 $\operatorname{Im}\mathscr{B}$ 与核 $\operatorname{Ker}\mathscr{B}$ 是线性变换 $\mathscr{A}$ 的不变子空间.

证明 设 $\alpha \in \operatorname{Im}\mathscr{B}$，则 $\alpha = \mathscr{B}(\beta)$，$\beta \in V$. 因此 $\mathscr{A}(\alpha) = \mathscr{A}\mathscr{B}(\beta)$. 因为 $\mathscr{A}$ 与 $\mathscr{B}$ 可交换，所以

$$\mathscr{A}(\alpha) = \mathscr{B}\mathscr{A}(\beta) = \mathscr{B}\big(\mathscr{A}(\beta)\big) \in \operatorname{Im}\mathscr{B},$$

即 $\operatorname{Im}\mathscr{B}$ 是 $\mathscr{A}$ 的不变子空间.

设 $\alpha \in \operatorname{Ker}\mathscr{B}$，则 $\mathscr{B}(\alpha) = 0$. 因此

$$\mathscr{B}\big(\mathscr{A}(\alpha)\big) = \mathscr{B}\mathscr{A}(\alpha) = \mathscr{A}\mathscr{B}(\alpha) = \mathscr{A}\big(\mathscr{B}(\alpha)\big) = \mathscr{A}(0) = 0,$$

故 $\mathscr{A}(\alpha) \in \operatorname{Ker}\mathscr{B}$. 所以 $\operatorname{Ker}\mathscr{B}$ 是 $\mathscr{A}$ 的不变子空间. ■

例 5.6.2 设 $\mathscr{D}: \mathbb{F}_n[x] \to \mathbb{F}_n[x]$ 是微商变换，$\mathbb{F}_n[x]$ 中由向量 $1, x, \ldots, x^{k-1}$ 生成的子空间记为 $\mathbb{F}_k[x]$，其中正整数 $k \geqslant 1$. 证明 $\mathbb{F}_k[x]$ 是 $\mathscr{D}$ 的不变子空间，并求 $\mathscr{D}\big|_{\mathbb{F}_k[x]}$ 与

$\mathscr{D}\big|_{\mathbb{F}_n[x]/\mathbb{F}_k[x]}$ 分别在 $\mathbb{F}_k$ 的基 $\{1, x, \ldots, x^{k-1}\}$ 与 $\mathbb{F}_n[x]/\mathbb{F}_k[x]$ 的基 $\{\overline{x^k}, \overline{x^{k+1}}, \ldots, \overline{x^{n-1}}\}$ 下的方阵.

证明　设 $f(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} \in \mathbb{F}_k[x]$,则

$$\mathscr{D}\big(f(x)\big) = f'(x) = a_1 + 2a_2 x + \cdots + (k-1)a_{k-1}x^{k-2} \in \mathbb{F}_k[x].$$

因此 $\mathbb{F}_k[x]$ 是 $\mathscr{D}$ 的不变子空间. 由于 $\mathscr{D}\big|_{\mathbb{F}_k[x]}(x^j) = \mathscr{D}(x^j) = jx^{j-1}, j = 0, \ldots, k-1$,故

$$\mathscr{D}\big|_{\mathbb{F}_k[x]}(1, x, \ldots, x^{k-1}) = (1, x, \ldots, x^{k-1})\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & k-1 \\ & & & 0 \end{pmatrix}.$$

所以 $\mathscr{D}\big|_{\mathbb{F}_k[x]}$ 在基 $(1, x, \ldots, x^{k-1})$ 下的方阵为

$$\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & k-1 \\ & & & 0 \end{pmatrix}_{k\times k}.$$

当 $k \leqslant j \leqslant n$ 时,$\mathscr{D}\big|_{\mathbb{F}_n[x]/\mathbb{F}_k[x]}(\overline{x^j}) = \overline{D(x^j)} = \overline{jx^{j-1}} = j\overline{x^{j-1}}$. 所以

$$\mathscr{D}\big|_{\mathbb{F}_n[x]/\mathbb{F}_k[x]}(\overline{x^k}, \overline{x^{k+1}}, \ldots, \overline{x^{n-1}}) = (\overline{x^k}, \overline{x^{k+1}}, \ldots, \overline{x^{n-1}})\begin{pmatrix} 0 & k+1 & & \\ & 0 & \ddots & \\ & & \ddots & n-1 \\ & & & 0 \end{pmatrix}.$$

因此 $\mathscr{D}\big|_{\mathbb{F}_n[x]/\mathbb{F}_k[x]}$ 在基 $\{\overline{x^k}, \overline{x^{k+1}}, \ldots, \overline{x^{n-1}}\}$ 下的方阵为

$$\begin{pmatrix} 0 & k+1 & & \\ & 0 & \ddots & \\ & & \ddots & n-1 \\ & & & 0 \end{pmatrix}_{(n-k)\times(n-k)}.$$ ■

注　$\mathbb{F}_n[x]$ 中由向量 $x^k, x^{k+1}, \ldots, x^{n-1}$ 生产的子空间记为 U. 显然 $\mathbb{F}_n[x] = \mathbb{F}_k[x] \oplus U$. 由于 $x^k \in U$,但 $\mathscr{D}(x^k) = kx^{k-1} \notin U$,因此 $\mathbb{F}_k[x]$ 的补 U 不是 $\mathscr{D}$ 的不变子空间.

习　题　5.6

1. 设 $\mathscr{A}: V \to V$ 与 $\mathscr{B}: V \to V$ 是线性变换,U 是 $\mathscr{A}$ 与 $\mathscr{B}$ 的不变子空间. 证明 U 是 $\mathscr{A} + \mathscr{B}$ 与 $\mathscr{A}\mathscr{B}$ 的不变子空间. 如果 $\mathscr{A}$ 可逆,则 U 也是 $\mathscr{A}^{-1}$ 的不变子空间.

2. 设 U 是线性变换 $\mathscr{A}: V \to V$ 的不变子空间. 证明 $\widetilde{U} = \{\alpha \in V \mid \mathscr{A}(\alpha) \in U\}$ 也是 $\mathscr{A}$ 的不变子空间.

3. 设 $\mathscr{A}: V \to V$ 是线性变换,$0 \neq \alpha \in V$. 证明,V 中由向量 $\alpha, \mathscr{A}(\alpha), \ldots, \mathscr{A}^k(\alpha), \ldots$ 生成的子空间 U 是 $\mathscr{A}$ 的不变子空间. 设 $\dim U = r$,证明 $\{\alpha, \mathscr{A}(\alpha), \ldots, \mathscr{A}^{r-1}(\alpha)\}$ 是 U 的基. 求 $\mathscr{A}\big|_U$ 在

这组基下的方阵.

4. 设 $\mathscr{A}: V \to V$ 是线性变换,$\lambda_0 \in \mathbb{F}$. 记

$$V_{\lambda_0}(\mathscr{A}) = \left\{\alpha \in V \,\middle|\, \text{存在某个正整数 } k, \text{使得 } (\mathscr{A} - \lambda_0 \mathscr{I})^k(\alpha) = 0\right\}.$$

证明 $V_{\lambda_0}(\mathscr{A})$ 是 V 的子空间,而且是 $\mathscr{A}$ 的不变子空间.

5. 设 V 是 n 维复线性空间,$\mathscr{A}: V \to V$ 是线性变换. 证明 V 的每个子空间都是 $\mathscr{A}$ 的不变子空间的充分必要条件是 $\mathscr{A}$ 为纯量变换,即 $\mathscr{A} = \lambda \mathscr{I}$,其中 $\lambda \in \mathbb{C}$.

6. 设 V 是 2 维复线性空间,线性变换 $\mathscr{A}: V \to V$ 在基 $\{\alpha_1, \alpha_2\}$ 下的方阵为

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

求线性变换 $\mathscr{A}$ 的所有不变子空间.

7. 设 V 是区间 $[0,1]$ 上所有连续实函数构成的实线性空间,U 是 V 中所有偶函数构成的子空间,W 是 V 中所有奇函数构成的子空间. 设 $\mathscr{A}: V \to V$ 是积分变换,即对任意 $f(x) \in V$,

$$\mathscr{A}\big(f(x)\big) = \int_0^x f(t)\,\mathrm{d}t.$$

U 与 W 是否是 $\mathscr{A}$ 的不变子空间?

8. 设 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 $\mathbb{F}$ 中两两不等的 n 个数,线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 $\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. 求 $\mathscr{A}$ 的不变子空间的个数.

§5.7 特征值与特征向量

在线性变换 $\mathscr{A}: V \to V$ 的所有不变子空间中,一维不变子空间是极为重要的. 设 U 是线性变换 $\mathscr{A}$ 的一维不变子空间. 显然 U 中任意一个非零向量 α 都可以构成 U 的基 $\{\alpha\}$. 由于 U 是 $\mathscr{A}$ 的不变子空间,因此 $\mathscr{A}(\alpha) \in U$. 所以 $\mathscr{A}(\alpha) = \lambda\alpha, \lambda \in \mathbb{F}$. 这就引出下面的定义.

定义 5.7.1 设 $\mathscr{A}: V \to V$ 是线性变换. 如果存在非零向量 $\alpha \in V$,以及纯量 $\lambda \in \mathbb{F}$,使得

$$\mathscr{A}(\alpha) = \lambda\alpha,$$

则 λ 称为线性变换 $\mathscr{A}$ 的**特征值**,α 称为线性变换 $\mathscr{A}$ 的属于特征值 λ 的**特征向量**.

定理 5.7.1 设线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A,向量 $\alpha \in V$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 x. 则 α 是线性变换 $\mathscr{A}$ 的属于特征值 λ 的特征向量之充分必要条件为 x 是齐次线性方程组 $(\lambda I_{(n)} - A)x = 0$ 的非零解.

证明 必要性 设 α 是线性变换 $\mathscr{A}$ 的属于特征值 λ 的特征向量,则 $\mathscr{A}(\alpha) = \lambda\alpha$.

记 $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$,即 $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)x$. 由于 $\mathscr{A}$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A,所以

$$\mathscr{A}(\alpha) = (\alpha_1, \alpha_2, \ldots, \alpha_n)Ax.$$

由于 $\mathscr{A}(\alpha) = \lambda\alpha$,因此 $Ax = \lambda x$. 即 $(\lambda I_{(n)} - A)x = 0$. 由于 α 是 $\mathscr{A}$ 的特征向量,故

$\alpha \neq 0$. 从而 $x \neq 0$. 因此 x 是齐次线性方程组 $(\lambda I_{(n)} - A)x = 0$ 的非零解.

充分性 设 x 是 $(\lambda I_{(n)} - A)x = 0$ 的非零解,则 $Ax = \lambda x$. 因此,

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)Ax = \lambda(\alpha_1, \alpha_2, \ldots, \alpha_n)x.$$

所以 $A(\alpha) = \lambda\alpha$. 由于 x 是非零的,因此 α 是非零向量. 所以 α 是 $\mathscr{A}$ 的属于特征值 λ 的特征向量. ■

定理 5.7.1 说明,求线性变换 $\mathscr{A}$ 的属于特征值 λ 的特征向量 α 的问题等价于求齐次线性方程组 $(\lambda I_{(n)} - A)x = 0$ 的非零解问题. 显然齐次线性方程组 $(\lambda I_{(n)} - A)x = 0$ 有非零解的充分必要条件是 $\det(\lambda I_{(n)} - A) = 0$. 这就引出下面的定义.

定义 5.7.2 设 $A \in \mathbb{F}^{n\times n}$,则多项式 $\varphi(\lambda) = \det(\lambda I_{(n)} - A)$ 称为方阵 A 的**特征多项式**. 特征多项式 $\varphi(\lambda)$ 的根称为方阵 A 的**特征值**. 而方程组 $Ax = \lambda x$ 的非零解 x 称为方阵 A 的属于特征值 λ 的**特征向量**.

按照定义 5.7.2,定理 5.7.1 可以改述为

定理 5.7.2 设线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A,向量 $\alpha \in V$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 x. 则

(1) 当且仅当 λ 是方阵 A 的特征值时,λ 是线性变换 $\mathscr{A}$ 的特征值;

(2) 当且仅当 x 是方阵 A 的属于特征值,λ 的特征向量时 α 是线性变换 $\mathscr{A}$ 的属于特征值 λ 的特征向量.

由定理 5.7.2 立即得到

定理 5.7.3 设数域 $\mathbb{F}$ 上 n 维线性空间 V 到自身的线性变换 $\mathscr{A}$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A. 则线性变换 $\mathscr{A}$ 具有一维不变子空间 U 的必要与充分条件是,方阵 A 的特征多项式 $\varphi(\lambda)$ 在数域 $\mathbb{F}$ 中有根 λ_0.

证明 证明不难,请读者自己完成. ■

我们知道,n 次复系数多项式一定有 n 个复根. 因此,对 n 维复线性空间 V,线性变换 $\mathscr{A}: V \to V$ 的一维不变子空间恒存在. 但是,由于 n 次实系数多项式不一定具有实根(尽管它的复根总存在),所以对 n 维实线性空间,线性变换 $\mathscr{A}: V \to V$ 的一维不变子空间并不是总存在的. 所以在讨论一维不变子空间时,一定要注意线性变换 V 的基域 $\mathbb{F}$.

现在讨论方阵 A 的特征多项式 $\varphi(\lambda) = \det(\lambda I_{(n)} - A)$ 的性质.

记 n 阶方阵 $A = (a_{ij})$,

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$$

表示取自方阵 A 的第 $i_1, i_2, \ldots, i_k$ 行与第 $j_1, j_2, \ldots, j_k$ 列组成的 k 阶子式,其中 $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$,$1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n$. 方阵 A 的特征多项式 $\varphi(\lambda)$ 的 n 个特征值分别记为 $\lambda_1, \lambda_2, \ldots, \lambda_n$.

性质 5.7.1

$$\varphi(\lambda) = \lambda^n - (a_{11} + \cdots + a_{nn})\lambda^{n-1} + \sum_{1\leqslant i_1<i_2\leqslant n} A\begin{pmatrix} i_1 & i_2 \\ i_1 & i_2 \end{pmatrix}\lambda^{n-2} + \cdots$$
$$+ (-1)^k \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}\lambda^{n-k} + \cdots + (-1)^n \det A.$$

证明 这是 §2.5 习题 6. 也可如下证明之. 容易看出,特征多项式 $\varphi(\lambda)$ 是 λ 的 n 次多项式,并且 n 次项系数为 1. 因此可设

$$\varphi(\lambda) = \varphi(0) + \frac{\varphi'(0)}{1!}\lambda + \frac{\varphi''(0)}{2!}\lambda^2 + \cdots + \frac{\varphi^{(n-1)}(0)}{(n-1)!}\lambda^{n-1} + \lambda^n.$$

然后对行列式 $\det(\lambda I_{(n)} - A)$ 求关于 λ 的各阶导数,从而求出

$$\frac{\varphi^{(k)}(0)}{k!},$$

$k = 0, 1, 2, \ldots, n-1$,即可证明性质 5.7.1. ■

性质 5.7.2 设 $\sigma_1(x_1, x_2, \ldots, x_n), \ldots, \sigma_n(x_1, x_2, \ldots, x_n)$ 是关于 $x_1, x_2, \ldots, x_n$ 的基本对称多项式. 则对任意 $k = 1, 2, \ldots, n$,有

$$\sigma_k(\lambda_1, \lambda_2, \ldots, \lambda_n) = \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}.$$

证明 由于 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 n 次多项式 $\varphi(\lambda)$ 的 n 个根,因此由 Viéta 定理和性质 5.7.1,即得性质 5.7.2. ■

由性质 5.7.2,有

$$\mathrm{Tr}\, A = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \qquad \det A = \lambda_1\lambda_2\cdots\lambda_n.$$

性质 5.7.3 设 A 是 n 阶复方阵. 则方阵 A 可逆的充分必要条件是方阵 A 的特征值全不为零.

证明 这是 $\det A = \lambda_1\lambda_2\cdots\lambda_n$ 的直接推论. ■

性质 5.7.4 设方阵 A 是准上三角的,即设

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

其中 A_{11} 与 A_{22} 是子方阵. 则方阵 A 的特征多项式等于 A_{11} 与 A_{22} 的特征多项式的乘积.

证明 由定义,方阵 A 的特征多项式 $\varphi(\lambda)$ 为

$$\varphi(\lambda) = \det(\lambda I_{(n)} - A) = \det\begin{pmatrix} \lambda I_{(k)} - A_{11} & -A_{12} \\ 0 & \lambda I_{(n-k)} - A_{22} \end{pmatrix}$$
$$= \det(\lambda I_{(k)} - A_{11})\det(\lambda I_{(n-k)} - A_{22}).$$

其中 k 是子方阵 A_{11} 的阶数. 显然,$\det(\lambda I_{(k)} - A_{11})$ 与 $\det(\lambda I_{(n-k)} - A_{22})$ 分别是子方阵 A_{11} 与 A_{22} 的特征多项式. ■

性质 5.7.5 设 A 是上三角方阵,则方阵 A 的对角元是方阵 A 的特征值.

性质 5.7.6 相似的方阵具有相同的特征多项式,从而具有相同的特征值.

证明　设方阵 A 与 B 相似,则存在可逆方阵 P,使得 $B=P^{-1}AP$. 因此,

$$\begin{aligned}\det(\lambda I_{(n)}-B)&=\det(\lambda I_{(n)}-P^{-1}AP)=\det P^{-1}(\lambda I_{(n)}-A)P\\&=\det P^{-1}\det(\lambda I_{(n)}-A)\det P=\det(\lambda I_{(n)}-A).\end{aligned}$$ ■

性质 5.7.6 表明,方阵的特征值是方阵在相似下的不变量. 但应当指出,它并不是方阵在相似下的全系不变量,即特征值相同的方阵并不一定相似. 例如方阵

$$A=\begin{pmatrix}0&0\\0&0\end{pmatrix}\text{ 与 }B=\begin{pmatrix}0&1\\0&0\end{pmatrix}$$

的特征值相同,但方阵 A 与 B 并不相似.

关于方阵的特征多项式,有下面的重要定理.

定理 5.7.4(Cayley-Hamilton 定理) 设 n 阶方阵 A 的特征多项式为

$$\varphi(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda_1+a_0,$$

则

$$\varphi(A)=A^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0I_{(n)}=0.$$

证明　设方阵 $\lambda I_{(n)}-A$ 的附属方阵为 B. 显然方阵 B 的元素都是关于 λ 的多项式,且其次数不超过 $n-1$. 因此可设 $B=\big(b_{ij}(\lambda)\big)$,其中

$$b_{ij}(\lambda)=b_{ij}^{(n-1)}\lambda^{n-1}+b_{ij}^{(n-2)}\lambda^{n-2}+\cdots+b_{ij}^{(1)}\lambda+b_{ij}^{(0)}.$$

所以方阵 B 可以表为

$$\begin{aligned}B&=\lambda^{n-1}\big(b_{ij}^{(n-1)}\big)+\lambda^{n-2}\big(b_{ij}^{(n-2)}\big)+\cdots+\lambda\big(b_{ij}^{(1)}\big)+\big(b_{ij}^{(0)}\big)\\&=\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\cdots+\lambda B_1+B_0,\end{aligned}$$

其中 $B_k=\big(b_{ij}^{(k)}\big),k=0,1,2,\ldots,n-1$. 所以

$$(\lambda I_{(n)}-A)B=\lambda^nB_{n-1}+\lambda^{n-1}(B_{n-2}-AB_{n-1})+\cdots+\lambda(B_0-AB_1)-AB_0.$$

另一方面,

$$(\lambda I_{(n)}-A)B=\big(\det(\lambda I_{(n)}-A)\big)I_{(n)}=\varphi(\lambda)I_{(n)}.$$

比较上面两式的两端系数,即得

$$\begin{aligned}B_{n-1}&=I_{(n)},\\B_{n-2}-AB_{n-1}&=a_{n-1}I_{(n)},\\&\cdots\cdots\\B_0-AB_1&=a_1I_{(n)},\\-AB_0&=a_0I_{(n)}.\end{aligned}$$

上述各式依次乘以方阵 $A^n,A^{n-1},\ldots,A,I_{(n)}$,并相加,即得到

$$\varphi(A)=A^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0I_{(n)}=0.$$ ■

定义 5.7.3　设方阵 $A\in\mathbb{F}^{n\times n}$,且非零多项式 $f(\lambda)\in\mathbb{F}[\lambda]$. 如果 $f(A)=0$,则 $f(\lambda)$ 称为方阵 A 的一个**化零多项式**.

利用定义 5.7.3，Cayley-Hamilton 定理可以叙述为：

方阵 A 的特征多项式是方阵 A 的一个化零多项式.

定义 5.7.4 方阵 A 的所有化零多项式中次数最小的首一多项式称为方阵 A 的**最小多项式**，记为 $d_A(\lambda)$，或 $d(\lambda)$.

考虑正整数集合

$$M = \left\{\deg f(\lambda) \,\middle|\, f(\lambda)\ \text{为方阵}\ A\ \text{的化零多项式}\right\}.$$

由于方阵 A 的特征多项式 $\varphi(\lambda)$ 是方阵 A 的一个化零多项式，因此集合 $M \neq \varnothing$. 所以集合 M 中存在最小的正整数 m，而且 m 是方阵 A 的某个化零多项式 $g(\lambda)$ 的次数. 于是 $d(\lambda) = a^{-1}g(\lambda)$ 是方阵 A 的最小多项式，其中 a 是多项式 $g(\lambda)$ 的首项系数. 这表明，任意一个方阵 A 的最小多项式均存在.

关于最小多项式，以下性质成立.

性质 5.7.7 方阵 A 的最小多项式 $d(\lambda)$ 整除方阵 A 的任一化零多项式 $f(\lambda)$. 特别 $d(\lambda)$ 整除方阵 A 的特征多项式 $\varphi(\lambda)$.

证明 由 Euclid 长除法，存在 $q(\lambda), r(\lambda) \in \mathbb{F}[\lambda]$，使得 $f(\lambda) = q(\lambda)d(\lambda) + r(\lambda)$，其中 $\deg r(\lambda) < \deg d(\lambda)$.

如果 $r(\lambda) \neq 0$，则 $f(A) = q(A)d(A) + r(A)$. 由于 $f(\lambda)$ 与 $d(\lambda)$ 均是方阵 A 的化零多项式，因此 $f(A) = d(A) = 0$. 所以 $r(A) = 0$. 这说明，$f(\lambda)$ 是方阵 A 的化零多项式，且 $\deg r(\lambda) < \deg d(\lambda)$. 这与 $d(\lambda)$ 关于次数的最小性相矛盾. 因此 $r(\lambda) = 0$，即 $d(\lambda)$ 整除方阵 A 的化零多项式 $f(\lambda)$. ■

性质 5.7.8 方阵 A 的最小多项式 $d(\lambda)$ 存在且唯一.

证明 前面已经证明方阵 A 的最小多项式的存在性. 现在证明唯一性.

设 $d_1(\lambda), d_2(\lambda)$ 是方阵 A 的最小多项式，则 $d_1(\lambda), d_2(\lambda)$ 是方阵 A 的化零多项式. 由性质 5.7.7，$d_1(\lambda) \mid d_2(\lambda)$，$d_2(\lambda) \mid d_1(\lambda)$. 而 $d_1(\lambda), d_2(\lambda)$ 是首一多项式，所以 $d_1(\lambda) = d_2(\lambda)$. ■

性质 5.7.9 相似的方阵具有相同的最小多项式.

证明 设方阵 A 与 B 相似，即存在可逆方阵 P，使得 $B = P^{-1}AP$. 设方阵 A 的最小多项式为

$$d_A(\lambda) = a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k,$$

则
$$d_A(A) = a_0I + a_1A + \cdots + a_{k-1}A^{k-1} + A^k = 0.$$

因此，
$$\begin{aligned} d_A(B) &= a_0I + a_1B + \cdots + a_{k-1}B^{k-1} + B^k \\ &= a_0P^{-1}P + a_1P^{-1}AP + \cdots + a_{k-1}P^{-1}A^{k-1}P + P^{-1}A^kP \\ &= P^{-1}\left(a_0I + a_1A + \cdots + a_{k-1}A^{k-1} + A^k\right)P \\ &= P^{-1}d_A(A)P = 0. \end{aligned}$$

所以 $d_A(\lambda)$ 是方阵 B 的化零多项式.

由性质 5.7.7,方阵 B 的最小多项式 $d_B(\lambda) \mid d_A(\lambda)$. 同样有,$d_A(\lambda) \mid d_B(\lambda)$. 由于 $d_A(\lambda)$ 与 $d_B(\lambda)$ 都是首一多项式,故 $d_A(\lambda) = d_B(\lambda)$. ■

性质 5.7.9 说明,方阵的最小多项式是方阵在相似下的不变量. 应当指出,方阵的最小多项式并不是方阵在相似下的全系不变量,即具有相同的最小多项式的方阵并不一定相似. 在习题中有这方面的例子.

定理 5.7.5　纯量 $x \in \mathbb{F}$ 为方阵 A 的最小多项式 $d(\lambda)$ 的根的充分必要条件是,c 是方阵 A 的特征值.

证明　设 $d(c) = 0$,则 $d(\lambda) = (\lambda - c)q(\lambda)$,其中 $\deg q(\lambda) < \deg d(\lambda)$.

由 $d(\lambda)$ 关于次数的最小性,$q(A) \neq 0$. 因此必有列向量 x,使得 $q(A)x \neq 0$. 取 $y = q(A)x$. 因为 $d(A) = (A - cI_{(n)})q(A) = 0$,所以

$$(A - cI_{(n)})y = (A - cI_{(n)})q(A)x = d(A)x = 0.$$

因此 $Ay = cy$. 这表明,c 方阵 A 的特征值.

反之,设 c 是方阵 A 的特征值,则存在列向量 $x \neq 0$,使得 $Ax = cx$. 于是 $A^j x = c^j x$. 因此,如果设 $d(x) = a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k$,则

$$\begin{aligned} d(A)x &= a_0 I_{(n)}x + a_1 Ax + \cdots + a_{k-1}A^{k-1}x + A^k x \\ &= a_0 x + a_1 cx + \cdots + a_{k-1}c^{k-1}x + c^k x \\ &= d(c)x. \end{aligned}$$

由于 $d(A) = 0$,故 $d(c)x = 0$. 但 $x \neq 0$,因此 $d(c) = 0$. 这表明,c 是方阵 A 的最小多项式的根. ■

现在转到线性变换 $\mathscr{A}: V \to V$. 我们知道,同一个线性变换 $\mathscr{A}$ 在不同基下的方阵是相似的. 而相似方阵的特征多项式相同. 于是可以引进如下的定义.

定义 5.7.5　设线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A,则方阵 A 的特征多项式称为线性变换 $\mathscr{A}$ 的特征多项式.

关于线性变换,相应的 Cayley-Hamilton 定理成立,即有

定理 5.7.6　设 $\varphi(\lambda)$ 是线性变换 $\mathscr{A}$ 的特征多项式,则 $\varphi(\mathscr{A}) = \mathscr{O}$.

定义 5.7.6　设 $\mathscr{A}: V \to V$ 是线性变换,且非零多项式 $f(\lambda) \in \mathbb{F}[\lambda]$. 如果 $f(\mathscr{A}) = 0$,则 $f((\lambda))$ 称为线性变换 $\mathscr{A}$ 的化零多项式.

于是定理 5.7.6 可以叙述为:

线性变换 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 是线性变换 $\mathscr{A}$ 的一个化零多项式.

定义 5.7.7　线性变换 $\mathscr{A}$ 的所有化零多项式中次数最小的首一多项式称为线性变换 $\mathscr{A}$ 的**最小多项式**.

和方阵的情形相仿,容易证明,线性变换 $\mathscr{A}$ 的最小多项式整除线性变换 $\mathscr{A}$ 的化零多项式. 而且线性变换 $\mathscr{A}$ 的最小多项式存在且唯一.

定理 5.7.7　设线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A,则

线性变换 $\mathscr{A}$ 的最小多项式等于方阵 A 的最小多项式.

证明 考虑线性变换 $d(\mathscr{A})\colon V\to V$. 显然, $d(\mathscr{A})$ 在基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的方阵为 $d(A)$. 由于 $d(\lambda)$ 是方阵 A 的最小多项式,因此 $d(A)=0$. 于是线性变换 $d(\mathscr{A})$ 是零变换. 所以 $d(\lambda)$ 是线性变换 A 的化零多项式. 因此线性变换 $\mathscr{A}$ 的最小多项式 $\widetilde{d}(\lambda)\mid d(\lambda)$.

反之,线性变换 $\widetilde{d}(\mathscr{A})\colon V\to V$ 在基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 下的方阵为 $\widetilde{d}(A)$. 由于 $\widetilde{d}(\mathscr{A})$ 是零变换,所以 $\widetilde{d}(A)=0$. 因此 $\widetilde{d}(\lambda)$ 是方阵 A 的化零多项式. 这表明,方阵 A 的最小多项式 $d(\lambda)\mid\widetilde{d}(\lambda)$.

由于 $\widetilde{d}(\lambda)$ 与 $d(\lambda)$ 都是首一多项式,所以 $\widetilde{d}(\lambda)=d(\lambda)$. ■

和定理 5.7.5 相平行的结论是

定理 5.7.8 纯量 $x\in\mathbb{F}$ 为线性变换 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 的根的充分必要条件是, c 是线性变换 $\mathscr{A}$ 的特征值.

例 5.7.1 设 A 与 B 是 n 阶复方阵,则方阵 AB 与 BA 的特征多项式相同.

证明 只需证明,

$$\det(\lambda I_{(n)}-AB)=\det(\lambda I_{(n)}-BA). \tag{5.7.1}$$

这是例 3.3.3. 这里另给一个证明.

设 $\operatorname{rank}A=r$,并且方阵 A 幂等,即 $A^2=A$. 由 §5.5 习题 12,存在可逆方阵 P,使得

$$A=P^{-1}\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}P.$$

因此,

$$\begin{aligned}\det(\lambda I_{(n)}-AB)&=\det\left(\lambda I_{(n)}-P^{-1}\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}PBP^{-1}P\right)\\&=\det\left(\lambda I_{(n)}-\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}PBP^{-1}\right),\\\det(\lambda I_{(n)}-BA)&=\det\left(\lambda I_{(n)}-P^{-1}PBP^{-1}\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}P\right)\\&=\det\left(\lambda I_{(n)}-PBP^{-1}\begin{pmatrix}I_{(r)}&0\\0&0\end{pmatrix}\right).\end{aligned}$$

记

$$PBP^{-1}=\begin{pmatrix}B_{11}&B_{12}\\B_{21}&B_{22}\end{pmatrix},$$

其中 B_{11} 是 r 阶子方阵. 则

$$\det(\lambda I_{(n)}-AB)=\det\left(\lambda I_{(n)}-\begin{pmatrix}B_{11}&B_{12}\\0&0\end{pmatrix}\right)=\lambda^{n-r}\det(\lambda I_{(r)}-B_{11}),$$

$$\det(\lambda I_{(n)} - BA) = \det\left(\lambda I_{(n)} - \begin{pmatrix} B_{11} & 0 \\ B_{21} & 0 \end{pmatrix}\right) = \lambda^{n-r}\det(\lambda I_{(r)} - B_{11}).$$

因此式 (5.7.1) 成立.

对任意方阵 A,由例 5.4.3,存在可逆方阵 P,使得 PA 为幂等方阵. 而

$$\det(\lambda I_{(n)} - AB) = \det(\lambda I_{(n)} - PA\cdot BP^{-1})$$

由上面的证明,

$$= \det(\lambda I_{(n)} - BP^{-1}\cdot PA) = \det(\lambda I_{(n)} - BA).$$

因此式 (5.7.1) 对任意方阵 A 成立. ■

注　一般地说,方阵 BA 与 AB 的最小多项式不一定相同. 例如,设

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}.$$

则

$$AB = 0, \qquad BA = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}.$$

方阵 AB 的特征多项式(也是方阵 BA)的特征多项式为 $\det(\lambda I_{(2)} - AB) = \lambda^2$. 因此方阵 AB 的最小多项式 $d_{AB}(\lambda)\mid\lambda^2$. 所以 $d_{AB}(\lambda) = \lambda$,或 λ^2. 将方阵 AB 代入 λ,即知 $d_{AB}(\lambda) = \lambda$. 同样,方阵 BA 的最小多项式 $d_{BA}(\lambda)\mid\lambda^2$,故 $d_{BA}(\lambda) = \lambda$,或 λ^2. 因为 $BA \neq 0$,故 $d_{BA}(\lambda) = \lambda^2$. 因此 AB 与 BA 的最小多项式并不相等.

作为方阵的特征多项式与化零多项式的一个简单应用,这里介绍求逆方阵的一种方法. 我们知道,可逆方阵 A 的行列式 $\det A \neq 0$,因此方阵 A 的特征多项式

$$\varphi(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

的常数项 $a_n = (-1)^n \det A \neq 0$. 反之亦然. 由 Cayley-Hamilton 定理,

$$\varphi(A) = A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI_{(n)} = 0.$$

所以,

$$A\left(-\frac{1}{a_n}\left(A^{n-1} + a_1A^{n-2} + \cdots + a_{n-1}I_{(n)}\right)\right) = I_{(n)}.$$

因此,

$$A^{-1} = -\frac{1}{a_n}\left(A^{n-1} + a_1A^{n-2} + \cdots + a_{n-1}I_{(n)}\right).$$

例 5.7.2　设

$$A = \begin{pmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{pmatrix}.$$

求 A^{-1}.

解　方阵 A 的特征多项式为

$$\varphi(\lambda) = \det(\lambda I_{(3)} - A) = \lambda^3 - 4\lambda^2 + 7\lambda - 10.$$

因此,

$$A^{-1} = \frac{1}{10}\left(A^2 - 4A + 7I_{(3)}\right) = \frac{1}{10}\begin{pmatrix} -5 & 2 & -16 \\ 0 & 2 & 4 \\ 5 & 0 & 10 \end{pmatrix}.$$ ∎

例 5.7.3 设 n 阶方阵 A 为

$$A = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}.$$

求 A^{-1}.

解 所有元素全为 1 的 n 阶方阵记为 $J_{(n)}$. 则 $A = J_{(n)} - I_{(n)}$. 显然, $J_{(n)}^2 = nJ_{(n)}$. 所以 $f(\lambda) = \lambda^2 - n\lambda$ 是方阵 $J_{(n)}$ 的化零多项式. 由 Euclid 长除法,

$$\lambda^2 - n\lambda = (\lambda - 1)\big(\lambda - (n-1)\big) - (n-1).$$

因此,

$$\big(J_{(n)} - I_{(n)}\big)\big(J_{(n)} - (n-1)I_{(n)}\big) - (n-1)I_{(n)} = 0.$$

所以,

$$A\left(\frac{1}{n-1}\big(J_{(n)} - (n-1)I_{(n)}\big)\right) = I_{(n)}.$$

因此,

$$A^{-1} = \frac{1}{n-1}\big(J_{(n)} - (n-1)I_{(n)}\big).$$ ∎

习 题 5.7

1. 求下列方阵的特征多项式,特征值,和属于每个特征值的特征向量:

(1) $\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$; (2) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$;

(3) $\begin{pmatrix} a_1^2 & a_1a_2 & \cdots & a_1a_n \\ a_2a_1 & a_2^2 & \cdots & a_2a_n \\ & \cdots\cdots\cdots & & \\ a_na_1 & a_na_2 & \cdots & a_n^2 \end{pmatrix}$; (4) $\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ & \cdots\cdots & & \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}$.

2. 设 n 阶可逆方阵 A 的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$. 求逆方阵 A^{-1} 的特征值.

3. 设 n 阶方阵 A 的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$, 且 $f(\lambda)$ 是关于 λ 的多项式. 证明方阵 $f(A)$ 的特征值为 $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$. 特别, 当 $f(\lambda) = \lambda^2$ 时, 方阵 A^2 的特征值为 ${\lambda^2}_1, {\lambda^2}_2, \ldots, {\lambda^2}_n$.

4. 证明,方阵 A 的属于不同特征值的特征向量线性无关.

5. 设 $\alpha_1, \alpha_2, \ldots, \alpha_m$ 是方阵 A 的分别属于特征值 $\lambda_1, \lambda_2, \ldots, \lambda_m$ 的特征向量, 并且 $\alpha_1 + \alpha_2 + \cdots + \alpha_m$ 也是方阵 A 的特征向量. 证明 $\lambda_1 = \lambda_2 = \cdots = \lambda_m$.

6. 满足 $A^k = 0$ 的方阵称为幂零方阵, 其中 k 是正整数. 证明, 方阵 A 为幂零的充分必要条件是, 方阵 A 的特征值全为零.

7. 设 A 与 B 为 n 阶复方阵, 且方阵 A 的特征多项式为 $\varphi(\lambda)$. 证明, 方阵 $\varphi(B)$ 可逆的充分必要条件是, 方阵 A 与 B 没有公共特征值.

8. 设 A 与 B 为 n 阶复方阵. 则关于未知方阵 X 的方阵方程 $AX = XB$ 只有零解的充分必要条件是,方阵 A 与 B 没有公共特征值.

9. 设 A, B 为 n 阶方阵. 定义映射 $\mathscr{P}_{A,B}: \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}$ 如下: 设 $X \in \mathbb{F}^{n\times n}$,则令 $\mathscr{P}_{A,B}(X) = AX - XB$. 显然 $\mathscr{P}_{A,B}$ 是 $\mathbb{F}^{n\times n}$ 到自身的线性变换. 证明,线性变换 $\mathscr{P}_{A,B}$ 可逆的充分必要条件是,方阵 A 与 B 没有公共的特征值.

10. 证明,方阵 A 的最小多项式为 $d(\lambda) = \lambda - a$ 的充分必要条件是,方阵 $A = aI_{(n)}$.

11. 证明,准对角方阵的最小多项式等于每个对角块的最小多项式的最小公倍式.

12. 求下列方阵的最小多项式.

(1) $\begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} 4 & -2 & 2 \\ -5 & 7 & -5 \\ -6 & 7 & -4 \end{pmatrix}$; (3) $\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$.

13. 求下列方阵的特征多项式与最小多项式.

(1) $\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ a_0 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}$; (2) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{pmatrix}$.

14. 举例说明,不相似的方阵可以具有相同的特征多项式与最小多项式.

15. 求 3 阶方阵 A,使得方阵 A 的最小多项式是 λ^2.

16. 取方阵 $A \in \mathbb{C}^{n\times n}$. 定义映射 $\mathscr{A}: \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ 如下: 设 $X \in \mathbb{C}^{n\times n}$,则令 $\mathscr{A}(X) = AX$. 证明线性变换 $\mathscr{A}$ 的最小多项式等于方阵 A 的最小多项式.

17. 设 $\mathscr{A}: V \to V$ 是线性变换,且 $U_1, U_2, \ldots, U_k$ 是线性变换 $\mathscr{A}$ 的不变子空间,

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_k.$$

证明线性变换 $\mathscr{A}$ 的最小多项式等于 $\mathscr{A}\big|_{V_1}, \mathscr{A}\big|_{V_2}, \ldots, \mathscr{A}\big|_{V_k}$ 的最小多项式的最小公倍式.

18. 设 n 阶复方阵 A 满足 $A^k = 0$, k 是正整数. 求方阵 $I_{(n)} - A$ 的逆方阵.

19. 设 n 阶复方阵 A 满足 $A^2 = aA$, $a \neq 1$,且 $\det A = 0$. 求方阵 $I_{(n)} - A$ 的逆方阵.

20. 设 n 阶复方阵 A 为

$$A = \begin{pmatrix} a_1(1-a_1) & -a_1a_2 & \cdots & -a_1a_n \\ -a_2a_1 & a_2(1-a_2) & \cdots & -a_2a_n \\ \cdots & \cdots & \cdots & \cdots \\ -a_na_1 & -a_na_2 & \cdots & a_n(1-a_n) \end{pmatrix}.$$

当 $a_1, a_2, \ldots, a_n$ 满足什么条件时,方阵 A 可逆? 并且 A 可逆时,求逆方阵 A^{-1}.

§5.8 特征子空间

本节继续讨论线性变换 $\mathscr{A}: V \to V$ 的特征值与特征向量.

定义 5.8.1 设 λ_0 是线性变换 $\mathscr{A}: V \to V$ 的特征值. 属于 λ_0 的所有特征值以及零向量构成的集合记为 V_{λ_0},即

$$V_{\lambda_0} = \left\{\alpha \in V \;\middle|\; \mathscr{A}(\alpha) = \lambda_0\alpha\right\}.$$

集合 V_{λ_0} 称为属于特征值 λ_0 的**特征子空间**.

容易验证,特征子空间 V_{λ_0} 确实是 V 的子空间. 不但如此,还可证明,

定理 5.8.1 特征子空间 V_{λ_0} 是线性变换 $\mathscr{A}$ 的不变子空间.

证明 设 $\alpha \in V_{\lambda_0}$,则 $\mathscr{A}(\alpha) = \lambda_0\alpha$. 因此

$$\mathscr{A}\big(\mathscr{A}(\alpha)\big) = \lambda_0\mathscr{A}(\alpha),$$

所以 $\mathscr{A}(\alpha) \in V_{\lambda_0}$. 因此 V_{λ_0} 是 $\mathscr{A}$ 的不变子空间. ■

定义 5.8.2 特征子空间 V_{λ_0} 的维数 m_0 称为线性变换 $\mathscr{A}$ 的特征值 λ_0 的**几何重数**. 而 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 的根 λ_0 的重数称为特征值 λ_0 的**代数重数**.

定理 5.8.2 线性变换 $\mathscr{A}$ 的特征值 λ_0 的几何重数不超过它的代数重数.

证明 设$\mathscr{A}$ 的特征值 λ_0 的几何重数为 m_0,即 $\dim V_{\lambda_0} = m_0$,并且 $\left\{\alpha_1, \alpha_2, \dots, \alpha_{m_0}\right\}$ 是 V_{λ_0} 的基,而 $\left\{\alpha_1, \dots, \alpha_{m_0}, \alpha_{m_0+1}, \dots, \alpha_n\right\}$ 为 V 的基. 由于 $\alpha_j \in V_{\lambda_0}$,故 $\mathscr{A}(\alpha_j) = \lambda_0\alpha_j$, $j = 1, 2, \dots, m_0$. 所以,

$$\mathscr{A}(\alpha_1, \dots, \alpha_{m_0}, \alpha_{m_0+1}, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_{m_0}, \alpha_{m_0+1}, \dots, \alpha_n)\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

其中 $A_{11} = \lambda_0 I_{(m_0)}$. 因此方阵 A 的特征多项式

$$\varphi(\lambda) = \det(\lambda I_{(n)} - A) = (\lambda - \lambda_0)^{m_0} \det(\lambda I_{(n-m_0)} - A_{22}).$$

由于线性变换 $\mathscr{A}$ 的特征多项式等于方阵 A 的特征多项式,所以特征值 λ_0 的几何重数不超过 λ_0 的代数重数. ■

现在假定线性变换 $\mathscr{A}$ 的全部互异的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_t$,它们的代数重数依次为 $e_1, e_2, \dots, e_t$,其中 $e_j \geqslant 1$. 且 $e_1 + e_2 + \cdots + e_t = n = \dim V$. 即设 $\mathscr{A}$ 的特征多项式为

$$\varphi(\lambda) = (\lambda - \lambda_1)^{e_1}(\lambda - \lambda_2)^{e_2}\cdots(\lambda - \lambda_t)^{e_t}.$$

线性变换 $\mathscr{A}$ 的属于特征值 $\lambda_1, \lambda_2, \dots, \lambda_t$ 的特征子空间依次记为 $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_t}$.

定理 5.8.3 特征子空间 $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_t}$ 的和 $V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_t}$ 是直和.

证明 设 $\alpha \in V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_t}$,则存在 $\alpha_j \in V_{\lambda_j}$,使得 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_t$. 假设另有 $\beta_j \in V_{\lambda_j}$,使得 $\alpha = \beta_1 + \beta_2 + \cdots + \beta_t$. 记 $\gamma_j = \alpha_j - \beta_j \in V_{\lambda_j}$. 则显然有

$$\gamma_1+\gamma_2+\cdots+\gamma_t=0. \tag{5.8.1}$$

由于 $\gamma_j\in V_{\lambda_j}$,故 $\mathscr{A}(\gamma_j)=\lambda_j\gamma_j$. 因此由 (5.8.1),

$$\mathscr{A}(\gamma_1+\gamma_2+\cdots+\gamma_t)=\lambda_1\gamma_1+\lambda_2\gamma_2+\cdots+\lambda_t\gamma_t=0.$$

再依次用线性变换 $\mathscr{A}^2,\ldots,\mathscr{A}^{t-1}$ 作用于式 (5.8.1),得到方程组

$$\begin{cases}\gamma_1+\quad\gamma_2+\cdots+\quad\gamma_t=0,\\ \lambda_1\gamma_1+\lambda_2\gamma_2+\cdots+\lambda_t\gamma_t=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \lambda_1^{t-1}\gamma_1+\lambda_2^{t-1}\gamma_2+\cdots+\lambda_t^{t-1}\gamma_t=0.\end{cases} \tag{5.8.2}$$

在 V 中取一组基 $\{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\}$. 设 γ_j 在这组基下的坐标为 $(x_{1j},x_{2j},\ldots,x_{nj})^{\mathrm{T}}$. 将式 (5.8.2) 写出第 i 个坐标分量形式,得到

$$\begin{cases}x_{i1}+\quad x_{i2}+\cdots+\quad x_{it}=0,\\ \lambda_1x_{i1}+\lambda_2x_{i2}+\cdots+\lambda_tx_{it}=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \lambda_1^{t-1}x_{i1}+\lambda_2^{t-1}x_{i2}+\cdots+\lambda_t^{t-1}x_{it}=0.\end{cases} \tag{5.8.3}$$

由于 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 两两不同,所以方程组 (5.8.3) 的系数行列式

$$\begin{vmatrix}1&1&\cdots&1\\ \lambda_1&\lambda_2&\cdots&\lambda_t\\ \vdots&\vdots&\ddots&\vdots\\ \lambda_1^{t-1}&\lambda_2^{t-1}&\cdots&\lambda_t^{t-1}\end{vmatrix}\neq 0.$$

因此诸 $x_{ij}=0$,即 γ_j 的坐标为零,所以 $\gamma_j=0$. 从而 $\alpha_j=\beta_j$. 这表明,和 $V_{\lambda_1}+V_{\lambda_2}+\cdots+V_{\lambda_t}$ 中每个向量 α 表成 $V_{\lambda_1},V_{\lambda_2},\ldots,V_{\lambda_t}$ 的向量之和的方式是唯一的. 因此

$$V_{\lambda_1}+V_{\lambda_2}+\cdots+V_{\lambda_t}=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_t}.$$ ■

定义 5.8.3　如果存在线性空间 V 的一组基,使得线性变换 $\mathscr{A}:V\to V$ 在这组基下的方阵是对角方阵,则线性变换 $\mathscr{A}$ 称为**可对角化**的.

定理 5.8.4　设 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 是线性变换 $\mathscr{A}:V\to V$ 的全部不同的特征值,则线性变换 $\mathscr{A}$ 可对角化当且仅当每个不同特征值 λ_j 的几何重数等于它的代数重数.

证明　设线性变换 $\mathscr{A}$ 是可对角化的,则存在 V 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$,使得

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A, \tag{5.8.4}$$

其中 A 是对角方阵,记 $A=\mathrm{diag}(a_1,a_2,\ldots,a_n)$.

由式 (5.8.4),$\mathscr{A}(\alpha_j)=a_j\alpha_j$,即 $a_1,a_2,\ldots,a_n$ 是线性变换 $\mathscr{A}$ 的特征值. 适当调整基向量 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 的次序,可设

$$A=\mathrm{diag}\big(\lambda_1I_{(e_1)},\lambda_2I_{(e_2)},\ldots,\lambda_tI_{(e_t)}\big).$$

于是线性变换 $\mathscr{A}$ 的特征多项式为

$$\varphi(\lambda)=\det(\lambda I_{(n)}-A)=(\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_t)^{e_t}.$$

即特征值 λ_j 的代数重数为 $e_j, j=1,2,\ldots,t$.

另一方面,由式 (5.8.4),对于 $1\leqslant i\leqslant e_j, 1\leqslant j\leqslant t$,有

$$\mathscr{A}(\alpha_{e_1+\cdots+e_{j-1}+i})=\lambda_j\alpha_{e_1+\cdots+e_{j-1}+i},$$

其中约定 $e_0=0$,因此 $\alpha_{e_1+\cdots+e_{j-1}+i}\in V_{\lambda_j}, i=1,2,\ldots,e_j$. 所以 $\dim V_{\lambda_j}\geqslant e_j$.

由定理 5.8.2, $\dim V_{\lambda_j}\leqslant e_j$. 从而 $e_j=\dim V_{\lambda_j}, j=1,2,\ldots,t$.

反之,设每个特征值 λ_j 的几何重数等于它的代数重数. 由定理 5.8.3,

$$V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_t}.$$

取特征子空间 V_{λ_j} 的基为 $\left\{\alpha_{e_1+\cdots+e_{j-1}+1},\ldots,\alpha_{e_1+\cdots+e_{j-1}+e_j}\right\}$,则 $\left\{\alpha_1,\alpha_2,\ldots,\alpha_n\right\}$ 是 V 的基. 于是

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\operatorname{diag}\left(\lambda_1 I_{(e_1)},\lambda_2 I_{(e_2)},\ldots,\lambda_t I_{(e_t)}\right).$$

即线性变换 $\mathscr{A}$ 是可以对角化的. ∎

定义 5.8.4 设 $\mathscr{A}: V\to V$ 是线性变换, $\left\{\alpha_1,\alpha_2,\ldots,\alpha_n\right\}$ 是 V 的基. 如果每个基向量 α_i 都是 $\mathscr{A}$ 的特征向量,则 $\left\{\alpha_1,\alpha_2,\ldots,\alpha_n\right\}$ 称为**完全特征向量组**.

定理 5.8.5 线性变换 $\mathscr{A}$ 可对角化的当且仅当 $\mathscr{A}$ 存在一组完全特征向量组.

证明 设线性变换 $\mathscr{A}$ 可对角化,则存在 V 的基 $\left\{\alpha_1,\alpha_2,\ldots,\alpha_n\right\}$,使得

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n).$$

因此 $\mathscr{A}(\alpha_j)=\lambda_j\alpha_j$,即每个基向量 α_j 都是 $\mathscr{A}$ 的特征向量.

反之,设 $\left\{\alpha_1,\alpha_2,\ldots,\alpha_n\right\}$ 是完全特征向量组,则 $\left\{\alpha_1,\alpha_2,\ldots,\alpha_n\right\}$ 是 V 的基,并且 $\mathscr{A}(\alpha_j)=\lambda_j\alpha_j$. 因此

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n).$$

所以线性变换 $\mathscr{A}$ 可对角化. ∎

例 5.8.1 证明,任意 n 阶复方阵 A 都相似于上三角方阵.

证明 对方阵 A 的阶数 n 用归纳法. 当 $n=1$ 时,1 阶方阵即是上三角方阵,因此结论成立. 假设结论对 $n-1$ 阶方阵成立. 下面证明结论对 n 阶方阵成立.

设 λ_1 是方阵 A 的特征值, $\alpha_1\in\mathbb{C}^n$ 是方阵 A 的属于特征值 λ_1 的特征向量. 设 $\left\{\alpha_1,\alpha_2,\ldots,\alpha_n\right\}$ 是 n 维复列向量空间 $\mathbb{C}^n$ 的基. 因为 $A\alpha_j\in\mathbb{C}^n, j\geqslant 2$,故

$$A\alpha_j=b_{j1}\alpha_1+b_{j2}\alpha_2+\cdots+b_{jn}\alpha_n.$$

又 $\mathscr{A}\alpha_1=\lambda_1\alpha_1$. 所以

$$A(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\begin{pmatrix}\lambda_1 & b_{21} & \cdots & b_{n1}\\ 0 & b_{22} & \cdots & b_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & b_{2n} & \cdots & b_{nn}\end{pmatrix}$$

$$=: (\alpha_1, \alpha_2, \ldots, \alpha_n)\begin{pmatrix} \lambda_1 & \beta \\ 0 & A_{22} \end{pmatrix}. \tag{5.8.5}$$

其中 $\beta = (b_{21}, b_{31}, \ldots, b_{n1})$.

由归纳假设,存在 $n-1$ 阶可逆方阵 Q_1,使得 $Q_1^{-1}A_{22}Q_1 = B_1$ 是上三角方阵. 把 $A_{22} = Q_1B_1Q_1^{-1}$ 代入式 (5.8.5),得到

$$\begin{aligned} A(\alpha_1, \alpha_2, \ldots, \alpha_n) &= (\alpha_1, \alpha_2, \ldots, \alpha_n)\begin{pmatrix} \lambda_1 & \beta \\ 0 & Q_1B_1Q_1^{-1} \end{pmatrix} \\ &= (\alpha_1, \alpha_2, \ldots, \alpha_n)\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}\begin{pmatrix} \lambda_1 & \beta Q_1 \\ 0 & B_1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{pmatrix}. \end{aligned}$$

因为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性无关,所以方阵 $P = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ 可逆. 因此

$$A = P\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}\begin{pmatrix} \lambda_1 & \beta Q_1 \\ 0 & B_1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}^{-1}P^{-1}.$$

由于 Q_1 可逆,所以方阵 $\operatorname{diag}(1, Q_1)$ 可逆,从而 $R = P\operatorname{diag}(1, Q_1)$ 可逆. 因此方阵 A 相似于上三角方阵

$$\begin{pmatrix} \lambda_1 & \beta Q_1 \\ 0 & B_1 \end{pmatrix}.$$

■

例 5.8.2　证明,秩为 r 的 n 阶幂等复方阵 A (即满足 $A^2 = A$)相似于对角方阵

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

先用矩阵方法证明之.

证法 1　因为 $\operatorname{rank} A = r$,故存在可逆方阵 P 与 Q,使得

$$A = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q.$$

由于 $A^2 = A$,故

$$P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}QP\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q = P\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q.$$

因此,

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}QP\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

设

$$R = QP = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

其中 R_{11} 为 r 阶子方阵. 则由上式得到,$R_{11} = I_{(r)}$. 因此,

$$Q = \begin{pmatrix} I_{(r)} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}P^{-1}.$$

所以,

$$A = P\begin{pmatrix} I_{(r)} & R_{12} \\ 0 & 0 \end{pmatrix}P^{-1} = P\begin{pmatrix} I_{(r)} & -R_{12} \\ 0 & I_{(n-r)} \end{pmatrix}\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_{(r)} & R_{12} \\ 0 & I_{(n-r)} \end{pmatrix}P^{-1}.$$

记

$$T = P\begin{pmatrix} I_{(r)} & -R_{12} \\ 0 & I_{(n-r)} \end{pmatrix}.$$

则 T 可逆,并且

$$A = T\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}T^{-1}.$$ ■

证法 2 因为方阵 A 满足 $A^2 = A$,所以 $f(\lambda) = \lambda^2 - \lambda$ 是方阵 A 的化零多项式. 因此方阵 A 的最小多项式 $d(\lambda)$ 的次数只能是 1 或 2. 并且由于 $d(\lambda) \mid f(\lambda)$,故 $d(\lambda)$ 只能是 λ,或 $\lambda - 1$,或 $\lambda(\lambda - 1)$.

如果 $d(\lambda) = \lambda$,则 $d(A) = A = 0$,即 A 为零方阵,所以结论成立; 如果 $d(\lambda) = \lambda - 1$,则 $d(A) = A - I_{(n)} = 0$,从而 $A = I_{(n)}$,所以结论也成立.

现在设 $d(\lambda) = \lambda(\lambda - 1)$. 因此方阵 A 的最小多项式的根为 0,1. 所以方阵 A 的全部不同的特征值为 0,1. 设它们的代数重数分别为 ℓ 与 k,$\ell + k = n$. 由例 5.8.1,存在可逆方阵 P,使得

$$P^{-1}AP = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

其中 A_{11} 是主对角元全为 1 的 k 阶上三角阵,A_{22} 是主对角元全为 0 的 ℓ 阶上三角阵.

由于 $A^2 = A$,所以 $A_{11} = I_{(k)}$,$A_{22} = 0$. 即

$$P^{-1}AP = \begin{pmatrix} I_{(k)} & A_{12} \\ 0 & 0 \end{pmatrix}.$$

于是,

$$\begin{pmatrix} I_{(k)} & A_{12} \\ 0 & I_{(\ell)} \end{pmatrix}P^{-1}AP\begin{pmatrix} I_{(k)} & -A_{12} \\ 0 & I_{(\ell)} \end{pmatrix} = \begin{pmatrix} I_{(k)} & 0 \\ 0 & 0 \end{pmatrix}.$$

容易证明,矩阵的秩是矩阵在相似下的不变量,因此 $k = r$. 这就证明了结论. ■

再用几何方法证明之.

证法 3 设 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 n 维复线性空间 V 的基. 由

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A \tag{5.8.6}$$

可以确定线性变换 V 上的线性变换 $\mathscr{A}$. 设 $\alpha \in V$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标为 x,即 $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)x$. 由式 (5.8.6),

$$\mathscr{A}(\alpha) = (\alpha_1, \alpha_2, \ldots, \alpha_n)Ax,$$
$$\mathscr{A}^2(\alpha) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A^2x.$$

由于 $A^2 = A$,故对任意 $\alpha \in V$,$\mathscr{A}^2(\alpha) = \mathscr{A}(\alpha)$. 因此线性变换 $\mathscr{A}$ 满足 $\mathscr{A}^2 = \mathscr{A}$,它称为幂等变换.

设 $\alpha \in \operatorname{Im}\mathscr{A} \cap \operatorname{Ker}\mathscr{A}$. 由于 $\alpha \in \operatorname{Im}\mathscr{A}$,故存在 $\beta \in V$,使得 $\alpha = \mathscr{A}(\beta)$; 由于 $\alpha \in \operatorname{Ker}\mathscr{A}$,故 $\mathscr{A}(\alpha) = 0$. 因为 $\mathscr{A}^2 = \mathscr{A}$,所以

$$\alpha = \mathscr{A}(\beta) = \mathscr{A}^2(\beta) = \mathscr{A}\big(\mathscr{A}(\beta)\big) = \mathscr{A}(\alpha) = 0.$$

因此 $\operatorname{Im}\mathscr{A} \cap \operatorname{Ker}\mathscr{A} = 0$. 这表明,$\operatorname{Im}\mathscr{A} + \operatorname{Ker}\mathscr{A} = \operatorname{Im}\mathscr{A} \oplus \operatorname{Ker}\mathscr{A}$.

由于 $\operatorname{rank} A = r$,故 $\dim \operatorname{Im}\mathscr{A} = r$,$\dim \operatorname{Ker}\mathscr{A} = n - r$. 于是

$$V = \operatorname{Im}\mathscr{A} \oplus \operatorname{Ker}\mathscr{A}.$$

设 $\{\beta_1, \beta_2, \ldots, \beta_r\}$ 是 $\operatorname{Im}\mathscr{A}$ 的基,$\{\beta_{r+1}, \beta_{r+2}, \ldots, \beta_n\}$ 是 $\operatorname{Ker}\mathscr{A}$ 的基,则

$$\{\beta_1, \beta_2, \ldots, \beta_r, \beta_{r+1}, \beta_{r+2}, \ldots, \beta_n\}$$

是 V 的基. 设 $\beta_j \in \operatorname{Im}\mathscr{A}$,则存在 $\gamma_j \in V$,使得 $\beta_j = \mathscr{A}(\gamma_j)$. 因为 $\mathscr{A}^2 = \mathscr{A}$,所以

$$\beta_j = \mathscr{A}(\gamma_j) = \mathscr{A}^2(\gamma_j) = \mathscr{A}\big(\mathscr{A}(\gamma_j)\big) = \mathscr{A}(\beta_j).$$

设 $\beta_j \in \operatorname{Ker}\mathscr{A}$,则显然 $\mathscr{A}(\beta_j) = 0$. 因此

$$\mathscr{A}(\beta_1, \beta_2, \ldots, \beta_n) = (\beta_1, \beta_2, \ldots, \beta_n)\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

由于同一个线性变换在不同基下的方阵相似,所以方阵 A 相似于

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

■

证法 4　仍如证法 3. 由式 (5.8.6) 确定的线性变换 $\mathscr{A}$ 满足 $\mathscr{A}^2 = \mathscr{A}$. 所以 $f(\lambda) = \lambda^2 - \lambda$ 是线性变换 $\mathscr{A}$ 的化零多项式. 因此线性变换 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 整除 $\lambda^2 - \lambda$. 这表明,$d(\lambda) = \lambda$,或 $\lambda - 1$,或 $\lambda(\lambda - 1)$.

由于最小多项式的根必是特征多项式的根,反之亦然,所以线性变换 $\mathscr{A}$ 的特征值只能是 0 或 1. 线性变换 $\mathscr{A}$ 的分别属于特征值 0 与 1 的特征子空间分别记为 V_0 与 V_1. 由定理 5.8.3,$V_0 + V_1 = V_0 \oplus V_1$.

设 $\alpha \in V$,则 $\alpha = \big(\alpha - \mathscr{A}(\alpha)\big) + \mathscr{A}(\alpha)$. 由于 $\mathscr{A}^2 = \mathscr{A}$,故

$$\mathscr{A}\big(\alpha - \mathscr{A}(\alpha)\big) = \mathscr{A}(\alpha) - \mathscr{A}^2(\alpha) = 0,$$

且

$$\mathscr{A}\big(\mathscr{A}(\alpha)\big) = \mathscr{A}(\alpha).$$

所以 $\alpha - \mathscr{A}(\alpha) \in V_0$,$\mathscr{A}(\alpha) \in V_1$,即 $\alpha \in V_0 + V_1$. 因此 $V = V_0 \oplus V_1$. 从而

$$\dim V_0 + \dim V_1 = n.$$

这就证明特征值 0 与 1 的几何重数分别等于它们的代数重数,由定理 5.8.4,存在 V 的基,使得线性变换在这组基下的方阵为对角方阵,而且对角元为线性变换 $\mathscr{A}$ 的特征值. 所以可设对角方阵为

$$\begin{pmatrix} I_{(k)} & 0 \\ 0 & 0 \end{pmatrix}.$$

因为同一个线性变换 $\mathscr{A}$ 在不同基下的方阵相似,所以方阵 A 相似于以上对角

阵. 由于相似的方阵的秩相等,所以 $k=r$. ■

例 5.8.3 设 E_{ij} 是第 i 行与第 j 列交叉位置上的元素为 1,其它元素为 0 的 n 阶方阵. 且设 n^2 个 n 阶非零的复方阵为 F_{ij} 满足,对于 $1\leqslant i,j,k,\ell\leqslant n$,

$$F_{ij}F_{k\ell}=\delta_{jk}F_{i\ell}, \tag{5.8.7}$$

其中 δ_{jk} 是 Kronecker 符号. 证明,存在 n 阶可逆复方阵 P,使得对任意 $1\leqslant i,j\leqslant n$,

$$P^{-1}F_{ij}P=E_{ij}.$$

证明 由式 (5.8.7),$F_{11}^2=F_{11}$,即 F_{11} 是幂等方阵. 因为 F_{11} 非零,因此由例 5.8.2,方阵 F_{11} 必有特征值 1. 设 $\alpha_1\in\mathbb{C}^n$ 是方阵 F_{11} 属于特征值 1 的特征向量,即 $F_{11}\alpha_1=\alpha_1$.

记 $\alpha_j=F_{j1}\alpha_1$, $j=1,2,\ldots,n$. 设 $\lambda_1,\lambda_2,\ldots,\lambda_n\in\mathbb{C}$ 使得

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_n=0.$$

则对任意 $j=1,2,\ldots,n$,

$$\begin{aligned}F_{1j}(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_n\alpha_n)&=\lambda_1F_{1j}\alpha_1+\lambda_2F_{1j}\alpha_2+\cdots+\lambda_nF_{1j}\alpha_n\\&=\lambda_1F_{1j}F_{11}\alpha_1+\lambda_2F_{1j}F_{21}\alpha_1+\cdots+\lambda_nF_{1j}F_{n1}\alpha_1\\&=\lambda_1\delta_{j1}F_{11}\alpha_1+\lambda_2\delta_{j2}F_{11}\alpha_1+\cdots+\lambda_n\delta_{jn}F_{11}\alpha_1\\&=\lambda_jF_{11}\alpha_1=\lambda_j\alpha_1=0.\end{aligned}$$

由于 $\alpha_1\neq 0$,所以 $\lambda_j=0$. 因此列向量 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 线性无关. 令 n 阶方阵

$$P=(\alpha_1,\alpha_2,\ldots,\alpha_n).$$

显然方阵 P 可逆,由式 (5.8.7),对任意 $k=1,2,\ldots,n$,

$$F_{ij}\alpha_k=F_{ij}F_{k1}\alpha_1=\delta_{jk}F_{i1}\alpha_1=\delta_{ij}\alpha_i.$$

所以

$$\begin{aligned}F_{ij}P&=F_{ij}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(F_{ij}\alpha_1,F_{ij}\alpha_2,\ldots,F_{ij}\alpha_n)\\&=(0,\ldots,0,\alpha_i,0,\ldots,0)=(\alpha_1,\alpha_2,\ldots,\alpha_n)E_{ij}\\&=PE_{ij}.\end{aligned}$$

这就证明,$P^{-1}F_{ij}P=E_{ij}$,$1\leqslant i,j\leqslant n$. ■

习 题 5.8

1. 设 n 阶复方阵 A 满足 $A^k=I_{(n)}$,k 为正整数. 证明方阵 A 相似于对角方阵.

2. 如果方阵 N 满足 $N^k=0$,k 为正整数,则方阵 N 称为幂零方阵. 使得 $N^k=0$ 的最小正整数 k 称为幂零方阵 N 的**幂零指数**. 证明幂零指数为 n 的 n 阶幂零复方阵 N 相似于

$$\begin{pmatrix}0&1&&\\&\ddots&\ddots&\\&&0&1\\&&&0\end{pmatrix}_{n\times n}.$$

3. 由于方阵 A 的迹 $\mathrm{Tr}\,A$ 是方阵在相似下的不变量,因此可以定义线性变换 $\mathscr{A}:V\to V$ 在 V 的某组基下的方阵 A 的 $\mathrm{Tr}\,A$ 为线性变换 $\mathscr{A}$ 的迹 $\mathrm{Tr}\,\mathscr{A}$. 证明,如果 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 满足 $\mathrm{Tr}\,\mathscr{A}=0$,则存在 V 的一组基,使得线性变换 $\mathscr{A}$ 在这组基下的方阵的主对角元都是零.

4. 设 3 阶实方阵 A 在实数域 $\mathbb{R}$ 上不相似于上三角阵，即不存在 3 阶可逆实方阵 P，使得 $P^{-1}AP$ 是上三角方阵. 证明方阵 A 在复数域上相似于对角方阵.

5. 设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 的 n 个特征值两两不等. 证明线性变换 $\mathscr{A}$ 可相似对角化.

6. 设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 可对角化，且 U 是线性变换 $\mathscr{A}$ 的不变子空间. 证明线性变换 $\mathscr{A}$ 在 U 上的限制 $\mathscr{A}\big|_U$ 也是可对角化的.

7. 取定 n 阶复方阵 $A \in \mathbb{C}^{n\times n}$，定义线性变换 $\mathscr{A}_1 : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ 与 $\mathscr{A}_2 : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ 如下：对任意 $X \in \mathbb{C}^{n\times n}$，令

$$\mathscr{A}_1(X) = AX, \qquad \mathscr{A}_2(X) = AX - XA.$$

如果方阵 A 可对角化，则线性变换 $\mathscr{A}_1$ 与 $\mathscr{A}_2$ 是否也可对角化?

8. 设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 可交换. 证明线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 具有公共特征向量. 进而证明，设 I 是下标集合，V 上的线性变换集合 $\left\{\mathscr{A}_\nu \,\middle|\, \nu \in I\right\}$ 中任意两个线性变换 $\mathscr{A}_{\nu_1}$ 与 $\mathscr{A}_{\nu_2}$ 可交换，则众线性变换 $\mathscr{A}_\nu$ 具有公共特征向量.

9. 设 n 阶复方阵 A 与 B 可交换. 证明，存在 n 阶可逆方阵 P，使得 $P^{-1}AP$ 与 $P^{-1}BP$ 都是上三角方阵，即方阵 A 与 B 可以同时相似于上三角形. 推广到任意多个两两可交换的方阵的情形.

§5.9　特征值的界

在 §5.7 已经讲过，方阵 A 的特征值是方阵 A 的特征多项式 $\varphi(\lambda)$ 的根，而方阵 A 的特征多项式 $\varphi(\lambda)$ 是方阵 $\lambda I_{(n)} - A$ 的行列式. 因此欲求 n 阶方阵 A 的特征值，必须先求出 n 阶方阵 $\lambda I_{(n)} - A$ 的行列式，即确定方阵 A 的特征多项式 $\varphi(\lambda)$，然后再求 n 次多项式 $\varphi(\lambda)$ 的根. 我们知道，计算 n 阶行列式的计算量是相当大的. 换句话说，要求出方阵 A 的特征多项式是相当困难的. 即便方阵 A 的特征多项式 $\varphi(\lambda)$ 已经求出，要求出 n 次多项式 $\varphi(\lambda)$ 的根式很困难的. 甚至当 $n \geqslant 5$ 时，一般 n 次多项式都不能用根号求得(这是法国著名数学家 E. Galois 于 1831 年所证明的). 因此人们转向特征值的界的估计.

本世纪头一个关于方阵的特征值的界是 Hirsch 于 1900 年给出的.

定理 5.9.1 (Hirsch) 设 $A = (a_{ij})$ 是 n 阶复方阵. 记

$$B = \frac{1}{2}\left(A + \overline{A}^{\mathrm{T}}\right) = (b_{k\ell}), \qquad C = \frac{1}{2\mathrm{i}}\left(A - \overline{A}^{\mathrm{T}}\right) = (c_{k\ell}),$$

其中 $\mathrm{i}^2 = -1$. 且记

$$M_A = \max\left\{|a_{k\ell}| \,\middle|\, 1 \leqslant k, \ell \leqslant n\right\},$$

其中 $|a_{k\ell}|$ 表示复数 $a_{k\ell}$ 的模.

设 λ_0 是方阵 A 的特征值，则

$$|\lambda_0| \leqslant nM_A, \quad |\mathrm{Re}\,\lambda_0| \leqslant nM_B, \quad |\mathrm{Im}\,\lambda_0| \leqslant nM_C,$$

其中 $\mathrm{Re}\,\lambda_0$ 与 $\mathrm{Im}\,\lambda_0$ 分别是复数 λ_0 的实部与虚部.

如果 A 是实方阵，则

$$|\mathrm{Im}\,\lambda_0| \leqslant M_C\left(\frac{n(n-1)}{2}\right)^{\frac{1}{2}}.$$

证明 设 $x=(x_1,x_2,\dots,x_n)^{\mathrm{T}}$ 是方阵 A 的属于特征值 λ_0 的特征向量,即

$$Ax=\lambda_0 x. \tag{5.9.1}$$

上式两端同时右乘以 $\overline{x}^{\mathrm{T}}$,则得到

$$\overline{x}^{\mathrm{T}}Ax=\lambda_0\overline{x}^{\mathrm{T}}x. \tag{5.9.2}$$

由于

$$\overline{x}^{\mathrm{T}}Ax=\sum_{1\leqslant k,\ell\leqslant n}a_{k\ell}\overline{x}_k x_\ell,\quad \lambda_0\overline{x}^{\mathrm{T}}x=\lambda_0\left(\sum_{k=1}^n|x_k|^2\right).$$

因此,

$$\lambda_0\left(\sum_{k=1}^n|x_k|^2\right)=\sum_{1\leqslant k,\ell\leqslant n}a_{k\ell}\overline{x}_k x_\ell.$$

所以,

$$\begin{aligned}|\lambda_0|\left(\sum_{k=1}^n|x_k|^2\right)&\leqslant\sum_{1\leqslant k,\ell\leqslant n}|a_{k\ell}|\cdot|\overline{x}_k|\cdot|x_\ell|\\&\leqslant M_A\sum_{1\leqslant k,\ell\leqslant n}|\overline{x}_k|\cdot|x_\ell|\leqslant M_A\left(\sum_{k=1}^n|x_k|\right)^2.\end{aligned}$$

由 Cauchy 不等式,

$$\left(\sum_{k=1}^n|x_k|\right)^2\leqslant n\left(\sum_{k=1}^n|x_k|^2\right).$$

因此

$$|\lambda_0|\left(\sum_{k=1}^n|x_k|^2\right)\leqslant nM_A\left(\sum_{k=1}^n|x_k|^2\right).$$

由于特征向量 x 非零,所以

$$\sum_{k=1}^n|x_k|^2>0.$$

于是得到,$|\lambda_0|\leqslant nM_A$.

取式 (5.9.2) 两端的共轭转置,得到

$$\overline{x}^{\mathrm{T}}\overline{A}^{\mathrm{T}}x=\overline{\lambda}_0\overline{x}^{\mathrm{T}}x. \tag{5.9.3}$$

式 (5.9.2) 与式 (5.9.3) 相加或相减,并乘以 $\dfrac{1}{2}$ 或 $\dfrac{1}{2\mathrm{i}}$,得到

$$(\mathrm{Re}\,\lambda_0)\overline{x}^{\mathrm{T}}x=\overline{x}^{\mathrm{T}}Bx, \tag{5.9.4}$$

$$(\mathrm{Im}\,\lambda_0)\overline{x}^{\mathrm{T}}x=\overline{x}^{\mathrm{T}}Cx. \tag{5.9.5}$$

把式 (5.9.4) 中的 $\mathrm{Re}\,\lambda_0$ 与 B 分别视为式 (5.9.2) 中的 λ_0 与 A,便得到

$$|\mathrm{Re}\,\lambda_0|\leqslant nM_B.$$

把式 (5.9.5) 中的 $\mathrm{Im}\,\lambda_0$ 与 C 分别视为式 (5.9.2) 中的 λ_0 与 A,便得到

$$|\mathrm{Im}\,\lambda_0|\leqslant nM_C.$$

当 $A=(a_{ij})$ 为实方阵时,$C=\dfrac{1}{2\mathrm{i}}(A-A^{\mathrm{T}})$. 因此,

$$C^{\mathrm{T}}=\frac{1}{2\mathrm{i}}(A^{\mathrm{T}}-A)=-\frac{1}{2\mathrm{i}}(A-A^{\mathrm{T}})=-C.$$

即 $c_{kk}=0, k=1,2,\dots,n$. 由式 (5.9.5) 得到,

$$\begin{aligned}|\operatorname{Im}\lambda_0|\Big(\sum_{k=1}^n |x_k|^2\Big) &= \Big|\sum_{1\leqslant k\neq \ell\leqslant n} c_{k\ell}\overline{x}_k x_\ell\Big| = \Big|\sum_{1\leqslant k<\ell\leqslant n} c_{k\ell}(\overline{x}_k x_\ell - x_k\overline{x}_\ell)\Big|\\ &\leqslant \sum_{1\leqslant k<\ell\leqslant n} |c_{k\ell}|\cdot|\overline{x}_k x_\ell - x_k\overline{x}_\ell| \leqslant M_C\Big(\sum_{1\leqslant k<\ell\leqslant n}\Big|\frac{\overline{x}_k x_\ell - x_k\overline{x}_\ell}{\mathrm{i}}\Big|\Big).\end{aligned}$$

由于

$$\overline{\Big(\frac{\overline{x}_k x_\ell - x_k\overline{x}_\ell}{\mathrm{i}}\Big)} = \frac{x_k\overline{x}_\ell - \overline{x}_k x_\ell}{-\mathrm{i}} = \frac{\overline{x}_k x_\ell - x_k\overline{x}_\ell}{\mathrm{i}},$$

所以 $\dfrac{\overline{x}_k x_\ell - x_k\overline{x}_\ell}{\mathrm{i}}$ 是实数, $1\leqslant k<\ell\leqslant n$. 因此由 Cauchy 不等式,

$$\Big(\sum_{1\leqslant k<\ell\leqslant n}\Big|\frac{\overline{x}_k x_\ell - x_k\overline{x}_\ell}{\mathrm{i}}\Big|\Big)^2 \leqslant \frac{n(n-1)}{2}\Big(\sum_{1\leqslant k<\ell\leqslant n}\Big|\frac{\overline{x}_k x_\ell - x_k\overline{x}_\ell}{\mathrm{i}}\Big|^2\Big).$$

因为

$$\begin{aligned}\Big|\frac{\overline{x}_k x_\ell - x_k\overline{x}_\ell}{\mathrm{i}}\Big|^2 &= |\overline{x}_k x_\ell - x_k\overline{x}_\ell|^2 = (\overline{x}_k x_\ell - x_k\overline{x}_\ell)\overline{(\overline{x}_k x_\ell - x_k\overline{x}_\ell)}\\ &= -(\overline{x}_k x_\ell - x_k\overline{x}_\ell)^2,\end{aligned}$$

所以,

$$|\operatorname{Im}\lambda_0|\Big(\sum_{k=1}^n |x_k|^2\Big) \leqslant M_C\Big(-\frac{n(n-1)}{2}\sum_{1\leqslant k<\ell\leqslant n}(\overline{x}_k x_\ell - x_k\overline{x}_\ell)^2\Big)^{\frac12}.$$

不难验证,

$$-\sum_{1\leqslant k<\ell\leqslant n}(\overline{x}_k x_\ell - x_k\overline{x}_\ell)^2 = \Big(\sum_{k=1}^n x_k\overline{x}_k\Big)^2 - \Big(\sum_{k=1}^n x_k^2\Big)\Big(\sum_{k=1}^n \overline{x}_k^2\Big) \leqslant \Big(\sum_{k=1}^n x_k\overline{x}_k\Big)^2.$$

因此,

$$|\operatorname{Im}\lambda_0|\Big(\sum_{k=1}^n |x_k|^2\Big) \leqslant M_C\Big(\frac{n(n-1)}{2}\Big)^{\frac12}\Big(\sum_{k=1}^n |x_k|^2\Big).$$

由此得到,

$$|\operatorname{Im}\lambda_0| \leqslant M_C\Big(\frac{n(n-1)}{2}\Big)^{\frac12}.$$ ■

适合 $\overline{H}^{\mathrm{T}}=H$ 的复方阵 H 称为 **Hermite 方阵**; 适合 $S^{\mathrm{T}}=S$ 的实方阵称为**实对称方阵**. 由 Hirsch 定理, 容易证明,

定理 5.9.2 Hermite 方阵与实对称方阵的特征值都是实数.

证明 设 H 是 Hermite 方阵, 则

$$C=\frac{1}{2\mathrm{i}}(H-\overline{H}^{\mathrm{T}})=\frac{1}{2\mathrm{i}}(H-H)=0.$$

因此 $M_C=0$. 于是 Hermite 方阵 H 的特征值 λ_0 满足

$$|\operatorname{Im}\lambda_0|\leqslant nM_C=0.$$

所以 $\operatorname{Im}\lambda_0=0$. 这就证明, λ_0 是实数.

对于实对称方阵 S, 证明完全相同. ■

满足 $\overline{K}^{\mathrm{T}}=-K$ 的复方阵 K 称为**斜 Hermite 方阵**; 满足 $K^{\mathrm{T}}=-K$ 的实方阵称为**实斜对称方阵**.

定理 5.9.3 斜 Hermite 方阵与实斜对称方阵的非零特征值为纯虚数.

证明 设 λ_0 是斜 Hermite 方阵 K 的非零特征值. 因为

$$B=\frac{1}{2}(K+\overline{K}^{\mathrm{T}})=\frac{1}{2}(K-K)=0,$$

所以 $M_B=0$. 因此

$$|\mathrm{Re}\,\lambda_0|\leqslant nM_B=0.$$

故 $\mathrm{Re}\,\lambda_0=0$,即 λ_0 是纯虚数.

对实斜对称方阵,证明完全相同. ■

在介绍关于特征值的另一个界之前,先证明有关主对角占优的一个结论. 它是 Levy 于 1881 年首先得到的,Desplanques 于 1887 年作了推广,1903 年 Hadamard 把它收入他所著的书中. 因此通常称为 Hadamard 定理.

n 阶复方阵 $A=(a_{ij})$ 的第 i 行上所有所有元素的模之和记为 R_i,即

$$R_i=\sum_{j=1}^{n}|a_{ij}|.$$

方阵 $A=(a_{ij})$ 的第 j 列上所有元素的模之和记为 T_j,即

$$T_j=\sum_{i=1}^{n}|a_{ij}|.$$

记 $P_i=R_i-|a_{ii}|, Q_j=T_j-|a_{jj}|$. 如果方阵 A 满足,对任意 i,

$$|a_{ii}|>P_i,$$

则方阵 A 称为**行主对角占优方阵**; 如果方阵 A 满足,对任意 j,

$$|a_{jj}|>Q_j,$$

则方阵 A 称为**列主对角占优方阵**.

关于主对角占优方阵,有

定理 5.9.4(Levy-Desplanques) 设 A 是行或列主对角占优方阵,则 A 的行列式 $\det A\neq 0$.

证明 设 $A=(a_{ij})$ 是行主对角占优方阵,且 $\det A=0$. 那么齐次线性方程组 $Ax=0$ 有非零解 $x=(x_1,x_2,\ldots,x_n)^{\mathrm{T}}$. 故 $\max\{|x_1|,\ldots,|x_n|\}=|x_k|>0$. 因此,

$$\sum_{j=1}^{n}a_{kj}x_j=0.$$

于是,

$$a_{kk}x_k=-\sum_{\substack{1\leqslant j\leqslant n\\ j\neq k}}a_{kj}x_j.$$

上式两端取模,得到,

$$|a_{kk}|\cdot|x_k|=\Big|\sum_{\substack{1\leqslant j\leqslant n\\ j\neq k}}a_{kj}x_j\Big|\leqslant\sum_{\substack{1\leqslant j\leqslant n\\ j\neq k}}|a_{kj}|\cdot|x_j|\leqslant|x_k|\cdot P_k.$$

因此 $|a_{kk}|\leqslant P_k$,这与 A 为行主对角占优方阵相矛盾.

如果 A 是列主对角占优方阵,注意到此时 A^{T} 为行主对角占优方阵,因此

$\det A = \det A^{\mathrm{T}} \neq 0.$ ■

利用 Levy-Desplanques 定理,立即得到 Gersgörin 1931 年所证明的圆盘定理.

定理 5.9.5(Gersgörin 圆盘定理) 任意 n 阶复方阵 $A = (a_{ij})$ 的特征值一定落在复平面上 n 个圆盘

$$|z - a_{ii}| \leqslant P_i$$

的并集内,其中 $i = 1, 2, \ldots, n$.

证明 设 λ_0 是方阵 A 的特征值,则 $\varphi(\lambda_0) = \det(\lambda_0 I_{(n)} - A) = 0$. 因此根据 Levy-Desplanques 定理,方阵 $\lambda_0 I_{(n)} - A$ 不是行主对角占优方阵. 所以至少有某个 i,使得

$$|\lambda_0 - a_{ii}| \leqslant P_i.$$ ■

Gersgörin 圆盘定理可以推广,先证明 Levy-Desplanques 定理的一个推广.

定理 5.9.6 设 $A = (a_{ij})$ 是 n 阶复方阵,且对任意 $1 \leqslant i \neq j \leqslant n$,

$$|a_{ii}| \cdot |a_{jj}| > P_i P_j, \tag{5.9.6}$$

则 $\det A \neq 0$.

证明 设 $\det A = 0$,则齐次线性方程组 $Ax = 0$ 有非零解

$$x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}.$$

设 x_r 与 x_s 满足($i = 1, \ldots, r-1, r, \ldots, n$):

$$|x_r| \geqslant |x_s| \geqslant |x_i|.$$

如果 $x_s = 0$,则 $x_i = 0$. 因此方程组 $Ax = 0$ 的第 r 个方程转化为 $a_{rr} x_r = 0$. 但 x 非零,所以 $x_r \neq 0$. 因此 $a_{rr} = 0$. 这与条件 (5.9.6) 矛盾. 从而 $x_s \neq 0$.

由齐次线性方程组 $Ax = 0$ 中第 r 与 s 个方程得到,

$$|a_{rr}| \cdot |x_r| = \left| -\sum_{\substack{1 \leqslant j \leqslant n \\ j \neq r}} a_{rj} x_j \right| \leqslant \sum_{\substack{1 \leqslant j \leqslant n \\ j \neq r}} |a_{rj}| \cdot |x_j| \leqslant |x_s| \cdot P_r,$$

$$|a_{ss}| \cdot |x_s| = \left| -\sum_{\substack{1 \leqslant j \leqslant n \\ j \neq s}} a_{sj} x_j \right| \leqslant \sum_{\substack{1 \leqslant j \leqslant n \\ j \neq s}} |a_{sj}| \cdot |x_j| \leqslant |x_r| \cdot P_s.$$

因此 $|a_{rr}| \cdot |a_{ss}| \leqslant P_r P_s$,与这与条件 (5.9.6) 矛盾. 这就证明,$\det A \neq 0$. ■

定理 5.9.7 任意 n 阶复方阵 $A = (a_{ij})$ 的特征值一定落在复平面上 $\dfrac{1}{2} n(n-1)$ 个 Cassini 卵形区域

$$|z - a_{ii}| \cdot |z - a_{jj}| \leqslant P_i P_j$$

的并集内,其中 $1 \leqslant i \neq j \leqslant n$.

证明 设 λ_0 是方阵 A 的特征值,则 $\varphi(\lambda_0) = \det(\lambda_0 I_{(n)} - A) = 0$. 由定理 5.9.6,不等式

$$|\lambda_0 - a_{ii}| \cdot |\lambda_0 - a_{jj}| > P_i P_j$$

不可能对所有的 i, j 成立. 因此必存在某对 i, j, $1 \leqslant i \neq j \leqslant n$,使得

$$|\lambda_0 - a_{ii}| \cdot |\lambda_0 - a_{jj}| \leqslant P_i P_j.$$ ■

应当指出,Hirsch 定理与 Gersgörin 圆盘定理不但本身在估计特征值时具有重要的意义,而且它们的证明也为处理有关特征值的问题提供典型的方法.

例 5.9.1 满足

$$\overline{U}^{\mathrm{T}}U=U\overline{U}^{\mathrm{T}}=I_{(n)}$$

的 n 阶复方阵 U 称为**酉方阵**; 满足

$$O^{\mathrm{T}}O=OO^{\mathrm{T}}=I_{(n)}$$

的 n 阶实方阵 O 称为**实正交方阵**. 证明,酉方阵与实正交方阵的特征值的模为 1.

证明 设 x 是酉方阵 U 的属于特征值 λ_0 的特征向量,即 $Ux=\lambda_0 x$. 两端取共轭转置,得到 $\overline{x}^{\mathrm{T}}\overline{U}^{\mathrm{T}}=\overline{\lambda}_0\overline{x}^{\mathrm{T}}$. 因此,

$$\overline{x}^{\mathrm{T}}\overline{U}^{\mathrm{T}}Ux=|\lambda_0|^2\overline{x}^{\mathrm{T}}x.$$

因为 U 为酉方阵,故 $\overline{U}^{\mathrm{T}}U=I_{(n)}$. 于是得到,

$$\left(|\lambda_0|^2-1\right)\overline{x}^{\mathrm{T}}x=0.$$

由于 x 非零,所以 $\overline{x}^{\mathrm{T}}x>0$,因此 $|\lambda_0|^2=1$. 所以酉方阵 U 的特征值的模为 1.

对实正交方阵,证明完全相同. ■

习 题 5.9

1. 设复方阵 $A=\mathrm{diag}(a_1,a_2,\dots,a_n)$,$U$ 是酉方阵. 证明方阵 UA 的特征值 λ_0 满足

$$\min\left\{|a_1|,|a_2|,\dots,|a_n|\right\}\leqslant|\lambda_0|\leqslant\max\left\{|a_1|,|a_2|,\dots,|a_n|\right\}.$$

2. 证明,酉方阵 U 的任意一个子方阵 U_1 的特征值的模不大于 1.

3. 设 A 是 n 阶方阵,M 是 k 阶方阵,$k\leqslant n$,且存在 $n\times k$ 列满秩矩阵 P,使得 $AP=PM$. 证明,方阵 M 的特征值一定是方阵 A 的特征值.

4. 设 O 是奇数阶实正交方阵,且 $\det O=1$. 证明方阵 O 具有特征值 1.

5. 证明,行列式为 -1 的实正交方阵具有特征值 -1.

6. 设 A 与 B 是 n 阶实正交方阵,且 $\det A=-\det B$. 证明,$\det(A+B)=0$.

7. 设 A 是 n 阶实方阵,且方阵 $B=\dfrac{1}{2}(A+A^{\mathrm{T}})$ 的最大与最小特征值分别为 μ_1 与 μ_n. 证明方阵 A 的特征值 λ_0 的实部 $\mathrm{Re}\,\lambda_0$ 满足 $\mu_n\leqslant\mathrm{Re}\,\lambda_0\leqslant\mu_1$.

8. 设 $A=(a_{ij})$ 是 n 阶复方阵,且

$$m_A=\min_{1\leqslant i\leqslant n}\left\{|a_{ii}|-\sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}}|a_{ij}|\right\}>0.$$

证明,$|\det A|\geqslant(m_A)^n$.

第六章

Jordan 标准形

- Jordan 标准形理论,也即方阵在相似下的分类理论,可以说是线性代数的最深刻的地方.
- 由于同一线性变换在不同基下的矩阵表示彼此相似的方阵,因此,方阵在相似问题下的分类问题也就是寻求线性变换的最简单的矩阵表示问题.
- 研究 Jordan 标准形,有两种途径:几何的与矩阵的. 本章继续沿着上一章的几何方法,首先在 §6.1 中阐明如何将线性空间分解成线性变换的根子空间的直和,然后在 §6.2 与 §6.3 中,先是对幂零线性变换,而后是对一般的线性变换,阐述如何将每个根子空间再分解成循环子空间的直和,并用初等因子组来刻画每个循环子空间的构造. 至此,寻求线性变换的最简单的矩阵表示问题,也即方阵在相似下的分类问题也已解决.
- 由于初等因子组的计算利用矩阵比较方便,也由于利用矩阵方法来表示方阵在相似下的分类问题具有同等重要的价值,所以 §6.4 与 §6.5 通过 λ 矩阵重新讨论了 Jordan 标准形理论,并得到计算初等因子组的有效方法.
- §6.6 通过例子说明,如何利用这一深刻的结果,也即方阵的 Jordan 标准形来解决线性代数的种种问题.
- 以上的讨论是在复数域上进行的. §6.7 讨论了实方阵在实数域上的相似分类问题.

§6.1 根子空间

设 V 是 n 维复线性空间,$\mathscr{A}: V \to V$ 是线性变换,λ_0 是线性变换 $\mathscr{A}$ 的特征值. 根据 §5.8,线性变换 $\mathscr{A}$ 的属于特征值 λ_0 的特征子空间为

$$V_{\lambda_0} = \left\{\alpha \in V \,\middle|\, \mathscr{A}(\alpha) = \lambda_0\alpha\right\} = \left\{\alpha \in V \,\middle|\, (\mathscr{A} - \lambda_0\mathscr{I})(\alpha) = 0\right\},$$

其中 $\mathscr{I}: V \to V$ 是单位变换. 也就是说,V_{λ_0} 是线性变换 $\mathscr{A} - \lambda_0\mathscr{I}$ 的核.

设 m 是正整数. 考虑线性变换 $(\mathscr{A} - \lambda_0\mathscr{I})^m$ 的核.

$$W_{\lambda_0}^{(m)} = \operatorname{Ker}(\mathscr{A} - \lambda_0\mathscr{I})^m = \left\{\alpha \in V \,\middle|\, (\mathscr{A} - \lambda_0\mathscr{I})^m(\alpha) = 0\right\}.$$

显然,$W_{\lambda_0}^{(m)}$ 是线性变换 $\mathscr{A}$ 的不变子空间,并且当 $k < \ell$ 时,$W_{\lambda_0}^{(k)} \subseteq W_{\lambda_0}^{(\ell)}$. 于是得到线性变换 $\mathscr{A}$ 的不变子空间的上升序列:

$$W_{\lambda_0}^{(1)} \subseteq W_{\lambda_0}^{(2)} \subseteq \cdots \subseteq W_{\lambda_0}^{(m)} \subseteq \cdots \subseteq V. \tag{6.1.1}$$

和例 5.4.2 的证明相仿,可以证明,对于线性变换 $\mathscr{A}$ 的不变子序列 (6.1.1),存在

某个正整数 k,使得

$$W_{\lambda_0}^{(1)} \subsetneqq W_{\lambda_0}^{(2)} \subsetneqq \cdots \subsetneqq W_{\lambda_0}^{(k-1)} \subsetneqq W_{\lambda_0}^{(k)} = W_{\lambda_0}^{(k+1)} = \cdots. \tag{6.1.2}$$

因此,

$$W_{\lambda_0} \coloneqq \bigcup_{m=1}^{\infty} W_{\lambda_0}^{(m)} = W_{\lambda_0}^{(k)}. \tag{6.1.3}$$

定义 6.1.1 设 λ_0 是线性变换 $\mathscr{A}: V \to V$ 的特征值. 则所有线性变换 $(\mathscr{A} - \lambda_0\mathscr{I})^m$ 的核 $\mathrm{Ker}(\mathscr{A}-\lambda_0\mathscr{I})^m$ 的并集称为线性变换 $\mathscr{A}$ 的属于特征值 λ_0 的**根子空间**,记为 W_{λ_0}.

根子空间 W_{λ_0} 中非零向量称为线性变换 $\mathscr{A}$ 的属于特征值 λ_0 的**根向量**.

由定义可以看出,根子空间 $W_{\lambda_0} = W_{\lambda_0}^{(k)}$,因此它是线性变换 $\mathscr{A}$ 的不变子空间. 设根向量 $\alpha \in W_{\lambda_0}$,则由 (6.1.2) 与 (6.1.3),存在正整数 ℓ,使得 $\alpha \in W_{\lambda_0}^{(\ell)}$,但 $\alpha \notin W_{\lambda_0}^{(\ell-1)}$,也即 $(\mathscr{A} - \lambda_0\mathscr{I})^{\ell}(\alpha) = 0$,但 $(\mathscr{A} - \lambda_0\mathscr{I})^{\ell-1}(\alpha) \neq 0$. 正整数 ℓ 称为**根向量 α 的次数**. 显然,1 次根向量即是特征向量,而且特征子空间 $V_{\lambda_0} = W_{\lambda_0}^{(1)} \subseteq W_{\lambda_0}$.

关于根子空间,下面的定理成立.

定理 6.1.1 设 λ_0 是线性变换 $\mathscr{A}: V \to V$ 的特征值,W_{λ_0} 是线性变换 $\mathscr{A}$ 的属于 λ_0 的根子空间,并且正整数 k 满足,

$$W_{\lambda_0}^{(k-1)} \subsetneqq W_{\lambda_0}^{(k)} = W_{\lambda_0},$$

则线性变换 $\mathscr{A}$ 在根子空间 W_{λ_0} 上的限制 $\mathscr{A}\big|_{W_{\lambda_0}}$ 的最小多项式为 $(\lambda - \lambda_0)^k$.

证明 因为 $W_{\lambda_0} = W_{\lambda_0}^{(k)}$,因此对任意 $\alpha \in W_{\lambda_0}$,均有

$$\alpha \in W_{\lambda_0}^{(k)} = \mathrm{Ker}(\mathscr{A} - \lambda_0\mathscr{I})^k,$$

即

$$(\mathscr{A} - \lambda_0\mathscr{I})^k\big|_{W_{\lambda_0}}(\alpha) = (\mathscr{A} - \lambda_0\mathscr{I})^k(\alpha) = 0.$$

所以 $(\lambda - \lambda_0)^k$ 是 $\mathscr{A}\big|_{W_{\lambda_0}}$ 的化零多项式. 于是 $\mathscr{A}\big|_{W_{\lambda_0}}$ 的最小多项式应为 $(\lambda - \lambda_0)^{\ell}$,$1 \leqslant \ell \leqslant k$.

如果 $\ell < k$,则 $\ell \leqslant k-1$. 因此对任意 $\alpha \in W_{\lambda_0}$,将有

$$(\mathscr{A} - \lambda_0\mathscr{I})^{k-1}(\alpha) = (\mathscr{A} - \lambda_0\mathscr{I})^{k-\ell-1}\big((\mathscr{A} - \lambda_0\mathscr{I})^{\ell}(\alpha)\big) = (\mathscr{A} - \lambda_0\mathscr{I})^{k-\ell-1}(0) = 0.$$

所以 $\alpha \in W_{\lambda_0}^{(k-1)}$,即 $W_{\lambda_0}^{(k)} = W_{\lambda_0} \subseteq W_{\lambda_0}^{(k-1)}$. 因此 $W_{\lambda_0}^{(k)} = W_{\lambda_0}^{(k-1)}$,矛盾.

这就说明,$\mathscr{A}\big|_{W_{\lambda_0}}$ 的最小多项式为 $(\lambda - \lambda_0)^k$. ■

定理 6.1.2 设 λ_1 与 λ_2 是线性变换 $\mathscr{A}: V \to V$ 的不同特征值,则根子空间 W_{λ_1} 与 W_{λ_2} 的和 $W_{\lambda_1} + W_{\lambda_2}$ 是直和.

证明 只需证明,$W_{\lambda_1} \cap W_{\lambda_2} = 0$.

设 $0 \neq \alpha \in W_{\lambda_1} \cap W_{\lambda_2}$. 则 $\alpha \in W_{\lambda_1}$. 因此存在正整数 k,使得 $\beta = (\mathscr{A} - \lambda_1\mathscr{I})^{k-1}(\alpha) \neq 0$,但是 $(\mathscr{A} - \lambda_1\mathscr{I})^k(\alpha) = (\mathscr{A} - \lambda_1\mathscr{I})(\beta) = 0$,即 $\mathscr{A}(\beta) = \lambda_1\beta$.

另一方面,$\alpha \in W_{\lambda_2}$,而 W_{λ_2} 是线性变换 $\mathscr{A}$ 的不变子空间,显然 W_{λ_2} 也是线

性变换 $\lambda_1\mathscr{I}$ 的不变子空间,因此 W_{λ_2} 是线性变换 $\mathscr{A}-\lambda_1\mathscr{I}$ 的不变子空间,所以 $\beta\in W_{\lambda_2}$. 于是存在正整数 ℓ,使得 $(\mathscr{A}-\lambda_2\mathscr{I})^{\ell-1}(\beta)\neq 0$,而 $(\mathscr{A}-\lambda_2\mathscr{I})^{\ell}(\beta)=0$. 由于 $\mathscr{A}(\beta)=\lambda_1\beta$,因此

$$\begin{aligned}(\mathscr{A}-\lambda_2\mathscr{I})^{\ell}(\beta)&=\sum_{j=0}^{\ell}(-1)^{\ell-j}\mathrm{C}_{\ell}^{j}\lambda_2^{\ell-j}\mathscr{A}^{j}(\beta)\\&=\Big(\sum_{j=0}^{\ell}(-1)^{\ell-j}\mathrm{C}_{\ell}^{j}\lambda_2^{\ell-j}\lambda_1^{j}\Big)\beta=(\lambda_1-\lambda_2)^{\ell}\beta=0.\end{aligned}$$

因为 $\beta\neq 0$,故 $\lambda_1=\lambda_2$,与假设矛盾. 因此 $W_{\lambda_1}\cap W_{\lambda_2}=0$. ■

下面是本节的主要定理.

定理 6.1.3(空间第一分解定理) 设 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 是线性变换 $\mathscr{A}:V\to V$ 的全部不同地特征值,它们的代数重数分别为 $e_1,e_2,\ldots,e_t$,$e_1,e_2,\ldots,e_t\geqslant 1$,$e_1+e_2+\cdots+e_t=n$,即线性变换 $\mathscr{A}$ 的特征多项式

$$\varphi(\lambda)=(\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_t)^{e_t},$$

则

$$V=W_{\lambda_1}\oplus W_{\lambda_2}\oplus\cdots\oplus W_{\lambda_t},$$

其中 W_{λ_j} 是线性变换 $\mathscr{A}$ 的属于特征值 λ_j 的根子空间. 并且线性变换 $\mathscr{A}$ 在 W_{λ_j} 上的限制 $\mathscr{A}\big|_{W_{\lambda_j}}$ 的特征多项式为 $\varphi_j(\lambda)=(\lambda-\lambda_j)^{e_j}$,$j=1,2,\ldots,t$.

证明 首先证明,$V=W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_t}$. 记

$$f_j(\lambda)=\frac{\varphi(\lambda)}{\varphi_j(\lambda)}=(\lambda-\lambda_1)^{e_1}\cdots(\lambda-\lambda_{j-1})^{e_{j-1}}(\lambda-\lambda_{j+1})^{e_{j+1}}\cdots(\lambda-\lambda_t)^{e_t},$$

其中 $j=1,2,\ldots,t$,并约定 $e_0=0$. 显然多项式 $f_1(\lambda),f_2(\lambda),\ldots,f_t(\lambda)$ 是互素的,因此存在多项式 $g_1(\lambda),g_2(\lambda),\ldots,g_t(\lambda)$,使得

$$g_1(\lambda)f_1(\lambda)+g_2(\lambda)f_2(\lambda)+\cdots+g_t(\lambda)f_t(\lambda)=1.$$

于是 $g_1(\mathscr{A})f_1(\mathscr{A})+g_2(\mathscr{A})f_2(\mathscr{A})+\cdots+g_t(\mathscr{A})f_t(\mathscr{A})=\mathscr{I}$. 设 $\alpha\in V$,则

$$\alpha=\mathscr{I}(\alpha)=g_1(\mathscr{A})f_1(\mathscr{A})(\alpha)+g_2(\mathscr{A})f_2(\mathscr{A})(\alpha)+\cdots+g_t(\mathscr{A})f_t(\mathscr{A})(\alpha).$$

记 $\alpha_j=g_j(\mathscr{A})f_j(\mathscr{A})(\alpha)$,$j=1,2,\ldots,t$. 因为

$$(\mathscr{A}-\lambda_j\mathscr{I})^{e_j}(\alpha_j)=\big(\varphi_j(\mathscr{A})g_j(\mathscr{A})f_j(\mathscr{A})\big)(\alpha)=g_j(\mathscr{A})\big(\varphi(\mathscr{A})(\alpha)\big).$$

而 $\varphi(\lambda)$ 是线性变换 $\mathscr{A}$ 的特征多项式,所以由 Cayley-Hamilton 定理,$\varphi(\mathscr{A})$ 是零变换. 因此 $(\mathscr{A}-\lambda_j\mathscr{I})^{e_j}(\alpha_j)=g_j(\mathscr{A})(0)=0$,即 $\alpha_j\in W_{\lambda_j}$.

这表明,$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_t\in W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_t}$. 于是

$$V=W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_t}.$$

其次证明,和 $W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_t}$ 是直和. 为此只需证明,对任意 ℓ,$1\leqslant\ell\leqslant t-1$,

$$\big(W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_\ell}\big)\cap W_{\lambda_{\ell+1}}=0. \tag{6.1.4}$$

对 ℓ 用归纳法. 当 $\ell=1$ 时,由定理 6.1.2,结论成立.

现在假设式 (6.1.4) 对 $\ell \leqslant m-1$ 成立. 下面证明式 (6.1.4) 对 m 成立.

设 $\alpha \neq 0, \alpha \in \left(W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_m}\right) \cap W_{\lambda_{m+1}}$. 由于 $\alpha \in W_{\lambda_{m+1}}$,故存在正整数 k,使得 $\beta=(\mathscr{A}-\lambda_{m+1}\mathscr{I})^{k-1}(\alpha) \neq 0$,但 $(\mathscr{A}-\lambda_{m+1}\mathscr{I})^k(\alpha)=0$. 因此 $\mathscr{A}(\beta)=\lambda_{m+1}\beta$.

另一方面,$\alpha \in W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_m}$. 由于 W_{λ_j} 是线性变换 $\mathscr{A}$ 的不变子空间,因此 $W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_t}$ 也是线性变换 $\mathscr{A}$ 的不变子空间. 显然 $W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_t}$ 也是线性变换 $\lambda_{m+1}\mathscr{I}$ 的不变子空间. 所以

$$\beta=(\mathscr{A}-\lambda_{m+1}\mathscr{I})^{k-1}(\alpha) \in W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_m},$$

而且

$$\mathscr{A}(\beta) \in W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_m}.$$

因此 $\beta=\beta_1+\beta_2+\cdots+\beta_m, \beta_j \in W_{\lambda_j}, j=1,2,\ldots,m$,且

$$\mathscr{A}(\beta)=\mathscr{A}(\beta_1)+\mathscr{A}(\beta_2)+\cdots+\mathscr{A}(\beta_m),$$

其中 $\mathscr{A}(\beta_j) \in W_{\lambda_j}$. 由于 $\mathscr{A}(\beta)=\lambda_{m+1}\beta$,故

$$\mathscr{A}(\beta)=\lambda_{m+1}\beta_1+\lambda_{m+1}\beta_2+\cdots+\lambda_{m+1}\beta_m.$$

由归纳假设,

$$W_{\lambda_1}+W_{\lambda_2}+\cdots+W_{\lambda_m}=W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_m},$$

因此 $\mathscr{A}(\beta)$ 分解为 $W_{\lambda_1}, W_{\lambda_2}, \ldots, W_{\lambda_m}$ 中向量之和的方式唯一,所以

$$\mathscr{A}(\beta_j)=\lambda_{m+1}\beta_j,$$

于是 $\beta_j \in W_{\lambda_j} \cap W_{\lambda_{m+1}}$. 由定理 6.1.2,$\beta_j=0$. 也即 $\beta=0$. 矛盾. 因此 $\alpha=0$. 这就证明,式 (6.1.4) 对 $\ell=m$ 成立.

最后证明,线性变换 $\mathscr{A}|_{W_{\lambda_j}}$ 的特征多项式为 $\varphi_j(\lambda)=(\lambda-\lambda_j)^{e_j}, j=1,2,\ldots,t$.

根据定理 6.1.1 以及定理 5.7.5,除 λ_j 外,线性变换 $\mathscr{A}|_{W_{\lambda_j}}$ 不可能有其它的特征值,所以线性变换 $\mathscr{A}|_{W_{\lambda_j}}$ 的特征多项式 $\varphi_j(\lambda)$ 应具有形式 $(\lambda-\lambda_j)^{m_j}$. 由于

$$V=W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_t},$$

因此

$$\mathscr{A}=\mathscr{A}|_{W_{\lambda_1}} \oplus \mathscr{A}|_{W_{\lambda_2}} \oplus \cdots \oplus \mathscr{A}|_{W_{\lambda_t}}.$$

于是线性变换 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)=\varphi_1(\lambda)\varphi_2(\lambda)\cdots\varphi_t(\lambda)$,即

$$\begin{aligned}\varphi(\lambda)&=(\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_t)^{e_t}\\&=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t}.\end{aligned}$$

所以 $\varphi_j(\lambda)=(\lambda-\lambda_j)^{e_j}, j=1,2,\ldots,t$. ■

由定理 6.1.3 的证明可以得到,线性变换 $\mathscr{A}$ 的属于特征值 λ_0 的根子空间 W_{λ_0} 的维数等于特征值 λ_0 的代数重数.

习 题 6.1

1. 举例说明,最小多项式相同的 4 阶幂零方阵 A 与 B 不一定相似.

2. 设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A，求线性变换 $\mathscr{A}$ 的特征值和根子空间.

(1) $A = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$;　　(2) $A = \begin{pmatrix} 0 & -2 & 3 & 2 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 2 & 1 \\ 1 & -1 & 0 & 1 \end{pmatrix}$.

3. 证明，n 维复线性空间 V 的线性变换 $\mathscr{A}$ 可对角化当且仅当所有根向量都是特征向量.

4. 证明，n 维复线性空间 V 的非零向量都是线性变换 $\mathscr{A}$ 的根向量当且仅当线性变换 $\mathscr{A}$ 的特征值都相等.

5. 证明，n 维复线性空间 V 的线性变换 $\mathscr{A}$ 的属于不同特征值的根向量线性无关.

6. 证明，3 阶复方阵 A 与 B 相似当且仅当方阵 A 与 B 具有相同的特征多项式与最小多项式.

7. (Fitting) 设 A 是属于 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换. 证明，存在线性变换 $\mathscr{A}$ 的不变子空间 V_1 和 V_2，使得 $V = V_1 \oplus V_2$，并且线性变换 $\mathscr{A}$ 在 V_1 上的限制 $\mathscr{A}\big|_{V_1}$ 是可逆的，而在 V_2 上的限制 $\mathscr{A}\big|_{V_2}$ 是幂零的. 简单地说，任意线性变换 $\mathscr{A}$ 都可以分解为可逆线性变换 $\mathscr{A}\big|_{V_1}$ 与幂零线性变换 $\mathscr{A}\big|_{V_2}$ 的直和. (注意，如果数域 $\mathbb{F}$ 是复数域，则本题可用第一分解定理加予证明，这里要求给出一个不用第一分解定理的证明)

8. 利用上题证明空间第一分解定理.

9. 设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 的最小多项式

$$d(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}\cdots(\lambda - \lambda_t)^{m_t},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_t$ 是线性变换 $\mathscr{A}$ 的全部不同特征值. 并设 W_j 是线性变换 $(\mathscr{A} - \lambda_j\mathscr{I})^{m_j}$ 的核. 证明，

(1) 根子空间 $W_{\lambda_j} = W_j$, $j = 1, 2, \ldots, t$;

(2) 线性变换 $\mathscr{A}$ 在 W_j 上的限制 $\mathscr{A}\big|_{W_j}$ 的最小多项式为 $(\lambda - \lambda_j)^{m_j}$, $j = 1, 2, \ldots, t$.

10. 设数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换 $\mathscr{A}$ 的最小多项式为 $d(\lambda) = p_1^{m_1}(\lambda)p_2^{m_2}(\lambda)\cdots p_t^{m_t}(\lambda)$，其中 $p_1(\lambda), p_2(\lambda), \ldots, p_t(\lambda)$ 是数域 $\mathbb{F}$ 上互不相同的首一多项式. 并设 W_j 是线性变换 $p_j^{m_j}(\mathscr{A})$ 的核. 证明，

(1) $V = W_1 \oplus W_2 \oplus \cdots \oplus W_t$;

(2) 线性变换 $\mathscr{A}$ 在 W_j 上的限制 $\mathscr{A}\big|_{W_j}$ 的最小多项式为 $p_j^{m_j}(\lambda)$, $j = 1, 2, \ldots, t$.

11. 设 3 维实向量空间 $\mathbb{R}^3$ 的线性变换 $\mathscr{A}$ 在标准基下的方阵为

$$\begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{pmatrix}.$$

记线性变换 $\mathscr{A}$ 的最小多项式 $d(\lambda) = p_1(\lambda)p_2(\lambda)$，其中 $p_1(\lambda), p_2(\lambda)$ 是实数域 $\mathbb{R}$ 上首一不可约多项式. 设 W_j 是线性变换 $p_j(\mathscr{A})$ 的核, $j = 1, 2$.

(1) 分别求子空间 W_1 与 W_2 的基;

(2) 求线性变换 $\mathscr{A}\big|_{W_1}$ 与 $\mathscr{A}\big|_{W_2}$ 分别在所求基下的方阵 A_j, $j = 1, 2$.

12. 设 3 维实向量空间 $\mathbb{R}^3$ 的线性变换 $\mathscr{A}$ 在标准基下的方阵为

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}.$$

求可对角化线性变换 $\mathscr{D}$ 与幂零变换 $\mathscr{N}$，使得 $\mathscr{A} = \mathscr{D} + \mathscr{N}$，且 $\mathscr{D}\mathscr{N} = \mathscr{N}\mathscr{D}$.

§6.2 循环子空间

设 V 是数域 $\mathbb{F}$ 上 n 维线性空间. 对于线性变换 $\mathscr{A}: V \to V$,除特征子空间与根子空间外,还有另一种不变子空间,这就是由向量生成的循环子空间.

定义 6.2.1 设 α_0 是 V 的非零向量,V 中的线性变换 $\mathscr{A}$ 的所有包含 α_0 的不变子空间的交称为**由向量 $\boldsymbol{\alpha_0}$ 生成的(相对线性变换 $\mathscr{A}$ 的)循环子空间**,记为 C_0.

由定义可以看出,C_0 是线性变换 $\mathscr{A}$ 的包含 α_0 的最小不变子空间. 即如果线性变换 $\mathscr{A}$ 的不变子空间 U 包含向量 α_0,则 U 也包含 C_0. 关于循环子空间 C_0,有

命题 6.2.1 由向量 α_0 生成的循环子空间 C_0 即是由向量 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^m(\alpha_0)$, $\dots$ 生成的子空间.

证明 由向量 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^m(\alpha_0), \dots$ 生成的子空间记为 U. 由于 C_0 是线性变换 $\mathscr{A}$ 的不变子空间,因此 $\mathscr{A}(\alpha_0), \mathscr{A}^2(\alpha_0), \dots, \mathscr{A}^m(\alpha_0), \dots \in C_0$. 所以 $U \subseteq C_0$.

反之,设 $\alpha \in U$,则存在 $a_1, a_2, \dots, a_k \in \mathbb{F}$,使得

$$\alpha = a_1\mathscr{A}^{m_1}(\alpha_0) + a_2\mathscr{A}^{m_2}(\alpha_0) + \cdots + a_k\mathscr{A}^{m_k}(\alpha_0).$$

因此

$$\mathscr{A}(\alpha) = a_1\mathscr{A}^{m_1+1}(\alpha_0) + a_2\mathscr{A}^{m_2+1}(\alpha_0) + \cdots + a_k\mathscr{A}^{m_k+1}(\alpha_0) \in U.$$

这表明,U 是线性空间 $\mathscr{A}$ 的不变子空间. 从而,$\alpha_0 \in U$. 因此 $C_0 \subseteq U$. ■

定义 6.2.2 设 $f(\lambda)$ 是数域 $\mathbb{F}$ 上关于未定元 λ 的非零多项式,$\alpha_0 \in V$. 如果 $f(\mathscr{A})(\alpha_0) = 0$,则 $f(\lambda)$ 称为**向量 $\boldsymbol{\alpha_0}$(相对线性变换 $\mathscr{A}$)的化零多项式**.

由定义可以看出,线性变换 $\mathscr{A}$ 的任意一个化零多项式都是向量 α_0(相对线性变换 $\mathscr{A}$)的化零多项式. 反之却不尽然.

定义 6.2.3 向量 α_0 的所有(相对线性变换 $\mathscr{A}$)化零多项式中次数最小的首一多项式称为**向量 $\boldsymbol{\alpha_0}$(相对线性变换 $\mathscr{A}$)的最小多项式**,记为 $d_{\alpha_0}(\lambda)$,或简记为 $d_0(\lambda)$.

向量 α_0 的所有化零多项式的次数集合记为 M. 由于线性变换 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 是 $\mathscr{A}$ 的化零多项式,因此 $d(\lambda)$ 是向量 α_0 的化零多项式,即 $\deg g(\lambda) \in M$,所以集合 M 非空.

这表明,集合 M 一定存在一个最小的数,记为 k. 于是存在向量 α_0 的一个次数为 k 的化零多项式 $g(\lambda)$. 多项式 $a^{-1}g(\lambda)$ 即是向量 α_0 的最小多项式,其中 a 是多项式 $g(\lambda)$ 的首项系数. 也就是说,向量 α_0 的最小多项式总是存在的.

容易证明,向量 α_0 的最小多项式 $d_0(\lambda)$ 整除向量 α_0 的任意一个化零多项式,而且向量 α_0 的最小多项式 $d_0(\lambda)$ 是唯一的.

因为由向量 α_0 生成的循环子空间 C_0 是线性变换 $\mathscr{A}$ 的不变子空间,所以可以

考虑线性变换 $\mathscr{A}$ 在 C_0 上的限制 $\mathscr{A}\big|_{C_0}$. 关于 $\mathscr{A}\big|_{C_0}$,有

命题 6.2.2　设 $\mathscr{A}: V \to V$ 是线性变换,C_0 是向量 α_0 生成的(相对 $\mathscr{A}$ 的)循环子空间,$d_0(\lambda)$ 是向量 α_0 的最小多项式,且 $\deg d_0(\lambda) = k$,则

(1) $\dim C_0 = k$,$\{\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)\}$是 C_0 的基,且 $\mathscr{A}\big|_{C_0}$ 在这组基下的方阵为

$$A = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 & 0 & -a_{k-2} \\ 0 & \cdots & 0 & 1 & -a_{k-1} \end{pmatrix}_{k\times k},$$

其中 $d_0(\lambda) = a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k$;

(2) $d_0(\lambda)$是 $\mathscr{A}\big|_{C_0}$ 的特征多项式与最小多项式.

证明　首先证明 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)$ 线性无关. 设有不全为零的 $b_0, b_1, \dots, b_{k-1} \in \mathbb{F}$,使得

$$b_0\alpha_0 + b_1\mathscr{A}(\alpha_0) + \cdots + b_{k-1}\mathscr{A}^{k-1}(\alpha_0) = 0.$$

$b_0, b_1, \dots, b_{k-1}$ 诸数中由后往前数头一个不为零的 b_ℓ 的下标记为 ℓ,即 $b_\ell \neq 0$,但 $b_{\ell+1} = b_{\ell+2} = \cdots = b_{k-1} = 0$,则

$$(b_0\mathscr{I} + b_2\mathscr{A}^2 + \cdots + b_\ell\mathscr{A}^\ell)(\alpha_0) = 0.$$

因此 $f(\lambda) = b_0 + b_1\lambda + \cdots + b_\ell\lambda^\ell$ 是向量 α_0 的一个化零多项式,且 $\deg f(\lambda) = \ell \leqslant k-1 < k$,这与 $d_0(\lambda)$ 关于次数的最小性矛盾.

其次证明,对任意 $\ell \geqslant k$,向量 $\mathscr{A}^\ell(\alpha_0)$ 可由 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)$ 线性表出. 事实上,由于 $d_0(\lambda)$ 是向量 α_0 的最小多项式,因此

$$d(\mathscr{A})(\alpha_0) = a_0\alpha_0 + a_1\mathscr{A}(\alpha_0) + \cdots + a_{k-1}\mathscr{A}^{k-1}(\alpha_0) + \mathscr{A}^k(\alpha_0) = 0.$$

所以

$$\mathscr{A}^k(\alpha_0) = -a_0\alpha_0 - a_1\mathscr{A}(\alpha_0) - \cdots - a_{k-1}\mathscr{A}^{k-1}(\alpha_0). \tag{6.2.1}$$

即 $\mathscr{A}^k(\alpha_0)$ 可由 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)$ 线性表出. 即结论对 $\ell = k$ 成立. 假设结论对 ℓ 成立. 即设$\mathscr{A}^\ell(\alpha_0) = b_0\alpha_0 + b_1\mathscr{A}(\alpha_0) + \cdots + b_{k-1}\mathscr{A}^{k-1}(\alpha_0)$,其中 $b_0, b_1, \dots, b_{k-1} \in \mathbb{F}$. 则

$$\mathscr{A}^{\ell+1}(\alpha_0) = b_0\mathscr{A}(\alpha_0) + b_1\mathscr{A}^2(\alpha_0) + \cdots + b_{k-2}\mathscr{A}^{k-1}(\alpha_0) + b_{k-1}\mathscr{A}^k(\alpha_0).$$

将式 (6.2.1) 代入上式,即知 $\mathscr{A}^{\ell+1}(\alpha_0)$ 确实可由 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)$ 线性表出.

现在设 $\alpha \in C_0$,则 α 是有限个形如 $\mathscr{A}^{m_1}(\alpha_0), \mathscr{A}^{m_2}(\alpha_0), \dots, \mathscr{A}^{m_\ell}(\alpha_0)$ 的线性组合,其中 $m_1, m_2, \dots, m_\ell$ 是非负整数. 由于每个 $\mathscr{A}^{m_i}(\alpha_0)$ 都可由 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)$ 线性表出,所以 α 也可由 $\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)$ 线性表出.

综上所述,$\{\alpha_0, \mathscr{A}(\alpha_0), \dots, \mathscr{A}^{k-1}(\alpha_0)\}$ 是 C_0 的基,并且 $\dim C_0 = k$. 由于

$$\begin{aligned} &\mathscr{A}\big|_{C_0}(\alpha_0) = \mathscr{A}(\alpha_0), \\ &\mathscr{A}\big|_{C_0}\big(\mathscr{A}(\alpha_0)\big) = \mathscr{A}^2(\alpha_0), \end{aligned}$$

$$\begin{aligned}
&\cdots\cdots\cdots\cdots\\
&\mathscr{A}\big|_{C_0}\big(\mathscr{A}^{k-2}(\alpha_0)\big) = \mathscr{A}^{k-1}(\alpha_0),\\
&\mathscr{A}\big|_{C_0}\big(\mathscr{A}^{k-1}(\alpha_0)\big) = \mathscr{A}^{k}(\alpha_0),\\
&\qquad\qquad\qquad\qquad = -a_0\alpha_0 - a_1\mathscr{A}(\alpha_0) - \cdots - a_{k-1}\mathscr{A}^{k-1}(\alpha_0),
\end{aligned}$$

所以

$$\mathscr{A}\big|_{C_0}\big(\alpha_0, \mathscr{A}(\alpha_0), \ldots, \mathscr{A}^{k-1}(\alpha_0)\big) = \big(\alpha_0, \mathscr{A}(\alpha_0), \ldots, \mathscr{A}^{k-1}(\alpha_0)\big)A.$$

这就证明了结论 (1).

由结论 (1) 容易算得,$\mathscr{A}\big|_{C_0}$ 的特征多项式为

$$\varphi(\lambda) = \det\big(\lambda I_{(k)} - A\big) = a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k = d_0(\lambda).$$

由于 $\mathscr{A}\big|_{C_0}$ 的最小多项式为 $d(\lambda)$ 整除 $\mathscr{A}\big|_{C_0}$ 的特征多项式 $d_0(\lambda)$,而 α_0 的最小多项式 $d_0(\lambda)$ 整除 $\mathscr{A}\big|_{C_0}$ 的化零多项式,特别,$d_0(\lambda)$整除 $\mathscr{A}\big|_{C_0}$ 的最小多项式 $d(\lambda)$,即 $d(\lambda) \mid d_0(\lambda)$,且 $d_0(\lambda) \mid d(\lambda)$. 由于 $d_0(\lambda)$ 与 $d(\lambda)$ 都是首一多项式,因此 $d(\lambda) = d_0(\lambda)$. ∎

现在讨论相对幂零变换的循环子空间.

我们知道,所谓线性变换 $\mathscr{A}: V \to V$ 是**幂零变换**,是指存在正整数 m,使得 $\mathscr{A}^m = 0$. 使 $\mathscr{A}^m = 0$ 的最小正整数 m 称为幂零变换 $\mathscr{A}$ 的**幂零指数**,而 $\mathscr{A}$ 也称为 m 次幂零变换. 显然,1 次幂零变换是零变换. 因此非零的幂零变换的幂零指数总是大于或等于 2.

定理 6.2.1 设 $\mathscr{A}: V \to V$ 是 m 次幂零变换,则

(1) 存在非零向量 $\alpha_1 \in V$,使得由 α_1 生成的循环子空间 C_1 的维数为 m,并且 $\mathscr{A}\big|_{C_1}$ 是 m 次幂零变换;

(2) 存在 $\mathscr{A}$ 的不变子空间 V_1,使得 $V = C_1 \oplus V_1$,并且 $\mathscr{A}\big|_{V_1}$ 也是幂零变换,它的幂零指数 $m_1 \leqslant m$.

证明 因为 $\mathscr{A}^m = 0$,$\mathscr{A}^{m-1} \neq 0$,故存在非零向量 $\alpha_1 \in V$,使得 $\mathscr{A}^{m-1}(\alpha_1) \neq 0$,但 $\mathscr{A}^m(\alpha_1) = 0$. 这表明 $f(\lambda) = \lambda^m$ 是 α_1 的化零多项式. 因此 α_1 的最小多项式 $d_1(\lambda) \big| \lambda^m$. 所以 $d_1(\lambda) = \lambda^\ell$,$1 \leqslant \ell \leqslant m$. 如果 $\ell \leqslant m-1$,则由 $\mathscr{A}^\ell(\alpha_1) = 0$ 得到,

$$\mathscr{A}^{m-1}(\alpha_1) = \mathscr{A}^{m-\ell-1}\big(\mathscr{A}^\ell(\alpha_1)\big) = \mathscr{A}^{m-\ell-1}(0) = 0,$$

这与 $\mathscr{A}^{m-1}(\alpha_1) \neq 0$ 矛盾. 因此 $\ell = m$. 即 $d_1(\lambda) = \lambda^m$. 由命题 6.2.2,$\dim C_1 = m$,且 $\mathscr{A}\big|_{C_1}$ 的最小多项式为 λ^m,也即 $\mathscr{A}\big|_{C_1}$ 是 m 次幂零变换. 结论 (1) 得证.

现在对幂零指数 m 用归纳法证明 (2).

当 $m = 1$ 时,$\mathscr{A}$ 是零变换. 因此对任意非零向量 $\alpha_1 \in V$,$\mathscr{A}(\alpha_1) = 0$,即由向量 α_1 生成的循环子空间 C_1 是 1 维的,而且 $\mathscr{A}\big|_{C_1}$ 是零变换. 由于 C_1 是 1 维的,故 $\{\alpha_1\}$ 是 C_1 的基. 设 $\{\alpha_2, \alpha_3, \ldots, \alpha_n\}$ 是 V 的基. V 中由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 生成的子空间记为 V_1. 显然,$V = C_1 \oplus V_1$,且 V_1 是 $\mathscr{A}$ 的不变子空间,而且 $\mathscr{A}\big|_{V_1}$ 是零变换,即 1 次幂零变换. 这就证明当 $m = 1$ 时结论 (2) 成立.

假设结论 (2) 对 $m-1$ 次幂零变换成立,下面证明结论 (2) 对 m 次幂零变换 $\mathscr{A}$ 也成立.

设 $U=\operatorname{Im}\mathscr{A}$,并记 $\beta_1=\mathscr{A}(\alpha_1)\in U$. 由于 $\operatorname{Im}\mathscr{A}$ 是 $\mathscr{A}$ 的不变子空间,因此 $\mathscr{A}\big|_U$ 有意义. 设 $\xi\in U$,则存在 $\eta\in V$,使得 $\xi=\mathscr{A}(\eta)$. 因此

$$(\mathscr{A}\big|_U)^{m-1}(\xi)=\mathscr{A}^{m-1}\big(\mathscr{A}(\eta)\big)=\mathscr{A}^m(\eta)=0,$$

并且

$$(\mathscr{A}\big|_U)^{m-2}(\beta_1)=\mathscr{A}^{m-2}\big(\mathscr{A}(\alpha_1)\big)=\mathscr{A}^{m-1}(\alpha_1)\neq 0.$$

所以 $(A\big|_U)^{m-2}\neq 0$,$(\mathscr{A}\big|_U)^{m-1}=0$,即 $\mathscr{A}\big|_U$ 是 $m-1$ 次幂零的.

在 U 中由向量 β_1 生成的(相对 $\mathscr{A}\big|_U$ 的)循环子空间 $\widetilde{C}_1$ 显然具有基

$$\left\{\beta_1,\mathscr{A}(\beta_1),\ldots,\mathscr{A}^{m-2}(\beta_1)\right\},$$

并且 $\mathscr{A}\big|_U$ 在 $\widetilde{C}_1$ 上的限制是 $m-1$ 次幂零的.

由归纳假设,存在 $\mathscr{A}\big|_U$ 的不变子空间 U_1,使得 $U=\widetilde{C}_1\oplus U_1$,而且 $\mathscr{A}\big|_U$ 在 U_1 上的限制是幂零的,其幂零指数不大于 $m-1$.

记 $\widetilde{V}_1=\left\{\alpha\in V\;\middle|\;\mathscr{A}(\alpha)\in U_1\right\}$. 先证明以下事实.

① $C_1\cap U_1=0$. 事实上,设 $\alpha\in C_1\cap U_1$. 由于 $\alpha\in C_1$,而 $\left\{\alpha_1,\mathscr{A}(\alpha_1),\ldots,\mathscr{A}^{m-1}(\alpha_1)\right\}$ 为 C_1 的基,所以 $\alpha=b_0\alpha_1+b_1\mathscr{A}(\alpha_1)+\cdots+b_{m-1}\mathscr{A}^{m-1}(\alpha_1)$. 因此

$$\mathscr{A}(\alpha)=b_0\beta_1+b_1\mathscr{A}(\beta_1)+\cdots+b_{m-2}\mathscr{A}^{m-2}(\beta_1)\in\widetilde{C}_1.$$

另一方面,由于 $\alpha\in U_1$,且 U_1 是 $\mathscr{A}\big|_U$ 的不变子空间,所以 $\mathscr{A}(\alpha)=\mathscr{A}\big|_U(\alpha)\in U_1$. 因此 $\mathscr{A}(\alpha)\in\widetilde{C}_1\cap U_1=0$,即

$$b_0\beta_1+b_1\mathscr{A}(\beta_1)+\cdots+b_{m-2}\mathscr{A}^{m-2}(\beta_1)=0,$$

而 $\left\{\beta_1,\mathscr{A}(\beta_1),\ldots,\mathscr{A}^{m-2}(\beta_1)\right\}$ 是 $\widetilde{C}_1$ 的基. 所以 $b_0=b_1=\cdots=b_{m-2}=0$. 于是

$$\alpha=b_{m-1}\mathscr{A}^{m-1}(\alpha_1)=b_{m-1}\mathscr{A}^{m-2}(\beta_1)=b_{m-1}\mathscr{A}\big(\mathscr{A}^{m-2}(\alpha_1)\big)\in\widetilde{C}_1\cap U_1=0.$$

因为 $\mathscr{A}^{m-2}(\beta_1)\neq 0$,故 $b_{m-1}=0$,所以 $\alpha=0$. 这就证明,$C_1\cap U_1=0$.

② $U_1\subseteq\widetilde{V}_1$. 事实上,由于 U_1 是 $\mathscr{A}\big|_U$ 的不变子空间,所以对任意 $\alpha\in U_1$,$\mathscr{A}\big|_U(\alpha)=\mathscr{A}(\alpha)\in U_1$. 因此 $U_1\subseteq\widetilde{V}_1$.

③ $V=C_1+\widetilde{V}_1$. 事实上,设 $\alpha\in V$,则 $\mathscr{A}(\alpha)\in U$. 由于 $U=\widetilde{C}_1\oplus U_1$,故 $\mathscr{A}(\alpha)=\xi+\eta$,其中 $\xi\in\widetilde{C}_1$,$\eta\in U_1$. 因为 $\xi\in\widetilde{C}_1$,故

$$\begin{aligned}\xi&=b_0\beta_1+b_1\mathscr{A}(\beta_1)+\cdots+b_{m-2}\mathscr{A}^{m-2}(\beta_1)\\&=\mathscr{A}\big(b_0\alpha_1+b_1\mathscr{A}(\alpha_1)+\cdots+b_{m-2}\mathscr{A}^{m-2}(\alpha_1)\big),\end{aligned}$$

其中 $b_0,b_1,\ldots,b_{m-2}\in\mathbb{F}$. 因此

$$\mathscr{A}\big(\alpha-b_0\alpha_1-b_1\mathscr{A}(\alpha_1)-\cdots-b_{m-2}\mathscr{A}^{m-2}(\alpha_1)\big)=\eta\in U_1,$$

故 $\alpha-b_0\alpha_1-b_1\mathscr{A}(\alpha_1)-\cdots-b_{m-2}\mathscr{A}^{m-2}(\alpha_1)=\zeta\in\widetilde{V}_1$. 也就是说,

$$\alpha=\big(b_0\alpha_1+b_1\mathscr{A}(\alpha_1)+\cdots+b_{m-2}\mathscr{A}^{m-2}(\alpha_1)\big)+\zeta,$$

其中 $b_0\alpha_1+b_1\mathscr{A}(\alpha_1)+\cdots+b_{m-2}\mathscr{A}^{m-2}(\alpha_1)\in C_1$. 这表明,$\alpha\in C_1+\widetilde{V}_1$. 因此 $V=C_1+\widetilde{V}_1$.

应当指出,$V=C_1+\widetilde{V}_1$ 并不一定是直和. 现在需要求出适合结论 (2) 的子空间 V_1. 因为 $U_1\cap(C_1\cap\widetilde{V}_1)=(U_1\cap C_1)\cap\widetilde{V}_1$,所以由 ① 可知

$$U_1\cap(C_1\cap\widetilde{V}_1)=0.$$

因此和 $U_1+(C_1\cap\widetilde{V}_1)$ 是直和. 由于 $U_1\subseteq\widetilde{V}_1, C_1\cap\widetilde{V}_1\subseteq\widetilde{V}_1$,因此

$$U_1\oplus(C_1\cap\widetilde{V}_1)\subseteq\widetilde{V}_1.$$

$U_1\oplus(C_1\cap\widetilde{V}_1)$ 在 $\widetilde{V}_1$ 中的补记为 W,即

$$W\oplus U_1\oplus(C_1\cap\widetilde{V}_1)=\widetilde{V}_1.$$

记 $V_1=W\oplus U_1$. 现在验证 V_1 是 $\mathscr{A}$ 的不变子空间,并且 $\mathscr{A}|_{V_1}$ 是幂零的,其幂零指数 $m_1\leqslant m$. 事实上,设 $\xi\in V_1\subseteq\widetilde{V}_1$,则

$$\mathscr{A}(\xi)\in U_1\subseteq W\oplus U_1=V_1.$$

所以 V_1 是 $\mathscr{A}$ 的不变子空间. 由于 $\xi\in V_1\subseteq V$,故 $\mathscr{A}^m(\xi)=0$. 所以 $\mathscr{A}|_{V_1}$ 是幂零的,它的幂零指数 $m_1\leqslant m$.

其次验证,$C_1\cap V_1=0$. 事实上,设 $\xi\in C_1\cap V_1$. 由于

$$C_1\cap V_1\subseteq V_1\subseteq\widetilde{V}_1,$$

所以 $\xi\in C_1\cap\widetilde{V}_1$. 而 $\xi\in V_1$,所以 $\xi\in V_1\cap(C_1\cap\widetilde{V}_1)$. 由于

$$\widetilde{V}_1=V_1\oplus(C_1\cap\widetilde{V}_1),$$

因此 $\xi=0$. 即 $C_1\cap V_1=0$.

最后验证,$V=C_1+V_1$. 事实上,设 $\xi\in V$. 由于 $V=C_1+\widetilde{V}_1$,所以 $\xi=\eta+\zeta$,其中 $\eta\in C_1,\zeta\in\widetilde{V}_1$. 但

$$\widetilde{V}_1=V_1\oplus(C_1\cap\widetilde{V}_1),$$

故 $\zeta=\zeta_1+\zeta_2$,其中 $\zeta_1\in V_1,\zeta_2\in(C_1\cap\widetilde{V}_1)$. 因此

$$\xi=\eta+\zeta_1+\zeta_2=(\eta+\zeta_2)+\zeta_1,$$

其中 $\eta+\zeta_2\in C_1,\zeta_1\in V_1$. 即 $\xi\in C_1+C_1$,所以 $V=C_1+V_1$.

这就证明,结论 (2) 对 m 次幂零变换 $\mathscr{A}$ 成立. ■

利用定理 6.2.1,并对线性空间 V 的维数用归纳法,即得下面的线性空间 V 分解为关于幂零变换 $\mathscr{A}$ 的循环子空间的直和的定理.

定理 6.2.2 设 $\mathscr{A}:V\to V$ 是 m 次幂零变换. 则

(1) 存在由非零向量 $\alpha_1,\alpha_2,\ldots,\alpha_k\in V$ 生成的循环子空间 $C_1,C_2,\ldots,C_k$,使得

$$V=C_1\oplus C_2\oplus\cdots\oplus C_k,$$

并且 $\mathscr{A}|_{C_j}$ 是 m_j 次幂零的,而 $m=m_1\geqslant m_2\geqslant\cdots\geqslant m_k$;

(2) $\left\{\alpha_j,\mathscr{A}(\alpha_j),\ldots,\mathscr{A}^{m_j-1}(\alpha_j)\right\}$ 是 C_j 的基,$j=1,2,\ldots,k$,它们的并构成 V 的基,而且幂零变换 $\mathscr{A}$ 在这组基下的方阵为

$$\operatorname{diag}\left(N_{(m_1)}^{\mathrm{T}}, N_{(m_2)}^{\mathrm{T}}, \ldots, N_{(m_k)}^{\mathrm{T}}\right), \tag{6.2.2}$$

其中 n 阶方阵 N_n^{T} 为

$$N_n^{\mathrm{T}} = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}_{n \times n}.$$

证明　结论 (1) 是定理 6.2.1 的自然推论. 这里只证明结论 (2).

由结论 (1), C_j 是由向量 α_j 生成的循环子空间, 且

$$V = C_1 \oplus C_2 \oplus \cdots \oplus C_k.$$

因此

$$\mathscr{A} = \mathscr{A}\big|_{C_1} \oplus \mathscr{A}\big|_{C_2} \oplus \cdots \oplus \mathscr{A}\big|_{C_k}.$$

由于 $\mathscr{A}\big|_{C_j}$ 是 m_j 次幂零的, 所以 $\mathscr{A}\big|_{C_j}$ 的最小多项式为 $d_j(\lambda) = \lambda^{m_j}$. 由命题 6.2.2, $\mathscr{A}\big|_{C_j}$ 在 C_j 的基 $\left\{\alpha_j, \mathscr{A}(\alpha_j), \ldots, \mathscr{A}^{m_j-1}(\alpha_j)\right\}$ 下的方阵为 $N_{(m_j)}^{\mathrm{T}}$. 因此, $\mathscr{A}$ 在给定的 V 的基下的方阵即为 (6.2.2). ■

定理 6.2.3　设 $\mathscr{A}: V \to V$ 是幂零变换, 而且 V 可按两种方式分解为相对 $\mathscr{A}$ 的循环子空间的直和, 即设

$$V = C_1 \oplus C_2 \oplus \cdots \oplus C_k = \widetilde{C}_1 \oplus \widetilde{C}_2 \oplus \cdots \oplus \widetilde{C}_\ell,$$

其中 C_j 是由向量 α_j 生成的循环子空间, $\mathscr{A}\big|_{C_j}$ 是 m_j 次幂零的, $m = m_1 \geqslant m_2 \geqslant \cdots \geqslant m_k$; 而 $\widetilde{C}_j$ 是由向量 $\widetilde{\alpha}_j$ 生成的循环子空间, $\mathscr{A}\big|_{\widetilde{C}_j}$ 是 $\widetilde{m}_i$ 次幂零的. 那么 $k = \ell$, 且可适当调整 $\widetilde{C}_i$ 的次序, 使得 $\widetilde{m}_i = m_i$.

定理 6.2.3 可以简述为: 对给定的幂零变换 $\mathscr{A}$, n 维线性空间 V 可以分解为相对 $\mathscr{A}$ 的循环子空间的直和, 而且分解中循环子空间的个数以及循环子空间的维数是由幂零变换 $\mathscr{A}$ 所唯一决定的.

证明　记 $\widetilde{m} = \max\left\{\widetilde{m}_1, \widetilde{m}_2, \ldots, \widetilde{m}_\ell\right\}$. 设 $\alpha \in V$, 则

$$\alpha = \xi_1 + \xi_2 + \cdots + \xi_\ell,$$

其中 $\xi_j \in \widetilde{C}_j$. 因此

$$\mathscr{A}^{\widetilde{m}}(\alpha) = \mathscr{A}^{\widetilde{m}}(\xi_1) + \mathscr{A}^{\widetilde{m}}(\xi_2) + \cdots + \mathscr{A}^{\widetilde{m}}(\xi_\ell).$$

但对任意 $j = 1, 2, \ldots, \ell$,

$$\mathscr{A}^{\widetilde{m}}(\xi_j) = \mathscr{A}^{\widetilde{m}-\widetilde{m}_j}\left(\mathscr{A}^{\widetilde{m}_j}(\xi_j)\right) = \mathscr{A}^{\widetilde{m}-\widetilde{m}_j}(0) = 0.$$

所以 $\mathscr{A}^{\widetilde{m}}(\alpha) = 0$, 即 $\mathscr{A}^{\widetilde{m}} = 0$. 另一方面, $\widetilde{m} = \widetilde{m}_i$. 由于 $\widetilde{m}_i$ 是 $\mathscr{A}\big|_{C_i}$ 的幂零指数, 不存在 $\xi \in \widetilde{C}_i \subseteq V$, 使得

$$\left(\mathscr{A}\big|_{\widetilde{C}_i}\right)^{\widetilde{m}-1}(\xi) = \mathscr{A}^{\widetilde{m}-1}(\xi) \neq 0.$$

因此 $\mathscr{A}^{\widetilde{m}-1} \neq 0$. 这表明, $\widetilde{m}$ 是 $\mathscr{A}$ 的幂零指数, 即 $\widetilde{m} = \widetilde{m}_i = m$. 把直和 $\widetilde{C}_1 \oplus \widetilde{C}_2 \oplus \cdots \oplus \widetilde{C}_\ell$

中的第一项 $\widetilde{C}_1$ 与第 i 项 $\widetilde{C}_i$ 对调后,可设 $\widetilde{C}_1$ 的维数为 $\widetilde{m}$.

记 $V_1 = C_2 \oplus C_2 \oplus \cdots \oplus C_k$, $\widetilde{V}_1 = \widetilde{C}_2 \oplus \widetilde{C}_2 \oplus \cdots \oplus \widetilde{C}_\ell$. 则

$$V = C_1 \oplus V_1 = \widetilde{C}_1 \oplus \widetilde{V}_1.$$

显然 V_1 与 $\widetilde{V}_1$ 都是 $\mathscr{A}$ 的不变子空间,而且 $\mathscr{A}|_{V_1}$ 与 $\mathscr{A}|_{\widetilde{V}_1}$ 都是幂零的,它们的幂零指数分别记为 r 与 $\widetilde{r}$. 于是

$$\mathscr{A}^r(V) = \mathscr{A}^r(C_1) \oplus \mathscr{A}^r(V_1) = \mathscr{A}^r(C_1) = \mathscr{A}^r(\widetilde{C}_1) \oplus \mathscr{A}^r(\widetilde{V}_1).$$

容易证明,

$$\dim\big(\mathscr{A}^r(C_1)\big) = m - r = \dim\big(\mathscr{A}^r(\widetilde{C}_1)\big).$$

因此由上式得到,$\dim\big(\mathscr{A}^r(\widetilde{V}_1)\big) = 0$,即 $\mathscr{A}^r(\widetilde{V}_1) = 0$. 所以 $(\mathscr{A}|_{\widetilde{V}_1})^r = 0$. 这表明,$\widetilde{r} \leqslant r$. 同理可证 $r \leqslant \widetilde{r}$. 即得 $r = \widetilde{r}$.

重复前面的证明过程,可以证明,$r = m_2$,并且存在某个 $\widetilde{m}_{i_2} = \widetilde{r} = m_2$,在直和 $\widetilde{V}_1 = \widetilde{C}_2 \oplus \cdots \oplus \widetilde{C}_\ell$ 中将第 2 项 $\widetilde{C}_2$ 与第 i_2 项 $\widetilde{C}_{i_2}$ 对调,即可设 $\widetilde{C}_2$ 的维数为 $\widetilde{m}_2 = m_2$. 如此继续即可证明 $\ell = k$,且 $\widetilde{m}_j = m_j$. ■

现在叙述与定理 6.2.2 和定理 6.2.3 相应的关于矩阵的定理. 设数域 $\mathbb{F}$ 上 n 阶方阵 A 满足 $A^m = 0$,其中 m 是某个正整数,则 A 称为幂零方阵. 使得 $A^m = 0$ 的最小正整数 m 称为方阵 A 的幂零指数. 关于幂零方阵,有

定理 6.2.4 设 A 是数域 $\mathbb{F}$ 上 n 阶幂零方阵,其幂零指数为 m,则方阵 A 在数域 $\mathbb{F}$ 上相似于 n 阶准对角方阵

$$J = \operatorname{diag}\big(N_{(m_1)}^{\mathrm{T}}, N_{(m_2)}^{\mathrm{T}}, \ldots, N_{(m_k)}^{\mathrm{T}}\big),$$

其中 $m = m_1 \geqslant m_2 \geqslant \cdots \geqslant m_k$.

证明 设 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的基. 则可以验证,由

$$\mathscr{A}(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1, \xi_2, \ldots, \xi_n)A$$

所确定的线性变换 $\mathscr{A}$ 是 m 次幂零的. 由定理 6.2.2,存在 V 的基 $\{\beta_1, \beta_2, \ldots, \beta_n\}$,使得

$$\mathscr{A}(\beta_1, \beta_2, \ldots, \beta_n) = (\beta_1, \beta_2, \ldots, \beta_n)\operatorname{diag}\big(N_{(m_1)}^{\mathrm{T}}, N_{(m_2)}^{\mathrm{T}}, \ldots, N_{(m_k)}^{\mathrm{T}}\big).$$

由于同一个线性变换在不同基下的方阵是相似的,所以定理 6.2.4 成立. ■

定理 6.2.5 设数域 $\mathbb{F}$ 上 n 阶幂零方阵 A 与 B 分别相似于下面的准对角方阵

$$J_A = \operatorname{diag}\big(N_{(m_1)}^{\mathrm{T}}, N_{(m_2)}^{\mathrm{T}}, \ldots, N_{(m_k)}^{\mathrm{T}}\big),$$
$$J_B = \operatorname{diag}\big(N_{(n_1)}^{\mathrm{T}}, N_{(n_2)}^{\mathrm{T}}, \ldots, N_{(n_\ell)}^{\mathrm{T}}\big),$$

其中 $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_k$, $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_\ell$. 则方阵相似当且仅当 $J_A = J_B$.

证明 充分性是显然的,这里只证必要性.

设 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的基,则由

$$\mathscr{A}(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1, \xi_2, \ldots, \xi_n)A$$

所确定的线性变换 $\mathscr{A}$ 是幂零的. 因为方阵 A 与方阵 B 相似,所以 $B = P^{-1}AP$,其中

P 是数域 $\mathbb{F}$ 上 n 阶可逆方阵. 记

$$(\beta_1,\beta_2,\ldots,\beta_n)=(\xi_1,\xi_2,\ldots,\xi_n)P.$$

则 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 是 V 的基,而且

$$\mathscr{A}(\beta_1,\beta_2,\ldots,\beta_n)=(\beta_1,\beta_2,\ldots,\beta_n)B.$$

由假设,$J_A=Q^{-1}AQ$,$J_B=R^{-1}BR$,其中 Q 与 R 是数域 $\mathbb{F}$ 上 n 阶方阵. 记

$$(\eta_1,\eta_2,\ldots,\eta_n)=(\xi_1,\xi_2,\ldots,\xi_n)Q,$$
$$(\zeta_1,\zeta_2,\ldots,\zeta_n)=(\beta_1,\beta_2,\ldots,\beta_n)R.$$

则

$$\mathscr{A}(\eta_1,\eta_2,\ldots,\eta_n)=(\eta_1,\eta_2,\ldots,\eta_n)J_A,$$
$$\mathscr{A}(\zeta_1,\zeta_2,\ldots,\zeta_n)=(\zeta_1,\zeta_2,\ldots,\zeta_n)J_B.$$

记 V 中由基向量 $\eta_{e_j+1},\eta_{e_j+2},\ldots,\eta_{e_j+m_j}$ 生成的子空间为 C_j,其中 $e_j=m_1+m_2+\cdots+m_{j-1}$,$j=1,2,\ldots,k$,且约定 $m_0=0$. 显然 $\dim C_j=m_j$,而且可以证明,C_j 是由向量 η_{e_j+1} 生成的循环子空间.

同样,V 中由基向量 $\zeta_{f_j+1},\zeta_{f_j+2},\ldots,\zeta_{f_j+n_j}$ 生成的子空间 $\widetilde{C}_j$ 是由向量 ζ_{f_j+1} 生成的循环子空间,其中 $f_j=n_1+n_2+\cdots+n_{j-1}$,$j=1,2,\ldots,\ell$,且 $n_0=0$. 于是

$$V=C_1\oplus C_2\oplus\cdots\oplus C_k=\widetilde{C}_1\oplus\widetilde{C}_2\oplus\cdots\oplus\widetilde{C}_\ell.$$

由定理 6.2.3,$k=\ell$,而且 $n_j=m_j$这表明,$N_{(n_j)}^{\mathrm{T}}=N_{(m_j)}^{\mathrm{T}}$,所以 $J_A=J_B$. ■

定理 6.2.4 和定理 6.2.5 表明,对于 n 阶幂零方阵,方阵在相似下的标准形问题已经基本解决. 即对幂零方阵而言,可以用准对角方阵 J 作为相似等价类的代表元,而方阵 J 中准对角块的个数以及对角块的大小是幂零方阵在相似下的全系不变量.

留下的问题是,对给定的幂零方阵 A,如何确定准对角方阵 J 中准对角块的个数以及准对角块的大小. 这一问题留待 §6.3 解决.

习　题　6.2

1. 设 3 维复向量空间 $\mathbb{C}^3$ 的线性变换 $\mathscr{A}$ 在 $\mathbb{C}^3$ 的标准基下的方阵为

$$A=\begin{pmatrix}1&\mathrm{i}&0\\-1&2&-\mathrm{i}\\0&1&1\end{pmatrix},$$

其中 $\mathrm{i}^2=-1$. 求向量 $\varepsilon_1=(1,0,0)$ 与 $\alpha=(1,0,\mathrm{i})$ 的最小多项式.

2. 设 $\mathscr{A}$ 是 n 维复线性空间 V 的 k 次幂零变换. 证明,由非零向量 $\alpha_0\in V$ 生成的循环子空间 C_0 的维数不超过 k.

3. 证明 6 阶幂零方阵 N_1 与 N_2 相似的必要与充分条件是,它们具有相同的秩和最小多项式.

4. 设 $\mathscr{A}$ 是 n 维复线性空间 V 的 k 次幂零变换,且 V 分解为循环子空间 $C_1,C_2,\ldots,C_k$ 的直和. $C_1,C_2,\ldots,C_k$ 中维数为 j 的子空间的个数记为 n_j,$j=1,2,\ldots,k$. 证明,

(1) $\dim\operatorname{Im}\mathscr{A}=n_2+2n_3+\cdots+(k-1)n_k$;

(2) $n_j=\dim\operatorname{Im}\mathscr{A}^{j+1}+\dim\operatorname{Im}\mathscr{A}^{j-1}-2\dim\operatorname{Im}\mathscr{A}^{j}$,$j=1,2,\ldots,k$;

(3) V 本身是非零向量 α 生成的循环子空间的充分必要条件是，$\dim \operatorname{Im} \mathscr{A}^j = n-j$，其中 $j=1,2,\ldots,n$.

5. 设 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换. 如果存在非零向量 $\alpha_0 \in V$，使得由向量 α_0 生成的循环子空间 $C_0 = V$，则称 $\mathscr{A}$ 为**循环变换**，向量 α_0 称为 $\mathscr{A}$ 的**循环向量**. 证明，$\mathscr{A}$ 为循环变换的充分必要条件是，存在 V 的基，使得 $\mathscr{A}$ 在这组基下的方阵为

$$A = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

并由此证明，$\mathscr{A}$ 为循环变换的充分必要条件是，$\mathscr{A}$ 的最小多项式等于 $\mathscr{A}$ 的特征多项式. 注：形如 A 的方阵称为**友方阵**.

6. 设 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 2 维向量空间 $\mathbb{F}^2$ 的线性变换，非零向量 $\alpha \in \mathbb{F}^2$ 不是 $\mathscr{A}$ 的特征向量. 证明 α 是 $\mathscr{A}$ 的循环向量.

7. 设 3 维实向量空间 $\mathbb{R}^3$ 的线性变换 $\mathscr{A}$ 在 $\mathbb{R}^3$ 的标准正交基下的方阵为 $A = \operatorname{diag}(2,2,-1)$. 证明 $\mathscr{A}$ 不具有循环向量. 求向量 $\alpha = (1,-1,3)$ 生成的循环子空间.

8. 证明，如果数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换 $\mathscr{A}$ 的二次幂 $\mathscr{A}^2$ 为循环变换，则 $\mathscr{A}$ 本身也是循环变换. 反之是否成立？

9. 设数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换 $\mathscr{A}$ 可对角化. 证明，

(1) 如果 $\mathscr{A}$ 是循环变换，则 $\mathscr{A}$ 的 n 个特征值两两不同；

(2) 如果 $\mathscr{A}$ 的 n 个特征值两两不同，且 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 $\mathscr{A}$ 的完全特征向量组，则 $\alpha_1+\alpha_2+\cdots+\alpha_n$ 是循环向量.

10. 设 $\mathscr{A}$ 与 $\mathscr{B}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的可交换的线性变换，且 $\mathscr{A}$ 是循环变换. 证明，存在多项式 $f(\lambda) \in \mathbb{F}[\lambda]$，使得 $\mathscr{B} = f(\mathscr{A})$.

11. 设 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换，而且 V 的任意一个与 $\mathscr{A}$ 可交换的线性变换 $\mathscr{B}$ 都可表为 $\mathscr{A}$ 的多项式. 证明 $\mathscr{A}$ 是循环变换.

12. 设 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 的线性变换. 证明，V 的每个非零向量都是 $\mathscr{A}$ 的循环向量的充分必要条件为，$\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 在 $\mathbb{F}$ 上不可约.

§6.3 Jordan 标准形

设 V 是 n 维复线性空间，$\mathscr{A}: V \to V$ 是线性变换，$\lambda_1,\lambda_2,\ldots,\lambda_t$ 是 $\mathscr{A}$ 的全部不同特征值，且 $\mathscr{A}$ 的特征多项式为

$$\varphi(\lambda) = (\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_t)^{e_t},$$

其中 $e_1,e_2,\ldots,e_t$ 是正整数，且 $e_1+e_2+\cdots+e_t = n$.

定理 6.3.1（空间第二分解定理） 设 W_{λ_j} 是线性变换 $\mathscr{A}: V \to V$ 的属于特征值 λ_j 的根子空间. 则

$$W_{\lambda_j} = C_{j1} \oplus C_{j2} \oplus \cdots \oplus \lambda_{jk_j},$$

其中 $C_{j\ell}$ 是相对 $(\mathscr{A} - \lambda_j\mathscr{I})\big|_{W_{\lambda_j}}$ 的循环子空间，其中 $\ell = 1,2,\ldots,k_j$，$j = 1,2,\ldots,t$.

简言之,$\mathscr{A}$ 的每个根子空间 W_{λ_j} 都可分解为循环子空间的直和.

证明　由 §6.1 可知,存在正整数 m_j,使得

$$W_{\lambda_j}=W_{\lambda_j}^{(m_j)}=\left\{\alpha\in V\,\middle|\,(\mathscr{A}-\lambda_j\mathscr{I})^{m_j}(\alpha)=0\right\},$$

并且存在 $\xi\in W_{\lambda_j}$,使得 $(\mathscr{A}-\lambda_j\mathscr{I})^{m_j-1}(\xi)\neq 0$.

这表明 $\mathscr{A}-\lambda_j\mathscr{I}$ 在 W_{λ_j} 上的限制 $(\mathscr{A}-\lambda_j\mathscr{I})|_{W_{\lambda_j}}$ 是 m_j 次幂零的.

由定理 6.2.2 可知,存在非零向量 $\xi_{j1},\xi_{j2},\ldots,\xi_{jk_j}\in W_{\lambda_j}$,它们生产的相对于 $(\mathscr{A}-\lambda_j\mathscr{I})|_{W_{\lambda_j}}$ 的循环子空间依次为 $C_{j1},C_{j2},\ldots,C_{jk_j}$,使得

$$W_{\lambda_j}=C_{j1}\oplus C_{j2}\oplus\cdots\oplus C_{jk_j}.\qquad\blacksquare$$

在定理 6.3.1 的证明已经指出,$(\mathscr{A}-\lambda_j\mathscr{I})|_{W_{\lambda_j}}$ 是 m_j 次幂零的. 因此 $(\mathscr{A}-\lambda_j\mathscr{I})|_{C_{j\ell}}$ 也是幂零的. 设 $\dim C_{j\ell}=m_{j\ell}$. 由定理 6.2.2,可设 $m_j=m_{j1}\geqslant m_{j2}\geqslant\cdots\geqslant m_{jk_j}$. 由定理 6.2.1 可知,$(\lambda-\lambda_j)^{m_{j\ell}}$ 是 $\mathscr{A}|_{C_{j\ell}}$ 的特征多项式与最小多项式. 由定理 6.2.2,根子空间 W_{λ_j} 在分解为循环子空间的直和时,循环子空间的个数 k_j 与循环子空间的维数 $m_{j1},m_{j2},\ldots,m_{jk_j}$ 是由 $(\mathscr{A}-\lambda_j\mathscr{I})|_{W_{\lambda_j}}$ 所唯一确定的,也即 $\mathscr{A}|_{W_{\lambda_j}}$ 所唯一确定. 因此,$(\lambda-\lambda_j)^{m_{j\ell}}$ 称为 $\mathscr{A}$ 的属于特征值 λ_j 的一个**初等因子**,而 $C_{j\ell}$ 称为相应于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的循环子空间. $\mathscr{A}$ 的初等因子的全体称为 $\mathscr{A}$ 的**初等因子组**.

由于 $C_{j\ell}$ 是相应于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的循环子空间,而 $C_{j\ell}$ 具有基

$$\left\{\alpha_{j\ell},(\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}),(\mathscr{A}-\lambda_j\mathscr{I})^2(\alpha_{j\ell}),\ldots,(\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right\},$$

所以 $(\mathscr{A}-\lambda_j\mathscr{I})|_{C_{j\ell}}$ 在这组基下的方阵为

$$\begin{aligned}&(\mathscr{A}-\lambda_j\mathscr{I})|_{C_{j\ell}}\left(\alpha_{j\ell},(\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}),\ldots,(\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right)\\&\qquad=\left(\alpha_{j\ell},(\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}),\ldots,(\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right)N_{(m_{j\ell})}^{\mathrm{T}}.\end{aligned}$$

由于 $(\mathscr{A}-\lambda_j\mathscr{I})|_{C_{j\ell}}=\mathscr{A}|_{C_{j\ell}}-\lambda_j\mathscr{I}|_{C_{j\ell}}$,且

$$\lambda_j\mathscr{I}|_{C_{j\ell}}(\alpha)=\lambda_j\alpha,\ \alpha\in C_{j\ell},$$

所以,

$$\begin{aligned}&\mathscr{A}|_{C_{j\ell}}\left(\alpha_{j\ell},(\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}),\ldots,(\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right)\\&\qquad=\left(\alpha_{j\ell},(\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}),\ldots,(\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right)\left(\lambda_jI_{(m_{j\ell})}+N_{(m_{j\ell})}^{\mathrm{T}}\right).\end{aligned}$$

记

$$J_{(m_{j\ell})}^{\mathrm{T}}(\lambda_j)=\lambda_jI_{(m_{j\ell})}+N_{(m_{j\ell})}^{\mathrm{T}}.$$

它称为属于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的 **Jordan 块**.

容易验证,属于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的 Jordan 块 $J_{(m_{j\ell})}^{\mathrm{T}}$ 的最小多项式与特征多项式都等于 $(\lambda-\lambda_j)^{m_{j\ell}}$.

设

$$(\lambda-\lambda_j)^{m_{j1}},(\lambda-\lambda_j)^{m_{j2}},\ldots,(\lambda-\lambda_j)^{m_{jk_j}}$$

是 $\mathscr{A}$ 的属于特征值 λ_j 的全部初等因子,其中 $m_{j1}\geqslant m_{j2}\geqslant\cdots\geqslant m_{jk_j}$. 它们所相应的

循环子空间依次为 $C_{j1}, C_{j2}, \ldots, C_{jk_j}$. 那么根据定理 6.3.1,线性变换 $\mathscr{A}$ 的属于特征值 λ_j 的根子空间 $W_{\lambda_j} = C_{j1} \oplus C_{j2} \oplus \cdots \oplus C_{jk_j}$. 因此这些 $C_{j\ell}$ 的基

$$\left\{\alpha_{j\ell}, (\mathscr{A} - \lambda_j\mathscr{I})(\alpha_{j\ell}), (\mathscr{A} - \lambda_j\mathscr{I})^2(\alpha_{j\ell}), \cdots, (\mathscr{A} - \lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right\}$$

合并后便得到根子空间 W_{λ_j} 的基,而且

$$\mathscr{A}\big|_{W_{\lambda_j}} = \mathscr{A}\big|_{C_{j1}} \oplus \mathscr{A}\big|_{C_{j2}} \oplus \cdots \oplus \mathscr{A}\big|_{C_{jk_j}}.$$

所以 $\mathscr{A}\big|_{W_{\lambda_j}}$ 在 W_{λ_j} 的这组基下的方阵为

$$\operatorname{diag}\left(J^{\mathrm{T}}_{(m_{j1})}(\lambda_j), J^{\mathrm{T}}_{(m_{j2})}(\lambda_j), \ldots, J^{\mathrm{T}}_{(m_{jk_j})}(\lambda_j)\right).$$

由此不难得到

定理 6.3.2 设 $\lambda_1, \lambda_2, \ldots, \lambda_t$ 是线性变换 $\mathscr{A}: V \to V$ 的全部不同特征值,且

$$\left.\begin{array}{llll} (\lambda-\lambda_1)^{m_{11}}, & (\lambda-\lambda_1)^{m_{12}}, & \ldots, & (\lambda-\lambda_1)^{m_{1k_1}}, \\ (\lambda-\lambda_2)^{m_{21}}, & (\lambda-\lambda_2)^{m_{22}}, & \ldots, & (\lambda-\lambda_2)^{m_{2k_2}}, \\ & \cdots\cdots\cdots\cdots & & \\ (\lambda-\lambda_t)^{m_{t1}}, & (\lambda-\lambda_t)^{m_{t2}}, & \ldots, & (\lambda-\lambda_t)^{m_{tk_t}}, \end{array}\right\} \tag{6.3.1}$$

是线性变换 $\mathscr{A}$ 的初等因子组,其中 $m_{j1} \geqslant m_{j2} \geqslant \cdots \geqslant m_{jk_j}$, $j = 1, 2, \ldots, t$.

并且设 $C_{j\ell}$是相应于 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的由向量 $\alpha_{j\ell}$ 生成的(相对于 $(\mathscr{A}-\lambda_j\mathscr{I})\big|_{C_{j\ell}}$ 的)循环子空间,$\ell = 1, 2, \ldots, k_j$,则

(1) $V = C_{11} \oplus \cdots C_{1k_1} \oplus \cdots\cdots \oplus C_{t1} \oplus \cdots \oplus C_{tk_t}$;

(2) 存在 V 的基,使得线性变换 $\mathscr{A}$ 在这组基下的方阵为

$$J = \operatorname{diag}\left(J^{\mathrm{T}}_{(m_{11})}(\lambda_1), \ldots, J^{\mathrm{T}}_{(m_{1k_1})}(\lambda_1), \ldots\ldots, J^{\mathrm{T}}_{(m_{t1})}(\lambda_t), \ldots, J^{\mathrm{T}}_{(m_{tk_t})}(\lambda_t)\right).$$

(3) 线性变换 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 等于 $\mathscr{A}$ 的所有初等因子的乘积;

(4) 线性变换 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 为

$$d(\lambda) = (\lambda-\lambda_1)^{m_{11}}(\lambda-\lambda_2)^{m_{21}}\cdots(\lambda-\lambda_t)^{m_{t1}}.$$

证明 (1) 由于 $(\lambda-\lambda_j)^{m_{j1}}, (\lambda-\lambda_j)^{m_{j2}}, \ldots, (\lambda-\lambda_j)^{m_{jk_j}}$是线性变换 $\mathscr{A}$ 的属于 λ_j 的全部初等因子,$C_{j1}, C_{j2}, \ldots, C_{jk_j}$ 分别是依次相对应的循环子空间,所以依据定理 6.3.1,线性变换 $\mathscr{A}$ 的属于特征值 λ_j 的根子空间为

$$W_{\lambda_j} = C_{j1} \oplus C_{j2} \oplus \cdots \oplus \lambda_{jk_j},$$

其中 $j = 1, 2, \ldots, t$. 由定理 6.1.3, $V = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_t}$. 于是

$$V = C_{11} \oplus \cdots C_{1k_1} \oplus \cdots\cdots \oplus C_{t1} \oplus \cdots \oplus C_{tk_t}.$$

(2) 由于 $C_{j\ell}$ 是由向量 $\alpha_{j\ell}$ 生成的循环子空间,所以

$$\left\{\alpha_{j\ell}, (\mathscr{A} - \lambda_j\mathscr{I})(\alpha_{j\ell}), (\mathscr{A} - \lambda_j\mathscr{I})^2(\alpha_{j\ell}), \cdots, (\mathscr{A} - \lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right\} \tag{$*$}$$

是 $C_{j\ell}$ 的基,$\ell = 1, 2, \ldots, k_j$, $j = 1, 2, \ldots, t$. 由结论 (1),将所有这些循环子空间 $C_{j\ell}$ 的基合并,即得 V 的一组基,而且

$$\mathscr{A} = \mathscr{A}\big|_{C_{11}} \oplus \cdots \oplus \mathscr{A}\big|_{C_{1k_1}} \oplus \cdots\cdots \oplus \mathscr{A}\big|_{C_{t1}} \oplus \cdots \oplus \mathscr{A}\big|_{C_{tk_t}}.$$

由于 $\mathscr{A}\big|_{C_{j\ell}}$ 在 $C_{j\ell}$ 的基 $(*)$ 下的方阵为 Jordan 块 $J^{\mathrm{T}}_{(m_{j\ell})}(\lambda_j)$，所以线性变换 $\mathscr{A}$ 在 V 所取得基下的方阵为 J.

(3) 只需注意到线性变换 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 等于 $\mathscr{A}$ 在 V 的基下的方阵 J 的特征多项式，而准对角方阵 J 的特征多项式显然等于所有对角块 $J^{\mathrm{T}}_{(m_{j\ell})}$ 的特征多项式 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的乘积，即知结论成立.

(4) 对于同一个 j，将所有 $C_{j\ell}$ 的基 $(*)$ 合并，即得根子空间 W_{λ_j} 的基. $\mathscr{A}\big|_{W_{\lambda_j}}$ 在 W_{λ_j} 的这组基下的方阵为

$$\operatorname{diag}\left(J^{\mathrm{T}}_{(m_{j1})}(\lambda_j), J^{\mathrm{T}}_{(m_{j2})}(\lambda_j), \ldots, J^{\mathrm{T}}_{(m_{jk_j})}(\lambda_j)\right). \tag{6.3.2}$$

由于线性变换 $\mathscr{A}\big|_{W_{\lambda_j}}$ 的最小多项式等于它在基下的方阵的最小多项式，所以 $\mathscr{A}\big|_{W_{\lambda_j}}$ 的最小多项式为 $(\lambda-\lambda_j)^{m_{j1}}$. 由于

$$\mathscr{A} = \mathscr{A}\big|_{W_{\lambda_1}} \oplus \mathscr{A}\big|_{W_{\lambda_2}} \oplus \cdots \oplus \mathscr{A}\big|_{W_{\lambda_t}},$$

且 $\lambda_1, \lambda_2, \ldots, \lambda_t$ 两两不同，所以线性变换 $\mathscr{A}$ 的最小多项式等于 $\mathscr{A}\big|_{W_{\lambda_1}}, \mathscr{A}\big|_{W_{\lambda_2}}, \ldots, \mathscr{A}\big|_{W_{\lambda_t}}$ 的最小多项式的乘积. ■

定理 6.3.2 结论 (1) 可以简述为，对于给定的线性变换 $\mathscr{A}: V \to V$，线性空间 V 可以分解为(相对 $\mathscr{A}-\lambda_1\mathscr{I}, \mathscr{A}-\lambda_2\mathscr{I}, \ldots, \mathscr{A}-\lambda_t\mathscr{I}$ 的)循环子空间的直和.

结论 (2) 中的方阵 J 称为 Jordan 方阵，或 **Jordan 标准形**. 因此结论 (2) 可叙述为，对给定的线性变换 $\mathscr{A}$，存在 V 的一组基，使得 $\mathscr{A}$ 在这组基下的方阵为 Jordan 标准形.

设 $\mathscr{A}$ 与 $\mathscr{B}$ 是 n 维复线性空间 V 的线性变换. 如果存在可逆线性变换 $\mathscr{P}: V \to V$，使得

$$\mathscr{B} = \mathscr{P}^{-1}\mathscr{A}\mathscr{P},$$

则称线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 是等价的，记作 $\mathscr{B} \sim \mathscr{A}$.

容易验证，线性空间 V 的线性变换之间的关系 $\sim$ 满足自反性、对称性与传递性. 于是关系 $\sim$ 是 V 的线性变换之间的一种等价关系.

定理 6.3.3　线性空间 V 的线性变换 $\mathscr{A}$ 和 $\mathscr{B}$ 等价的充分必要条件是，$\mathscr{A}$ 与 $\mathscr{B}$ 具有相同的初等因子组.

证明　必要性　设线性变换 $\mathscr{A}$ 的初等因子组为 (6.3.1)，则由定理 6.3.2，

$$V = C_{11} \oplus \cdots C_{1k_1} \oplus \cdots\cdots \oplus C_{t1} \oplus \cdots C_{tk_t},$$

其中 $C_{j\ell}$是相应于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的循环子空间. 设

$$\left\{\alpha_{j\ell}, (\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}), \ldots, (\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right\},$$

是 $C_{j\ell}$ 的基.

因为线性变换 $\mathscr{A}$ 和 $\mathscr{B}$ 等价，所以存在可逆线性变换 $\mathscr{P}$，使得 $\mathscr{B} = \mathscr{P}^{-1}\mathscr{A}\mathscr{P}$. 记

$\widetilde{C}_{j\ell} = \mathscr{P}^{-1}(C_{j\ell})$. 由于 $\mathscr{P}$ 可逆,故 $\mathscr{P}^{-1}$ 可逆,因此 $\dim \widetilde{C}_{j\ell} = \dim C_{j\ell} = m_{j\ell}$. 而且

$$\left\{\mathscr{P}^{-1}(\alpha_{j\ell}), \mathscr{P}^{-1}(\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}), \ldots, \mathscr{P}^{-1}(\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right\},$$

是 $\widetilde{C}_{j\ell}$ 的基. 因为

$$\begin{aligned}\mathscr{P}^{-1}(\mathscr{A}-\lambda_j\mathscr{I})^i(\alpha_{j\ell}) &= \mathscr{P}^{-1}(\mathscr{A}-\lambda_j\mathscr{I})^i\mathscr{P}\cdot\left(\mathscr{P}^{-1}(\alpha_{j\ell})\right)\\ &= \left(\mathscr{P}^{-1}(\mathscr{A}-\lambda_j\mathscr{I})\mathscr{P}\right)^i\left(\mathscr{P}^{-1}(\alpha_{j\ell})\right)\\ &= (\mathscr{B}-\lambda_j\mathscr{I})^i\left(\mathscr{P}^{-1}(\alpha_{j\ell})\right),\end{aligned}$$

所以 $\widetilde{C}_{j\ell}$ 是由向量 $\mathscr{P}^{-1}(\alpha_{j\ell})$ 生成的循环子空间,而且 $(\lambda-\lambda_j)^{m_{j\ell}}$ 是线性变换 $\mathscr{B}$ 的初等因子,其中 $\ell = 1, 2, \ldots, k_j, j = 1, 2, \ldots, t$. 这就证明,线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 具有相同的初等因子组.

充分性 设线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 的初等因子组都是 (6.3.1). 则由定理 6.3.2,

$$\begin{aligned}V &= C_{11} \oplus \cdots C_{1k_1} \oplus \cdots\cdots \oplus C_{t1} \oplus \cdots C_{tk_t}\\ &= \widetilde{C}_{11} \oplus \cdots \widetilde{C}_{1k_1} \oplus \cdots\cdots \oplus \widetilde{C}_{t1} \oplus \cdots \widetilde{C}_{tk_t},\end{aligned}$$

其中 $C_{j\ell}$ 与 $\widetilde{C}_{j\ell}$ 分别是线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 的属于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的循环子空间.

设 $C_{j\ell}$ 与 $\widetilde{C}_{j\ell}$ 分别具有基

$$\left\{\alpha_{j\ell}, (\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}), (\mathscr{A}-\lambda_j\mathscr{I})^2(\alpha_{j\ell}), \ldots, (\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\right\}$$

与

$$\left\{\widetilde{\alpha}_{j\ell}, (\mathscr{B}-\lambda_j\mathscr{I})(\widetilde{\alpha}_{j\ell}), (\mathscr{B}-\lambda_j\mathscr{I})^2(\widetilde{\alpha}_{j\ell}), \ldots, (\mathscr{B}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\widetilde{\alpha}_{j\ell})\right\}.$$

依次将 $C_{11}, \ldots, C_{1k_1}, \ldots, C_{t1}, \ldots, C_{tk_t}$ 的基合并后得到 V 的基记为 $\{\beta_1, \beta_2, \ldots, \beta_n\}$. 而依次将 $\widetilde{C}_{11}, \ldots \widetilde{C}_{1k_1}, \ldots, \widetilde{C}_{t1}, \ldots \widetilde{C}_{tk_t}$ 的基合并后得到 V 的基记为 $\{\gamma_1, \gamma_2, \ldots, \gamma_n\}$. 记

$$(\beta_1, \beta_2, \ldots, \beta_n) = (\gamma_1, \gamma_2, \ldots, \gamma_n)P,$$

其中 P 是 n 阶可逆复方阵. 由

$$\mathscr{P}(\gamma_1, \gamma_2, \ldots, \gamma_n) = (\gamma_1, \gamma_2, \ldots, \gamma_n)P$$

所确定的线性变换 $\mathscr{P}$ 显然可逆,并且

$$\mathscr{P}(\gamma_1, \gamma_2, \ldots, \gamma_n) = (\beta_1, \beta_2, \ldots, \beta_n).$$

由定理 6.3.2,

$$\begin{aligned}\mathscr{A}\mathscr{P}(\gamma_1, \gamma_2, \ldots, \gamma_n) &= \mathscr{A}(\beta_1, \beta_2, \ldots, \beta_n)\\ &= (\beta_1, \beta_2, \ldots, \beta_n)J\\ &= (\gamma_1, \gamma_2, \ldots, \gamma_n)PJ,\\ \mathscr{P}\mathscr{B}(\gamma_1, \gamma_2, \ldots, \gamma_n) &= \mathscr{P}\left((\gamma_1, \gamma_2, \ldots, \gamma_n)J\right)\\ &= (\gamma_1, \gamma_2, \ldots, \gamma_n)PJ,\end{aligned}$$

其中

$$J = \operatorname{diag}\left(J^{\mathrm{T}}_{(m_{11})}(\lambda_1), \ldots, J^{\mathrm{T}}_{(m_{1k_1})}(\lambda_1), \ldots\ldots, J^{\mathrm{T}}_{(m_{t1})}(\lambda_t), \ldots, J^{\mathrm{T}}_{(m_{tk_t})}(\lambda_t)\right).$$

于是 $\mathscr{P}\mathscr{B} = \mathscr{A}\mathscr{P}$. 所以 $\mathscr{B} = \mathscr{P}^{-1}\mathscr{A}\mathscr{P}$. ■

现在转到 n 阶复方阵在相似下的标准形. 先给出下面的定义.

定义 6.3.1 设 A 是 n 阶复方阵, $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维复线性空间 V 的基. 由

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A$$

所确定的线性变换 $\mathscr{A}\colon V\to V$ 的初等因子组称为**方阵 A 的初等因子组**. 方阵 A 的初等因子组中的成员 $(\lambda-\lambda_j)^{m_{j\ell}}$ 称为方阵 A 的属于特征值 λ_j 的**初等因子**.

应当指出,方阵的初等因子的定义既和线性空间 V 本身无关,也和线性空间 V 的基选取无关,也就是说,方阵 A 的初等因子时方阵 A 本身具有的特性.

事实上,设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 分别是同一个 n 维复线性空间 V 的基,方阵 A 由这两组基所确定的线性变换分别记为 $\mathscr{A}$ 与 $\mathscr{B}$,即

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A,$$
$$\mathscr{B}(\beta_1,\beta_2,\ldots,\beta_n)=(\beta_1,\beta_2,\ldots,\beta_n)A.$$

由于

$$(\beta_1,\beta_2,\ldots,\beta_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)P,$$

其中 P 是 n 阶可逆复方阵,所以由

$$\mathscr{P}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)P=(\beta_1,\beta_2,\ldots,\beta_n)$$

所确定的线性变换 $\mathscr{P}$ 可逆,并且

$$\begin{aligned}\mathscr{P}^{-1}\mathscr{B}\mathscr{P}(\alpha_1,\alpha_2,\ldots,\alpha_n)&=\mathscr{P}^{-1}\mathscr{B}(\beta_1,\beta_2,\ldots,\beta_n)\\&=\mathscr{P}^{-1}\big((\beta_1,\beta_2,\ldots,\beta_n)A\big)\\&=\mathscr{P}^{-1}\big((\beta_1,\beta_2,\ldots,\beta_n)\big)A\\&=(\alpha_1,\alpha_2,\ldots,\alpha_n)A.\end{aligned}$$

因此 $\mathscr{P}^{-1}\mathscr{B}\mathscr{P}=\mathscr{A}$. 这表明,线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 是等价的. 由定理 6.3.3,线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 具有相同的初等因子组,即方阵 A 的初等因子的定义与线性空间 V 的基的选取无关.

其次,设 V 与 $\widetilde{V}$ 都是 n 维复线性空间,向量集合 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n\}$ 分别是 V 与 $\widetilde{V}$ 的基. 设 $\mathscr{A}$ 与 $\widetilde{\mathscr{A}}$ 分别是方阵 A 在线性空间 V 与 $\widetilde{V}$ 中由所取定的基确定的线性变换,即

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A,$$
$$\widetilde{\mathscr{A}}(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)=(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\ldots,\widetilde{\alpha}_n)A.$$

设线性变换 $\mathscr{A}\colon V\to V$ 的初等因子组为 (6.3.1),则由定理 6.3.2,

$$V=C_{11}\oplus\cdots C_{1k_1}\oplus\cdots\cdots\oplus C_{t1}\oplus\cdots C_{tk_t},$$

其中 $C_{j\ell}$是属于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$ 的循环子空间,而且 $C_{j\ell}$ 具有基

$$\{\alpha_{j\ell},(\mathscr{A}-\lambda_j\mathscr{I})(\alpha_{j\ell}),(\mathscr{A}-\lambda_j\mathscr{I})^2(\alpha_{j\ell}),\cdots,(\mathscr{A}-\lambda_j\mathscr{I})^{m_{j\ell}-1}(\alpha_{j\ell})\},$$

其中 $\ell=1,2,\ldots,k_j$, $j=1,2,\ldots,t$.

定义 V 到 $\widetilde{V}$ 的映射 σ 如下：设 $\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n\in V$，则令

$$\sigma(\alpha)=a_1\widetilde{\alpha}_1+a_2\widetilde{\alpha}_2+\cdots+a_n\widetilde{\alpha}_n\in\widetilde{V}.$$

容易证明，σ 是 V 到 $\widetilde{V}$ 上的同构映射，而且对任意 $\alpha\in V$，$\sigma\big(\mathscr{A}(\alpha)\big)=\widetilde{\mathscr{A}}\big(\sigma(\alpha)\big)$. 记 $\sigma(C_{j\ell})=\widetilde{C}_{j\ell}$. 因为 $\sigma\colon V\to\widetilde{V}$ 为同构映射，所以 $\dim\widetilde{C}_{j\ell}=\dim C_{j\ell}$，而且

$$\Big\{\sigma(\alpha_{j\ell}),(\widetilde{\mathscr{A}}-\lambda_j\widetilde{\mathscr{I}})\big(\sigma(\alpha_{j\ell})\big),(\widetilde{\mathscr{A}}-\lambda_j\widetilde{\mathscr{I}})^2\big(\sigma(\alpha_{j\ell})\big),\ldots,(\widetilde{\mathscr{A}}-\lambda_j\widetilde{\mathscr{I}})^{m_{j\ell}-1}\big(\sigma(\alpha_{j\ell})\big)\Big\}$$

是 $\widetilde{C}_{j\ell}$ 的基，其中 $\widetilde{\mathscr{I}}$ 是 $\widetilde{V}$ 的单位变换.

这表明，$\widetilde{C}_{j\ell}$ 是由向量 $\sigma(\alpha_{j\ell})$ 生成的循环子空间，而且 $(\lambda-\lambda_j)^{m_{j\ell}}$ 是线性变换 $\widetilde{\mathscr{A}}$ 的初等因子. 所以方阵 A 的初等因子的定义与线性空间 V 的选取无关.

定理 6.3.4 设 n 维复方阵 A 的初等因子组为 (6.3.1)，则方阵 A 相似于如下的 Jordan 标准形：

$$J=\operatorname{diag}\Big(J^{\mathrm{T}}_{(m_{11})}(\lambda_1),\ldots,J^{\mathrm{T}}_{(m_{1k_1})}(\lambda_1),\ldots\ldots,J^{\mathrm{T}}_{(m_{t1})}(\lambda_t),\ldots,J^{\mathrm{T}}_{(m_{tk_t})}(\lambda_t)\Big),$$

其中 $J^{\mathrm{T}}_{(m_{j\ell})}$是属于初等因子 $(\lambda-\lambda_j)^{m_{j\ell}}$的 Jordan 块，$\ell=1,2,\ldots,k_j$，$j=1,2,\ldots,t$.

证明 设 V 是 n 维复线性空间，由方阵 A 在 V 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 所确定的线性变换记为 $\mathscr{A}$，即

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A.$$

由定义，线性变换 $\mathscr{A}$ 的初等因子组也为 (6.3.1). 由定理 6.3.2，存在 V 的基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$，使得

$$\mathscr{A}(\beta_1,\beta_2,\ldots,\beta_n)=(\beta_1,\beta_2,\ldots,\beta_n)J.$$

由于同一个线性变换在不同基下的方阵相似，所以方阵 A 相似于方阵 J. ∎

定理 6.3.5 设 A 与 B 是 n 阶复方阵. 则方阵 A 与 B 相似的充分必要条件是，方阵 A 与 B 的初等因子组相同.

证明 **必要性** 设方阵 A 与 B 相似，则存在 n 阶可逆复方阵 P，使得 $B=P^{-1}AP$. 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维复线性空间 V 的基. 由

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A$$

便确定一个线性变换 $\mathscr{A}$. 因为方阵 P 可逆，所以由

$$(\beta_1,\beta_2,\ldots,\beta_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)P$$

所确定的 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 是 V 的基. 于是

$$\mathscr{A}(\beta_1,\beta_2,\ldots,\beta_n)=(\beta_1,\beta_2,\ldots,\beta_n)B,$$

即 A 与 B 是同一个线性变换 A 在不同基下的方阵. 由定义 6.3.1，方阵 A 与 B 的初等因子组相同.

充分性 设方阵 A 与 B 的初等因子组都是 (6.3.1). 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维复线性空间 V 的基，A 与 B 在这组基下所确定的线性变换分别记为 $\mathscr{A}$ 与 $\mathscr{B}$，即

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A,$$

$$\mathscr{B}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)B.$$

由假设,线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 具有相同的初等因子组. 由定理 6.3.3,存在可逆线性变换 $\mathscr{P}$,使得 $\mathscr{B}=\mathscr{P}^{-1}\mathscr{A}\mathscr{P}$. 设

$$\mathscr{P}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)P,$$

其中 P 是 n 阶可逆复方阵. 于是,

$$\mathscr{P}^{-1}\mathscr{A}\mathscr{P}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)P^{-1}AP.$$

由于 $\mathscr{B}=\mathscr{P}^{-1}\mathscr{A}\mathscr{P}$,所以 $B=P^{-1}AP$,即方阵 A 与 B 相似. ■

定理 6.3.4 与定理 6.3.5 回答了方阵在相似下的标准形理论的两个基本问题,尚未解决的问题是,给定 n 阶复方阵 A,如何求出方阵 A 的初等因子组? 也就是说,给定 n 维复线性空间 V 的线性变换 $\mathscr{A}$,如何求线性变换 $\mathscr{A}$ 的初等因子组? 本节习题 1 给出了求初等因子组的一种方法. 在以后诸节中将另给一种方法.

习 题 6.3

1. 设 A 是 n 阶复方阵,$\lambda_1,\lambda_2,\dots,\lambda_t$ 是方阵 A 的所有不同特征值. 证明

(1) 存在正整数 m,使得

$$\operatorname{rank}A^m=\operatorname{rank}A^{m+1}=\operatorname{rank}A^{m+2}=\cdots;$$

(2) 设 m_j 是使

$$\operatorname{rank}(A-\lambda_j\mathscr{I})^{m_j}=\operatorname{rank}(A-\lambda_j\mathscr{I})^{m_j+1}=\operatorname{rank}(A-\lambda_j\mathscr{I})^{m_j+2}=\cdots$$

的最小正整数,则方阵 A 的最小多项式 $d(\lambda)$ 为

$$d(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t};$$

(3) 设 $(\lambda-\lambda_j)^\ell$ 是方阵 A 的属于特征值 λ_j 的初等因子,则 $\ell\leqslant m_j$;

(4) 设方阵 A 的初等因子组为

$$\begin{array}{llll}(\lambda-\lambda_1)^{m_{11}}, & (\lambda-\lambda_1)^{m_{12}}, & \dots, & (\lambda-\lambda_1)^{m_{1k_1}},\\ (\lambda-\lambda_2)^{m_{21}}, & (\lambda-\lambda_2)^{m_{22}}, & \dots, & (\lambda-\lambda_2)^{m_{2k_2}},\\ & \cdots\cdots\cdots\cdots & & \\ (\lambda-\lambda_t)^{m_{t1}}, & (\lambda-\lambda_t)^{m_{t2}}, & \dots, & (\lambda-\lambda_t)^{m_{tk_t}},\end{array}$$

其中属于特征值 λ_j 且次数为 ℓ 的初等因子 $(\lambda-\lambda_j)^\ell$ 的个数记为 $n_{j\ell}$,当 $(\lambda-\lambda_j)^\ell$ 不是方阵 A 的初等因子时,$n_{j\ell}=0$. 则

$$n_{j\ell}=\operatorname{rank}(A-\lambda_j\mathscr{I})^{\ell+1}+\operatorname{rank}(A-\lambda_j\mathscr{I})^{\ell-1}-2\operatorname{rank}(A-\lambda_j\mathscr{I})^\ell,$$

其中 $1\leqslant\ell\leqslant m_j$, $j=1,2,\dots,t$.

注 习题 1 建议采用如下步骤求方阵 A 的 Jordan 标准形:

(1) 求出方阵 A 的特征多项式 $\varphi(\lambda)=\det(\lambda\mathscr{I}-A)$,并求出方阵 A 的全部不同特征值 $\lambda_1,\lambda_2,\dots,\lambda_t$;

(2) 对每个特征值 λ_j,由

$$\operatorname{rank}(A-\lambda_j\mathscr{I})^{m_j-1}>\operatorname{rank}(A-\lambda_j\mathscr{I})^{m_j}=\operatorname{rank}(A-\lambda_j\mathscr{I})^{m_j+1}=\cdots$$

求出 m_j;

(3) 对每个 $\ell, 1 \leqslant \ell \leqslant m_j$, 计算

$$n_{j\ell} = \operatorname{rank}(A - \lambda_j \mathscr{I})^{\ell+1} + \operatorname{rank}(A - \lambda_j \mathscr{I})^{\ell-1} - 2 \operatorname{rank}(A - \lambda_j \mathscr{I})^{\ell},$$

由此确定 $(\lambda - \lambda_j)^\ell$ 是否是方阵 A 的初等因子, 以及初等因子 $(\lambda - \lambda_j)^\ell$ 在方阵 A 的初等因子组中出现的次数;

(4) 根据 (3) 中所确定的方阵 A 的初等因子组写出方阵 A 的 Jordan 标准形.

2. 利用习题 1 的方法, 求出下列方阵 A 的 Jordan 标准形.

(1) $A = \begin{pmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{pmatrix}$; (2) $A = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}$;

(3) $A = \begin{pmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix}$; (4) $A = \begin{pmatrix} 3 & 7 & -3 \\ -2 & -5 & 2 \\ -4 & -10 & -3 \end{pmatrix}$;

(5) $A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix}$; (6) $A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & 0 & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}_{n \times n}$.

3. 设 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 与做小多项式 $d(\lambda)$ 分别为

$$\varphi(\lambda) = (\lambda - \lambda_1)^{e_1} (\lambda - \lambda_2)^{e_2} \cdots (\lambda - \lambda_t)^{e_t},$$
$$d(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_t)^{m_t}.$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_t$ 是方阵 A 的全部不同特征值. 证明线性变换 $(\mathscr{A} - \lambda_j \mathscr{I})^{e_j}$ 与 $(\mathscr{A} - \lambda_j \mathscr{I})^{m_j}$ 的核相等.

§6.4 λ 矩阵的相抵

设 $\mathbb{F}$ 是数域, λ 是未定元. 数域 $\mathbb{F}$ 上所有关于未定元 λ 的多项式集合记为 $\mathbb{F}[\lambda]$. 取 $m \times n$ 个多项式 $a_{ij}(\lambda) \in \mathbb{F}[\lambda], 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$, 则

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix}$$

称为数域 $\mathbb{F}$ 上的 $m \times n$ 阶 λ 矩阵, 简称 **λ 矩阵**. 也就是说, 所谓 λ 矩阵是元素为 λ 的多项式的矩阵. 数域 $\mathbb{F}$ 上所有 $m \times n$ 阶 λ 矩阵的集合记为 $\big(\mathbb{F}[\lambda]\big)^{m \times n}$. λ 矩阵的加法, 纯量与 λ 矩阵的乘法, λ 矩阵与 λ 矩阵的乘法, 以及 λ 方阵的行列式的定义和通常数域 $\mathbb{F}$ 上方阵相同. 但应注意, λ 方阵的行列式是关于未定元 λ 的多项式.

和通常矩阵一样可以引进 λ 矩阵的秩的概念. 所谓 λ 矩阵 $A(\lambda)$ 的秩是指 λ 矩阵 $A(\lambda)$ 中非零子式的最高阶数, 仍记为 $\operatorname{rank} A(\lambda)$. 如果 n 阶方阵 $A(\lambda)$ 的秩为 n,

则 λ 方阵 $A(\lambda)$ 称为满秩的,否则 λ 方阵 $A(\lambda)$ 称为降秩的. 很显然,λ 方阵 $A(\lambda)$ 满秩的充分必要条件是 $\det A(\lambda) \neq 0$.

对 n 阶 λ 方阵 $A(\lambda) \in \left(\mathbb{F}[\lambda]\right)^{n\times n}$,如果存在 n 阶 λ 方阵 $B(\lambda) \in \left(\mathbb{F}[\lambda]\right)^{n\times n}$,使得

$$A(\lambda)B(\lambda) = I_{(n)} = B(\lambda)A(\lambda),$$

其中 $I_{(n)}$ 是 n 阶单位方阵,则 λ 方阵 $A(\lambda)$ 称为可逆的; 而 $B(\lambda)$ 称为 $A(\lambda)$ 的逆方阵,并记为 $A(\lambda)^{-1}$.

定理 6.4.1 n 阶 λ 方阵 $A(\lambda)$ 可逆的充分必要条件是,λ 方阵 $A(\lambda)$ 的行列式是数域 $\mathbb{F}$ 中非零的数.

证明 设 n 阶 λ 方阵 $A(\lambda)$ 可逆,则存在 n 阶 λ 方阵 $B(\lambda)$,使得 $A(\lambda)B(\lambda) = I_{(n)} = B(\lambda)A(\lambda)$. 两端取行列式,即得到 $\det A(\lambda)\cdot\det B(\lambda) = 1$. 其中 $\det A(\lambda)$ 与 $\det B(\lambda)$ 是数域 $\mathbb{F}$ 上关于 λ 的多项式. 比较两端多项式的系数可知,$\det A(\lambda)$ 与 $\det B(\lambda)$ 是零次多项式. 因此 $\det A(\lambda)$ 是数域 $\mathbb{F}$ 中非零的数.

反之,设 $\det A(\lambda)$ 是数域 $\mathbb{F}$ 中非零的数. 记 $\det A(\lambda) = a$. λ 方阵 $A(\lambda)$ 的附属方阵记为 $A^*(\lambda)$. 由于方阵 $A^*(\lambda)$ 中的元素是 λ 方阵 $A(\lambda)$ 的 $n-1$ 阶子式,而 $A(\lambda)$ 的 $n-1$ 阶子式是数域 $\mathbb{F}$ 上关于 λ 的多项式,因此 $A^*(\lambda)$ 是数域 $\mathbb{F}$ 上的 n 阶 λ 方阵. 所以 $a^{-1}A^*(\lambda)$ 是数域 $\mathbb{F}$ 上 n 阶 λ 方阵,并且

$$A(\lambda)\left(a^{-1}A^*(\lambda)\right) = I_{(n)} = \left(a^{-1}A^*(\lambda)\right)A(\lambda).$$

即 λ 方阵 $A(\lambda)$ 可逆,而且 $A(\lambda) = a^{-1}A^*(\lambda)$. ∎

由定理 6.4.1 可以看出,如果 λ 方阵 $A(\lambda)$ 可逆,则 $A(\lambda)$ 是满秩的. 反之则不然. 这可通常方阵是不同的.

对 λ 矩阵,同样有所谓行或列的**初等变换**. 对调 λ 矩阵 $A(\lambda)$ 的某两行(或列); λ 矩阵 $A(\lambda)$ 的某一行(或列)遍乘数域 $\mathbb{F}$ 上某个非零多项式并加到 $A(\lambda)$ 的某一行(或列); 以及 λ 矩阵 $A(\lambda)$ 的某一行(或列)遍乘以数域 $\mathbb{F}$ 中非零的数,依次称为对 λ 矩阵 $A(\lambda)$ 施行行(或列)的**第一、第二和第三种初等 $\boldsymbol{\lambda}$ 变换**. 记

$$\begin{aligned} P_{ij} &= I_{(n)} + E_{ij} + E_{ji} - E_{ii} - E_{jj}, \\ Q_{ij}\left(f(\lambda)\right) &= I_{(n)} + f(\lambda)E_{ij}, \\ P_i(a) &= I_{(n)} + (a-1)E_{ii}. \end{aligned}$$

其中 E_{ij} 是仅在 $(i,\ j)$ 位置上元素为 1 而其它元素为 0 的 n 阶方阵; $f(\lambda)$ 是数域 $\mathbb{F}$ 上关于 λ 的多项式.

P_{ij}、$Q_{ij}\left(f(\lambda)\right)$ 和 $P_i(a)$ 依次称为第一、第二和第三种初等 λ 方阵. 分别用第一、第二和第三种初等 λ 方阵左乘 $n\times m$ 阶 λ 矩阵 $A(\lambda)$,相当于分别对 λ 矩阵 $A(\lambda)$ 施行第一、第二和第三种行的初等 λ 变换. 而右乘于 $n\times m$ 阶 λ 矩阵 $A(\lambda)$,则相当于 $A(\lambda)$ 的列变换.

定理 6.4.2 设 $m\times n$ 阶 λ 矩阵 $A(\lambda)$ 的秩为 $\operatorname{rank} A(\lambda) = r$. 则 $A(\lambda)$ 可以经过有

限次行或列的初等 λ 变换化为以下形式:

$$B(\lambda)=\begin{pmatrix} D(\lambda) & 0 \\ 0 & 0 \end{pmatrix},$$

其中 $D(\lambda)=\mathrm{diag}\big(d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)\big)$ 是 n 阶对角 λ 方阵,$d_1(\lambda),d_2(\lambda),\ldots,d_k(\lambda)$ 是数域 $\mathbb{F}$ 上首一多项式,并且依次一个整除另一个.

证明 对零矩阵 $A(\lambda)$,定理 6.4.2 成立. 因此设 $A(\lambda)=\big(a_{ij}(\lambda)\big)\neq 0$.

于是存在某个元素 $a_{ij}(\lambda)\neq 0$. 对调 λ 矩阵 $A(\lambda)$ 的第 1 行与第 i 行,第 1 列与第 j 列. 则元素 $a_{ij}(\lambda)$ 掉到 $(1,1)$ 位置上,所以不妨设 $a_{11}(\lambda)\neq 0$.

首先证明,λ 矩阵 $A(\lambda)$ 可以经有限次初等 λ 变换化为 λ 矩阵 $C(\lambda)=\big(c_{ij}(\lambda)\big)$,其中 $c_{11}(\lambda)\neq 0$,且 $c_{11}(\lambda)$ 整除每个 $c_{ij}(\lambda)$. 为此对次数 $\deg a_{11}(\lambda)$ 用归纳法.

当 $\deg a_{11}(\lambda)=0$,即 $a_{11}(\lambda)$ 为数域 $\mathbb{F}$ 中非零的数时,显然 $a_{11}(\lambda)$ 整除每个 $a_{ij}(\lambda)$. 因此结论对 $\deg a_{11}(\lambda)=0$ 成立.

假设结论对 $\deg a_{11}(\lambda)\leqslant t-1$ 成立,下面证明结论对 $\deg a_{11}(\lambda)=t$ 成立. 如果 $a_{11}(\lambda)$ 整除每个 $a_{ij}(\lambda)$,则结论已经成立. 因此不妨设存在某个 $a_{ij}(\lambda)$,使得 $a_{11}(\lambda)$ 不整除 $a_{ij}(\lambda)$.

情形 1 $i=1$,即 $a_{11}(\lambda)$ 不整除 $a_{1j}(\lambda)$.

由 Euclid 长除法,存在 $q(\lambda),r(\lambda)\in\mathbb{F}[\lambda]$,使得

$$a_{1j}(\lambda)=q(\lambda)a_{11}(\lambda)+r(\lambda),$$

其中 $0\leqslant\deg r(\lambda)<t$. 用 $-q(\lambda)$ 遍乘矩阵 $A(\lambda)$ 的第 1 列,并加到第 j 列,得到一个 λ 矩阵,它的 $(1,j)$ 位置上的元素为 $r(\lambda)$. 再对调第 1 列与第 j 列,得到的 λ 矩阵中 $(1,1)$ 位置上的元素为 $r(\lambda)$. 对这个 λ 矩阵用归纳假设,即知结论成立.

情形 2 $j=1$,即 $a_{11}(\lambda)$ 不整除 $a_{i1}(\lambda)$.

此情形的证明同情形 1,只需将情形 1 的证明中将列改成行即可.

情形 3 $a_{11}(\lambda)$ 整除每个 a_{1j} 与每个 a_{i1},$j=2,\ldots,n$,$i=2,\ldots,m$,而且 $a_{11}(\lambda)$ 不整除某个 $a_{k\ell}(\lambda)$,$k\neq 1$,$\ell\neq 1$.

此时 $a_{1j}(\lambda)=q_j(\lambda)a_{11}(\lambda)$,$q_j(\lambda)\in\mathbb{F}[\lambda]$. 用 $-q_j(\lambda)$ 遍乘 λ 矩阵 $A(\lambda)$ 的第 1 列,并加到第 j 列,得到如下形式的 λ 矩阵

$$\begin{pmatrix} a_{11}(\lambda) & 0 & 0 & \cdots & 0 \\ a_{21}(\lambda) & \widetilde{a}_{22}(\lambda) & \widetilde{a}_{23}(\lambda) & \cdots & \widetilde{a}_{2n}(\lambda) \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1}(\lambda) & \widetilde{a}_{m2}(\lambda) & \widetilde{a}_{m3}(\lambda) & \cdots & \widetilde{a}_{mn}(\lambda) \end{pmatrix}.$$

设 $a_{i1}(\lambda)=p_i(\lambda)a_{11}(\lambda)$,$p_i(\lambda)\in\mathbb{F}[\lambda]$. 用 $-p_i(\lambda)$ 遍乘上述矩阵的第一行,并加到第 i 行,得

$$\begin{pmatrix} a_{11}(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & \widetilde{a}_{22}(\lambda) & \widetilde{a}_{23}(\lambda) & \cdots & \widetilde{a}_{2n}(\lambda) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \widetilde{a}_{m2}(\lambda) & \widetilde{a}_{m3}(\lambda) & \cdots & \widetilde{a}_{mn}(\lambda) \end{pmatrix}. \tag{6.4.1}$$

如果 $a_{11}(\lambda)$ 整除每个 $\widetilde{a}_{ij}(\lambda)$，则结论已成立. 因此设 $a_{11}(\lambda)$ 不整除某个 $\widetilde{a}_{k\ell}(\lambda)$. 将 λ 矩阵 (6.4.1) 的第 ℓ 列加到第 1 列，得到

$$\begin{pmatrix} a_{11}(\lambda) & 0 & 0 & \cdots & 0 \\ \widetilde{a}_{2\ell}(\lambda) & \widetilde{a}_{22}(\lambda) & \widetilde{a}_{23}(\lambda) & \cdots & \widetilde{a}_{2n}(\lambda) \\ \vdots & \vdots & \vdots & & \vdots \\ \widetilde{a}_{k\ell}(\lambda) & \widetilde{a}_{k2}(\lambda) & \widetilde{a}_{k3}(\lambda) & \cdots & \widetilde{a}_{kn}(\lambda) \\ \vdots & \vdots & \vdots & & \vdots \\ \widetilde{a}_{m\ell}(\lambda) & \widetilde{a}_{m2}(\lambda) & \widetilde{a}_{m3}(\lambda) & \cdots & \widetilde{a}_{mn}(\lambda) \end{pmatrix}.$$

于是化为情形 2. 所以结论成立.

现在设 λ 矩阵 $A(\lambda)$ 已经通过有限次初等 λ 变换化为 λ 矩阵 $C(\lambda)$. 因为 $c_{11}(\lambda)$ 整除每个 $c_{i1(\lambda)}$，故可设 $c_{i1}(\lambda)=f_i(\lambda)c_{11}(\lambda)$，$f_i(\lambda)\in\mathbb{F}[\lambda]$. 用 $-f_i(\lambda)$ 遍乘 λ 矩阵 $C(\lambda)$ 的第 1 行，并加到第 i 行，得到 λ 矩阵

$$\begin{pmatrix} c_{11}(\lambda) & c_{12}(\lambda) & \cdots & c_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & b_{m2}(\lambda) & \cdots & b_{mn}(\lambda) \end{pmatrix}, \tag{6.4.2}$$

其中 $b_{k\ell}(\lambda)=c_{k\ell}(\lambda)-f_k(\lambda)c_{11}(\lambda)$. 由于 $c_{k\ell}(\lambda)$ 整除每个 $c_{ij}(\lambda)$，所以 $c_{11}(\lambda)$ 整除每个 $b_{k\ell}$. 再对矩阵 (6.4.2) 作初等 λ 变换，即得到

$$\begin{pmatrix} c_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & b_{m2}(\lambda) & \cdots & b_{mn}(\lambda) \end{pmatrix}, \tag{6.4.3}$$

其中 $c_{11}(\lambda)$ 整除每个 $b_{k\ell}(\lambda)$.

记 $A_1(\lambda)=\big(b_{k\ell}(\lambda)\big)$. 对 λ 矩阵 $A_1(\lambda)$ 重复上述过程，则 λ 矩阵 $A_1(\lambda)$ 可经有限次初等 λ 变换化为

$$\begin{pmatrix} c_{22}(\lambda) & 0 \\ 0 & A_2(\lambda) \end{pmatrix}.$$

其中 $c_{22}(\lambda)$ 整除 λ 矩阵 $A_2(\lambda)$ 的每个元素. 于是 λ 矩阵 (6.4.3) 可化为

$$\begin{pmatrix} c_{11}(\lambda) & 0 & 0 \\ 0 & c_{22}(\lambda) & 0 \\ 0 & 0 & A_2(\lambda) \end{pmatrix}. \tag{6.4.4}$$

注意，由 λ 矩阵 (6.4.3) 化为 λ 矩阵 (6.4.4)，只需对 λ 矩阵 $A_1(\lambda)$ 施行初等 λ 变

换. 由于 $c_{11}(\lambda)$ 整除 λ 矩阵 $A_1(\lambda)$ 的每个元素,因此对换 $A_1(\lambda)$ 的某两行(或列),得到的 λ 矩阵中每个元素都能被 $c_{11}(\lambda)$ 整除.

对于用非零多项式 $f(\lambda)\in\mathbb{F}[\lambda]$ 遍乘 λ 矩阵 $A_1(\lambda)$ 的第 i 列,并加到第 j 列后得到的 λ 矩阵,它的第 j 列元素为 $b_{kj}(\lambda)+f(\lambda)b_{kj}(\lambda)$,显然它们都能被 $c_{11}(\lambda)$ 整除,至于第 j 列以外的元素,则是 λ 矩阵 $A_1(\lambda)$ 的元素未经变动而得到的,当然也能被 $c_{11}(\lambda)$ 整除.

对于用非零的数遍乘 λ 矩阵 $A_1(\lambda)$ 的某一行(或列)而得到的 λ 矩阵,它的每个元素显然都能被 $c_{11}(\lambda)$ 整除. 因此 $c_{11}(\lambda)$ 整除 $c_{22}(\lambda)$ 以及 $A_2(\lambda)$ 的每个元素,而且 $c_{22}(\lambda)$ 整除 $A_2(\lambda)$ 的么个元素.

重复上面的过程,λ 矩阵 $A(\lambda)$ 可以经有限次初等 λ 变换化为以下的形式:

$$\begin{pmatrix}\operatorname{diag}\big(c_{11}(\lambda),c_{22}(\lambda),\dots,c_{kk}(\lambda)\big) & 0\\ 0 & 0\end{pmatrix}\tag{6.4.5}$$

其中非零多项式 $c_{11}(\lambda),c_{22}(\lambda),\dots,c_{kk}(\lambda)$ 依次一个整除另一个.

显然初等 λ 变换不改变 λ 矩阵 $A(\lambda)$ 的秩,所以 $k=\operatorname{rank}A(\lambda)=r$.

设 $c_{ii}(\lambda)$ 的首项系数为 b_i,则用 b_i^{-1} 遍乘 λ 矩阵 (6.4.5) 的第 i 行. 于是 λ 矩阵 $A(\lambda)$ 便化为 λ 矩阵 $B(\lambda)$. ■

由于对 λ 矩阵 $A(\lambda)$ 施行行或列的初等 λ 变换相当于用初等 λ 方阵左乘或右乘 λ 矩阵 $A(\lambda)$,因此定理 6.4.2 也可以叙述为:

定理 6.4.3　设数域 $\mathbb{F}$ 上 $m\times n$ 阶 λ 矩阵 $A(\lambda)$ 的秩为 r. 则存在数域 $\mathbb{F}$ 上 m 阶初等 λ 方阵 $P_1(\lambda),P_2(\lambda),\dots,P_s(\lambda)$ 和属于 $\mathbb{F}$ 上的 n 阶初等 λ 方阵 $Q_1(\lambda),Q_2(\lambda),\dots,Q_t(\lambda)$,使得

$$P_s(\lambda)\cdots,P_1(\lambda)A(\lambda)Q_1(\lambda)\cdots Q_t(\lambda)=\begin{pmatrix}D(\lambda)&0\\0&0\end{pmatrix},$$

其中 $D(\lambda)=\operatorname{diag}\big(d_1(\lambda),d_2(\lambda),\dots,d_r(\lambda)\big)$,而 $d_1(\lambda),d_2(\lambda),\dots,d_r(\lambda)\in\mathbb{F}[\lambda]$ 是依次一个整除另一个的首一多项式.

由定理 6.4.3 立即得到

定理 6.4.4　λ 矩阵 $A(\lambda)$ 可逆的充分必要条件是,λ 矩阵 $A(\lambda)$ 可以表为有限多个初等 λ 方阵的乘积.

证明　设 n 阶 λ 方阵 $A(\lambda)$ 可逆,则 $\operatorname{rank}A(\lambda)=n$. 根据定理 6.4.3,存在 n 阶初等 λ 方阵 $P_1(\lambda),\dots,P_s(\lambda)$ 和 $Q_1(\lambda)\cdots,Q_t(\lambda)$,使得

$$\begin{aligned}P_s(\lambda)\cdots,P_1(\lambda)A(\lambda)Q_1(\lambda)\cdots Q_t(\lambda)&=B(\lambda)\\&=\operatorname{diag}\big(d_1(\lambda),\dots,d_n(\lambda)\big).\end{aligned}$$

显然,初等 λ 方阵可逆,而且可逆 λ 方阵的乘积仍可逆,所以 λ 方阵 $B(\lambda)$ 可逆. 由定理 6.4.1,$\det B(\lambda)=d_1(\lambda)\cdots d_n(\lambda)$ 是数域 $\mathbb{F}$ 中非零的数,即 $\det B(\lambda)$ 是零次多项式,因此 $d_1(\lambda),\dots,d_n(\lambda)$ 都是零次多项式. 由于 $d_1(\lambda),\dots,d_n(\lambda)$ 都是首一多项式,

所以 $d_1(\lambda)=\cdots=d_n(\lambda)=1$. 于是 $B(\lambda)=I_{(n)}$,而且

$$A(\lambda)=P_1(\lambda)^{-1}\cdots P_s(\lambda)^{-1}Q_t(\lambda)^{-1}\cdots Q_1(\lambda)^{-1}.$$

显然,初等 λ 方阵的逆方阵仍是同一类型的初等 λ 方阵,所以 λ 方阵 $A(\lambda)$ 可以表为有限多个初等 λ 方阵的乘积.

反之,设 λ 矩阵 $A(\lambda)$ 是有限多个初等 λ 方阵的乘积. 由于初等 λ 方阵可逆,而且可逆 λ 方阵的乘积仍可逆,所以 λ 矩阵 $A(\lambda)$ 可逆. ■

定义 6.4.1　设 $A(\lambda)$ 与 $B(\lambda)$ 是数域 $\mathbb{F}$ 上的 $m\times n$ 阶 λ 矩阵. 如果 λ 矩阵 $A(\lambda)$ 可经有限次行或列的初等 λ 变换化为 λ 矩阵 $B(\lambda)$,也即存在数域 $\mathbb{F}$ 上 m 阶可逆 λ 方阵 $P(\lambda)$ 与数域 $\mathbb{F}$ 上 n 阶可逆 λ 方阵 $Q(\lambda)$,使得

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda),$$

则称 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 是**相抵**的.

容易验证,数域上 $m\times n$ 阶 λ 矩阵之间的相抵关系满足自反性、对称性与传递性,因此 λ 矩阵之间的相抵关系是等价关系. 于是,根据 λ 矩阵之间的相抵关系可以把数域 $\mathbb{F}$ 上所有 $m\times n$ 阶 λ 矩阵分类: 彼此相抵的 λ 矩阵归于同一类,而不相抵的 λ 矩阵归不同的类.

和通常数域 $\mathbb{F}$ 上矩阵的相抵相同,两个基本问题是: λ 矩阵在相抵下的标准形是什么? λ 矩阵在相抵下的全系不变量是什么?

先考虑 λ 矩阵在相抵下的全系不变量. 容易看出,相抵的 λ 矩阵具有相同的秩. 也就是说,λ 矩阵的秩是 λ 矩阵在相抵下的不变量. 但是,光有 λ 矩阵的秩,还不足以判断两个 λ 矩阵是否相抵. 例如,λ 矩阵

$$A(\lambda)=\begin{pmatrix}1&0\\0&1\end{pmatrix},\qquad B(\lambda)=\begin{pmatrix}1&0\\0&\lambda\end{pmatrix},$$

它们的秩都是 2. 如果 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵,即

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)=P(\lambda)Q(\lambda),$$

其中 $P(\lambda)$ 与 $Q(\lambda)$ 是可逆的 2 阶 λ 矩阵,则 $B(\lambda)$ 是可逆 λ 方阵. 但 $\det B(\lambda)=\lambda$,它不是零次多项式. 由定理 6.4.1,λ 方阵 $B(\lambda)$ 不可逆,矛盾. 因此 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 并不相抵.

这说明,λ 矩阵的秩不足构成 λ 矩阵在相抵下的全系不变量. 所以必须寻求 λ 在相抵下的其它不变量.

定义 6.4.2　设 $A(\lambda)\in(\mathbb{F}[\lambda])^{m\times n}$. $A(\lambda)$ 中所有 k 阶非零子式的最大公因子称为 $A(\lambda)$ 的 ***k* 阶行列式因子**,记为 $D_k(\lambda)$. 如果 λ 矩阵 $A(\lambda)$ 的所有 k 阶子式全为零,则约定 $D_k(\lambda)=0$.

容易看出,如果 $\operatorname{rank}A(\lambda)=r$,则 $D_{r+1}(\lambda)=\cdots=D_s(\lambda)=0$,其中 $s=\min\{m,n\}$,而且当 $1\leqslant k\leqslant r$ 时,$D_1(\lambda),\ldots,D_r(\lambda)$ 都是非零多项式,同时依次一个整除一个.

定理 6.4.5 $m\times n$ 阶 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵的充分必要条件是，它们的行列式因子相同. 简单地说，λ 矩阵的行列式因子是 λ 矩阵在相抵下的全系不变量.

证明 必要性 设 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵，那么 λ 矩阵 $B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$，其中 $P(\lambda)$ 与 $Q(\lambda)$ 分别是 m 阶与 n 阶可逆 λ 矩阵.

根据 Binet-Cauchy 公式，$B(\lambda)$ 的 k 阶子式

$$\begin{aligned}B(\lambda)\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k\end{pmatrix} &= \sum_{1\leqslant \ell_1<\cdots<\ell_k\leqslant m} P(\lambda)\begin{pmatrix} i_1 & \cdots & i_k \\ \ell_1 & \cdots & \ell_k\end{pmatrix}\Big(\big(A(\lambda)Q(\lambda)\big)\begin{pmatrix} \ell_1 & \cdots & \ell_k \\ j_1 & \cdots & j_k\end{pmatrix}\Big)\\ &= \sum_{1\leqslant \ell_1<\cdots<\ell_k\leqslant m}\ \sum_{1\leqslant t_1<\cdots<t_k\leqslant n} P(\lambda)\begin{pmatrix} i_1 & \cdots & i_k \\ \ell_1 & \cdots & \ell_k\end{pmatrix}A(\lambda)\begin{pmatrix} \ell_1 & \cdots & \ell_k \\ t_1 & \cdots & t_k\end{pmatrix}Q(\lambda)\begin{pmatrix} t_1 & \cdots & t_k \\ j_1 & \cdots & j_k\end{pmatrix}.\end{aligned}$$

如果 $A(\lambda)$ 的 k 阶行列式因子 $D_k(\lambda)=0$，则 $A(\lambda)$ 的每个 k 阶子式都为零. 由上式，$B(\lambda)$ 的每个 k 阶子式都为零. 所以 $B(\lambda)$ 的 k 阶行列式因子 $\widetilde{D}_k(\lambda)=0$.

如果 $D_k(\lambda)\neq 0$，则 $D_k(\lambda)$ 整除 $A(\lambda)$ 的每个 k 阶子式，由上式，$D_k(\lambda)$ 整除 $B(\lambda)$ 的每个 k 阶子式，所以 $D_k(\lambda)$ 整除 $\widetilde{D}_k(\lambda)$.

由于相抵关系满足对称性，因此 $\widetilde{D}_k(\lambda)$ 整除 $D_k(\lambda)$. 作为最大公因子 $D_k(\lambda)$ 与 $\widetilde{D}(\lambda)$ 都是首一多项式，所以 $D_k(\lambda)=\widetilde{D}_k(\lambda)$. 这就证明，相抵的 λ 矩阵具有相同的行列式因子.

充分性 设 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 的行列式因子相同，且设 $\operatorname{rank} A(\lambda)=r$. 由定理 6.4.2，$\lambda$ 矩阵 $A(\lambda)$ 相抵于如下的 λ 矩阵：

$$C(\lambda)=\begin{pmatrix}\operatorname{diag}\big(d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)\big) & 0\\ 0 & 0\end{pmatrix}.$$

$C(\lambda)$ 的 k 阶行列式因子记为 $D_k(\lambda)$. 易知，$D_1(\lambda)=d_1(\lambda)$，$D_2(\lambda)=d_1(\lambda)d_2(\lambda),\ldots,$ $D_r(\lambda)=d_1(\lambda)\cdots d_r(\lambda)$，$D_{r+1}(\lambda)=\cdots=D_s(\lambda)=0$，其中 $s=\min\{m,\ n\}$. 因此，

$$d_1(\lambda)=D_1(\lambda),d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)},\ldots,d_r(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}.$$

同样，λ 矩阵 $B(\lambda)$ 相抵于如下的 λ 矩阵：

$$G(\lambda)=\begin{pmatrix}\operatorname{diag}\big(\widetilde{d}_1(\lambda),\widetilde{d}_2(\lambda),\ldots,\widetilde{d}_{\widetilde{r}}(\lambda)\big) & 0\\ 0 & 0\end{pmatrix}.$$

λ 矩阵 $G(\lambda)$ 的 k 阶行列式因子记为 $\widetilde{D}_k(\lambda)$，则

$$\widetilde{d}_1(\lambda)=\widetilde{D}_1(\lambda),\widetilde{d}_2(\lambda)=\frac{\widetilde{D}_2(\lambda)}{\widetilde{D}_1(\lambda)},\ldots,\widetilde{d}_{\widetilde{r}}(\lambda)=\frac{\widetilde{D}_{\widetilde{r}}(\lambda)}{\widetilde{D}_{\widetilde{r}-1}(\lambda)}.$$

根据必要性，λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 的 k 阶行列式分别为 $D_k(\lambda)$ 与 $\widetilde{D}_k(\lambda)$. 由假设，$r=\widetilde{r}$，并且 $D_k=\widetilde{D}_k$. 所以 $d_k(\lambda)=\widetilde{d}_k(\lambda)$. 因此 $C(\lambda)=G(\lambda)$. 再由相抵关系的传递性即知 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵. ■

受定理 6.4.5 的充分性证明的启发，可以引进如下的定义.

定义 6.4.3 设 $A(\lambda)\in\big(\mathbb{F}[\lambda]\big)^{m\times n}$ 的秩为 r，$A(\lambda)$ 的 k 阶行列式因子记为 $D_k(\lambda)$.

则称
$$d_k(\lambda) := \frac{D_k(\lambda)}{D_{k-1}(\lambda)}$$
为 $A(\lambda)$ 的**不变因子**,其中 $k=1,2,\ldots,r$,并约定 $D_0(\lambda)=1$.

定理 6.4.6　设 r 是 $m\times n$ 阶 λ 矩阵 $A(\lambda)$ 的秩,$d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)$ 是 $A(\lambda)$ 的不变因子,则 λ 矩阵 $A(\lambda)$ 相抵于如下的 Smith 标准形:
$$\begin{pmatrix}\operatorname{diag}\big(d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)\big) & 0\\ 0 & 0\end{pmatrix}. \tag{6.4.6}$$
并且 λ 矩阵的不变因子是 λ 矩阵在相抵下的全系不变量.

证明　由定理 6.4.2,λ 矩阵 $A(\lambda)$ 相抵于如下的 λ 矩阵:
$$B(\lambda)=\begin{pmatrix}\operatorname{diag}\big(\widetilde{d}_1(\lambda),\widetilde{d}_2(\lambda),\ldots,\widetilde{d}_r(\lambda)\big) & 0\\ 0 & 0\end{pmatrix},$$
其中 $\widetilde{d}_1(\lambda),\widetilde{d}_2(\lambda),\ldots,\widetilde{d}_r(\lambda)$ 是依次一个整除另一个的首一多项式.

容易算出,λ 矩阵 $B(\lambda)$ 的行列式因子 $D_1(\lambda)=\widetilde{d}_1(\lambda)$, $\widetilde{D}_2(\lambda)=\widetilde{d}_1(\lambda)\widetilde{d}_2(\lambda),\ldots$, $D_r(\lambda)=\widetilde{d}_1(\lambda)\cdots\widetilde{d}_r(\lambda)$, $D_{r+1}(\lambda)=\cdots=D_s(\lambda)=0$,其中 $s=\min\{m,n\}$. 于是,
$$\widetilde{d}_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)}.$$
由定理 6.4.5,$D_1(\lambda),D_2(\lambda),\ldots,D_s(\lambda)$ 是 λ 矩阵 $A(\lambda)$ 的行列式因子. 故
$$\widetilde{d}_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)}=d_k(\lambda),$$
其中 $k=1,\ldots,r$. 这就证明,λ 矩阵 $A(\lambda)$ 相抵于 Smith 标准形 (6.4.6).

由定理 6.4.5,相抵的 λ 矩阵的行列式因子相同,因此不变因子也相同.

反之,如果两个 λ 矩阵的不变因子相同,则它们的行列式因子也相同. 由定理 6.4.5 可知,这两个 λ 矩阵相抵. 所以 λ 矩阵的不变因子是 λ 矩阵在相抵下的全系不变量. ∎

定理 6.4.5 与定理 6.4.6 解决了 λ 矩阵在相抵下的标准形理论的两个基本问题. 下面将给出复数域上 λ 矩阵在相抵下的全系不变量.

设 $A(\lambda)$ 是复数域上 $m\times n$ 阶 λ 矩阵,$\operatorname{rank}A(\lambda)=r$,且 $d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)$ 是 $A(\lambda)$ 的不变因子. 由于复系数多项式可以分解为一次因子的乘积,因此可设
$$\begin{aligned}
d_1(\lambda)&=(\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_t)^{e_{1t}},\\
d_2(\lambda)&=(\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdots(\lambda-\lambda_t)^{e_{2t}},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
d_r(\lambda)&=(\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_t)^{e_{rt}}.
\end{aligned}$$
其中 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 是两两不同地复数,e_{ij} 是非负整数,$i=1,\ldots,r$, $j=1,\ldots,t$.

由于 $d_i(\lambda)$ 整除 $d_{i+1}(\lambda)$,所以 $0\leqslant e_{1j}\leqslant e_{2j}\leqslant\cdots\leqslant e_{rj}$. 当 $e_{j\ell}>0$ 时,因子 $(\lambda-\lambda_j)^{e_{j\ell}}$ 称为矩阵 $A(\lambda)$ 的属于 λ_j 的初等因子,$A(\lambda)$ 的初等因子的全体称为 $A(\lambda)$ 的初等因

子组.

定理 6.4.7 在复数域上的 λ 矩阵的秩和初等因子组是 λ 矩阵在相抵下的全系不变量.

证明 设复数域上 $m\times n$ 阶 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵,则 $\operatorname{rank} A(\lambda)=\operatorname{rank} B(\lambda)$,并且 $A(\lambda)$ 和 $B(\lambda)$ 具有相同的不变因子. 由初等因子组的定义可以看出,λ 矩阵的初等因子组由它的不变因子唯一确定,因此 $A(\lambda)$ 与 $B(\lambda)$ 具有相同的初等因子组.

反之设 $\operatorname{rank} A(\lambda)=\operatorname{rank} B(\lambda)=r$,且

$$\begin{array}{cccc}
(\lambda-\lambda_1)^{m_{11}}, & (\lambda-\lambda_1)^{m_{12}}, & \ldots, & (\lambda-\lambda_1)^{m_{1k_1}}, \\
(\lambda-\lambda_2)^{m_{21}}, & (\lambda-\lambda_2)^{m_{22}}, & \ldots, & (\lambda-\lambda_2)^{m_{2k_2}}, \\
\multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\
(\lambda-\lambda_t)^{m_{t1}}, & (\lambda-\lambda_t)^{m_{t2}}, & \ldots, & (\lambda-\lambda_t)^{m_{tk_t}},
\end{array}$$

是 λ 矩阵 $A(\lambda)$ 的初等因子组,其中 $m_{j1}\geqslant m_{j2}\geqslant\cdots\geqslant m_{jk_j}>0$, $j=1,2,\ldots,t$.

由于 λ 矩阵 $A(\lambda)$ 的不变因子 $d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)$ 依次一个整除另一个,而且 $A(\lambda)$ 的每个初等因子必定是 $A(\lambda)$ 的某个不变因子的一个因子,因此 $d_r(\lambda)$ 是初等因子组中分别属于 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 的最高次幂的初等因子的乘积,即

$$d_r(\lambda)=(\lambda-\lambda_1)^{m_{11}}(\lambda-\lambda_2)^{m_{21}}\cdots(\lambda-\lambda_t)^{m_{t1}}.$$

从 $A(\lambda)$ 的初等因子组中去掉初等因子 $(\lambda-\lambda_1)^{m_{11}},(\lambda-\lambda_2)^{m_{21}},\ldots,(\lambda-\lambda_t)^{m_{t1}}$.

同理,不变因子 $d_{r-1}(\lambda)$ 是余下的初等因子中分别属于 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 的最高次幂的初等因子的乘积,即

$$d_{r-1}(\lambda)=(\lambda-\lambda_1)^{m_{12}}(\lambda-\lambda_2)^{m_{22}}\cdots(\lambda-\lambda_t)^{m_{t2}}.$$

当然,如果在余下的初等因子中不出现属于 λ_j 的初等因子,则令 $(\lambda-\lambda_j)^{m_{j2}}=1$,即 $m_{j2}=0$.

如此继续,即可确定 λ 矩阵 $A(\lambda)$ 的所有不变因子 $d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)$. 这表明,$\lambda$ 矩阵 $A(\lambda)$ 的不变因子由 $A(\lambda)$ 的秩和初等因子组唯一确定. 由于 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 的秩相等,初等因子组相同,所以 $A(\lambda)$ 与 $B(\lambda)$ 具有相同的不变因子. 由定理 6.4.6,λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵. ■

应当指出,定理 6.4.7 中关于 λ 矩阵的秩的条件是不能去掉的. 即是说,光有初等因子组,还不足以构成 λ 矩阵在相抵下的全系不变量. 请读者自己举例说明,具有相同的初等因子组的两个 $m\times n$ 阶 λ 矩阵并不一定相抵.

习 题 6.4

1. 求下列 λ 矩阵的 Smith 标准形,并求出它们的行列式因子,不变因子和初等因子组.

(1) $\begin{pmatrix} \lambda(\lambda+1) & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & (\lambda+1)^2 \end{pmatrix}$; (2) $\begin{pmatrix} \lambda(\lambda-1) & 0 & 0 \\ 0 & \lambda(\lambda-2) & 0 \\ 0 & 0 & (\lambda-1)(\lambda-2) \end{pmatrix}$;

(3) $\begin{pmatrix} \lambda-2 & -1 & 0 \\ 0 & \lambda-2 & -1 \\ 0 & 0 & \lambda-2 \end{pmatrix}$;　　(4) $\begin{pmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{pmatrix}$;

(5) $\begin{pmatrix} \lambda & -1 & 0 & 0 & 0 \\ 0 & \lambda & -1 & 0 & 0 \\ 0 & 0 & \lambda & -1 & 0 \\ 0 & 0 & 0 & \lambda & -1 \\ 1 & 2 & 3 & 4 & 5+\lambda \end{pmatrix}$;　　(6) $\begin{pmatrix} \lambda & 1 & \cdots & 1 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}_{n\times n}$.

2. 证明,任意一个满秩 λ 方阵 $A(\lambda)$ 都可表为 $A(\lambda)=P(\lambda)Q(\lambda)$,其中 $P(\lambda)$ 是可逆 λ 方阵,$Q(\lambda)$ 是上三角 λ 方阵,而且它的对角元都是首一多项式,对角线以上的元素都是次数小于同一列的对角元的次数的多项式. 并证明,这种表法唯一.

3. 证明,对 n 阶 Hermite 方阵 H_1 和 H_2,λ 方阵 $H_1+\lambda H_2$ 的不变因子都是实系数多项式.

§6.5　Jordan 标准形的求法

本节将利用 λ 矩阵在相抵下的标准形理论,重新处理复方阵在相似下的标准形问题,所说的方阵都指复方阵. 先证明一个引理.

引理 6.5.1　任意一个 n 阶 λ 方阵都可以表为矩阵系数的多项式. 具体地说,对给定的 n 阶非零的复 λ 矩阵 $A(\lambda)=\big(a_{ij}(\lambda)\big)$,恒存在 n 阶复方阵 $A_0,A_1,\ldots,A_m$,使得

$$A(\lambda)=A_m\lambda^m+A_{m-1}\lambda^{m-1}+\cdots+A_1\lambda+A_0, \tag{6.5.1}$$

其中 $A_m\neq 0$.

证明　对于 λ 矩阵 $A(\lambda)=\big(a_{ij}(\lambda)\big)$,每个元素 $a_{ij}(\lambda)$ 都是复系数多项式,记为

$$a_{ij}(\lambda)=a_{ij}^{(m)}\lambda^m+a_{ij}^{(m-1)}\lambda^{m-1}+\cdots+a_{ij}^{(1)}\lambda+a_{ij}^{(0)},$$

其中 $1\leqslant i,j\leqslant n$,并且存在某一对 i,j,使得 $a_{ij}^{(m)}\neq 0$. 因此,

$$A(\lambda)=\big(a_{ij}^{(m)}\big)\lambda^m+\big(a_{ij}^{(m-1)}\big)\lambda^{m-1}+\cdots+\big(a_{ij}^{(1)}\big)\lambda+\big(a_{ij}^{(0)}\big).$$

记 $A_k=\big(a_{ij}^{(k)}\big)$,$k=0,1,2,\ldots,m$. 则上式即为 (6.5.1). ■

下面的定理给出方阵的相似与 λ 方阵的相抵之间的重要关系.

定理 6.5.1　n 阶复方阵 A 与 B 相似的充分必要条件是,λ 方阵 $\lambda I_{(n)}-A$ 与 $\lambda I_{(n)}-B$ 相抵.

证明　设方阵 A 与 B 相似,即 $B=P^{-1}AP$,其中 P 是某个 n 阶可逆复方阵. 则

$$\lambda I_{(n)}-B=P^{-1}(\lambda I_{(n)}-A)P.$$

方阵 P 当然是可逆的 λ 方阵. 因此 λ 矩阵 $\lambda I_{(n)}-A$ 与 $\lambda I_{(n)}-B$ 相抵.

反之,设 λ 方阵 $\lambda I_{(n)}-A$ 与 $\lambda I_{(n)}-B$ 相抵,即

$$\lambda I_{(n)}-B=P(\lambda)(\lambda I_{(n)}-A)Q(\lambda),$$

其中 $P(\lambda)$ 与 $Q(\lambda)$ 是 n 阶可逆的 λ 方阵. 由引理 6.5.1,可设

$$Q(\lambda)=Q_k\lambda^k+Q_{k-1}\lambda^{k-1}+\cdots+Q_1\lambda+Q_0,$$
$$Q(\lambda)^{-1}=R_m\lambda^m+R_{m-1}\lambda^{m-1}+\cdots+R_1\lambda+R_0,$$

记

$$W=Q(B)=Q_kB^k+Q_{k-1}B^{k-1}+\cdots+Q_1B+Q_0.$$

因为 $Q(\lambda)^{-1}\cdot Q(\lambda)=I_{(n)}$,所以

$$R_mQ(\lambda)\lambda^m+\cdots+R_1Q(\lambda)\lambda+R_0Q(\lambda)=I_{(n)}.$$

因此

$$R_mWB^m+\cdots+R_1WB+R_0W=I_{(n)}. \tag{6.5.2}$$

但是,由于 $\lambda I_{(n)}-B=P(\lambda)(\lambda_{(n)}-A)Q(\lambda)$,所以

$$P(\lambda)^{-1}(\lambda I_{(n)}-B)=(\lambda I_{(n)}-A)Q(\lambda)=Q(\lambda)\lambda-AQ(\lambda).$$

用方阵 B 代换 λ,得到 $Q(B)B=AQ(B)$. 即 $WB=AW$. 由此得到,

$$WB^2=A^2W,WB^3=A^3W,\ldots,WB^m=A^mW.$$

于是式 (6.5.2) 化为

$$(R_mA^m+\cdots+R_1A+R_0)W=I_{(n)}.$$

这表明,方阵 W 可逆,并且 $B=W^{-1}AW$. ■

对于 n 阶复方阵 A,λ 方阵 $\lambda I_{(n)}-A$ 称为方阵 A 的**特征方阵**.

显然特征方阵 $\lambda I_{(n)}-A$ 总是满秩的. 因此由定理 6.5.1 与定理 6.4.7,立即得到

定理 6.5.2 n 阶复方阵 A 与 B 相似当且仅当它们的特征方阵 $\lambda I_{(n)}-A$ 与 $\lambda I_{(n)}-B$ 的初等因子组相同.

也就是说,复方阵的特征方阵的初等因子组是复方阵在相似下的全系不变量.

利用复方阵 A 的特征方阵 $\lambda I_{(n)}-A$ 的初等因子组,可以给出复方阵 A 在相似下的 Jordan 标准形. 为此需要

引理 6.5.2 准对角 λ 矩阵 $A(\lambda)=\operatorname{diag}\big(A_1(\lambda),\ A_2(\lambda)\big)$ 的初等因子组由 $A_1(\lambda)$ 与 $A_2(\lambda)$ 的初等因子组合并而成.

证明 设 $d_1(\lambda),\ldots,d_r(\lambda)$ 和 $d_{r+1}(\lambda),\ldots,d_{r+s}(\lambda)$ 分别是 λ 矩阵 $A_1(\lambda)$ 和 $A_2(\lambda)$ 的不变因子. 则存在可逆 λ 方阵 $P_1(\lambda),P_2(\lambda)$ 和 $Q_1(\lambda),Q_2(\lambda)$,使得

$$P_1(\lambda)A_1(\lambda)Q_1(\lambda)=\begin{pmatrix}\operatorname{diag}\big(d_1(\lambda),\ldots,d_r(\lambda)\big) & 0\\ 0 & 0\end{pmatrix},$$

$$P_2(\lambda)A_2(\lambda)Q_2(\lambda)=\begin{pmatrix}\operatorname{diag}\big(d_{r+1}(\lambda),\ldots,d_{r+s}(\lambda)\big) & 0\\ 0 & 0\end{pmatrix}.$$

记

$$P(\lambda)=\operatorname{diag}\big(P_1(\lambda),\ P_2(\lambda)\big),\quad Q(\lambda)=\operatorname{diag}\big(Q_1(\lambda),\ Q_2(\lambda)\big).$$

显然 $P(\lambda)$ 和 $Q(\lambda)$ 都可逆,并且

$$P(\lambda)A(\lambda)Q(\lambda) = \operatorname{diag}\big(D_1, 0, D_2, 0\big),$$

其中

$$D_1 = \operatorname{diag}\big(d_1(\lambda), \ldots, d_r(\lambda)\big), \quad D_2 = \operatorname{diag}\big(d_{r+1}(\lambda), \ldots, d_{r+s}(\lambda)\big).$$

记右端的 λ 矩阵为 $B(\lambda)$. 于是 λ 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵,因此它们的初等因子组相同.

下面求 λ 矩阵 $B(\lambda)$ 的初等因子组. 记

$$\begin{gathered}
d_1(\lambda) = (\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_k)^{e_{1k}},\\
\cdots\cdots\cdots\cdots\cdots\cdots\\
d_r(\lambda) = (\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_k)^{e_{rk}},\\
d_{r+1}(\lambda) = (\lambda-\lambda_1)^{e_{r+1,1}}(\lambda-\lambda_2)^{e_{r+1,2}}\cdots(\lambda-\lambda_k)^{e_{r+1,k}},\\
\cdots\cdots\cdots\cdots\cdots\cdots\\
d_1(\lambda) = (\lambda-\lambda_1)^{e_{r+s,1}}(\lambda-\lambda_2)^{e_{r+s,2}}\cdots(\lambda-\lambda_k)^{e_{r+s,k}},
\end{gathered}$$

其中 e_{ij} 是非负整数,$\lambda_1, \lambda_2, \ldots, \lambda_k$ 是两两不同地复数. 由于不变因子依次一个整除另一个,所以 $0 \leqslant e_{1j} \leqslant e_{2j} \leqslant \cdots \leqslant e_{rj}, 0 \leqslant e_{r+1,j} \leqslant e_{r+2,j} \leqslant \cdots \leqslant e_{r+s,j}, j = 1, 2, \ldots, k$.

显然,λ 矩阵 $A_1(\lambda)$ 的初等因子组为 $(\lambda-\lambda_j)^{e_{ij}}(1 \leqslant i \leqslant r)$,而 λ 矩阵 $A_2(\lambda)$ 的初等因子组为 $(\lambda-\lambda_j)^{e_{r+i,j}}(1 \leqslant i \leqslant s)$.

对于每个 j,将指数 $e_{1j}, \ldots, e_{rj}, e_{r+1,j}, \ldots, e_{r+s,j}$ 重排为 $e'_{1j}, e'_{2j}, \ldots, e'_{r+s,j}$,使得

$$0 \leqslant e'_{1j} \leqslant e'_{2j} \leqslant \cdots \leqslant e'_{r+s,j}.$$

容易看出,λ 矩阵 $B(\lambda)$ 的 $r+s$ 阶行列式因子 $D_{r+s}(\lambda)$ 为

$$\begin{aligned}
D_{r+s}(\lambda) &= d_1(\lambda)\cdots d_r(\lambda)d_{r+1}(\lambda)\cdots d_{r+s}(\lambda)\\
&= (\lambda-\lambda_1)^{e'_{11}+\cdots+e'_{r+s,1}}(\lambda-\lambda_2)^{e'_{12}+\cdots+e'_{r+s,2}}\cdots(\lambda-\lambda_k)^{e'_{1k}+\cdots+e'_{r+s,k}}.
\end{aligned}$$

$B(\lambda)$ 的所有非零 $r+s-1$ 阶子式为 $\dfrac{D_{r+s}(\lambda)}{d_i(\lambda)}, i = 1, 2, \ldots, r+s$,而

$$\frac{D_{r+s}(\lambda)}{d_i(\lambda)} = (\lambda-\lambda_1)^{e'_{11}+\cdots+e'_{r+s,1}-e_{i1}}\cdots(\lambda-\lambda_k)^{e'_{1k}+\cdots+e'_{r+s,k}-e_{ik}}.$$

由于 $e'_{r+s,j} \geqslant e_{ij}$,因此

$$e'_{1j}+\cdots+e'_{r+s,j}-e_{ij} \geqslant e'_{1j}+\cdots+e'_{r+s-1,j}.$$

并且等式当且仅当 $e_{ij} = e'_{r+s,j}$ 时成立.

所以,多项式 $\dfrac{D_{r+s}(\lambda)}{d_1(\lambda)}, \ldots, \dfrac{D_{r+s}(\lambda)}{d_{r+s}(\lambda)}$ 的最大公因子,也即 $B(\lambda)$ 的 $r+s-1$ 阶行列式因子 $D_{r+s-1}(\lambda)$ 为

$$D_{r+s-1}(\lambda) = (\lambda-\lambda_1)^{e'_{11}+\cdots+e'_{r+s-1,1}}\cdots(\lambda-\lambda_k)^{e'_{1k}+\cdots+e'_{r+s-1,k}}.$$

与 $B(\lambda)$ 的第 $r+s$ 个不变因子为

$$\widetilde{d}_{r+s}(\lambda) = (\lambda-\lambda_1)^{e'_{r+s,1}}\cdots(\lambda-\lambda_k)^{e'_{r+s,k}}.$$

同理可得,$B(\lambda)$ 的第 ℓ 个不变因子为 $(\ell=1,2,\ldots,r+s)$:

$$\widetilde{d}_\ell(\lambda)=(\lambda-\lambda_1)^{e'_{\ell 1}}\cdots(\lambda-\lambda_k)^{e'_{\ell k}}.$$

因此 $B(\lambda)$ 的初等因子组为

$$\left\{(\lambda-\lambda_j)^{e'_{\ell j}}\;\middle|\; e'_{\ell j}>0, 1\leqslant \ell\leqslant r+s, 1\leqslant j\leqslant k\right\}.$$

由于 $e'_{1j},e'_{2j},\ldots,e'_{r+s,j}$ 是 $e_{1,j},e_{2j},\ldots,e_{r+s,j}$ 的重排,$j=1,2,\ldots,k$,所以 λ 矩阵 $B(\lambda)$ 的初等因子组是 $A_1(\lambda)$ 与 $A_2(\lambda)$ 的初等因子组合并而成的. ■

定理 6.5.3 设 n 阶复方阵 A 的特征方阵 $\lambda I_{(n)}-A$ 的初等因子组为

$$\left\{\begin{array}{llll}(\lambda-\lambda_1)^{m_{11}}, & (\lambda-\lambda_1)^{m_{12}}, & \ldots, & (\lambda-\lambda_1)^{m_{1k_1}},\\ (\lambda-\lambda_2)^{m_{21}}, & (\lambda-\lambda_2)^{m_{22}}, & \ldots, & (\lambda-\lambda_2)^{m_{2k_2}},\\ & \cdots\cdots\cdots\cdots\cdots\cdots & & \\ (\lambda-\lambda_t)^{m_{t1}}, & (\lambda-\lambda_t)^{m_{t2}}, & \ldots, & (\lambda-\lambda_t)^{m_{tk_t}},\end{array}\right. \tag{6.5.3}$$

其中 $m_{j1}\geqslant m_{j2}\geqslant\cdots\geqslant m_{jk_j}>0$, $j=1,2,\ldots,t$; 且 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 是两两不同地复数. 则复方阵 A 相似于如下的 Jordan 标准形:

$$J=\operatorname{diag}\left(J^{\mathrm T}_{(m_{11})}(\lambda_1),\ldots,J^{\mathrm T}_{(m_{1k_1})}(\lambda_1),\ldots\ldots,J^{\mathrm T}_{(m_{t1})}(\lambda_t),\ldots,J^{\mathrm T}_{(m_{tk_t})}(\lambda_t)\right),$$

其中对任意 $\ell=1,2,\ldots,k_j$, $j=1,2,\ldots,t$,

$$J^{\mathrm T}_{(m_{j\ell})}(\lambda_j)=\lambda_j I_{(m_{j\ell})}+N^{\mathrm T}_{(m_{j\ell})}.$$

证明 先计算准对角方阵 J 的对角块 $J^{\mathrm T}_{(m_{j\ell})}(\lambda_j)$ 的特征方阵 $\lambda I_{(m_{j\ell})}-J^{\mathrm T}_{(m_{j\ell})}(\lambda_j)$ 的初等因子组. 显然

$$\lambda I_{(m_{j\ell})}-J^{\mathrm T}_{(m_{j\ell})}(\lambda_j)=\begin{pmatrix}\lambda-\lambda_j & 0 & \cdots & \cdots & 0\\ -1 & \lambda-\lambda_j & \ddots & & \vdots\\ 0 & -1 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & -1 & \lambda-\lambda_j\end{pmatrix}_{m_{j\ell}\times m_{j\ell}}.$$

容易求出,它的行列式因子为

$$D_1(\lambda)=\cdots=D_{m_{j\ell}-1}(\lambda)=1, D_{m_{j\ell}}(\lambda)=(\lambda-\lambda_j)^{m_{j\ell}}.$$

所以,它的不变因子为

$$d_1(\lambda)=\cdots=d_{m_{j\ell}-1}(\lambda)=1, d_{m_{j\ell}}(\lambda)=(\lambda-\lambda_j)^{m_{j\ell}}.$$

于是特征方阵 $\lambda I_{(m_{j\ell})}-J^{\mathrm T}_{(m_{j\ell})}(\lambda_j)$ 的初等因子组为 $(\lambda-\lambda_j)^{m_{j\ell}}$.

由引理 6.5.2,J 的初等因子组即为 (6.5.3). 由定理 6.5.2,方阵 A 与 J 相似. ■

至此我们重新用 λ 矩阵在相抵下的标准形理论处理了复方阵在相似下的标准形问题. 现在的问题是,复方阵 A 的特征方阵 $\lambda I_{(n)}-A$ 的初等因子组和 §6.3 所定义的复方阵 A 的初等因子组有何联系? 对此有

定理 6.5.4 n 阶复方阵 A 的特征方阵 $\lambda I_{(n)}-A$ 的初等因子组即是 A 的初等因子组.

证明 设 n 阶复方阵 A 的特征方阵 $\lambda I_{(n)}-A$ 的初等因子组为

$$(\lambda-\lambda_1)^{e_1},(\lambda-\lambda_2)^{e_2},\dots,(\lambda-\lambda_t)^{e_t}.$$

这里复数 $\lambda_1,\lambda_2,\dots,\lambda_t$ 允许相同. 则由定理 6.5.3 得

$$P^{-1}AP=J=\operatorname{diag}\big(J_{(e_1)}^{\mathrm{T}}(\lambda_1),J_{(e_2)}^{\mathrm{T}}(\lambda_2),\dots,J_{(e_t)}^{\mathrm{T}}(\lambda_t)\big),$$

其中 P 为 n 阶可逆复方阵. 取 n 维复线性空间 V 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$. 则由

$$\begin{aligned}\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n)&=(\alpha_1,\alpha_2,\dots,\alpha_n)A,\\ \mathscr{J}(\alpha_1,\alpha_2,\dots,\alpha_n)&=(\alpha_1,\alpha_2,\dots,\alpha_n)J,\\ \mathscr{P}(\alpha_1,\alpha_2,\dots,\alpha_n)&=(\alpha_1,\alpha_2,\dots,\alpha_n)P\end{aligned}\tag{6.5.4}$$

便确定了 V 的线性变换 $\mathscr{A},\mathscr{J}$ 与可逆线性变换 $\mathscr{P}$. 由于

$$\begin{aligned}\mathscr{P}^{-1}\mathscr{A}\mathscr{P}(\alpha_1,\alpha_2,\dots,\alpha_n)&=(\alpha_1,\alpha_2,\dots,\alpha_n)P^{-1}AP\\&=(\alpha_1,\alpha_2,\dots,\alpha_n)J,\end{aligned}$$

所以 $\mathscr{J}=\mathscr{P}^{-1}\mathscr{A}\mathscr{P}$,即线性变换 $\mathscr{A}$ 与 $\mathscr{J}$ 等价. 由定理 6.3.3,线性变换 $\mathscr{A}$ 与 $\mathscr{J}$ 的初等因子组相同. 现在确定线性变换 $\mathscr{J}$ 的初等因子组. 由于

$$J_{(e_j)}^{\mathrm{T}}(\lambda_j)=\lambda_jI_{(e_j)}+N_{(e_j)}^{\mathrm{T}},$$

所以由 (6.5.4) 得

$$\begin{aligned}\mathscr{J}(\alpha_{f_{j-1}+1})&=\lambda_j\alpha_{f_{j-1}+1}+\alpha_{f_{j-1}+2};\\ \mathscr{J}(\alpha_{f_{j-1}+2})&=\lambda_j\alpha_{f_{j-1}+2}+\alpha_{f_{j-1}+3};\\ &\cdots\cdots\cdots\cdots\\ \mathscr{J}(\alpha_{f_{j-1}+e_j-1})&=\lambda_j\alpha_{f_{j-1}+e_j-1}+\alpha_{f_j};\\ \mathscr{J}(\alpha_{f_j})&=\lambda_j\alpha_{f_j},\end{aligned}$$

其中 $f_j=e_1+e_2+\cdots+e_j$, $j=1,2,\dots,e_t$. 因此,

$$\begin{aligned}\alpha_{f_{j-1}+2}&=(\mathscr{J}-\lambda_j\mathscr{I})(\alpha_{f_{j-1}+1});\\ \alpha_{f_{j-1}+3}&=(\mathscr{J}-\lambda_j\mathscr{I})^2(\alpha_{f_{j-1}+1});\\ &\cdots\cdots\cdots\cdots\\ \alpha_{f_{j-1}+e_j-1}&=(\mathscr{J}-\lambda_j\mathscr{I})^{e_j-2}(\alpha_{f_{j-1}+1});\\ \alpha_{f_j}&=(\mathscr{J}-\lambda_j\mathscr{I})^{e_j-1}(\alpha_{f_{j-1}+1}),\end{aligned}$$

其中 $\mathscr{I}$ 是 V 的单位变换.

记 V 中由向量 $\alpha_{f_{j-1}+1},\alpha_{f_{j-1}+2},\dots,\alpha_{f_{j-1}+e_j-1},\alpha_{f_j}$ 生成的子空间为 C_j,则 C_j 是由向量 $\alpha_{f_{j-1}+1}$ 生成的循环子空间,并且

$$V=C_1\oplus C_2\oplus\cdots\oplus C_t.$$

容易算得,$\mathscr{J}\big|_{C_j}$ 的最小多项式为 $(\lambda-\lambda_j)^{e_j}$. 这就证明,线性变换 $\mathscr{J}$ 的初等因子组也是

$$(\lambda-\lambda_1)^{e_1},(\lambda-\lambda_1)^{e_2},\dots,(\lambda-\lambda_t)^{e_t}.$$

■

最后总结一下求复方阵 A 的 Jordan 标准形的过程如下：

(1) 对特征方阵 $\lambda I_{(n)}-A$ 施行行或列的初等 λ 变换，把 $\lambda I_{(n)}-A$ 化为 Smith 标准形，由此求得不变因子 $d_1(\lambda), d_2(\lambda), \ldots, d_n(\lambda)$；

(2) 把不变因子 $d_1(\lambda), d_2(\lambda), \ldots, d_n(\lambda)$ 分解为一次因子的乘积，求出初等因子组

$$(\lambda-\lambda_1)^{e_1}, \quad (\lambda-\lambda_2)^{e_2},, \quad \ldots, \quad (\lambda-\lambda_t)^{e_t};$$

(3) 写出每个初等因子 $(\lambda-\lambda_j)^{e_j}$ 相应的 Jordan 块 $\lambda_j I_{(e_j)}+N_{(e_j)}^{\mathrm{T}}$，$j=1,2,\ldots,t$，然后将它们组成准对角方阵，即得方阵 A 的 Jordan 标准形 J.

使得 $P^{-1}AP$ 为 Jordan 标准形 J 的可逆方阵 P 称为**过渡方阵**.

过渡方阵 P 的求法如下（请读者自己证明，如此求得的方阵 P 确实是过渡方阵）：

(1) 对特征方阵 $\lambda I_{(n)}-A$ 施行行或列的初等 λ 变换，把 $\lambda I_{(n)}-A$ 化为 $\lambda I_{(n)}-J$，便求得 n 阶可逆 λ 方阵 $P(\lambda)$ 和 $Q(\lambda)$，使得

$$P(\lambda)(\lambda I_{(n)}-A)Q(\lambda)=\lambda I_{(n)}-J;$$

(2) 通过行的初等 λ 变换，把 $n\times 2n$ 阶 λ 矩阵 $\big(Q(\lambda),\ I_{(n)}\big)$ 化为 $\big(I_{(n)},\ R(\lambda)\big)$，即可求出 $Q(\lambda)$ 的逆 $Q(\lambda)^{-1}=R(\lambda)$；

(3) 把 $R(\lambda)$ 写成矩阵系数的多项式，即

$$R(\lambda)=R_m\lambda^m+R_{m-1}\lambda^{m-1}+R_1\lambda+R_0.$$

用方阵 A 代换其中的 λ，得

$$R(A)=R_mA^m+R_{m-1}A^{m-1}+R_1\mathscr{A}+R_0I_{(n)}.$$

方阵 $R(A)$ 即是所求的过渡方阵 P.

习　题　6.5

1. 求下列方阵的 Jordan 标准形.

(1) $\begin{pmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{pmatrix}$;　　(2) $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$;

(3) $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$;　　(4) $\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ \alpha & 0 & \alpha \end{pmatrix}, \alpha\neq 0$;

(5) $\begin{pmatrix} 4 & 6 & -15 \\ 3 & 4 & -12 \\ 2 & 3 & -8 \end{pmatrix}$;　　(6) $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$;

(7) $\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 1 & -1 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}_{n\times n}$;

(8) $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}_{n\times n}$;

(9) $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}_{n\times n}^2$;

(10) $\begin{pmatrix} a & 0 & 1 & 0 & \cdots & 0 \\ 0 & a & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 & 1 \\ \vdots & & & \ddots & a & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & a \end{pmatrix}_{n\times n}, n \geq 3$;

(11) $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}_{n\times n}$;

(12) $\begin{pmatrix} 0 & \cdots & 0 & a_{1n} \\ \vdots & \iddots & \iddots & 0 \\ 0 & \iddots & \iddots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{pmatrix}_{n\times n}$.

2. 设 n 阶方阵 J 为

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix},$$

$f(\lambda)$ 是关于 λ 的多项式. 证明:

$$f(J) = \begin{pmatrix} f(\lambda) & \dfrac{f'(\lambda)}{1!} & \dfrac{f''(\lambda)}{2!} & \cdots & \dfrac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & f(\lambda) & \dfrac{f'(\lambda)}{1!} & \cdots & \dfrac{f^{(n-2)}(\lambda)}{(n-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & f(\lambda) & \dfrac{f'(\lambda)}{1!} \\ 0 & \cdots & \cdots & 0 & f(\lambda) \end{pmatrix}.$$

3. 证明,一组两两可交换的可对角化方阵可以用同一个可逆方阵相似于对角形.

4. 证明,方阵 A 的最小多项式 $d(\lambda)$ 的某次幂能被 A 的特征多项式 $\varphi(\lambda)$ 所整除.

5. 证明,方阵 A 相似于对角形的充分必要条件为,它的初等因子都是一次的.

6. 设 A 是可逆方阵. 证明 AB 和 BA 相似. 当 A 不可逆时,结论是否仍成立?

7. 设方阵 A 的特征值全是 1. 证明方阵 A 的任意次幂都与 A 相似.

8. 设 A 是 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 在 V 的某组基下的方阵. 证明 $\mathscr{A}$ 是循环变换的充分必要条件为,方阵 A 的特征方阵 $\lambda I_{(n)} - A$ 的 $n-1$ 阶行列式因子 $D_{n-1}(\lambda) = 1$.

§6.6 一些例子

本节通过一些例子来说明 Jordan 标准形的应用.

例 6.6.1 n 阶复方阵 A 和它的转置 A^{T} 相似.

证法 1 容易看出,特征方阵 $\lambda I_{(n)}-A$ 和 $\lambda I_{(n)}-A^{\mathrm{T}}$ 的行列式因子相同. 因此特征方阵 $\lambda I_{(n)}-A$ 和 $\lambda I_{(n)}-A^{\mathrm{T}}$ 相抵. 所以方阵 A 和 A^{T} 相似. ∎

证法 2 设方阵 A 的初等因子组为 $(\lambda-\lambda_1)^{e_1},(\lambda-\lambda_2)^{e_2},\ldots,(\lambda-\lambda_t)^{e_t}$,则方阵 A 相似于 Jordan 标准形 J,即

$$P^{-1}AP=J=\mathrm{diag}\left(\lambda_1 I_{(e_1)}+N_{(e_1)}^{\mathrm{T}},\lambda_2 I_{(e_2)}+N_{(e_2)}^{\mathrm{T}},\ldots,\lambda_t I_{(e_t)}+N_{(e_t)}^{\mathrm{T}}\right),$$

其中 P 是某个 n 阶阶可逆复方阵. 于是

$$P^{\mathrm{T}}A^{\mathrm{T}}(P^{\mathrm{T}})^{-1}=J^{\mathrm{T}}=\mathrm{diag}\left(\lambda_1 I_{(e_1)}+N_{(e_1)},\lambda_2 I_{(e_2)}+N_{(e_2)},\ldots,\lambda_t I_{(e_t)}+N_{(e_t)}\right).$$

记

$$S_j=\begin{pmatrix} & & 1\\ & \iddots & \\ 1 & & \end{pmatrix}_{e_j\times e_j}.$$

则

$$S_j^{-1}\left(\lambda_j I_{(e_j)}+N_{(e_j)}^{\mathrm{T}}\right)S_j=\lambda_j I_{(e_j)}+N_{(e_j)},$$

记 $S=\mathrm{diag}(S_1,S_2,\ldots,S_t)$,则

$$S^{-1}P^{-1}APS=J^{\mathrm{T}}=P^{\mathrm{T}}A^{\mathrm{T}}(P^{\mathrm{T}})^{-1}.$$

所以 $(PSP^{\mathrm{T}})^{-1}A(PSP^{\mathrm{T}})=A^{\mathrm{T}}$,其中方阵 PSP^{T} 显然可逆. ∎

注 1 证法 2 不但证明方阵 A 和它的转置 A^{T} 相似,而且还指出可以选取可逆对称方阵 P,使得 $P^{-1}AP=A^{\mathrm{T}}$.

注 2 由例 6.6.1,属于初等因子 $(\lambda-\lambda_j)^{e_j}$ 的 Jordan 块 $\lambda_j I_{(e_j)}+N_{(e_j)}^{\mathrm{T}}$ 和 $\lambda_j I_{(e_j)}+N_{(e_j)}$ 相似,Jordan 标准形

$$J=\mathrm{diag}\left(\lambda_1 I_{(e_1)}+N_{(e_1)}^{\mathrm{T}},\lambda_2 I_{(e_2)}+N_{(e_2)}^{\mathrm{T}},\ldots,\lambda_t I_{(e_t)}+N_{(e_t)}^{\mathrm{T}}\right)$$

和

$$\mathrm{diag}\left(\lambda_1 I_{(e_1)}+N_{(e_1)},\lambda_2 I_{(e_2)}+N_{(e_2)},\ldots,\lambda_t I_{(e_t)}+N_{(e_t)}\right) \tag{6.6.1}$$

相似. 所以 $\lambda_j I_{(e_j)}+N_{(e_j)}$ 也称为属于 $(\lambda-\lambda_j)^{e_j}$ 的 Jordan 块,而形如 (6.6.1) 的上三角阵也称为 Jordan 标准形. 以后我们说方阵 A 的 Jordan 标准形都指形如 (6.6.1) 的上三角方阵.

例 6.6.2 任意一个 n 阶复方阵 A 都可以表示为两个 n 阶对称方阵的乘积,而且可以指定其中一个是非奇异的.

证明　由例 6.6.1，存在 n 阶可逆对称复方阵 S，使得 $S^{-1}A^{\mathrm{T}}S = A$.

记 $S_1 = S^{-1}, S_2 = A^{\mathrm{T}}S$. 则 $A = S_1S_2$. 由于 $A^{\mathrm{T}}S = SA$，因此

$$S_2^{\mathrm{T}} = (A^{\mathrm{T}}S)^{\mathrm{T}} = (SA)^{\mathrm{T}} = A^{\mathrm{T}}S = S_2,$$

所以 S_2 是对称的. 显然 $S_1 = S^{-1}$ 是对称而且可逆的. 记 $\widetilde{S}_1 = S^{-1}A^{\mathrm{T}}, \widetilde{S}_2 = S$，则 $A = \widetilde{S}_1\widetilde{S}_2$. 由于

$$S^{-1}A^{\mathrm{T}} = AS^{-1},$$

所以

$$\widetilde{S}_1^{\mathrm{T}} = (S^{-1}A^{\mathrm{T}})^{\mathrm{T}} = (AS^{-1})^{\mathrm{T}} = S^{-1}A' = \widetilde{S}_1,$$

即方阵 $\widetilde{S}_1$ 是对称的. 显然 $\widetilde{S}_2 = S$ 是可逆对称的. ■

例 6.6.3　任意一个 n 阶复方阵 A 都相似于这样的上三角阵，它的非零的非对角元都是任意小的正数.

证明　设 $(\lambda-\lambda_1)^{e_1}, (\lambda-\lambda_2)^{e_2}, \dots, (\lambda-\lambda_t)^{e_t}$ 是方阵 A 的初等因子组，则存在可逆方阵 P，使得

$$P^{-1}AP = J = \operatorname{diag}\big(\lambda_1 I_{(e_1)} + N_{(e_1)}, \lambda_2 I_{(e_2)} + N_{(e_2)}, \dots, \lambda_t I_{(e_t)} + N_{(e_t)}\big).$$

设 ε 是任意小的正数，且

$$E_{(e_j)} = \operatorname{diag}\big(\varepsilon, \varepsilon^2, \dots, \varepsilon^{e_j}\big).$$

则

$$E_{(e_j)}^{-1}\big(\lambda_j I_{(e_j)} + N_{(e_j)}\big)E_{(e_j)} = \begin{pmatrix} \lambda_j & \varepsilon & & \\ & \ddots & \ddots & \\ & & \ddots & \varepsilon \\ & & & \lambda_j \end{pmatrix}_{e_j \times e_j}.$$

记

$$E = \operatorname{diag}\big(E_{(e_1)}, E_{(e_2)}, \dots, E_{(e_t)}\big).$$

则

$$(PE)^{-1}A(PE) = \operatorname{diag}\big(\lambda_1 I_{(e_1)} + \varepsilon N_{(e_1)}, \lambda_2 I_{(e_2)} + \varepsilon N_{(e_2)}, \dots, \lambda_t I_{(e_t)} + \varepsilon N_{(e_t)}\big). \quad ■$$

注　由例 6.6.1，若把例 6.6.3 中上三角方阵改成为下三角方阵，结论仍成立.

例 6.6.4　n 阶复方阵 A 的最小多项式 $d(\lambda)$ 等于 A 的特征方阵 $\lambda I_{(n)} - A$ 的第 n 个不变因子 $d_n(\lambda)$.

证明　设方阵 A 的初等因子组为

$$\begin{array}{cccc}
(\lambda-\lambda_1)^{m_{11}}, & (\lambda-\lambda_1)^{m_{12}}, & \dots, & (\lambda-\lambda_1)^{m_{1k_1}}, \\
(\lambda-\lambda_2)^{m_{21}}, & (\lambda-\lambda_2)^{m_{22}}, & \dots, & (\lambda-\lambda_2)^{m_{2k_2}}, \\
& \cdots\cdots\cdots\cdots\cdots\cdots & & \\
(\lambda-\lambda_t)^{m_{t1}}, & (\lambda-\lambda_t)^{m_{t2}}, & \dots, & (\lambda-\lambda_t)^{m_{tk_t}},
\end{array}$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_t$ 是两两不同的复数，$m_{j1} \geqslant m_{j2} \geqslant \cdots \geqslant m_{jk_j}, j = 1, 2, \dots, t$.

那么特征方阵 $\lambda I_{(n)} - A$ 的第 n 个不变因子为

$$d_n(\lambda) = (\lambda - \lambda_1)^{m_{11}}(\lambda - \lambda_2)^{m_{21}}\cdots(\lambda - \lambda_t)^{m_{t1}}.$$

记属于 $(\lambda - \lambda_j)^{m_{j\ell}}$ 的 Jordan 块为

$$J_{j\ell} = \lambda_j I_{(m_{j\ell})} + N_{(m_{j\ell})},$$

其中 $\ell = 1, 2, \ldots, k_j$，$j = 1, 2, \ldots, t$. 则存在可逆方阵 P，使得

$$A = P^{-1}JP = P^{-1}\operatorname{diag}\big(J_{11}, \ldots, J_{1k_1}, \ldots\ldots, J_{t1}, \ldots, J_{tk_t}\big)P.$$

容易验证，如果 $f(\lambda)$ 是复系数多项式，且 $A = P^{-1}BP$，则 $f(A) = P^{-1}f(B)P$. 而且如果方阵 $A = \operatorname{diag}\big(A_1, A_2, \ldots, A_k\big)$，则

$$f(A) = \operatorname{diag}\big(f(A_1), f(A_2), \ldots, f(A_k)\big).$$

因此

$$d_n(J) = \operatorname{diag}\big(d_n(J_{11}), \ldots, d_n(J_{1k_1}), \ldots\ldots, d_n(J_{t1}), \ldots, d_n(J_{tk_t})\big).$$

由于

$$d_n(J_{j\ell}) = (J_{j\ell} - \lambda_1 I_{(m_{j\ell})})^{m_{11}}(J_{j\ell} - \lambda_2 I_{(m_{j\ell})})^{m_{21}}\cdots(J_{j\ell} - \lambda_t I_{(m_{j\ell})})^{m_{t1}},$$

并且

$$(J_{j\ell} - \lambda_j I_{(m_{j\ell})})^{m_{j1}} = (N_{(m_{j\ell})})^{m_{j1}} = 0,$$

所以 $d_n(J_{j\ell}) = 0$，其中 $\ell = 1, 2, \ldots, k_j$，$j = 1, 2, \ldots, t$. 于是 $d_n(J) = 0$. 因此

$$d_n(A) = P^{-1}d_n(J)P = 0.$$

这就证明，$d_n(\lambda)$是方阵 A 的化零多项式. 因为 $d(\lambda)$ 是方阵 A 的最小多项式，所以 $d(\lambda)$ 整除 $d_n(\lambda)$. 而 $d(\lambda)$ 又是首一多项式，因此可设

$$d(\lambda) = (\lambda - \lambda_1)^{e_1}(\lambda - \lambda_2)^{e_2}\cdots(\lambda - \lambda_t)^{e_t},$$

其中 $e_j \leqslant m_{j1}$，$j = 1, 2, \ldots, t$. 如果有某个 $e_j < m_{j1}$，则

$$d(J) = \operatorname{diag}\big(d(J_{11}), \ldots, d(J_{1k_1}), \ldots\ldots, d(J_{t1}), \ldots, d(J_{tk_t})\big),$$

其中

$$d(J_{j1}) = (J_{j1} - \lambda_1 I_{(m_{j1})})^{e_1}(J_{j1} - \lambda_2 I_{(m_{j1})})^{e_2}\cdots(J_{j1} - \lambda_t I_{(m_{j1})})^{e_t}.$$

显然，当 $i \neq j$ 时，

$$(J_{j1} - \lambda_i I_{(m_{j1})})^{e_i}$$

是上三角阵，而且对角元素都是 $(\lambda_j - \lambda_i)^{e_i} \neq 0$，所以 $(J_{j1} - \lambda_i I_{(m_{j1})})^{e_i}$ 可逆.

当 $i = j$ 时，由于 $e_j < m_{j1}$，所以

$$(J_{j1} - \lambda_j I_{(m_{j1})})^{e_j} = (N_{(m_{j1})})^{e_j} \neq 0.$$

因此 $d(J_{j1}) \neq 0$，从而 $d(J) \neq 0$. 于是

$$d(A) = P^{-1}d(J)P \neq 0,$$

这与最小多项式是方阵 A 的化零多项式相矛盾.

所以 $e_j = m_{j1}$，$j = 1, 2, \ldots, t$，也即 $d(\lambda) = d_n(\lambda)$. ■

例 6.6.5 n 阶复方阵 A 相似于对角形的当且仅当 A 的最小多项式没有重根.

证明　设方阵 A 相似于对角形

$$J=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n).$$

则方阵 J 即是方阵 A 的 Jordan 标准形，所以方阵 A 的初等因子组为

$$\lambda-\lambda_1,\lambda-\lambda_2,\ldots,\lambda-\lambda_n.$$

设 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 是 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 中全部不同的复数，则方阵 A 的最小多项式（也即特征方阵 $\lambda I_{(n)}-A$ 的第 n 个不变因子）为

$$d(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_t).$$

因此 $d(\lambda)$ 没有重根.

反之，设方阵 A 的最小多项式 $d(\lambda)$ 没有重根，则特征方阵 $\lambda I_{(n)}-A$ 的第 n 个不变因子没有重根. 所以方阵 A 的初等因子都是一次的. 因此属于每个初等因子的 Jordan 块都是一阶方阵. 这表明，方阵 A 的 Jordan 标准形是对角方阵. ■

例 6.6.6　设 n 阶复方阵 A 的最小多项式 $d(\lambda)$ 的次数为 s，且 $B=(b_{ij})$ 是 s 阶方阵，其中 $b_{ij}=\operatorname{Tr}A^{i+j}$. 则方阵 A 相似于对角形的充分必要条件是 $\det B\neq 0$.

证明　设 $\lambda_1,\ldots,\lambda_t$ 是 A 的全部不同特征值，它们的代数重数依次为 $e_1,\ldots,e_t$.

必要性　因为方阵 A 相似于对角形，所以方阵 A 的最小多项式 $d(\lambda)$ 没有重根，因此 $\deg d(\lambda)=s=t$. 我们知道方阵 A^{i+j} 的特征值为 $\lambda_1^{i+j},\lambda_2^{i+j},\ldots,\lambda_t^{i+j}$，而且它们的代数重数依次为 $e_1,e_2,\ldots,e_t$，所以

$$\operatorname{Tr}A^{i+j}=e_1\lambda_1^{i+j}+e_2\lambda_2^{i+j}+\cdots+e_s\lambda_t^{i+j}.$$

于是

$$\begin{aligned}\det B&=\det\left(\operatorname{Tr}A^{i+j}\right)=\det\left(e_1\lambda_1^{i+j}+e_2\lambda_2^{i+j}+\cdots+e_s\lambda_t^{i+j}\right)\\&=\det\begin{pmatrix}e_1\lambda_1&e_2\lambda_2&\cdots&e_t\lambda_t\\e_1\lambda_1^2&e_2\lambda_2^2&\cdots&e_t\lambda_t^2\\\vdots&\vdots&\ddots&\vdots\\e_1\lambda_1^t&e_2\lambda_2^t&\cdots&e_t\lambda_t^t\end{pmatrix}\begin{pmatrix}\lambda_1&\lambda_1^2&\cdots&\lambda_1^t\\\lambda_2&\lambda_2^2&\cdots&\lambda_2^t\\\vdots&\vdots&\ddots&\vdots\\\lambda_t&\lambda_t^2&\cdots&\lambda_t^t\end{pmatrix}\\&=e_1e_2\cdots e_t\lambda_1^2\lambda_2^2\cdots\lambda_t^2\prod_{1\leqslant i<j\leqslant t}(\lambda_i-\lambda_j)^2\neq 0.\end{aligned}$$

充分性　设 $\det B\neq 0$. 如果方阵 A 不相似于对角形，则方阵 A 的最小多项式 $d(\lambda)$ 有重根，因此 $t<s$. 所以

$$\det B=\det\left(\operatorname{Tr}A^{i+j}\right)=\det\left(e_1\lambda_1^{i+j}+e_2\lambda_2^{i+j}+\cdots+e_s\lambda_t^{i+j}\right)$$

$$
\begin{aligned}
&= \det \begin{pmatrix} e_1\lambda_1 & e_2\lambda_2 & \cdots & e_t\lambda_t & 0 & \cdots & 0 \\ e_1\lambda_1^2 & e_2\lambda_2^2 & \cdots & e_t\lambda_t^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ e_1\lambda_1^t & e_2\lambda_2^t & \cdots & e_t\lambda_t^t & \underbrace{0 \quad \cdots \quad 0}_{s-t} \end{pmatrix} \begin{pmatrix} \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^t \\ \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^t \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_t & \lambda_t^2 & \cdots & \lambda_t^t \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \\
&= 0.
\end{aligned}
$$

矛盾. 这就证明,方阵 A 相似于对角形. ■

例 6.6.7 如果 n 阶复方阵 R 相似于如下的准对角方阵

$$
\operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \underbrace{1, 1, \ldots, 1}_{n-2\text{个}}\right),
$$

则方阵 R 称为**反射**. 证明,如果对合矩阵 A (即满足 $A^2 = I_{(n)}$)不是纯量方阵,则方阵 A 可以分解为有限个反射的乘积.

证明 因为方阵 A 是对合方阵,即 $A^2 = I_{(n)}$,所以 $f(\lambda) = \lambda^2 - 1$ 是方阵 A 的化零多项式. 由于方阵 A 的最小多项式 $d(\lambda)$ 整除方阵 A 的化零多项式,因此 $d(\lambda) = \lambda - 1$,或 $\lambda + 1$,或 $\lambda^2 - 1$. 如果 $d(\lambda) = \lambda - 1$,或者 $\lambda + 1$,则 $A = I_{(n)}$,或 $A = -I_{(n)}$,即方阵 A 为纯量方阵,不可能. 所以 $d(\lambda) = \lambda^2 - 1$.

于是方阵 A 的最小多项式没有重根. 因此方阵 A 相似于对角方阵,其对角元应为方阵 A 的特征值. 由于方阵 A 的特征值应是最小多项式的根,所以 A 的特征值只能是 1 或者 −1. 于是方阵 A 相似于对角形

$$
\operatorname{diag}\left(-I_{(s)},\ I_{(n-s)}\right).
$$

注意 $1 \leqslant s \leqslant n - 1$,否则方阵为纯量方阵. 这表明,存在 n 阶可逆复方阵 P,使得

$$
\begin{aligned}
A &= P^{-1} \operatorname{diag}\big(\underbrace{-1, -1, \ldots, -1}_{s\text{个}}, 1, 1, \ldots, 1\big) P \\
&= P^{-1} \operatorname{diag}(-1, 1, \ldots, 1) P \cdot P^{-1} \operatorname{diag}(1, -1, 1, \ldots, 1) P \times \cdots \\
&\quad \times P^{-1} \operatorname{diag}(1, \ldots, 1, \underset{s}{-1}, 1, \ldots, 1) P.
\end{aligned}
$$

记

$$
R_i = P^{-1} \operatorname{diag}(1, \ldots, 1, \underset{i}{-1}, 1, \ldots, 1) P,
$$

其中 $i = 1, 2, \ldots, s$. 取

$$
Q_i = \operatorname{diag}\left(1, \ldots, 1, \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \underset{i+2}{1}, 1, \ldots, 1\right),
$$

其中 $i = 1, 2, \ldots, s$. 则当 $1 \leqslant i \leqslant s$ 时

$$
\operatorname{diag}(1, \ldots, 1, \underset{i}{-1}, 1, \ldots, 1) = Q_i \operatorname{diag}\left(1, \ldots, 1, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \underset{i+2}{1}, 1, \ldots, 1\right) Q_i^{-1}.
$$

所以 R_i 是反射，$i=1,2,\ldots,s$. 于是方阵 A 是反射 $R_1,R_2,\ldots,R_s$ 的乘积. ■

例 6.6.8　设 A 是 $m\times n$ 复矩阵，B 和 C 分别是 n 阶和 m 阶复方阵，方阵 B 和 C 没有公共的特征值，而且 $AB=CA$. 证明 $A=0$.

证法 1　先考虑方阵 B 和 C 都是 Jordan 块的情形，即设

$$B=\begin{pmatrix}\lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & & & \lambda_1\end{pmatrix}_{n\times n},\quad C=\begin{pmatrix}\lambda_2 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & & & \lambda_2\end{pmatrix}_{m\times m}.$$

其中 $\lambda_1\neq\lambda_2$. 设 $A=\left(a_{ij}\right)_{m\times n}$，则由 $AB=CA$ 得到，

$$\begin{pmatrix}\lambda_1a_{11} & a_{11}+\lambda_1a_{12} & \cdots & a_{1,n-1}+\lambda_1a_{1n}\\ \lambda_1a_{21} & a_{21}+\lambda_1a_{22} & \cdots & a_{2,n-1}+\lambda_1a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \lambda_1a_{m1} & a_{11}+\lambda_1a_{m2} & \cdots & a_{m,n-1}+\lambda_1a_{mn}\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_2a_{11}+a_{21} & \lambda_2a_{12}+a_{22} & \cdots & \lambda_2a_{1n}+a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \lambda_2a_{m-2,1}+a_{m-1,1} & \lambda_2a_{m-2,2}+a_{m-1,2} & \cdots & \lambda_2a_{m-2,n}+a_{m-2,n}\\ \lambda_2a_{m1} & \lambda_2a_{m2} & \cdots & \lambda_2a_{mn}\end{pmatrix}.$$

比较上式等号两端矩阵的元素，由比较第 m 行上第一个元素得到 $\lambda_1a_{m1}=\lambda_2a_{m1}$，因为 $\lambda_1\neq\lambda_2$，故 $a_{m1}=0$；再比较第 m 上的第 2 个元素得到 $a_{m2}=0$；……；再比较第 m 行上第 n 个元素得到 $a_{mn}=0$. 然后再比较第 $m-1$ 行上的元素，得到

$$a_{m-1,1}=a_{m-1,2}=\cdots=a_{m-1,n}=0.$$

如此继续，即得 $A=0$. 现在考虑一般情形. 设

$$P^{-1}BP=J=\operatorname{diag}\bigl(J_1,J_2,\ldots,J_k\bigr),$$
$$Q^{-1}CQ=\widetilde{J}=\operatorname{diag}\bigl(\widetilde{J}_1,\widetilde{J}_2,\ldots,\widetilde{J}_\ell\bigr),$$

其中 P 和 Q 分别是 n 阶和 m 阶可逆复方阵，$J_1,J_2,\ldots,J_k$ 和 $\widetilde{J}_1,\widetilde{J}_2,\ldots,\widetilde{J}_\ell$ 都是 Jordan 块.

记 $\widetilde{A}=Q^{-1}AP$，则由 $AB=CA$ 得到，$\widetilde{A}J=\widetilde{J}\widetilde{A}$. 将方阵 $\widetilde{A}$ 分块为 $\widetilde{A}=\left(A_{ij}\right)$，使得等式 $\widetilde{A}J=\widetilde{J}\widetilde{A}$ 能按分块进行运算. 因此由 $\widetilde{A}J=\widetilde{J}\widetilde{A}$ 得到，对任意 $1\leqslant i\leqslant\ell$，$1\leqslant j\leqslant k$，

$$A_{ij}J_j=\widetilde{J}_iA_{ij}.$$

因为方阵 B 和 C 没有公共特征值，所以方阵 J_j 和 $\widetilde{J}_i$ 也没有公共特征值. 因此由上一段的证明，$A_{ij}=0$，所以 $A=0$. ■

证法 2　设

$$P^{-1}BP=J=\operatorname{diag}\bigl(J_1,J_2,\ldots,J_k\bigr),$$

其中 P 是 n 阶可逆复方阵，$J_1,J_2,\ldots,J_k$ 是 Jordan 块.

现在设 $\varphi(\lambda)$ 是方阵 C 的特征多项式,那么

$$\varphi(B) = P\varphi(J)P^{-1} = P\,\mathrm{diag}\big(\varphi(J_1), \varphi(J_2), \ldots, \varphi(J_k)\big)P^{-1}.$$

设

$$\varphi(\lambda) = (\lambda - \lambda_1)^{e_1}(\lambda - \lambda_2)^{e_2}\cdots(\lambda - \lambda_t)^{e_t},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_t$ 是方阵 C 的全部不同特征值,而

$$J_i = \mu_i I_{(n_i)} + N_{(n_i)},$$

其中 μ_i 是方阵 B 的特征值. 则

$$\varphi(J_i) = \big((\mu_i - \lambda_1)I_{(n_i)} + N_{(n_i)}\big)^{e_1}\big((\mu_i - \lambda_2)I_{(n_i)} + N_{(n_i)}\big)^{e_2}\cdots\big((\mu_i - \lambda_t)I_{(n_i)} + N_{(n_i)}\big)^{e_t}.$$

显然,每个

$$(\mu_i - \lambda_j)I_{(n_i)} + N_{(n_i)}$$

都是上三角方阵,所以 $\varphi(J_i)$ 是上三角方阵,而且对角元为

$$(\mu_i - \lambda_1)^{e_1}(\mu_i - \lambda_2)^{e_2}\cdots(\mu_i - \lambda_t)^{e_t}.$$

因为方阵 B 与 C 没有公共特征值,所以 $\mu_i \neq \lambda_j$. 因此上三角方阵 $\varphi(J_i)$ 的对角元不为零,所以 $\varphi(J_i)$ 可逆. 于是方阵 $\mathrm{diag}\big(\varphi(J_1), \varphi(J_2), \ldots, \varphi(J_k)\big)$ 可逆,从而方阵 $\varphi(B)$ 可逆. 由条件 $AB = CA$ 得到 $AB^\ell = C^\ell A$. 所以 $A\varphi(B) = \varphi(C)A$.

由于 $\varphi(\lambda)$ 是方阵 C 的特征多项式,所以由 Cayley-Hamilton 定理,$\varphi(C) = 0$,即 $A\varphi(B) = 0$. 因为方阵 $\varphi(B)$ 可逆,所以 $A = 0$. ■

例 6.6.9 求出和 n 阶复方阵 A 可交换的所有 n 阶复方阵 B.

解 设 $\lambda_1, \lambda_2, \ldots, \lambda_t$ 是方阵 A 的全部不同特征值,并且 A 的初等因子组为

$$\begin{array}{cccc}
(\lambda - \lambda_1)^{m_{11}}, & (\lambda - \lambda_1)^{m_{12}}, & \ldots, & (\lambda - \lambda_1)^{m_{1k_1}}, \\
(\lambda - \lambda_2)^{m_{21}}, & (\lambda - \lambda_2)^{m_{22}}, & \ldots, & (\lambda - \lambda_2)^{m_{2k_2}}, \\
\multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\
(\lambda - \lambda_t)^{m_{t1}}, & (\lambda - \lambda_t)^{m_{t2}}, & \ldots, & (\lambda - \lambda_t)^{m_{tk_t}},
\end{array}$$

则

$$P^{-1}AP = J = \mathrm{diag}\big(J_{11}, J_{12}, \ldots, J_{1k_1}, \ldots\ldots, J_{t1}, J_{t2}, \ldots, J_{tk_t}\big),$$

其中 P 是某个 n 阶可逆复方阵,$J_{j\ell}$是属于初等因子 $(\lambda - \lambda_j)^{m_{j\ell}}$ 的 Jordan 块.

记

$$J_j = \mathrm{diag}\big(J_{j1}, J_{j2}, \ldots, J_{jk_j}\big),$$

其中 $j = 1, 2, \ldots, t$. 则

$$P^{-1}AP = \widetilde{J} = \mathrm{diag}(J_1, J_2, \ldots, J_t).$$

记 $\widetilde{B} = P^{-1}BP$,并按 $\widetilde{J} = \mathrm{diag}(J_1, J_2, \ldots, J_t)$ 的分块方式将 $\widetilde{B}$ 分块为 $\widetilde{B} = \big(B_{ij}\big)$,其中 B_{ij} 是 $(m_{i1} + \cdots + m_{ik_i}) \times (m_{j1} + \cdots + m_{jk_j})$ 阶子矩阵.

由于 $AB = BA$,所以 $\widetilde{B}\widetilde{J} = \widetilde{J}\widetilde{B}$. 于是 $B_{ij}J_j = J_iB_{ij}$,$1 \leqslant i, j \leqslant t$.

显然当 $i \neq j$ 时,J_i 和 J_j 没有公共特征值. 由例 6.6.8,$B_{ij} = 0$. 于是

$$\widetilde{B} = \mathrm{diag}(B_{11}, B_{22}, \ldots, B_{tt}),$$

而且对于 $i = 1, 2, \ldots, t$,

$$B_{ii}J_i = J_iB_{ii}.$$

按方阵$J_i = \mathrm{diag}\big(J_{j1}, J_{j2}, \ldots, J_{jk_i}\big)$ 的分块方式将 B_{ii} 分块为 $B_{ii} = \big(B_{k\ell}^{(i)}\big)$,其中 $B_{k\ell}^{(i)}$ 是 $m_{ik} \times m_{i\ell}$ 阶子矩阵. 因为 $B_{ii}J_i = J_iB_{ii}$,所以

$$B_{k\ell}^{(i)}J_{i\ell} = J_{ik}B_{k\ell}^{(i)}, 1 \leqslant k, \ell \leqslant k_i.$$

于是问题就化为: 已知

$$J_{(p)} = \lambda_0 I_{(p)} + N_{(p)}, \quad J_{(q)} = \lambda_0 I_{(q)} + N_{(q)},$$

求 $p \times q$ 阶矩阵 $C = \big(c_{ij}\big)$,使得 $CJ_{(p)} = J_{(q)}C$.

因为

$$C(\lambda_0 I_{(q)} + N_{(q)}) = (\lambda_0 I_{(p)} + N_{(p)})C,$$

所以

$$CN_{(q)} = N_{(p)}C.$$

当 $p = q$ 时,由 $CN_{(q)} = N_{(p)}C$ 得到:

$$\begin{pmatrix} 0 & c_{11} & c_{12} & \cdots & c_{1,p-1} \\ 0 & c_{21} & c_{22} & \cdots & c_{2,p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & c_{p1} & c_{p2} & \cdots & c_{p,p-1} \end{pmatrix} = \begin{pmatrix} c_{21} & c_{22} & \cdots & c_{2p} \\ c_{31} & c_{32} & \cdots & c_{3p} \\ \vdots & \vdots & & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

比较等号两端矩阵的元素,得到

$$C = \begin{pmatrix} c_1 & c_2 & \cdots & c_p \\ 0 & c_1 & \cdots & c_{p-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & c_1 \end{pmatrix},$$

其中 $c_i = c_{1i}, i = 1, 2, \ldots, p$. 当 $p > q$ 时,由 $CN_{(q)} = N_{(p)}C$ 得到

$$C = \begin{pmatrix} \begin{pmatrix} c_1 & c_2 & \cdots & c_p \\ 0 & c_1 & \cdots & c_{p-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & c_1 \end{pmatrix} \\ \\ 0_{(p-q)\times q} \end{pmatrix};$$

当 $p < q$ 时,则

$$C = \begin{pmatrix} 0_{(p-q)\times q} & \begin{pmatrix} c_1 & c_2 & \cdots & c_p \\ 0 & c_1 & \cdots & c_{p-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & c_1 \end{pmatrix} \end{pmatrix}.$$

这样一来就可以定出和方阵 J_i 可交换的方阵 $B_i, i = 1, 2, \ldots, t$,从而定出和方阵 A 交换的方阵 B. ■

例 6.6.10 如果 n 阶复方阵 A 的特征多项式等于它的最小多项式,则 A 称为**单纯方阵**. 设方阵 B 和单纯方阵 A 可交换,则方阵 B 可表为方阵 A 的多项式.

证明 设方阵 A 的特征多项式 $\varphi(\lambda)$ 为

$$\varphi(\lambda) = (\lambda - \lambda_1)^{e_1}(\lambda - \lambda_2)^{e_2}\cdots(\lambda - \lambda_t)^{e_t},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_t$ 是方阵 A 的全部不同特征值.

因为方阵 A 是单纯的,所以 A 的最小多项式 $d(\lambda) = \varphi(\lambda)$.

由于 $d(\lambda)$ 是特征方阵 $\lambda I_{(n)} - A$ 的第 n 个不变因子 $d_n(\lambda)$,所以 $\deg d_n(\lambda) = n$,即特征方阵 $\lambda I_{(n)} - A$ 的前 $n-1$ 个不变因子都是 1. 因此方阵 A 的初等因子组为

$$(\lambda - \lambda_1)^{e_1}, (\lambda - \lambda_2)^{e_2}, \ldots, (\lambda - \lambda_t)^{e_t}.$$

于是

$$P^{-1}AP = J = \mathrm{diag}(J_1, J_2, \ldots, J_t),$$

其中 P 是某个 n 阶可逆复方阵,

$$J_j = \lambda_j I_{(e_j)} + N_{(e_j)}$$

是属于初等因子的 Jordan 块, $j = 1, 2, \ldots, t$.

因为方阵 B 和 A 可交换,所以方阵 $\widetilde{B} = P^{-1}BP$ 和 J 可交换. 由例 6.6.9,

$$\widetilde{B} = P^{-1}BP = \mathrm{diag}(B_1, B_2, \ldots, B_t),$$

其中对任意 $j = 1, 2, \ldots, t$,

$$B_j = b_0^{(j)} I_{(e_j)} + b_1^{(j)} N_{(e_j)} + b_2^{(j)} N_{(e_j)}^2 + \cdots + b_{e_j-1}^{(j)} N_{(e_j)}^{e_j-1}.$$

现在确定一个复系数多项式 $f(\lambda)$,使得 $\widetilde{B} = f(J)$. 为此,设 $e = \max\{e_1, e_2, \ldots, e_t\}$,且设 $\mu_0^{(j)}, \mu_1^{(j)}, \ldots, \mu_{e_j-1}^{(j)}$ 是待定系数, $j = 1, 2, \ldots, t$. 记

$$\begin{aligned}
h_j(\lambda) &= \mu_0^{(j)} + \frac{\mu_1^{(j)}}{1!}(\lambda - \lambda_j) + \frac{\mu_2^{(j)}}{2!}(\lambda - \lambda_j)^2 + \cdots + \frac{\mu_{e_j-1}^{(j)}}{(e_j-1)!}(\lambda - \lambda_j)^{e_j-1},\\
g_j(\lambda) &= \frac{\varphi(\lambda)}{(\lambda - \lambda_j)^{e_j}} = (\lambda - \lambda_1)^{e-1}\cdots(\lambda - \lambda_{j-1})^{e_j-1}(\lambda - \lambda_{j+1})^{e_j+1}\cdots(\lambda - \lambda_t)^{e_t},\\
f_j(\lambda) &= h_j(\lambda) g_j(\lambda),
\end{aligned}$$

其中 $j = 1, 2, \ldots, t$. 显然对于 $i = 1, 2, \ldots, j-1, j+1, \ldots, t, k = 0, 1, 2, \ldots, e_j - 1$,

$$f_j^{(k)}(\lambda_i) = 0.$$

而且 $g_j(\lambda_j) \neq 0$. 由于 $f_j(\lambda) = h_j(\lambda) g_j(\lambda)$,所以

$$\begin{aligned}
f_j^{(k)}(\lambda_j) &= h_j^{(k)}(\lambda_j) g_j(\lambda_j) + \mathrm{C}_k^1 h_j^{(k-1)}(\lambda_j) g_j'(\lambda_j) + \cdots + \mathrm{C}_k^k h_j(\lambda_j) g_j^{(k)}(\lambda_j)\\
&= \mu_k^{(j)} g_j(\lambda_j) + \mathrm{C}_k^1 \mu_{k-1}^{(j)} g_j'(\lambda_j) + \cdots + \mathrm{C}_k^k \mu_0^{(j)} g_j^{(k)}(\lambda_j).
\end{aligned}$$

令 $\dfrac{1}{k!}f_j^{(k)}(\lambda_j)=b_k^{(j)}$. 于是由

$$\begin{cases}\mu_0^{(j)}=\dfrac{b_0^{(j)}}{g_j(\lambda_j)},\\ \mu_k^{(j)}=\dfrac{k!b_k^{(j)}-\mathrm{C}_k^1\mu_{k-1}^{(j)}g_j'(\lambda_j)-\cdots-\mathrm{C}_k^k\mu_0^{(j)}g_j^{(k)}(\lambda_j)}{g_j(\lambda_j)}\end{cases}$$

即可确定待定常数 $\mu_0^{(j)},\mu_1^{(j)},\ldots,\mu_{e_j-1}^{(j)},j=1,2,\ldots,t$.

容易验证,当 $i\neq j$ 时,$g_j(J_i)=0$,因此 $f_j(J_i)=0$; 当 $i=j$ 时,

$$\begin{aligned}f_j(J_j)&=f_j(\lambda_j)I_{(e_j)}+\frac{f_j'(\lambda_j)}{1!}N_{(e_j)}+\frac{f_j''(\lambda_j)}{2!}N_{(e_j)}^2+\cdots+\frac{f_j^{(e_j-1)}(\lambda_j)}{(e_j-1)!}N_{(e_j)}^{e_j-1}\\&=b_0^{(j)}I_{(e_j)}+b_1^{(j)}N_{(e_j)}+b_2^{(j)}N_{(e_j)}^2+\cdots+b_{e_j-1}^{(j)}N_{(e_j)}^{e_j-1}\\&=B_j.\end{aligned}$$

记 $f(\lambda)=f_1(\lambda)+f_2(\lambda)+\cdots+f_t(\lambda)$,则 $f(J_j)=B_j$. 所以

$$f(J)=\mathrm{diag}\big(f(J_1),f(J_2),\ldots,f(J_t)\big)=\mathrm{diag}\big(B_1,B_2,\ldots,B_t\big)=\widetilde{B}.$$

因此

$$f(A)=Pf(J)P^{-1}=P\widetilde{B}P^{-1}=B.$$ ■

注 例 6.6.10 中构造多项式 $f(\lambda)$ 的方法具有一般意义,应予重视. 例 6.6.10 的结论对任意非单纯的复方阵 A 不再成立. 但可以证明,若方阵 C 和每一个与方阵 A 可交换的方阵 B 都可交换,则 C 可以表为 A 的多项式. 请读者自证之.

例 6.6.11 设 A 是 n 阶可逆复方阵,则存在可逆方阵 B,使得 $B^2=A$.

证法 1 先考虑最简单情形,即设方阵 A 是 Jordan 块,也即设

$$A=\lambda_0I_{(n)}+N_{(n)}.$$

取 n 阶方阵 $D=\sqrt{\lambda_0}I_{(n)}+N_{(n)}$. 则 $D^2=\lambda_0I_{(n)}+2\sqrt{\lambda_0}N_{(n)}+N_{(n)}^2$. 方阵 D^2 的特征多项式为

$$\varphi(\lambda)=\det\big(\lambda I_{(n)}-D^2\big)=(\lambda-\lambda_0)^n.$$

因为 D^2 的最小多形式 $d(\lambda)$ 整除特征多项式,所以 $d(\lambda)=(\lambda-\lambda_0)^\ell,1\leqslant\ell\leqslant n$.

如果 $\ell<n$,则

$$\begin{aligned}d(D^2)&=\big(D^2-\lambda_0I_{(n)}\big)^\ell=\big(D-\sqrt{\lambda_0}I_{(n)}\big)^\ell\big(D+\sqrt{\lambda_0}I_{(n)}\big)^\ell\\&=N_{(n)}^\ell\big(2\sqrt{\lambda_0}I_{(n)}+N_{(n)}\big)^\ell=0.\end{aligned}$$

因为方阵 A 可逆,所以 A 的特征值 $\lambda_0\neq0$,因此方阵 $2\sqrt{\lambda_0}I_{(n)}+N_{(n)}$ 可逆. 于是由上式 $N_{(n)}^\ell=0$. 但是,当 $\ell<n$ 时 $N_{(n)}^\ell\neq0$. 这就导出矛盾. 所以 $\ell=n$.

由于 $\deg d(\lambda)=n$,所以方阵 D^2 的前 $n-1$ 个不变因子全为 1. 因此方阵 D^2 的初等因子组为 $(\lambda-\lambda_0)^n$. 显然方阵 A 的初等因子也为 $(\lambda-\lambda_0)^n$. 所以方阵 A 与 D^2 相似,即存在 n 阶可逆方阵 P,使得

$$A=P^{-1}D^2P=(P^{-1}DP)^2.$$

取 $B=P^{-1}DP$,即得所欲证的结论.

现在转到一般情形. 设

$$A=P^{-1}\operatorname{diag}\left(J_1,J_2,\ldots,J_t\right)P,$$

其中 P 是某个可逆方阵,$J_1,J_2,\ldots,J_t$ 都是可逆的 Jordan 块.

由上一段的证明,存在可逆方阵 B_j,使得 $J_j=B_j^2, j=1,2,\ldots,t$. 于是

$$A=P^{-1}\operatorname{diag}\left(B_1^2,B_2^2,\ldots,B_t^2\right)P=\left(P^{-1}\operatorname{diag}(B_1,B_2,\ldots,B_t)P\right)^2,$$

其中 $P^{-1}\operatorname{diag}(B_1,B_2,\ldots,B_t)P$ 可逆. ■

证法 2 先考虑方阵 A 是 Jordan 块的情形,也即设

$$A=\lambda_0 I_{(n)}+N_{(n)},\lambda_0\neq 0.$$

把方阵看成复数,于是问题转化为,已知复数 x,求复数 y,使得

$$\lambda_0+x=\lambda_0\left(1+\frac{x}{\lambda_0}\right)=y^2.$$

令 $y=\sqrt{\lambda_0}\,z$,则

$$z=\left(1+\frac{x}{\lambda_0}\right)^{\frac{1}{2}}=1+\sum_{k=1}^{\infty}\frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-k+1\right)}{k!}\lambda_0^{-k}x^k.$$

再把复数视为方阵,即用 $N_{(n)}$ 代换 x. 因为 $N_{(n)}^n=0$,所以构造方阵 C 如下:

$$C=I_{(n)}+\sum_{k=1}^{n-1}\frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-k+1\right)}{k!}\lambda_0^{-k}N_{(n)}^k.$$

显然方阵 C 可逆. 可以验证,$C^2=I_{(n)}+\dfrac{1}{\lambda_0}N_{(n)}$. 令 $B=\sqrt{\lambda_0}C$,则方阵 B 可逆,且

$$B^2=\lambda_0 I_{(n)}+N_{(n)}=A.$$

因此结论对一个只由 Jordan 块构成的方阵 A 成立. 然后再按证法 1 的一般情形证明之. ■

注 例 6.6.11 可以推广为: 设 p 是正整数,A 是 n 阶可逆方阵. 则存在 n 阶可逆复方阵 B,使得 $A=B^p$. 当应当指出,关于方阵 A 可逆的条件不能省略. 请读者自己给出例子.

习 题 6.6

1. 设 $i_1i_2\cdots i_n$ 是自然数 $1,2,\ldots,n$ 的一个排列. 把 n 阶单位方阵 $I_{(n)}$ 的第 $1,2,\ldots,n$ 行依次调到第 $i_1,i_2,\ldots,i_n$ 行,得到的方阵称为**置换方阵**. 证明,置换方阵相似于对角形.

2. 证明,所有 n 阶循环方阵

$$\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1}\\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix}$$

可以经同一个可逆方阵 P 化为标准形.

3. 设 A 和 B 是 n 阶方阵,且方阵 $\mathrm{diag}(A, A)$ 和 $\mathrm{diag}(B, B)$ 相似. 证明,方阵 A 和 B 相似.

4. 已知 5 阶方阵 A 的特征多项式 $\varphi(\lambda)$ 和 $d(\lambda)$ 分别为

$$\varphi(\lambda) = (\lambda-2)^3(\lambda+7)^2, \quad d(\lambda) = (\lambda-2)^2(\lambda+7).$$

求方阵 A 的 Jordan 标准形.

5. 设 A 是 n 阶可逆矩阵,且 A 和 A^k 相似,k 是正整数. 证明,方阵 A 的特征值都是单位根.

6. 设方阵 A 和任意一个可逆方阵都可交换. 证明,A 是纯量方阵.

7. 设方阵 C 和每一个与方阵 A 可交换的多项式都可交换. 证明,方阵 C 可以表示为方阵 A 的多项式.

8. 设 n 阶方阵 A 不可逆. 证明,$\mathrm{rank}\,A = \mathrm{rank}\,A^2$ 的充分必要条件是,方阵 A 的属于特征值 0 的初等因子都是一次的.

9. (Weyr) 证明,n阶复方阵 A 与 B 相似当且仅当对于每个复数 a 和每个正整数 k,

$$\mathrm{rank}\big(aI_{(n)}-A\big)^k = \mathrm{rank}\big(aI_{(n)}-B\big)^k.$$

§6.7　实方阵的实相似

一般地说,对于数域 $\mathbb{F}$ 上 n 阶方阵 A 和 B,如果存在数域 $\mathbb{F}$ 上 n 阶可逆方阵 P,使得 $B = P^{-1}AP$,则称方阵 A 和 B 在**数域 $\mathbb{F}$ 上相似**. 特别,如果 n 阶实方阵 A 和 B 在实数域 $\mathbb{R}$ 上相似,则称方阵 A 和 B **实相似**. 本节将讨论实方阵在实相似下的标准形问题.

显然,如果实方阵 A 和 B 实相似,则方阵 A 和 B 相似. 反之,如果实方阵 A 和 B 相似,方阵 A 和 B 是否实相似? 对此,有

定理 6.7.1　n 阶实方阵 A 和 B 相似的充分必要条件是,方阵 A 和 B 相似.

证明　仅证充分性. 设方阵 A 和 B 相似,则存在 n 阶可逆复方阵 P,使得 $B = P^{-1}AP$. 将方阵 P 分为实部和虚部,即记 $P = R + \mathrm{i}Q$,其中 $\mathrm{i}^2 = -1$,R 和 Q 都是 n 阶实方阵. 如果 Q 是零方阵,结论已然成立. 因此可设实方阵 $Q \neq 0$. 于是由 $B = P^{-1}AP$ 得到,

$$(R+\mathrm{i}Q)B = A(R+\mathrm{i}Q),$$

比较两端方阵的实部和虚部,得到

$$RB = AR, \quad QB = AQ.$$

于是对任意实数 λ,$A(R+\lambda Q) = (R+\lambda Q)B$. 记

$$f(\lambda) = \det(R+\lambda Q).$$

显然 $f(\lambda)$ 是关于 λ 的实系数多项式,当然也是关于 λ 的复系数多项式. 因为

$$f(\mathrm{i}) = \det(R+\mathrm{i}Q) = P \neq 0,$$

所以 $f(\lambda)$ 是非零多项式. 因此 $f(\lambda)$ 至多有有限个实根. 于是存在实数 λ_0,使得

$$f(\lambda_0) = \det(R+\lambda_0 Q) \neq 0.$$

这表明实方阵 $R+\lambda_0 Q$ 可逆,而且

$$B=(R+\lambda_0 Q)^{-1}A(R+\lambda_0 Q).$$

也即方阵 A 和 B 实相似. ■

定理 6.7.1 表明,把实方阵看成复方阵,则实方阵在相似下的全系不变量也就是实方阵在实相似下的全系不变量. 因此,实方阵在实相似下的标准形理论的两个基本问题中尚待解决的是:寻求实方阵的实相似的等价类的代表元.

当然,实相似等价类的代表元应当是实方阵. 所以我们不能用实方阵 A 在相似下的 Jordan 标准形 J 直接作为实方阵在实相似下的标准形. 尽管如此,我们还是从实方阵 A 在相似下的标准形 J 出发,构造一个和方阵 J 相似的实方阵 L. 由于实方阵 A 和 J 相似,且方阵 J 和 L 相似,所以由定理 6.7.1,实方阵 A 和 L 实相似. 于是即可求得实方阵 A 在实相似的标准形. 下面将给出具体构造方法.

容易看出,实方阵 A 的行列式因子是实系数多项式,所以它的不变因子也是实系数多项式. 但是,实系数多项式的复根是共轭成对出现的,所以如果虚部不为零的复数 τ 是实方阵 A 的某个不变因子的 k 重根,则它的共轭复数 $\overline{\tau}$ 也是实方阵 A 的这个不变因子的 k 重根. 也就是说,如果 $(\lambda-\tau)^k$ 是实方阵 A 的一个初等因子,则 $(\lambda-\overline{\tau})^k$ 也是 A 的初等因子. 所以可设实方阵 A 的初等因子组为

$$\begin{gathered}\lambda^{e_1},\lambda^{e_2},\dots,\lambda^{e_s},(\lambda-\lambda_1)^{f_1},(\lambda-\lambda_2)^{f_2},\dots,(\lambda-\lambda_t)^{f_t},\\(\lambda-\tau_1)^{g_1},(\lambda-\tau_2)^{g_2},\dots,(\lambda-\tau_k)^{g_k},(\lambda-\overline{\tau}_1)^{g_1},(\lambda-\overline{\tau}_2)^{g_2},\dots,(\lambda-\overline{\tau}_k)^{g_k}.\end{gathered}$$

其中 $\lambda_1,\lambda_2,\dots,\lambda_t$ 是非零实数,$\tau_1,\tau_2,\dots,\tau_k$ 是虚部不为零的复数.

对于实方阵 A 的初等因子 λ^{e_j},属于 λ^{e_j} 的 Jordan 块数 $N_{(e_j)}$,它是实方阵.

对于实方阵 A 的形如 $(\lambda-\lambda_0)^f$ 的初等因子,其中 λ_0 是非零实数,属于 $(\lambda-\lambda_0)^f$ 的 Jordan 块为 $\lambda_0 I_{(t)}+N_{(t)}$. 由于

$$\operatorname{diag}\left(1,\lambda_0^{-1},\dots,\lambda_0^{-(f-1)}\right)(\lambda_0 I_{(f)}+N_{(f)})\operatorname{diag}\left(1,\lambda_0,\dots,\lambda_0^{f-1}\right)=\lambda_0(I_{(f)}+N_{(f)}),$$

所以 $\lambda_0 I_{(f)}+N_{(f)}$ 和 $\lambda_0 M_{(f)}=\lambda_0(I_{(f)}+N_{(f)})$ 相似.

对于实方阵 A 的形如 $(\lambda-\tau)^g$ 和 $(\lambda-\overline{\tau}_0)^g$ 的初等因子,其中 τ 的虚部不为零,属于 $(\lambda-\tau)^g$ 和 $(\lambda-\overline{\tau}_0)^g$ 的 Jordan 块分别为 $\tau I_{(g)}+N_{(g)}$ 和 $\overline{\tau}I_{(g)}+N_{(g)}$.

由上一段可知,

$$\operatorname{diag}\left(\tau I_{(g)}+N_{(g)},\ \overline{\tau}I_{(g)}+N_{(g)}\right)$$

相似于 $\operatorname{diag}\left(\tau M_{(g)},\ \overline{\tau}M_{(g)}\right)$.

记 $\tau=|\tau|(\cos\theta+\mathrm{i}\sin\theta)$,其中 $|\tau|$ 是复数 τ 的模,$\mathrm{i}^2=-1$,且记

$$L_{(g)}=|\tau|\begin{pmatrix}\cos\theta M_{(g)} & \sin\theta M_{(g)}\\ -\sin\theta M_{(g)} & \cos\theta M_{(g)}\end{pmatrix}.$$

则

$$\begin{pmatrix}I_{(g)} & -\mathrm{i}I_{(g)}\\ I_{(g)} & I_{(g)}\end{pmatrix}\cdot L_{(g)}\cdot\left(\frac{1}{2}\begin{pmatrix}I_{(g)} & I_{(g)}\\ \mathrm{i}I_{(g)} & -\mathrm{i}I_{(g)}\end{pmatrix}\right)=\operatorname{diag}\left(\tau M_{(g)},\ \overline{\tau}M_{(g)}\right).$$

故 $\mathrm{diag}\big(\tau M_{(g)},\ \overline{\tau}M_{(g)}\big)$ 相似于 $L_{(g)}$，所以$\mathrm{diag}\big(\tau I_{(g)}+N_{(g)},\ \overline{\tau}I_{(g)}+N_{(g)}\big)$ 和 $L_{(g)}$ 相似.

定理 6.7.2 设 n 阶实方阵 A 的初等因子组为

$$\begin{gathered}\lambda^{e_1},\lambda^{e_2},\ldots,\lambda^{e_s},(\lambda-\lambda_1)^{f_1},(\lambda-\lambda_2)^{f_2},\ldots,(\lambda-\lambda_t)^{f_t},\\(\lambda-\tau_1)^{g_1},(\lambda-\overline{\tau}_1)^{g_1},\ldots,(\lambda-\tau_k)^{g_k},(\lambda-\overline{\tau}_k)^{g_k},\end{gathered}\tag{6.7.1}$$

其中 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 是非零实数，$\tau_1,\tau_2,\ldots,\tau_k$ 是虚部不为零的复数. 则方阵 A 实相似于如下的标准形：

$$\mathrm{diag}\big(N_{e_1},\ldots,N_{(e_s)},\lambda_1M_{(f_1)},\ldots,\lambda_tM_{(f_t)},L_{(g_1)},\ldots,L_{(g_k)}\big).\tag{6.7.2}$$

其中 $M_{(f_j)}=I_{(f_j)}+N_{(f_j)}$，$j=1,2,\ldots,t$；而 $\tau_i=|\tau_i|(\cos\theta_i+\mathrm{i}\sin\theta_i)$，$i=1,2,\ldots,k$，

$$L_{(g_i)}=|\tau_i|\begin{pmatrix}\cos\theta_iM_{(g_i)} & \sin\theta_iM_{(g_i)}\\-\sin\theta_iM_{(g_i)} & \cos\theta_iM_{(g_i)}\end{pmatrix}.$$

证明 根据定理 6.5.3，实方阵 A 相似于如下的 Jordan 标准形 J：

$$J=\mathrm{diag}\big(N_{e_1},\ldots,N_{e_s},J_{11},\ldots,J_{1t},J_{21},\overline{J}_{21},\ldots,J_{2k},\overline{J}_{2k}\big),$$

其中 J_{1j} 是属于 $(\lambda-\lambda_j)^{f_j}$ 的 Jordan 块，$j=1,2,\ldots,t$；J_{2j} 和 $\overline{J}_{2j}$ 分别是属于 $(\lambda-\tau_j)^{g_j}$ 和 $(\lambda-\overline{\tau}_j)^{g_j}$ 的 Jordan 块，$j=1,2,\ldots,k$.

由上一段的讨论，J_{1j} 相似于 $\lambda_jM_{(f_j)}$，$\mathrm{diag}\big(J_{2j},\ \overline{J}_{2j}\big)$ 相似于 $L_{(g_j)}$.

显然，如果准对角方阵 $\mathrm{diag}(A_1,\ A_2)$ 和 $\mathrm{diag}(B_1,\ B_2)$ 的对角块 A_j 和 B_j 相似 ($j=1,2$)，则 $\mathrm{diag}(A_1,\ A_2)$ 和 $\mathrm{diag}(B_1,\ B_2)$ 相似. 因此，方阵 J 相似于方阵 (6.7.2). 由于方阵 A 和 J 相似，所以方阵 A 和 (6.7.2) 相似. 由定理 6.7.1，方阵 A 和方阵 (6.7.2) 实相似. ■

例 6.7.1 设 A 是 $2n$ 阶实方阵，且 $A^2+I_{(2n)}=0$. 证明，存在 $2n$ 阶实方阵 P，使得

$$P^{-1}AP=\begin{pmatrix}0 & I_{(n)}\\-I_{(n)} & 0\end{pmatrix}.$$

证明 显然 $f(\lambda)=\lambda^2+1$ 是实方阵 A 的化零多项式. 因此方阵 A 的最小多项式 $d(\lambda)$ 只能是 $\lambda+\mathrm{i}$, $\lambda-\mathrm{i}$ 或 $\lambda^2=1$，其中 $\mathrm{i}^2=-1$. 由于 A 是实方阵，所以 $d(\lambda)$ 是实系数多项式. 因此 $d(\lambda)=\lambda^2+1$. 因为方阵的最小多项式的根式方阵的特征值，所以方阵 A 的特征值是 i 和 $-\mathrm{i}$，而且成对出现. 因此方阵具有 n 个特征值 i 和 n 个特征值 $-\mathrm{i}$. 因为方阵 A 的最小多项式没有重根，所以存在 $2n$ 阶可逆复方阵 Q，使得

$$Q^{-1}AQ=\begin{pmatrix}\mathrm{i}I_{(n)} & 0\\0 & -\mathrm{i}I_{(n)}\end{pmatrix}.$$

但是

$$\frac{1}{2}\begin{pmatrix}I_{(n)} & I_{(n)}\\\mathrm{i}I_{(n)} & -\mathrm{i}I_{(n)}\end{pmatrix}Q^{-1}AQ\begin{pmatrix}I_{(n)} & -\mathrm{i}I_{(n)}\\I_{(n)} & \mathrm{i}I_{(n)}\end{pmatrix}=\begin{pmatrix}0 & I_{(n)}\\-I_{(n)} & 0\end{pmatrix}.$$

这表明实方阵 A 和实方阵

$$\begin{pmatrix}0 & I_{(n)}\\-I_{(n)} & 0\end{pmatrix}$$

相似. 再根据定理 6.7.1 可知结论成立. ■

习 题 6.7

1. 证明,方阵

$$A=\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{pmatrix}$$

相似于对角形,但不实相似于对角形.

2. 设 $2k$ 维实线性空间 V 的线性变换 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 为 $d(\lambda)=(\lambda^2+a\lambda+b)^k$,其中 a,b 为实数,且 $a^2-4b<0$. 证明,存在 V 的基,使得线性变换 $\mathscr{A}$ 在这组基下的方阵为

$$\begin{pmatrix}0&1&&&&&&&\\-b&-a&1&&&&&&\\&0&0&1&&&&&\\&&-b&-a&1&&&&\\&&&\ddots&\ddots&\ddots&&&\\&&&&0&0&1&&\\&&&&&-b&-a&1&\\&&&&&&0&0&1\\&&&&&&&-b&-a\end{pmatrix}.$$

3. 证明,任意一个实方阵 A 都相似于准对角形,其对角块具有如下形式:

$$\begin{pmatrix}\lambda_0&1&&&\\&\lambda_0&1&&\\&&\ddots&\ddots&\\&&&\lambda_0&1\\&&&&\lambda_0\end{pmatrix},$$

或者

$$\begin{pmatrix}0&1&1&&&&&&\\-b&-a&0&1&&&&&\\&0&0&1&1&&&&\\&&-b&-a&0&1&&&\\&&&\ddots&\ddots&\ddots&\ddots&&\\&&&&0&0&1&1&\\&&&&&-b&-a&0&1\\&&&&&&0&0&1\\&&&&&&&-b&-a\end{pmatrix},$$

其中 λ_0, a 和 b 都是实数,且 $a^2<4b$.

4. 设 A 是数域 $\mathbb{F}$ 上 n 阶方阵. 视 A 为复方阵,它的特征多项式和最小多项式分别记为 $\varphi(\lambda)$ 和 $d(\lambda)$. 证明,作为数域 $\mathbb{F}$ 上 n 阶方阵,A 的特征多项式和最小多项式仍分别为 $\varphi(\lambda)$ 和 $d(\lambda)$.

5. 设 A 是数域 $\mathbb{F}$ 上 n 阶方阵. 视 A 为复方阵,它的行列式因子和不变因子分别记为 $D_1(\lambda),\ldots,D_n(\lambda)$ 和 $d_1(\lambda),\ldots,d_n(\lambda)$. 证明,作为数域 $\mathbb{F}$ 上 n 阶方阵,A 的行列式因子和不变因子仍分别是 $D_1(\lambda),\ldots,D_n(\lambda)$ 和 $d_1(\lambda),\ldots,d_n(\lambda)$.

6. 设 A 和 B 是数域 $\mathbb{F}$ 上 n 阶方阵. 证明,方阵 A 和 B 在数域 $\mathbb{F}$ 上相似的充分必要条件是,方阵 A 和 B 相似.

第七章

Euclid 空间

- 前面几章讨论的线性空间实际上只有代数结构. 为了使线性空间更像通常的三维空间,有必要在线性空间中增加几何结构.
- 我们首先限制在实线性空间. §7.1 定义了实线性空间中向量之间的内积,这实质上便在实线性空间中引进了距离的概念,从而赋予实线性空间一种几何结构. 这样的空间即是 Euclid 空间.
- 由于 Euclid 空间比线性空间具有更强的结果,基的选择应当与这种结构相适应,所以在 §7.2 中通过正交性的讨论引进了标准正交基,这就是 Euclid 空间所应当考虑的基,也是通常三维空间的笛卡尔坐标系的直接推广.
- 于是,Euclid 空间的每个线性变换在不同的标准正交基下的矩阵表示是彼此正交相似的方阵. 这就自然产生了方阵在正交相似下的分类.
- §7.3 与 §7.4 利用内积引进了规范变换与规范方阵的概念. 并解决了规范方阵在正交相似下的分类问题,得出了规范方阵的正交相似等价类完全由方阵的特征值所刻画的结论.
- 将这一结果应用于更为特殊的一些具有明显几何意义的变换以及相应的方阵上,§7.5 与 §7.6 解决了正交方阵、对称方阵以及斜对称方阵在正交相似下的分类问题.
- 由于正定对称方阵的特殊重要性以及应用的广泛性,§7.7 专门讨论了正定对称方阵,并通过矩阵的奇异值的概念,解决了矩阵在正交相似下的分类问题.
- §7.8 用纯矩阵的方法讨论了一般方阵的正交相似问题. 随后,§7.9 通过例子说明了这一理论的各种应用.

§7.1 内 积

在空间解析几何里,我们知道,在空间 V 中建立直角坐标系后,向量 α 便唯一地确定一个坐标 (x_1, x_2, x_3). 在空间 V 中向量的长度和向量间的夹角可以用向量的纯量积表示. 设向量 α 和 β 的坐标分别是 (x_1, x_2, x_3) 与 (y_1, y_2, y_3),则向量 α 与 β 的纯量积为

$$(\alpha, \beta) = x_1y_1 + x_2y_2 + x_3y_3.$$

于是向量 α 的长度为

$$\|\alpha\| = \sqrt{(\alpha, \alpha)},$$

向量 α 与 β 的夹角 θ 的余弦为

$$\cos\theta = \frac{(\alpha, \beta)}{\sqrt{(\alpha, \alpha)(\beta, \beta)}}.$$

为了在 n 维实线性空间中引进向量的长度与向量间的夹角等概念. 我们先分析一下上述向量的纯量积 (α,β) 的特性.

首先,它是定义在空间 V 上的一个二元实函数,即对任意一对向量 α,β,可以确定唯一一个实数 $x_1y_1+x_2y_2+x_3y_3$ 与向量对 α,β 相对应.

其次,向量的纯量积 (α,β) 具有对称性,即对任意 $\alpha,\beta\in V$,均有 $(\alpha,\beta)=(\beta,\alpha)$;而且对空间 V 中任意非零向量 α,均有 $(\alpha,\alpha)>0$,纯量积 (α,β) 的这一性质称为恒正性.

最后,纯量积 (α,β) 满足: 对任意 $\alpha_1,\alpha_2\in V$,均有 $(\alpha_1+\alpha_2,\beta)=(\alpha_1,\beta)+(\alpha_2,\beta)$,而且对任意 $\lambda\in\mathbb{R},\alpha\in V$,均有 $(\lambda\alpha,\beta)=\lambda(\alpha,\beta)$,其中 $\mathbb{R}$ 表示实数域. 也就是说,如果将纯量积 (α,β) 中的向量 β 视为不变,则纯量积 (α,β) 是关于向量元 α 的线性函数. 简单地说,纯量积 (α,β) 关于向量元 α 是线性的. 同样,纯量积 (α,β) 关于向量元 β 也是线性的. 即是说,纯量积 (α,β) 是双线性的.

现在将向量的纯量积 (α,β) 推广到实线性空间.

定义 7.1.1　设实线性空间 V 上二元实函数 (α,β) 满足下面三个性质:

(1) 对称性　对任意 $\alpha,\beta\in V,(\beta,\alpha)=(\alpha,\beta)$;

(2) 恒正性　对任意 $\alpha\in V,\alpha\neq 0,(\alpha,\alpha)>0$;

(3) 双线性性　对任意 $\alpha_1,\alpha_2,\alpha,\beta_1,\beta_2,\beta\in V,\lambda,\mu\in\mathbb{R}$,均有

$$(\alpha_1+\alpha_2,\beta)=(\alpha_1,\beta)+(\alpha_2,\beta),\qquad (\lambda\alpha,\beta)=\lambda(\alpha,\beta),$$
$$(\alpha,\beta_1+\beta_2)=(\alpha,\beta_1)+(\alpha,\beta_2),\qquad (\alpha,\mu\beta)=\mu(\alpha,\beta),$$

则二元实函数 (α,β) 称为实线性空间 V 的一个**内积**.

应当指出,在定义 7.1.1 中,内积 (α,β) 关于向量元 β 的线性性质可以由内积 (α,β) 对称性以及关于向量元 α 的线性性质导出. 这里为清楚起见,还是将它列入定义中. 另外,对实线性空间 V 而言,内积 (α,β) 并不是唯一的. 只要满足对称性、恒正性以及双线性性的二元实函数都是实线性空间 V 的一个内积.

现在给出内积的基本性质.

命题 7.1.1　设 (α,β) 是实线性空间 V 的内积,则 $(0,\beta)=(\alpha,0)=0$.

证明　由双线性性得到,$(0,\beta)=(0\cdot\alpha,\beta)=0(\alpha,\beta)=0$.

由对称性得到,$(\alpha,0)=(0,\alpha)=0$. ■

命题 7.1.2　设 (α,β) 是实线性空间 V 的内积,向量 $\alpha_1,\alpha_2,\ldots,\alpha_p,\beta_1,\beta_2,\ldots,\beta_q\in V$,并且 $\lambda_1,\lambda_2,\ldots,\lambda_p,\mu_1,\mu_2,\ldots,\mu_q\in\mathbb{R}$,则

$$\left(\sum_{i=1}^{p}\lambda_i\alpha_i,\ \sum_{j=1}^{q}\mu_j\beta_j\right)=\sum_{i=1}^{p}\sum_{j=1}^{q}\lambda_i\mu_j(\alpha_i,\beta_j).$$

证明　首先用归纳法证明,

$$\left(\sum_{i=1}^{p}\lambda_i\alpha_i,\ \beta\right)=\sum_{i=1}^{p}\lambda_i(\alpha_i,\beta).\tag{7.1.1}$$

当 $p=1$ 时式 (7.1.1) 左端为 $(\lambda_1\alpha_1,\beta)$,因此由内积 (α,β) 关于向量元 α 的线性性质,$(\lambda_1\alpha_1,\beta)=\lambda_1(\alpha_1,\beta)$,即式 (7.1.1) 当 $p=1$ 时成立. 现在设式 (7.1.1) 对 $p-1$ 成立. 于是由内积 (α,β) 关于向量元 α 的线性性质得到,

$$\begin{aligned}\left(\sum_{i=1}^{p}\lambda_i\alpha_i,\ \beta\right)&=\left(\sum_{i=1}^{p-1}\lambda_i\alpha_i+\lambda_p\alpha_p,\ \beta\right)=\left(\sum_{i=1}^{p-1}\lambda_i\alpha_i,\ \beta\right)+(\lambda_p\alpha_p,\beta)\\&=\left(\sum_{i=1}^{p-1}\lambda_i\alpha_i,\ \beta\right)+\lambda_p(\alpha_p,\beta)\end{aligned}$$

由归纳假设,

$$=\sum_{i=1}^{p-1}\lambda_i(\alpha_i,\beta)+\lambda_p(\alpha_p,\beta)=\sum_{i=1}^{p}\lambda_i(\alpha_i,\beta),$$

所以式 (7.1.1) 成立.

由内积 (α,β) 的对称性得

$$\left(\alpha,\ \sum_{j=1}^{q}\mu_j\beta_j\right)=\left(\sum_{j=1}^{q}\mu_j\beta_j,\ \alpha\right)=\sum_{j=1}^{q}\mu_j(\beta_j,\alpha)=\sum_{j=1}^{q}\mu_j(\alpha,\beta_j).$$

于是

$$\left(\sum_{i=1}^{p}\lambda_i\alpha_i,\ \beta\right)=\sum_{i=1}^{p}\lambda_i\left(\alpha_i,\ \sum_{j=1}^{q}\mu_j\beta_j\right)=\sum_{i=1}^{p}\sum_{j=1}^{q}\lambda_i\mu_j(\alpha_i,\beta_j).$$ ■

命题 7.1.3(Cauchy-Schwarz 不等式) 设 (α,β) 是实线性空间 V 的一个内积,则对任意 $\alpha,\beta\in V$,均有

$$(\alpha,\beta)^2\leqslant(\alpha,\alpha)(\beta,\beta),\tag{7.1.2}$$

等式当且仅当向量 α 与 β 线性相关时成立.

证明 当 $\alpha=0$ 时,由命题 7.1.1,$(0,\beta)=0$,$(0,0)=0$. 因此式 (7.1.2) 成等式,故当 $\alpha=0$ 时式 (7.1.2) 成立.

现在设 $\alpha\neq0$. 设 $t\in\mathbb{R}$,并记

$$f(t)=(t\alpha+\beta,t\alpha+\beta),$$

显然 $f(t)\geqslant0$. 由于

$$f(t)=t^2(\alpha,\alpha)+2t(\alpha,\beta)+(\beta,\beta),$$

所以 $f(t)$ 是关于 t 的二次多项式. 因此二次多项式 $f(t)$ 的判别式

$$(\alpha,\beta)^2-(\alpha,\alpha)(\beta,\beta)\leqslant0,$$

其中等式当且仅当方程 $f(t)=0$ 具有唯一解时成立. 即当 $\alpha\neq0$ 时,式 (7.1.2) 也成立.

设向量 α 与 β 线性相关,则存在不全为零的 $\lambda,\mu\in\mathbb{R}$,使得 $\lambda\alpha+\mu\beta=0$. 因为 $\alpha\neq0$,故 $\mu\neq0$. 于是 $\dfrac{\lambda}{\mu}\alpha+\beta=0$. 所以

$$f\left(\frac{\lambda}{\mu}\right)=\left(\frac{\lambda}{\mu}\alpha+\beta,\ \frac{\lambda}{\mu}\alpha+\beta\right)=0.$$

即 $\dfrac{\lambda}{\mu}$ 是方程 $f(t)=0$ 的解. 容易验证 $\dfrac{\lambda}{\mu}$ 是 $f(t)=0$ 的唯一解.

反之,设 t_0 是 $f(t)=0$ 的唯一解,则

$$f(t_0)=(t_0\alpha+\beta,t_0\alpha+\beta)=0.$$

由内积 (α,β) 的恒正性得到,$t_0\alpha+\beta=0$. 于是向量 α 与 β 线性相关. 这就证明,式 (7.1.2) 中等式当且仅当向量 α 与 β 线性相关时成立. ■

下面给出 n 维实线性空间 V 的内积的方阵表示. 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维实线性空间 V 的一组基,(α,β) 是 V 的一个内积,其中 $\alpha,\beta\in V$. 则

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n,\quad \beta_1=y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n.$$

因此

$$(\alpha,\beta)=\left(\sum_{i=1}^n x_i\alpha_i,\ \sum_{j=1}^n y_j\alpha_j\right)=\sum_{i=1}^n\sum_{j=1}^n x_iy_j(\alpha_i,\alpha_j).$$

记 $x=(x_1,x_2,\ldots,x_n)$,$y=(y_1,y_2,\ldots,y_n)$,且记 n 阶方阵

$$G=\begin{pmatrix}(\alpha_1,\alpha_1)&(\alpha_1,\alpha_2)&\cdots&(\alpha_1,\alpha_n)\\(\alpha_2,\alpha_1)&(\alpha_2,\alpha_2)&\cdots&(\alpha_2,\alpha_n)\\\vdots&\vdots&\ddots&\vdots\\(\alpha_n,\alpha_1)&(\alpha_n,\alpha_2)&\cdots&(\alpha_n,\alpha_n)\end{pmatrix},$$

则

$$(\alpha,\beta)=xGy^{\mathrm{T}}. \tag{7.1.3}$$

方阵 $G=\big((\alpha_i,\alpha_j)\big)_{n\times n}$ 称为内积 (α,β) 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的 **Gram 方阵.**

Gram 方阵 $G=\big((\alpha_i,\alpha_j)\big)$ 具有以下的性质.

(1) 方阵 G 是对称方阵,即 $G^{\mathrm{T}}=G$.

事实上,方阵 G 的 (i,j) 位置上的元素为 (α_i,α_j). 由内积 (α,β) 的对称性,

$$(\alpha_i,\alpha_j)=(\alpha_j,\alpha_i),$$

因此方阵 G 的 (i,j) 位置上的元素等于 (j,i) 位置上的元素,所以 $G^{\mathrm{T}}=G$.

(2) 正定性,即对任意 $x\in\mathbb{R}^n$,$xGx^{\mathrm{T}}\geqslant 0$,其中等式当且仅当 $x=0$ 时成立.

事实上对任意 $x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$,$\alpha=x_1\alpha_1+\cdots+x_n\alpha_n\in V$,由式 (7.1.3) 和内积的恒正性,$(\alpha,\alpha)=xGx^{\mathrm{T}}\geqslant 0$,而且当且仅当 $\alpha=0$,即 $x=0$ 时等式成立.

通常,满足 $S^{\mathrm{T}}=S$ 的 n 阶方阵 S 称为**对称方阵**. 设 S 是 n 阶对称方阵,如果对任意 $x\in\mathbb{R}^n$,$xSx^{\mathrm{T}}\geqslant 0$,且等式当且仅当 $x=0$ 时成立,则 S 称为**正定对称方阵**.

综上所得

定理 7.1.1 设 (α,β) 是 n 维实线性空间 V 的内积,$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 V 的一组基,向量 α 与 β 在这组基下的坐标分别为 x 与 y,而且内积 (α,β) 在这组基下的 Gram 方阵为 G,则 G 是正定对称方阵,而且 $(\alpha,\beta)=xGy^{\mathrm{T}}$.

定理 7.1.2 设 S 是 n 阶正定对称方阵,$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维实线性空间 V 的基,向量 α 与 β 在这组基下的坐标分别为 x 与 y,则由 $(\alpha,\beta)=xSy^{\mathrm{T}}$ 所定义的二元实函数 (α,β) 是 V 的一个内积,而且在这组基下的 Gram 方阵即为 S.

证明 对于任意 $\alpha,\beta\in V$,

$$(\beta,\alpha)=ySx^{\mathrm{T}}=(xSy^{\mathrm{T}})^{\mathrm{T}}=xSy^{\mathrm{T}}=(\alpha,\beta),$$

即二元实函数 (α,β) 是对称的.

其次对任意 $\alpha\in V$, $(\alpha,\alpha)=xSx^{\mathrm{T}}$. 因为对称方阵 S 是正定的,所以 $(\alpha,\alpha)\geqslant 0$. 并且当且仅当 $x=0$,即 $\alpha=0$ 时, $(\alpha,\alpha)=0$. 因此 (α,β) 是恒正的.

最后设 $\alpha,\widetilde{\alpha}\in V$ 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标分别为 x 与 $\widetilde{x}$, $\lambda,\mu\mathbb{R}$,则

$$\begin{aligned}(\lambda\alpha+\mu\widetilde{\alpha},\beta)&=(\lambda x+\mu\widetilde{x})Sy^{\mathrm{T}}=\lambda(xSy^{\mathrm{T}})+\mu(\widetilde{x}Sy^{\mathrm{T}})\\&=\lambda(\alpha,\beta)+\mu(\widetilde{\alpha},\beta).\end{aligned}$$

因此二元实函数 (α,β) 关于向量元 α 是线性的.

由于二元实函数 (α,β) 是对称的,因此 (α,β) 关于向量元 β 是线性的,即二元实函数 (α,β) 是双线性的. 这就证明, (α,β) 是 V 的一个内积.

记 $S=(s_{ij})_{n\times n}$. 基向量 α_i 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为

$$\varepsilon_i=(0,\ldots,0,\underset{i}{1},0,\ldots,0).$$

因此对 $1\leqslant i,j\leqslant n$,

$$(\alpha_i,\alpha_j)=\varepsilon_iS\varepsilon_j^{\mathrm{T}}=s_{ij}.$$

于是 $S=(s_{ij})=\big((\alpha_i,\alpha_j)\big)$,即方阵 S 是内积 (α,β) 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的 Gram 方阵. ■

定理 7.1.1 表明,在 n 维实线性空间 V 中取定一组基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$, V 中的一个内积 (α,β) 便由 (7.1.3) 确定一个 Gram 方阵 G,而 G 是正定对称方阵. 于是可以建立由 V 的所有内积的集合到所有 n 阶正定对称方阵的集合的映射 σ:

$$\sigma\big((\alpha,\beta)\big)=G.$$

容易验证,映射 σ 是单射. 定理 定理 7.1.2 表明,映射 σ 是满射. 因此 V 的所有内积的集合便和所有 n 阶正定对称方阵集合存在一个一一对应.

现在转到 n 维实线性空间 V 的一个内积在 V 的不同基下的方阵表示的关系. 为此先引进

定义 7.1.2 设 S_1 和 S_2 是 n 阶实对称方阵. 如果存在 n 阶可逆实方阵 P,使得

$$S_2=P^{\mathrm{T}}S_1P,$$

则称对称方阵 S_1 与 S_2 是**相合**的.

定理 7.1.3 设 n 维实线性空间 V 的内积 (α,β) 在 V 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 下的 Gram 方阵分别为 G_1 与 G_2,而且

$$(\beta_1,\beta_2,\ldots,\beta_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)P,$$

其中 P 是 n 阶可逆方阵. 则 $G_2=P^{\mathrm{T}}G_1P$,即方阵 G_1 与 G_2 是相合的.

证明 记 $P=(p_{ij})_{n\times n}$,则对于 $j=1,2,\ldots,n$,

$$\beta_j = \sum_{k=1}^{n} p_{kj}\alpha_j.$$

记 $G_1 = (a_{ij})_{n\times n}$, $G_2 = (b_{ij})_{n\times n}$, 其中 $a_{ij} = (\alpha_i, \alpha_j)$, $b_{ij} = (\beta_i, \beta_j)$. 因此,

$$\begin{aligned} b_{ij} = (\beta_i, \beta_j) &= \left(\sum_{k=1}^{n} p_{ki}\alpha_k, \sum_{\ell=1}^{n} p_{\ell j}\alpha_\ell\right) \\ &= \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_{ki}p_{\ell j}(\alpha_k, \alpha_\ell) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_{ki}p_{\ell j}a_{k\ell} \\ &= (p_{1i}, p_{2i}, \ldots, p_{ni})G_1(p_{1j}, p_{2j}, \ldots, p_{nj})^{\mathrm{T}}. \end{aligned}$$

这表明, b_{ij} 是方阵 $P^{\mathrm{T}}G_1P$ 的 (i,j) 位置上的元素, 所以 $G_2 = P^{\mathrm{T}}G_1P$. ■

实对称方阵之间的相合关系式一种重要的关系. 以后将详加讨论.

下面给出 Euclid 空间的定义.

定义 7.1.3　实线性空间 V 连同一个取定的内积 (α, β) 一起成为 **Euclid 空间**.

应当指出, 对 Euclid 空间 V_1 与 V_2, 如果作为线性空间, V_1 与 V_2 是不同的空间, 则 Euclid 空间 V_1 与 V_2 是不同的. 甚至作为线性空间, V_1 与 V_2 是同一个空间, 但确定的内积不同时, 则 Euclid 空间 V_1 与 V_2 也是不同的. 因此 Euclid 空间的定义是和实线性空间 V 以及 V 的内积 (α, β) 的选取密切联系的. 正因为如此, Euclid 空间也称为**内积空间**. 对于定义 Euclid 空间 V 的内积 (α, β) 前面给出的命题 7.1.1, 命题 7.1.2 与命题 7.1.3 当然成立.

在 Euclid 空间 V 中, 利用内积 (α, β) 的恒正性, 可以定义向量 $\alpha \in V$ 的**范数**(或**长度**) $\|\alpha\|$ 为

$$\|\alpha\| = \sqrt{(\alpha, \alpha)}.$$

范数为 1 的向量称为**单位向量**. 对非零向量 $\alpha \in V$, 令

$$\xi = \frac{\alpha}{\|\alpha\|},$$

则 ξ 是单位向量.

利用向量的范数, 命题 7.1.3 中的 Cauchy-Schwarz 不等式 (7.1.2) 可以改写成:

$$\left|\frac{(\alpha, \beta)}{\|\alpha\| \cdot \|\beta\|}\right| \leqslant 1,$$

其中 α 和 β 是非零向量. 于是, 对于非零向量 α 与 β, 可以定义它们的夹角 θ 为

$$\theta = \arccos \frac{(\alpha, \beta)}{\sqrt{(\alpha, \alpha)(\beta, \beta)}}.$$

如果向量 α 与 β 的夹角 $\theta = \dfrac{\pi}{2}$, 也即 $(\alpha, \beta) = 0$, 则称向量 α 和 β 是**正交的**, 记作

$$\alpha \perp \beta.$$

当 $\alpha = 0$ 时, 由于 $(0, \beta) = 0$, 故也称零向量和 β 正交.

关于向量的范数, 有

命题 7.1.4　对任意 $\alpha, \beta \in V$,

$$\|\alpha+\beta\| \leqslant \|\alpha\|+\|\beta\|. \tag{7.1.4}$$

证明 由范数的定义,

$$\begin{aligned}\|\alpha+\beta\|^2 &= (\alpha+\beta,\alpha+\beta) = (\alpha,\alpha)+2(\alpha,\beta)+(\beta,\beta)\\ &= \|\alpha\|^2+2\|\alpha\|\|\beta\|\cos\theta+\|\beta\|^2.\end{aligned}$$

根据 Cauchy-Schwarz 不等式,

$$\|\alpha+\beta\|^2 \leqslant \|\alpha\|^2+2\|\alpha\|\|\beta\|+\|\beta\|^2 = (\|\alpha\|+\|\beta\|)^2.$$ ∎

式 (7.1.4) 称为**向量范数的三角形不等式.**

最后给出 Euclid 空间的一些例子.

例 7.1.1 设 $\mathbb{R}^n$ 是所有有序 n 元实数组 $(x_1,x_2,\dots,x_n)$ 构成的实线性空间. 设 $\alpha=(x_1,x_2,\dots,x_n)$, $\beta=(y_1,y_2,\dots,y_n)\in\mathbb{R}^n$. 定义 $\mathbb{R}^n$ 上二元实函数 (α,β) 为

$$(\alpha,\beta)=x_1y_1+x_2y_2+\cdots+x_ny_n.$$

容易验证,二元实函数 (α,β) 满足: 对称性,恒正性与双线性性. 因此 (α,β) 是实线性空间 $\mathbb{R}^n$ 的一个内积,它称为 $\mathbb{R}^n$ 的**标准内积**. 实线性空间 $\mathbb{R}^n$ 连同标准内积 (α,β) 一起构成一个 Euclid 空间.

例 7.1.2 设 $\mathbb{R}^2$ 是所有 2 维实向量 (x_1,x_2) 构成的 2 维实向量空间. 设 $\alpha=(x_1,x_2)$, $\beta=(y_1,y_2)\in\mathbb{R}^2$. 定义 2 维实向量空间 $\mathbb{R}^2$ 上的二元实函数 (α,β) 为

$$(\alpha,\beta)=x_1y_1-x_2y_1-x_1y_2+4x_2y_2.$$

容易验证,二元实函数 (α,β) 满足: 对称性、恒正性和双线性性. 所以二元实函数 (α,β) 是 2 维实向量空间 $\mathbb{R}^2$ 的一个内积. 2 维实向量空间 $\mathbb{R}^2$ 连同内积 (α,β) 一起便构成一个 Euclid 空间.

例 7.1.3 设 $\mathbb{R}^{n\times n}$ 是所有 n 阶实方阵构成的实线性空间. 定义实线性空间 $\mathbb{R}^{n\times n}$ 上的二元实函数 (A,B) 如下: 设 $A,B\in\mathbb{R}^{n\times n}$,则

$$(A,B)=\operatorname{Tr}AB^{\mathrm{T}}.$$

证明实线性空间 $\mathbb{R}^{n\times n}$ 连同 (A,B) 一起构成一个 Euclid 空间.

证明 只需证明,二元实函数 (A,B) 是 $\mathbb{R}^{n\times n}$ 的一个内积. 为此只需验证,二元实函数 (A,B) 满足对称性,恒正性和双线性性.

设 $A,B\in\mathbb{R}^{n\times n}$,则

$$(B,A)=\operatorname{Tr}BA^{\mathrm{T}}=\operatorname{Tr}(AB^{\mathrm{T}})^{\mathrm{T}}=\operatorname{Tr}AB^{\mathrm{T}}=(A,B).$$

所以二元实函数 (A,B) 满足对称性. 设 $A=(a_{ij})\in\mathbb{R}^{n\times n}$, $A\neq 0$,则

$$(A,A)=\operatorname{Tr}AA^{\mathrm{T}}=\sum_{1\leqslant i,j\leqslant n}a_{ij}^2>0.$$

因此二元实函数 (A,B) 是恒正的. 设 $A_1,A_2\in\mathbb{R}^{n\times n}$, $\lambda\in\mathbb{R}$,则

$$(A_1+A_2,B)=\operatorname{Tr}(A_1+A_2)B^{\mathrm{T}}=\operatorname{Tr}(A_1B^{\mathrm{T}}+A_2B^{\mathrm{T}})$$

$$\begin{aligned} &= \operatorname{Tr} A_1 B^{\mathrm{T}} + \operatorname{Tr} A_2 B^{\mathrm{T}} = (A_1, B) + (A_2, B), \\ (\lambda A, B) &= \operatorname{Tr}(\lambda A) B^{\mathrm{T}} = \operatorname{Tr} \lambda (A B^{\mathrm{T}}) = \lambda \operatorname{Tr} A B^{\mathrm{T}} = \lambda (A, B). \end{aligned}$$

因此二元实函数 (A,B) 关于向量元 A 是线性的.

由对称性,二元实函数 (A,B) 关于向量元 B 也是线性的. 因此二元实函数 (A,B) 是双线性的. 所以 (A,B) 是实线性空间 $\mathbb{R}^{n\times n}$ 的一个内积. ■

例 7.1.4　设 $\mathbb{R}^n$ 是所有 n 维实向量 $(x_1,x_2,\ldots,x_n)$ 构成的实线性空间,P 是 n 阶可逆实方阵. 定义实线性空间 $\mathbb{R}^n$ 上二元实函数 (α,β) 如下:设 $\alpha,\beta\in\mathbb{R}^n$,则

$$(\alpha,\beta) = \alpha P P^{\mathrm{T}} \beta^{\mathrm{T}}.$$

验证二元实函数 (α,β) 是实线性空间 $\mathbb{R}^n$ 的一个内积.

证明　设 $\alpha,\beta\in\mathbb{R}^n$,则

$$(\beta,\alpha) = \beta P P^{\mathrm{T}} \alpha^{\mathrm{T}} = (\alpha P P^{\mathrm{T}} \beta^{\mathrm{T}})^{\mathrm{T}} = \alpha P P^{\mathrm{T}} \beta^{\mathrm{T}} = (\alpha,\beta),$$

即二元实函数 (α,β) 是实对称的.

设 $\alpha\in\mathbb{R}^n$,$\alpha\neq 0$,则方阵 P 可逆,故 $\alpha P\neq 0$. 记 $\alpha P=(y_1,y_2,\ldots,y_n)$,于是

$$(\alpha,\alpha) = \alpha P P^{\mathrm{T}} \alpha^{\mathrm{T}} = (\alpha P)(\alpha P)^{\mathrm{T}} = y_1^2 + y_2^2 + \cdots + y_n^2 > 0.$$

因此二元实函数 (α,β) 是恒正的.

设 $\alpha_1,\alpha_2\in\mathbb{R}^n$,$\lambda\in\mathbb{R}$,则

$$\begin{aligned} (\alpha_1+\alpha_2,\beta) &= (\alpha_1+\alpha_2) P P^{\mathrm{T}} \beta^{\mathrm{T}} = \alpha_1 P P^{\mathrm{T}} \beta^{\mathrm{T}} + \alpha_2 P P^{\mathrm{T}} \beta^{\mathrm{T}} = (\alpha_1,\beta) + (\alpha_2,\beta), \\ (\lambda\alpha,\beta) &= (\lambda\alpha) P P^{\mathrm{T}} \beta^{\mathrm{T}} = \lambda(\alpha P P^{\mathrm{T}} \beta^{\mathrm{T}}) = \lambda(\alpha,\beta). \end{aligned}$$

因此二元实函数 (α,β) 关于向量元 α 是线性的.

由对称性,二元实函数 (α,β) 关于向量元 β 也是线性的. 所以 (α,β) 是双线性的. 因此 (α,β) 是实线性空间 $\mathbb{R}^n$ 的一个内积. ■

注　在例 7.1.4 中取 n 阶可逆实方阵 P 为 n 阶单位方阵 $I_{(n)}$,取得实线性空间 $\mathbb{R}^n$ 的标准内积.

例 7.1.5　设 L_2 是区间 $[0,1]$ 上所有连续实函数 $f(x)$ 构成的实线性空间. 定义实线性空间 L_2 上二元实函数 $\big(f(x),g(x)\big)$ 为:设 $f(x),g(x)\in L_2$,则

$$\big(f(x),g(x)\big) = \int_0^1 f(x)g(x)\,\mathrm{d}x.$$

则二元实函数 $\big(f(x)g(x)\big)$ 是线性空间 L_2 的一个内积. 实线性空间 L_2 连同内积 $\big(f(x),g(x)\big)$ 一起构成一个 Euclid 空间.

例 7.1.6　取实线性空间 L_2 同例 7.1.5. 定义实线性空间 L_2 的变换 $\mathscr{A}$ 如下:设 $f(x)\in L_2$,则令

$$\big(\mathscr{A}(f)\big)(x) = x f(x).$$

容易验证,$\mathscr{A}$ 是 L_2 的线性变换. 定义实线性空间 L_2 上二元实函数 $P_A\big(f(x),g(x)\big)$ 如下:设 $f(x),g(x)\in L_2$,则

$$P_A\big(f(x),g(x)\big)=\int_0^1\big(\mathscr{A}(f)\big)(x)\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x.$$

验证 $P_A\big(f(x),g(x)\big)$ 是实线性空间 L_2 的内积.

证明 设 $f(x),g(x)\in L_2$,则

$$\begin{aligned}P_A\big(g(x),f(x)\big)&=\int_0^1\big(\mathscr{A}(g)\big)(x)\big(\mathscr{A}(f)\big)(x)\,\mathrm{d}x=\int_0^1\big(\mathscr{A}(f)\big)(x)\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x\\&=P_A\big(f(x),g(x)\big),\end{aligned}$$

因此二元实函数 $P_A\big(f(x),g(x)\big)$ 是对称的.

设 $f(x)\in L_2$,且 $\mathscr{A}(f)=0$,即实函数 $f(x)$ 在变换 $\mathscr{A}$ 下的象为零函数. 因此对任意 $x\in[0,1]$,

$$\big(\mathscr{A}(f)\big)(x)=xf(x)=0.$$

因为 $x\not\equiv 0$,所以 $f(x)\equiv 0$,即 $f(x)$ 是零函数. 因此 $\mathrm{Ker}\,\mathscr{A}=0$. 这表明,线性变换 $\mathscr{A}$ 是可逆的. 现在设 $f(x)\in L_2$, $f(x)\not\equiv 0$,则 $\big(\mathscr{A}(f)\big)(x)\not\equiv 0$. 因此

$$P_A\big(f(x),f(x)\big)=\int_0^1\big(\mathscr{A}(f)(x)\big)^2\,\mathrm{d}x>0.$$

所以二元实函数 $P_A\big(f(x),g(x)\big)$ 是恒正的.

设 $f_1(x),f_2(x)\in L_2$, $\lambda\in\mathbb{R}$,则

$$\begin{aligned}\big(\mathscr{A}(f_1+f_2)\big)(x)&=x(f_1+f_2)(x)=xf_1(x)+xf_2(x)\\&=\big(\mathscr{A}(f_1)\big)(x)+\big(\mathscr{A}(f_2)\big)(x),\\ \big(\mathscr{A}(\lambda f)\big)(x)&=x(\lambda f)(x)=\lambda\big(xf(x)\big)=\big(\lambda\mathscr{A}(f)\big)(x).\end{aligned}$$

因此

$$\begin{aligned}P_A\big(f_1(x)+f_2(x),g(x)\big)&=\int_0^1\big(\mathscr{A}(f_1+f_2)\big)(x)\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x\\&=\int_0^1\Big(\big(\mathscr{A}(f_1)\big)(x)+\big(\mathscr{A}(f_2)\big)(x)\Big)\cdot\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x\\&=\int_0^1\big(\mathscr{A}(f_1)\big)(x)\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x\\&\quad+\int_0^1\big(\mathscr{A}(f_2)\big)(x)\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x\\&=P_A\big(f_1(x),g(x)\big)+P_A\big(f_2(x),g(x)\big),\\ P_A\big(\lambda f(x),g(x)\big)&=\int_0^1\big(\mathscr{A}(\lambda f)\big)(x)\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x\\&=\int_0^1\lambda\big(\mathscr{A}(f)\big)(x)\big(\mathscr{A}(g)\big)(x)\,\mathrm{d}x\\&=\lambda P_A\big(f(x),g(x)\big).\end{aligned}$$

所以二元实函数 $P_A\big(f(x),g(x)\big)$ 关于向量元 $f(x)$ 是线性的.

根据对称性,二元实函数 $P_A\big(f(x),g(x)\big)$ 关于向量元 $g(x)$ 也是线性的. 从而 $P_A\big(f(x),g(x)\big)$ 是双线性的. 这就证明,二元实函数 $P_A\big(f(x),g(x)\big)$ 是实线性空间 L_2 的一个内积. ∎

习 题 7.1

1. 设 $\mathbb{R}^2$ 是所有 2 维实行向量的集合连同标准内积构成的 2 维 Euclid 空间, $A = (a_{ij})$ 是 2 阶实对称方阵. 对于 $x, y \in \mathbb{R}^2$, 定义

$$f_A(x, y) = xAy^{\mathrm{T}}.$$

证明, 二元实函数 $f_A(x, y)$ 是内积的充分必要条件为 $a_{11} > 0$, $a_{22} > 0$, 且 $\det A > 0$.

2. $\mathbb{R}^2$ 的意义同习题 1. 设 $x = (x_1, x_2)$, $y = (y_1, y_2)$, $\mathbb{R}^2$ 的标准内积即为

$$(x, y) = x_1y_1 + x_2y_2.$$

定义 $\mathbb{R}^2$ 的线性变换 $\mathscr{A}$ 如下: 对于 $x = (x_1, x_2) \in \mathbb{R}^2$, 令 $\mathscr{A}(x) = (-x_2, x_1)$. 构造 $\mathbb{R}^2$ 的一个内积 $[x, y]$, 使得对任意 $x \in \mathbb{R}^2$, 均有 $[x, \mathscr{A}(x)] = 0$.

3. 所有收敛的实数序列 $x = (x_1, x_2, \ldots)$ 的集合记为 V. 定义 V 中序列 $x = (x_1, x_2, \ldots)$ 与 $y = (y_1, y_2, \ldots)$ 的加法为

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots);$$

定义纯量 $\lambda \in \mathbb{R}$ 与序列 $x = (x_1, x_2, \ldots) \in V$ 的乘积为

$$\lambda x = (\lambda x_1, \lambda x_2, \ldots).$$

于是 V 在如此的加法和纯量与序列的乘法下成为实线性空间. 证明, V 上二元实函数

$$(x, y) = \sum_{k=1}^{\infty} x_k y_k$$

是 V 的一个内积, 其中 $x = (x_1, x_2, \ldots)$, $y = (y_1, y_2, \ldots) \in V$.

4. 设 V_1 和 V_2 是 Euclid 空间. 记 $V_1 \times V_2 = \{(\alpha, \beta) \mid \alpha \in V_1, \beta \in V_2\}$. 在 $V_1 \times V_2$ 中定义加法: 对任意 $(\alpha_1, \beta_1), (\alpha_2, \beta_2) \in V_1 \times V_2$, 令

$$(\alpha_1, \beta_1) + (\alpha_2, \beta_2) = (\alpha_1 + \alpha_2, \beta_1 + \beta_2);$$

定义纯量 $\lambda \in \mathbb{R}$ 与向量 $(\alpha, \beta) \in V_1 \times V_2$ 的乘积为

$$\lambda(\alpha, \beta) = (\lambda\alpha, \lambda\beta).$$

于是 $V_1 \times V_2$ 在如此的加法和纯量与向量的乘法下成为实线性空间. 定义 $V_1 \times V_2$ 上二元实函数为

$$[(\alpha_1, \beta_1), (\alpha_2, \beta_2)] = [\alpha_1, \alpha_2] + [\beta_1, \beta_2],$$

其中 $\alpha_1, \alpha_2 \in V_1$, $\beta_1, \beta_2 \in V_2$, 且 $[\alpha_1, \alpha_2]$ 与 $[\beta_1, \beta_2]$ 分别是 V_1 与 V_2 的内积.

证明, 二元实函数 $[(\alpha_1, \beta_1), (\alpha_2, \beta_2)]$ 是 $V_1 \times V_2$ 的一个内积.

5. 证明:

(1) n 维 Euclid 空间 V 中向量 α 与 β 正交的充分必要条件是

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2;$$

(2) 设 $\alpha, \beta \in V$, 且 $\|\alpha\| = \|\beta\|$, 则向量 $\alpha + \beta$ 与 $\alpha - \beta$ 正交;

(3) (平行四边形法则) 设 $\alpha, \beta \in V$, 则

$$\|\alpha + \beta\|^2 + \|\alpha - \beta\|^2 = 2\|\alpha\|^2 + 2\|\beta\|^2.$$

6. 证明, n 维 Euclid 空间 V 中向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 线性无关的充分必要条件是, 它们的 Gram 方阵 $G(\alpha_1, \alpha_2, \ldots, \alpha_n) = \big((\alpha_i, \alpha_j)\big)_{n\times n}$ 可逆, 其中 (α, β) 是 V 的内积.

§7.2 正 交 性

先讨论 Euclid 空间 V 中向量之间的正交性.

定理 7.2.1 Euclid 空间中 k 个两两正交的非零向量 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 线性无关.

证明 设 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$, 且 $\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_k\alpha_k = 0$. 则由命题 7.1.1 和命题 7.1.2,

$$\left(\sum_{i=1}^{k} \lambda_i\alpha_i,\ \alpha_j\right) = \sum_{i=1}^{k} \lambda_i(\alpha_i, \alpha_j) = 0.$$

因为 $\alpha_i \perp \alpha_j, i \neq j$, 所以 $(\alpha_i, \alpha_j) = 0, i \neq j$, 因此上式可化为 $\lambda_j(\alpha_j, \alpha_j) = 0$. 因为向量 α_j 非零, 所以 $(\alpha_j, \alpha_j) > 0$, 因此 $\lambda_j = 0$. 这就证明, 向量 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 线性无关. ■

定理 7.2.1 的逆命题并不成立, 即 Euclid 空间 V 中 k 个线性无关的向量并不一定两两正交. 另外, 定理 7.2.1 表明, n 维 Euclid 空间 V 中两两正交的非零向量组中所含向量的个数不超过 n. 问题是: n 维 Euclid 空间 V 中是否存在 n 个两两正交的非零向量构成的向量组? 对此有

定理 7.2.2 设 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 n 维 Euclid 空间 V 的一组基, 则存在两两正交的非零向量 $\beta_1, \beta_2, \ldots, \beta_n \in V$, 使得对每个正整数 k, $\{\beta_1, \beta_2, \ldots, \beta_k\}$ 是 V 中向量 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 生成的子空间 V_k 的一组基.

证明 首先, 取 $\beta_1 = \alpha_1$. 首先 $\{\beta_1\}$ 是子空间 V_1 的一组基, 因此结论对 $k = 1$ 成立. 现在假设存在两两正交的非零向量组 $\beta_1, \beta_2, \ldots, \beta_{k-1}$, 使得 $\{\beta_1, \beta_2, \ldots, \beta_{k-1}\}$ 是 V_{k-1} 的一组基. 下面寻求向量 β_k. 由于 $\beta_k \in V_k$, 但 $\beta_k \notin V_{k-1}$, 因此可设

$$\beta_k = \alpha_k + \lambda_{k-1}\beta_{k-1} + \cdots + \lambda_1\beta_1,$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ 是待定常数.

由于向量 β_k 应和 $\beta_1, \beta_2, \ldots, \beta_{k-1}$ 正交, 因此令 $(\beta_k, \beta_j) = 0$. 所以

$$(\beta_k, \beta_j) = \left(\alpha_k + \sum_{i=1}^{k-1} \lambda_i\beta_i,\ \beta_j\right) = (\alpha_k, \beta_j) + \sum_{i=1}^{k-1} \lambda_i(\beta_i, \beta_j) = 0.$$

由归纳假设, 向量 $\beta_1, \beta_2, \ldots, \beta_{k-1}$ 两两正交, 因此上式化为,

$$\lambda_j = -\frac{(\alpha_k, \beta_j)}{(\beta_j, \beta_j)}.$$

于是取

$$\beta_k = \alpha_k - \frac{(\alpha_k, \beta_{k-1})}{(\beta_{k-1}, \beta_{k-1})}\beta_{k-1} - \frac{(\alpha_k, \beta_{k-2})}{(\beta_{k-2}, \beta_{k-2})}\beta_{k-2} - \cdots - \frac{(\alpha_k, \beta_1)}{(\beta_1, \beta_1)}\beta_1,$$

则 $\beta_1, \beta_2, \ldots, \beta_k$ 是两两正交的, 而且 $\{\beta_1, \beta_2, \ldots, \beta_k\}$ 是子空间 V_k 的一组基. ■

定理 7.2.2 证明中由 n 维 Euclid 空间 V 的一组基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 得到两两正交的向量组 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 的过程成为 **Gram-Schmidt 正交化**. 当然, $\{\beta_1, \beta_2, \ldots,$

$\beta_n\}$ 是 Euclid 空间 V 的一组基，它称为 V 的**正交基**.

定义 7.2.1 设 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 是 Euclid 空间 V 的正交基. 如果每个向量 β_j 都是单位向量，则 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 称为 V 的一组**标准正交基**.

由定理 7.2.2 立即得到下面的定理.

定理 7.2.3 n 维 Euclid 空间 V 具有标准正交基.

证明 因为 V 是 n 维实线性空间，所以 V 具有一组基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$. 由定理 7.2.2，存在 V 的一组正交基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$. 记

$$\xi_j=\frac{\beta_j}{\|\beta_j\|},$$

则 $\|\xi_j\|=1, j=1,2,\ldots,n$. 于是 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的一组标准正交基. ■

注 对 n 维 Euclid 空间 V 的一组基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$，可以用 Gram-Schmidt 正交化过程得到 V 的正交基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$，然后再将每个向量 β_j 单位化，即令 $\xi_j=\dfrac{\beta_j}{\|\beta_j\|}$，便得到 V 的标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$. 由正交基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 得到标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 的过程称为对基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$ 施行**单位化**.

定理 7.2.4 n 维 Euclid 空间 V 中任意一组两两正交的单位向量组 $\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$ 都可以扩充成 V 的一组标准正交基 $\{\alpha_1,\alpha_2,\ldots,\alpha_m,\alpha_{m+1},\ldots,\alpha_n\}$.

证明 由定理 7.2.1，$m\leqslant n$. 现在对 $n-m=k$ 用归纳法.

显然当 $k=0$ 时，$m=n$，则 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 本身即为 V 的一组标准正交基. 所以结论对 $k=0$ 成立.

设结论对 $n-m=k$ 成立. 下面证明结论对 $n-m=k+1$ 成立.

此时 $m<n$. 因此 $\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$ 不是 V 的基. 记 V 中由向量 $\alpha_1,\alpha_2,\ldots,\alpha_m$ 生成的子空间记为 V_m，则存在 $\beta_0\in V$，但 $\beta_0\notin V_m$. 令

$$\xi_{m+1}=\beta_0+\lambda_m\alpha_m+\cdots+\lambda_1\alpha_1,$$

其中 $\lambda_1,\lambda_2,\ldots,\lambda_m$ 是待定常数. 设 $(\xi_{m+1},\alpha_j)=0$，则

$$(\beta_0+\lambda_m\alpha_m+\cdots+\lambda_1\alpha_1,\alpha_j)=(\beta_0,\alpha_j)+\lambda_j(\alpha_j,\alpha_j)=(\beta_0,\alpha_j)+\lambda_j=0.$$

因此 $\lambda_j=-(\beta_0,\alpha_j)$. 即取

$$\xi_{m+1}=\beta_0-(\beta_0,\alpha_m)\alpha_m-\cdots-(\beta_0,\alpha_1)\alpha_1,$$

并令

$$\alpha_{m+1}=\frac{\xi_{m+1}}{\|\xi_{m+1}\|},$$

则 $\alpha_1,\alpha_2,\ldots,\alpha_{m+1}$ 是 V 中一组两两正交的单位向量.

由于 $n-(m+1)=(n-m)-1=k$，所以由归纳假设，$\{\alpha_1,\alpha_2,\ldots,\alpha_m,\alpha_{m+1}\}$ 可以扩充成 V 的标准正交基

$$\{\alpha_1,\alpha_2,\ldots,\alpha_m,\alpha_{m+1},\ldots,\alpha_n\},$$

因此结论对 $n-m=k+1$ 成立. ■

下面的定理给出 n 维 Euclid 空间 V 中两组标准正交基的联系.

定理 7.2.5 设 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 与 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 n 维 Euclid 空间 V 的两组标准正交基,则由基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 到基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 的过渡方阵 P 是正交方阵,即方阵 P 满足

$$PP^{\mathrm{T}}=I_{(n)}=P^{\mathrm{T}}P.$$

证明 由假设,

$$(\eta_1,\eta_2,\ldots,\eta_n)=(\xi_1,\xi_2,\ldots,\xi_n)P.$$

记 $P=(p_{ij})_{n\times n}$,则,

$$\eta_j=\sum_{i=1}^n p_{ij}\xi_i.$$

由于 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 和 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 都是标准正交基,因此,

$$(\eta_i,\eta_j)=\delta_{ij}=(\xi_i,\xi_j).$$

所以

$$\begin{aligned}(\eta_i,\eta_j)&=\left(\sum_{k=1}^n p_{ki}\xi_k,\ \sum_{\ell=1}^n p_{\ell j}\xi_\ell\right)=\sum_{k=1}^n\sum_{\ell=1}^n p_{ki}p_{\ell j}(\xi_k,\xi_\ell)\\&=\sum_{k=1}^n\sum_{\ell=1}^n p_{ki}p_{\ell j}\delta_{k\ell}=\sum_{k=1}^n p_{ki}p_{kj}=\delta_{ij}.\end{aligned}$$

写成矩阵形式,即得 $P^{\mathrm{T}}P=I_{(n)}$. 因此 $P^{-1}=P^{\mathrm{T}}$. 所以 $PP^{\mathrm{T}}=PP^{-1}=I_{(n)}$. ■

定理 7.2.6 设 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 n 维 Euclid 空间 V 的一组标准正交基,P 是 n 阶正交方阵,则由

$$(\eta_1,\eta_2,\ldots,\eta_n)=(\xi_1,\xi_2,\ldots,\xi_n)P \tag{7.2.1}$$

确定的向量组 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 V 的标准正交基.

证明 记 $P=(p_{ij})_{n\times n}$. 因为 P 是正交方阵,所以 $P^{\mathrm{T}}P=I_{(n)}$. 因此

$$\sum_{k=1}^p p_{ki}p_{kj}=\delta_{ij}.$$

由于 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的标准正交基,所以 $(\xi_k,\xi_\ell)=\delta_{k\ell}$. 由式 (7.2.1) 得到,

$$\eta_j=\sum_{k=1}^n p_{kj}\xi_k.$$

因此对 $1\leqslant i,j\leqslant n$,

$$\begin{aligned}(\eta_i,\eta_j)&=\left(\sum_{k=1}^n p_{ki}\xi_k,\ \sum_{\ell=1}^n p_{\ell j}\xi_\ell\right)=\sum_{k=1}^n\sum_{\ell=1}^n p_{ki}p_{\ell j}(\xi_k,\xi_\ell)\\&=\sum_{k=1}^n\sum_{\ell=1}^n p_{ki}p_{\ell j}\delta_{k\ell}=\sum_{k=1}^n p_{ki}p_{kj}=\delta_{ij}.\end{aligned}$$

这就证明,$\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 V 的一组标准正交基. ■

定理 7.2.5 和定理 7.2.6 提到的正交方阵是一类正要的方阵. 在讨论 Euclid 空间时经常要遇到它. 容易看出,如果 O 是 n 阶正交方阵,则 $O^{\mathrm{T}}O=I_{(n)}=OO^{\mathrm{T}}$,因此

它的逆方阵 O^{-1} 是方阵 O 的转置. 而且由

$$(O^{\mathrm{T}})^{\mathrm{T}}O^{\mathrm{T}} = OO^{\mathrm{T}} = I_{(n)} = O^{\mathrm{T}}O = O^{\mathrm{T}}(O^{\mathrm{T}})^{\mathrm{T}}$$

可知,正交方阵 O 的转置 O^{T} 也是正交方阵,也即正交方阵 O 的逆方阵仍是正交方阵. 另外,如果 O_1 和 O_2 是正交方阵,则乘积 O_1O_2 仍是正交方阵.

对于 n 阶正交方阵 O,将方阵 O 按行分块为 $O = (\xi_1, \xi_2, \ldots, \xi_n)^{\mathrm{T}}$. 则 $\xi_1, \xi_2, \ldots, \xi_n$ 是 n 维实的行向量. 由于 $OO^{\mathrm{T}} = I_{(n)}$,因此 $\xi_i\xi_j^{\mathrm{T}} = \delta_{ij}$. 这表明,如果取 n 维实行向量空间 $\mathbb{R}^n$ 的内积为标准内积,则正交方阵 O 的 n 个行向量 $\xi_1, \xi_2, \ldots, \xi_n$ 构成 n 维 Euclid 空间 $\mathbb{R}^n$ 的一组标准正交基. 反之亦然.

同样,如果将正交方阵 O 按列分块,则方阵 O 的 n 个列向量 $\eta_1, \eta_2, \ldots, \eta_n$ 是 n 维列向量空间 $\mathbb{R}^n$ 的一组标准正交基. 反之亦然.

记所有 n 阶实正交方阵方阵的集合为 $\mathrm{O}_n(\mathbb{R})$. 定理 7.2.5 表明,在 n 维 Euclid 空间 V 中取定一组标准正交基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$, V 中的标准正交基 $\{\eta_1, \eta_2, \ldots, \eta_n\}$ 便由

$$(\eta_1, \eta_2, \ldots, \eta_n) = (\xi_1, \xi_2, \ldots, \xi_n)O$$

确定一个 n 阶正交方阵 $O \in \mathrm{O}_n(\mathbb{R})$. 于是可以建立 V 的所有标准正交基的集合到 $\mathrm{O}_n(\mathbb{R})$ 的映射 σ: 令

$$\sigma(\eta_1, \eta_2, \ldots, \eta_n) = O.$$

定理 7.2.6 表明,映射 σ 是满射. 容易验证,映射 σ 是单射. 所以 n 维 Euclid 空间 V 中所有标准正交基集合便和所有 n 阶实正交方阵集合 $\mathrm{O}_n(\mathbb{R})$ 建立了一一对应.

Euclid 空间 V 中向量间的正交性可以推广.

定义 7.2.2　设 U 是 Euclid 空间 V 的子空间, $\beta \in V$. 如果对任意 $\alpha \in U$, $(\alpha, \beta) = 0$,则称向量 β 和子空间 U 正交. V 中所有与子空间 U 正交的向量集合称为子空间 U 的正交补,记为 $U^{\perp}$.

由于零向量和任意向量都**正交**,因此 $0 \in U^{\perp}$. 所以 $U^{\perp} \neq \varnothing$. 其次设 $\beta_1, \beta_2 \in U^{\perp}$,则对任意 $\alpha \in U$, $(\alpha, \beta_1) = (\alpha, \beta_2) = 0$,因此

$$(\alpha, \beta_1 + \beta_2) = (\alpha, \beta_1) + (\alpha, \beta_2) = 0.$$

所以 $\beta_1 + \beta_2 \in U^{\perp}$. 最后设 $\lambda \in \mathbb{R}$, $\beta \in U^{\perp}$,则对任意 $\alpha \in U$,

$$(\alpha, \lambda\beta) = \lambda(\alpha, \beta) = 0.$$

所以 $\lambda\beta \in U^{\perp}$. 这表明,子空间 U 的正交补 $U^{\perp}$ 是 V 的子空间. 并且显然 $(U^{\perp})^{\perp} = U$.

定理 7.2.7　设 U 是 n 维 Euclid 空间 V 的子空间,则 $V = U \oplus U^{\perp}$.

证明　设 $\dim U = k$.

将 V 的内积 (α, β) 限定在子空间 U 上,即限定向量元 α 与 β 只取 U 中的向量,则 (α, β) 是子空间 U 的内积. 于是子空间 U 连同内积 (α, β) 一起是一个 k 维 Euclid 空间.

由定理 7.2.3，U 具有标准正交基 $\{\xi_1, \xi_2, \ldots, \xi_k\}$. 显然，$\xi_1, \xi_2, \ldots, \xi_k$ 是 V 中一组两两正交的单位向量. 由定理 7.2.4，它们可以扩充成 V 的标准正交基

$$\{\xi_1, \xi_2, \ldots, \xi_k, \xi_{k+1}, \xi_{k+2}, \ldots, \xi_n\}.$$

V 中由向量 $\xi_{k+1}, \xi_{k+2}, \ldots, \xi_n$ 生成的子空间记为 W. 显然 $V = U \oplus W$.

设 $\alpha \in U, \beta \in W$，则 $\alpha = a_1\xi_1 + a_2\xi_2 + \cdots + a_k\xi_k, \beta = b_{k+1}\xi_{k+1} + \cdots + b_n\xi_n$. 因此

$$(\alpha, \beta) = \left(\sum_{i=1}^{k} a_i\xi_i, \sum_{j=k+1}^{n} b_j\xi_j\right) = \sum_{i=1}^{k}\sum_{j=k+1}^{n} a_i b_j(\xi_i, \xi_j) = 0.$$

即 $\beta \subseteq U^{\perp}$. 因此 $W \subseteq U^{\perp}$.

反之设 $\beta \in U^{\perp}$，则 $\beta = c_1\xi_1 + c_2\xi_2 + \cdots + c_n\xi_n$，并且对 $1 \leqslant i \leqslant k, (\xi_i, \beta) = 0$. 所以

$$(\xi_i, \beta) = \left(\xi_i, \sum_{j=1}^{n} c_j\xi_j\right) = \sum_{j=1}^{n} c_j(\xi_i, \xi_j) = c_i = 0,$$

因此 $\beta = c_{k+1}\xi_{k+1} + \cdots + c_n\xi_n \in W$. 即 $U^{\perp} \subseteq W$. 从而 $U^{\perp} = W$.

综上所述，$V = U \oplus U^{\perp}$. ■

应当指出，定理 7.2.7 表明，子空间 U 和它的正交补 $U^{\perp}$ 的交 $U \cap U^{\perp} = 0$.

最后用几个例子结束本节.

例 7.2.1 任意一个 n 阶可逆实方阵 A 都可以表为一个实正交方阵 O 和一个对角元全为正数的上三角方阵 T 的乘积，即 $A = OT$. 并且这种表法唯一.

证明 设 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 n 维 Euclid 空间 V 的一组标准正交基. 因为方阵 A 可逆，所以由

$$(\beta_1, \beta_2, \ldots, \beta_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$$

所确定的向量组 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 是 V 的一组基.

对向量 $\beta_1, \beta_2, \ldots, \beta_n$ 施行 Gram-Schmidt 正交化，即令

$$\begin{aligned}
\xi_1 &= \beta_1, \\
\xi_2 &= \beta_2 - \frac{(\beta_2, \xi_1)}{(\xi_1, \xi_1)}\xi_1, \\
&\cdots\cdots\cdots \\
\xi_n &= \beta_n - \frac{(\beta_n, \xi_{n-1})}{(\xi_{n-1}, \xi_{n-1})}\xi_{n-1} - \cdots - \frac{(\beta_n, \xi_1)}{(\xi_1, \xi_1)}\xi_1,
\end{aligned}$$

则向量 $\xi_1, \xi_2, \ldots, \xi_n$ 两两正交，而且都是非零向量. 将上式写成矩阵的形式即得

$$(\beta_1, \beta_2, \ldots, \beta_n) = (\xi_1, \xi_2, \ldots, \xi_n)\begin{pmatrix} 1 & \dfrac{(\beta_2, \xi_1)}{(\xi_1, \xi_1)} & \cdots & \dfrac{(\beta_n, \xi_1)}{(\xi_1, \xi_1)} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \dfrac{(\beta_n, \xi_{n-1})}{(\xi_{n-1}, \xi_{n-1})} \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

再把向量 $\xi_1, \xi_2, \ldots, \xi_n$ 单位化，即令

$$\eta_1=\frac{\xi_1}{\|\xi_1\|},\quad \eta_2=\frac{\xi_2}{\|\xi_2\|},\quad \cdots\cdots,\quad \eta_n=\frac{\xi_n}{\|\xi_n\|}.$$

写成矩阵形式即为

$$(\xi_1,\xi_2,\ldots,\xi_n)=(\eta_1,\eta_2,\ldots,\eta_n)\operatorname{diag}\big(\|\xi_1\|,\|\xi_2\|,\ldots,\|\xi_n\|\big),$$

其中 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 V 的一组标准正交基. 于是

$$(\beta_1,\beta_2,\ldots,\beta_n)=(\eta_1,\eta_2,\ldots,\eta_n)T,$$

其中

$$T=\operatorname{diag}\big(\|\xi_1\|,\|\xi_2\|,\ldots,\|\xi_n\|\big)\begin{pmatrix}1 & \dfrac{(\beta_2,\xi_1)}{(\xi_1,\xi_1)} & \cdots & \dfrac{(\beta_n,\xi_1)}{(\xi_1,\xi_1)}\\ 0 & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \dfrac{(\beta_n,\xi_{n-1})}{(\xi_{n-1},\xi_{n-1})}\\ 0 & \cdots & 0 & 1\end{pmatrix}.$$

显然方阵 T 是上三角的,而且每个对角元 $\|\xi_i\|$ 都是正数.

由于 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 V 的标准正交基,故由定理 7.2.5,

$$(\eta_1,\eta_2,\ldots,\eta_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)O,$$

其中 O 是正交方阵. 因此

$$(\beta_1,\beta_2,\ldots,\beta_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)OT.$$

由于 V 中由一组基到另一组基的过渡矩阵是唯一的,所以 $A=OT$.

现在证明表法唯一. 设 $A=OT=O_1T_1$,其中 O_1 是正交方阵,T_1 是对角元全为正数的上三角阵. 则 $O_1^{\mathrm T}O=T_1T^{-1}$. 记 $C=O_1^{\mathrm T}O$,由于 O_1 是正交方阵,故 $O_1^{\mathrm T}$ 是正交方阵,又正交方阵的乘积是正交方阵,因此 C 是正交方阵.

另一方面,T 是上三角方阵,它的逆方阵 T^{-1} 也是上三角方阵. 又上三角方阵的乘积是上三角方阵,所以 C 是上三角的正交方阵,从而 C 是对角方阵,其对角元为 1 或 -1. 由于 $T_1=CT$,且 T 与 T_1 的对角元全为正数,所以 C 的对角元只能是 1,即 $C=I_{(n)}$,从而 $O_1=O$,$T_1=T$. ■

例 7.2.2　设 n 阶实正交方阵 O 的第 $i_1,i_2,\ldots,i_r$ 行与第 $j_1,j_2,\ldots,j_r$ 列上的第 r 阶子式为 $N=O\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}$,其中 $1\leqslant i_1<i_2\cdots<i_r\leqslant n$,$1\leqslant j_1<j_2\cdots<j_r\leqslant n$,且 r 阶子式 N 的代数余子式为 M,则 $N=M\det O$.

证明　先证明 N 是由方阵 O 的西北角构成的情形. 将方阵 O 按前 r 行与前 r 列分块,即设

$$O=\begin{pmatrix}O_{11} & O_{12}\\ O_{21} & O_{22}\end{pmatrix},$$

其中 O_{11} 是 r 阶方阵. 因为方阵 O 是正交的,所以

$$OO^{\mathrm T}=\begin{pmatrix}O_{11} & O_{12}\\ O_{21} & O_{22}\end{pmatrix}\begin{pmatrix}O_{11}^{\mathrm T} & O_{21}^{\mathrm T}\\ O_{12}^{\mathrm T} & O_{22}^{\mathrm T}\end{pmatrix}=I_{(n)},$$

即得到
$$O_{11}O_{11}^{\mathrm{T}}+O_{12}O_{12}^{\mathrm{T}}=I_{(r)},$$
$$O_{11}O_{21}^{\mathrm{T}}+O_{12}O_{22}^{\mathrm{T}}=0,$$
$$O_{21}O_{21}^{\mathrm{T}}+O_{22}O_{22}^{\mathrm{T}}=I_{(n-r)}.$$
因此
$$\begin{pmatrix}O_{11}&O_{12}\\O_{21}&O_{22}\end{pmatrix}\begin{pmatrix}I_{(r)}&O_{21}^{\mathrm{T}}\\0&O_{22}^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}O_{11}&0\\O_{21}&I_{(n-r)}\end{pmatrix}.$$
两端取行列式即得到
$$\det O\det O_{22}^{\mathrm{T}}=\det O_{11}.$$
于是 $N=M\det O$.

再讨论一般情形. 此时 $N=O\left(\begin{smallmatrix}i_1&i_2&\cdots&i_r\\j_1&j_2&\cdots&j_r\end{smallmatrix}\right)$ 将方阵 O 的第 $i_1,i_2,\ldots,i_r$ 行依次经相邻两行的对换调换到第 $1,2,\ldots,r$ 行,并将第 $j_1,j_2,\ldots,j_r$ 列依次经相邻两列的对换调换到第 $1,2,\ldots,r$ 列,得到的方阵记为 O_1. 由于对方阵 O 施行两行互换或两列互换相当于用初等置换方阵左乘或右乘于方阵,而初等置换方阵是正交方阵,两个正交方阵的乘积仍是正交方阵,所以 O_1 是正交方阵. 由行列式性质可知,
$$\det O_1=(-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r}\det O.$$
方阵 O_1 中前 r 行与前 r 列构成的 r 阶子式记为 N_1,显然 $N_1=N$. N_1 在方阵 O_1 中的余子式记为 M_1. 由于 M 是 N 在方阵 O 中的代数余子式,所以
$$M=(-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r}M_1.$$
于是由上段证明
$$N=N_1=M_1\det O_1=(-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r}M_1\det O,$$
即
$$N=M\det O.$$
■

习 题 7.2

1. (Bessel 不等式) 设 $\alpha_1,\alpha_2,\ldots,\alpha_k$ 是 n 维 Euclid 空间 V 的一组两两正交的单位向量,$\alpha\in V$,并记 $a_i=(\alpha,\alpha_i)$, $i=1,2,\ldots,k$. 证明
$$\sum_{i=1}^{k}|a_i|^2\leqslant\|\alpha\|^2,$$
而且向量
$$\beta=\alpha-\sum_{i=1}^{k}a_i\alpha_i$$
与每个向量 α_i 都正交,$i=1,2,\ldots,k$.

2. 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的一组向量. 证明,下面的命题等价:

(1) $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 V 的标准正交基;

(2) (Parseval 等式) 对任意 $\alpha,\beta\in V$,
$$(\alpha,\beta)=\sum_{i=1}^{n}(\alpha,\alpha_i)(\alpha_i,\beta);$$

(3) 对任意 $\alpha\in V$,
$$\|\alpha\|^2=\sum_{i=1}^{n}|(\alpha,\alpha_i)|^2.$$

3. 所有次数小于 $n+1$ 的实系数多项式 $f(x)$ 的集合 $\mathbb{R}_{n+1}[x]$ 连同内积

$$\big(f(x),g(x)\big)=\int_0^1 f(x)g(x)\,\mathrm{d}x$$

一起构成 Euclid 空间，其中 $f(x),g(x)\in\mathbb{R}_{n+1}[x]$. 设 $f_j(x)=x^j$, $j=0,1,2,\ldots,n-1$. 求 $\mathbb{R}_{n+1}[x]$ 中与多项式 $f_0(x),f_1(x),\ldots,f_{n-1}(x)$ 都正交的多项式.

4. $\mathbb{R}_{n+1}[x]$ 的意义同习题 3. 利用 Gram-Schmidt 正交化，由 $\mathbb{R}_{n+1}[x]$ 的基 $\{1,x,\ldots,x^n\}$ 求出 $\mathbb{R}_{n+1}[x]$ 的一组标准正交基.

5. 设 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 n 维 Euclid 空间 V 的一组标准正交基，向量 $\alpha_j\in V$ 在这组基下的坐标为 $x_j=(x_{j1},x_{j2},\ldots,x_{jn})^{\mathrm{T}}$, $j=1,2,\ldots,n$. 向量组 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 的 Gram 方阵为 $G(\alpha_1,\alpha_2,\ldots,\alpha_n)=\big((\alpha_i,\alpha_j)\big)_{n\times n}$. 证明，

$$\det G(\alpha_1,\alpha_2,\ldots,\alpha_n)=\big(\det(x_{ij})\big)^2.$$

6. 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的一组基. 对 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 施行 Gram-Schmidt 正交化得到的正交向量组记为 $\{\beta_1,\beta_2,\ldots,\beta_n\}$. 证明，对于 $j=1,2,\ldots,n$,

$$\|\beta_j\|^2=\frac{\det(\alpha_1,\alpha_2,\ldots,\alpha_j)}{\det(\alpha_1,\alpha_2,\ldots,\alpha_{j-1})},$$

其中约定零个向量的 Gram 方阵的行列式为 1.

7. 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的一组向量. 证明，

$$\det(\alpha_1,\alpha_2,\ldots,\alpha_n)\leqslant\|\alpha_1\|^2\|\alpha_2\|^2\cdots\|\alpha_n\|^2,$$

等式当且仅当 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 两两正交或其中含有零向量时成立. 由此证明，如果 $A=(a_{ij})$ 是 n 阶实方阵，则

$$(\det A)^2\leqslant\prod_{i=1}^n\sum_{j=1}^n a_{ij}^2.$$

8. 举例说明，方阵 A 的行向量两两正交，它的列向量并不一定两两正交.

9. 设 O 是 n 阶正交方阵. 而方阵 $A=\mathrm{diag}(a_1,a_2,\ldots,a_n)$. 证明，方阵 OA 的特征值 λ_0 满足 $m\leqslant|\lambda_0|\leqslant M$，其中

$$m=\min\big\{|a_j|\,\big|\,1\leqslant j\leqslant n\big\},\quad M=\max\big\{|a_j|\,\big|\,1\leqslant j\leqslant n\big\}.$$

10. 证明，正交方阵 O 的任意一个子方阵的特征值的绝对值小于或等于 1.

11. 证明，如果 n 阶正交方阵 O 的行列式为 1，则方阵 O 可以表为有限个形如

$$O_{jk}=I_{(n)}+(\cos\theta-1)(E_{jj}+E_{kk})+\sin\theta(E_{jk}-E_{kj})$$

的方阵的乘积，其中 E_{st} 是 (s,t) 位置上的元素为 1 而其它元素都为零的 n 阶方阵，并且 $1\leqslant j<k\leqslant n$. 如果 n 阶正交方阵 O 的行列式为 -1，则还应该添加上方阵 $\mathrm{diag}(\underbrace{1,\ldots,1}_{n-1\text{ 个}},-1)$.

12. 设 $\alpha=\beta+\mathrm{i}\gamma$ 是正交方阵 O 的属于特征值 λ 的特征向量，其中 β 与 γ 是实向量，$\mathrm{i}^2=-1$. 证明，$|\lambda|=1$，而且当 $\lambda\notin\mathbb{R}$ 时，实向量 β 与 γ 正交，且范数相等.

13. 设 λ 是 n 阶斜对称实方阵 K 的非零特征值，$\alpha=\beta+\mathrm{i}\gamma$ 是属于 λ 的特征向量，其中 β 与 γ 是实向量. 证明 λ 是纯虚数，而且实向量 β 与 γ 正交，范数相等.

14. 设 U 与 W 是 n 维 Euclid 空间 V 的子空间. 证明，

$$(U+W)^\perp=U^\perp\cap W^\perp,\quad (U\cap W)^\perp=W^\perp+W^\perp.$$

15. 设 $\mathbb{R}^4$ 是所有 4 维实行向量空间连同标准内积一起构成的 Euclid 空间. $\mathbb{R}^4$ 中由向量 $\alpha=(1,0,-1,1)$ 与 $\beta=(2,3,-1,2)$ 生成的的子空间记为 U. 求正交补 $U^\perp$ 的一组标准正交基.

16. 设 $\mathbb{R}_4[x]$ 是所有次数小于 4 的实系数多项式集合连同内积

$$\big(f(x),g(x)\big)=\int_0^1 f(x)g(x)\,\mathrm{d}x$$

一起构成的 Euclid 空间,其中 $f(x), g(x) \in \mathbb{R}_4[x]$.

求 $\mathbb{R}_4[x]$ 中由零次多项式生成的子空间 U 的正交补 $U^{\perp}$.

17. 设 A 是秩为 r 的 n 阶实方阵. 证明,存在 n 阶正交方阵 O 和 n 阶置换方阵,是的 $A = PTO$,其中

$$T = \begin{pmatrix} \begin{pmatrix} t_{11} & 0 & \cdots & 0 \\ t_{21} & t_{22} & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ t_{r1} & t_{r2} & \cdots & t_{rr} \end{pmatrix} & O_{r\times(n-r)} \\ T_{(n-r)\times r} & O_{(n-r)\times(n-r)} \end{pmatrix},$$

并且对角元 $t_{11}, t_{22}, \ldots, t_{rr}$ 都是正数.

18. 设 $\mathbb{R}^{n\times n}$ 是所有 n 阶实方阵集合连同内积 $(X, Y) = \operatorname{Tr} XY^{\mathrm{T}}$ 构成的 Euclid 空间,其中 $X, Y \in \mathbb{R}^{n\times n}$,求 $\mathbb{R}^{n\times n}$ 中由纯量方阵生成的子空间 U 的正交补 $U^{\perp}$.

19. 设 V_1 和 V_2 是有限维 Euclid 空间. 记 $V_1 \times V_2 = \{(\alpha, \beta) \mid \alpha \in V_1, \beta \in V_2\}$. 在 $V_1 \times V_2$ 中定义加法:对 $(\alpha_1, \beta_1), (\alpha_2, \beta_2) \in V_1 \times V_2$,令

$$(\alpha_1, \beta_1) + (\alpha_2, \beta_2) = (\alpha_1 + \alpha_2, \beta_1 + \beta_2);$$

定义纯量 $\lambda \in \mathbb{R}$ 与向量 $(\alpha, \beta) \in V_1 \times V_2$ 的乘积为

$$\lambda(\alpha, \beta) = (\lambda\alpha, \lambda\beta).$$

于是 $V_1 \times V_2$ 在如此的加法和纯量与向量的乘法下成为实线性空间. 设 $f_1(\alpha, \beta)$ 与 $f_2(\gamma, \delta)$ 分别是 V_1 与 V_2 的内积. 证明,$V_1 \times V_2$ 具有唯一一个内积 $f\big((\alpha_1, \beta_1), (\alpha_2, \beta_2)\big)$,它满足:

(1) $V_2 = V_1^{\perp}$;

(2) 当 $\alpha, \beta \in V_1$ 时,$f\big((\alpha, 0), (\beta, 0)\big) = f_1(\alpha, \beta)$;当 $\alpha, \beta \in V_2$ 时,$f\big((0, \alpha), (0, \beta)\big) = f_2(\alpha, \beta)$.

§7.3 线性函数与伴随变换

先讨论 Euclid 空间 V 上的线性函数.

定义 7.3.1 设 $f(\alpha)$ 是 Euclid 空间 V 上的实函数,其中 $\alpha \in V$. 如果对任意 $\lambda_1, \lambda_2 \in \mathbb{R}, \alpha_1, \alpha_2 \in V$,

$$f(\lambda_1\alpha_1 + \lambda_2\alpha_2) = \lambda_1 f(\alpha_1) + \lambda_2 f(\alpha_2),$$

则 $f(\alpha)$ 称为 V 上的**线性函数**.

例如,设 (α, β) 是 Euclid 空间 V 的内积,取定向量 $\beta \in V$,则 (α, β) 是 V 上的一个线性函数,记为 $f_\beta(\alpha)$.

容易看出,如果 $f(\alpha)$ 是 V 上的线性函数,则 $f(0) = 0$,其中左端的 0 维 V 中的零向量,右端的 0 为实数 0. 其次,对任意 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}, \alpha_1, \alpha_2, \ldots, \alpha_k \in V$,有

$$f\Big(\sum_{j=1}^{k} \lambda_j \alpha_j\Big) = \sum_{j=1}^{k} \lambda_j f(\alpha_j).$$

Euclid 空间 V 上所有线性函数的集合记为 V^*. 在 V^* 中定义加法如下:设 $f_1, f_2 \in V^*$,即 $f_1(\alpha)$ 与 $f_2(\alpha)$ 是 V 上的线性函数,则 f_1 与 f_2 的和 $f_1 + f_2$ 定义为:对

任意 $\alpha\in V$,令

$$(f_1+f_2)(\alpha)=f_1(\alpha)+f_2(\alpha).$$

容易验证,线性函数 f_1 与 f_2 的和 f_1+f_2 是线性函数.

事实上,对于任意 $\lambda_1,\lambda_2\in\mathbb{R},\alpha_1,\alpha_2\in V$,

$$\begin{aligned}(f_1+f_2)(\lambda_1\alpha_1+\lambda_2\alpha_2)&=f_1(\lambda_1\alpha_1+\lambda_2\alpha_2)+f_2(\lambda_1\alpha_1+\lambda_2\alpha_2)\\&=\lambda_1f_1(\alpha_1)+\lambda_2f_1(\alpha_2)+\lambda_1f_2(\alpha_1)+\lambda_2f_2(\alpha_2)\\&=\lambda_1\big(f_1(\alpha_1)+f_2(\alpha_1)\big)+\lambda_2\big(f_1(\alpha_2)+f_2(\alpha_2)\big)\\&=\lambda_1(f_1+f_2)(\alpha_1)+\lambda_2(f_1+f_2)(\alpha_2).\end{aligned}$$

所以 f_1+f_2 是 V 上的线性函数,即 $f_1+f_2\in V^*$.

其次定义纯量 $\lambda\in\mathbb{R}$ 与 $f\in V^*$ 的乘积 λf 如下:对任意 $\alpha\in V$,令

$$(\lambda f)(\alpha)=\lambda f(\alpha).$$

容易验证,λf 仍是 V 上的线性函数. 事实上,对于任意 $\lambda_1,\lambda_2\in\mathbb{R},\alpha_1,\alpha_2\in V$,

$$\begin{aligned}(\lambda f)(\lambda_1\alpha_1+\lambda_2\alpha_2)&=\lambda f(\lambda_1\alpha_1+\lambda_2\alpha_2)=\lambda\lambda_1f(\alpha_1)+\lambda\lambda_2f(\alpha_2)\\&=\lambda_1\lambda f(\alpha_1)+\lambda_2\lambda f(\alpha_2)=\lambda_1(\lambda f)(\alpha_1)+\lambda_2(\lambda f)(\alpha_2).\end{aligned}$$

所以 λf 是线性函数,即 $\lambda f\in V^*$.

容易验证,集合 V^* 在上述加法与乘法下满足线性空间的八条定理. 因此集合 V^* 是一个实线性空间,它称为 Euclid 空间 V 的**对偶空间**.

利用 Euclid 空间 V 的内积 (α,β),可以建立 V 到它的对偶空间 V^* 的映射 σ 如下:对于 $\beta\in V$,$f_\beta(\alpha)$ 是 V 上的线性函数,即 $f_\beta\in V^*$,于是令 $\sigma(\beta)=f_\beta$.

定理 7.3.1　n 维 Euclid 空间 V 到它的对偶空间 V^* 的映射 σ 是同构映射,从而 V 与 V^* 同构.

证明　首先证明,σ 是单射. 事实上,设 $\beta_1,\beta_2\in V$,且 $\sigma(\beta_1)=\sigma(\beta_2)$,即 $f_{\beta_1}=f_{\beta_2}$. 于是对任意 $\alpha\in V$,$f_{\beta_1}(\alpha)=f_{\beta_2}(\alpha)$,即 $(\alpha,\beta_1)=(\alpha,\beta_2)$. 因此 $(\alpha,\beta_1-\beta_2)=0$. 其中向量 α 是任意的,故可取 $\alpha=\beta_1-\beta_2$. 所以 $(\beta_1-\beta_2,\beta_1-\beta_2)=0$. 因此 $\beta_1-\beta_2=0$,即 $\beta_1=\beta_2$. 这表明,σ 是单射.

其次证明,σ 是满射. 事实上,设 $f\in V^*$,即 $f(\alpha)$ 是 V 上的线性函数. 设 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 是 V 的标准正交基,且 $\alpha=x_1\xi_1+x_2\xi_2+\cdots+x_n\xi_n$,则

$$f(\alpha)=f\Big(\sum_{j=1}^n x_j\xi_j\Big)=\sum_{j=1}^n x_jf(\xi_j).$$

另一方面,取 $\beta=f(\xi_1)\xi_1+f(\xi_2)\xi_2+\cdots+f(\xi_n)\xi_n\in V$,则

$$\begin{aligned}f_\beta(\alpha)&=(\alpha,\beta)=\Big(\sum_{j=1}^n x_j\xi_j,\ \sum_{k=1}^n f(\xi_k)\xi_k\Big)=\sum_{j=1}^n\sum_{k=1}^n x_jf(\xi_k)(\xi_j,\xi_k)\\&=\sum_{j=1}^n\sum_{k=1}^n x_jf(\xi_k)\delta_{jk}=\sum_{j=1}^n x_jf(\xi_j)=f(\alpha).\end{aligned}$$

由 α 的任意性得到,$f=f_\beta$. 即存在向量 $\beta\in V$,使得 $\sigma(\beta)=f_\beta=f$. 故 σ 是满射.

又，映射 σ 是保加法的. 事实上，设 $\beta_1,\beta_2\in V$，则 $\sigma(\beta_1+\beta_2)=f_{\beta_1+\beta_2}$. 而

$$\begin{aligned}f_{\beta_1+\beta_2}(\alpha)&=(\alpha,\beta_1+\beta_2)=(\alpha,\beta_1)+(\alpha,\beta_2)=f_{\beta_1}(\alpha)+f_{\beta_2}(\alpha)\\&=(f_{\beta_1}+f_{\beta_2})(\alpha),\end{aligned}$$

因此 $f_{\beta_1+\beta_2}=f_{\beta_1}+f_{\beta_2}$. 所以 $\sigma(\beta_1+\beta_2)=\sigma(\beta_1)+\sigma(\beta_2)$.

最后，映射 σ 是保乘法的. 事实上，设 $\lambda\in\mathbb{R},\beta\in V$，则 $\sigma(\lambda\beta)=f_{\lambda\beta}$. 由于

$$f_{\lambda\beta}(\alpha)=(\alpha,\lambda\beta)=\lambda(\alpha,\beta)=\lambda f_\beta(\alpha)=(\lambda f_\beta)(\alpha),$$

所以 $f_{\lambda\beta}=\lambda f_\beta$，即 $\sigma(\lambda\beta)=\lambda\sigma(\beta)$.

于是映射 σ 是 V 到 V^* 上的同构映射，从而 V 与它的对偶空间 V^* 同构. ■

定理 7.3.2 设 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 是 n 维 Euclid 空间 V 的一组基，则 $\{f_{\beta_1},f_{\beta_2},\dots,f_{\beta_n}\}$ 是对偶空间 V^* 的一组基. 它称为 $\beta_1,\beta_2,\dots,\beta_n$ 的**对偶基**.

证明 由定理 7.3.1，$\dim V^*=\dim V=n$. 因此只需证明，$f_{\beta_1},f_{\beta_2},\dots,f_{\beta_n}$ 线性无关. 设 $\lambda_1,\lambda_2,\dots,\lambda_n\in\mathbb{R}$，且 $\lambda_1 f_{\beta_1}+\lambda_2 f_{\beta_2}+\cdots+f_n f_{\beta_n}=0$. 由定理 7.3.1 的证明可得

$$f_{\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_n\beta_n}=\lambda_1 f_{\beta_1}+\lambda_2 f_{\beta_2}+\cdots+\lambda_n f_{\beta_n}=0,$$

因此对任意 $\alpha\in V$，有

$$f_{\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_n\beta_n}(\alpha)=\big(\alpha,\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_n\beta_n\big)=0.$$

由 α 的任意性，可取 $\alpha=\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_n\beta_n$. 于是

$$\big(\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_n\beta_n,\ \lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_n\beta_n\big)=0.$$

因此 $\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_n\beta_n=0$. 由于 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 是 V 的基，所以

$$\lambda_1=\lambda_2=\cdots=\lambda_n=0.$$

这表明，$f_{\beta_1},f_{\beta_2},\dots,f_{\beta_n}$ 线性无关. ■

现在转到 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$.

定理 7.3.3 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换，(α,β) 是 V 的内积. 则对给定的向量 $\beta\in V$，存在唯一的向量 $\overline{\beta}\in V$，使得对任意 $\alpha\in V$，均有 $\big(\mathscr{A}(\alpha),\beta\big)=(\alpha,\overline{\beta})$.

证明 记 $f(\alpha)=\big(\mathscr{A}(\alpha),\beta\big)$. 显然 $f(\alpha)$ 是 V 上的一个实函数. 因为 $\mathscr{A}$ 是 V 的线性变换，所以对任意 $\lambda_1,\lambda_2\in\mathbb{R},\alpha_1,\alpha_2\in V$，

$$\begin{aligned}f(\lambda_1\alpha_1+\lambda_2\alpha_2)&=\big(\mathscr{A}(\lambda_1\alpha_1+\lambda_2\alpha_2),\beta\big)=\big(\lambda_1\mathscr{A}(\alpha_1)+\lambda_2\mathscr{A}(\alpha_2),\beta\big)\\&=\lambda_1\big(\mathscr{A}(\alpha_1),\beta\big)+\lambda_2\big(\mathscr{A}(\alpha_2),\beta\big)=\lambda_1 f(\alpha_1)+\lambda_2 f(\alpha_2).\end{aligned}$$

所以 $f(\alpha)$ 是 V 上的线性函数.

由定理 7.3.1，存在唯一 $\overline{\beta}\in V$，使得 $f=f_{\overline{\beta}}$. 即 $\big(\mathscr{A}(\alpha),\beta\big)=(\alpha,\overline{\beta})$. ■

根据定理 7.3.3，可以引进如下定义.

定义 7.3.2 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换. 定义 V 到自身的变换 $\mathscr{A}^*$

如下：设 $\beta\in V$，则存在唯一 $\overline{\beta}\in V$，使得 $\big(\mathscr{A}(\alpha),\beta\big)=(\alpha,\overline{\beta})$，于是令 $\mathscr{A}^*(\beta)=\overline{\beta}$. 变换 $\mathscr{A}^*$ 称为 $\mathscr{A}$ 的**伴随变换**.

定理 7.3.4　线性变换 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$ 是线性变换.

证明　设 $\lambda_,\lambda_2\in\mathbb{R},\beta_1,\beta_2\in V$. 记

$$\mathscr{A}^*(\lambda_1\beta_1+\lambda_2\beta_2)=\overline{\lambda_1\beta_1+\lambda_2\beta_2}.$$

则由定义

$$\begin{aligned}\big(\alpha,\overline{\lambda_1\beta_1+\lambda_2\beta_2}\big)&=\big(\mathscr{A}(\alpha),\lambda_1\beta_1+\lambda_2\beta_2\big)=\lambda_1\big(\mathscr{A}(\alpha),\beta_1\big)+\lambda_2\big(\mathscr{A}(\alpha),\beta_2\big)\\&=\lambda_1(\alpha,\overline{\beta}_1)+\lambda_2(\alpha,\overline{\beta}_2)=\big(\alpha,\lambda_1\overline{\beta}_1+\lambda_2\overline{\beta}_2\big).\end{aligned}$$

即

$$\big(\alpha,\ \overline{\lambda_1\beta_1+\lambda_2\beta_2}-(\lambda_1\overline{\beta}_1+\lambda_2\overline{\beta}_2)\big)=0.$$

由 α 的任意性得到，$\overline{\lambda_1\beta_1+\lambda_2\beta_2}=\lambda_1\overline{\beta}_1+\lambda_2\overline{\beta}_2$，即

$$\mathscr{A}^*(\lambda_1\beta_1+\lambda_2\beta_2)=\lambda_1\mathscr{A}^*(\beta_1)+\lambda_2\mathscr{A}^*(\beta_2).$$ ■

下面给出伴随变换的一些例子.

例 7.3.1　设 V 是所有 n 维实行向量集合连同标准内积一起构成的 n 维 Euclid 空间. 设 A 是 n 阶实方阵. 则由 $\mathscr{A}(x)=xA,x\in V$ 便确定 V 的一个线性变换 $\mathscr{A}$，求 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$.

解　n 维 Euclid 空间 V 的内积 (x,y) 为

$$(x,y)=x_1y_1+x_2y_2+\cdots+x_ny_n=xy^{\mathrm{T}}.$$

其中 $x=(x_1,x_2,\dots,x_n),y=(y_1,y_2,\dots,y_n)\in V$. 因此对给定的 $y\in V$，

$$\big(\mathscr{A}(x),y\big)=(xA,y)=xAy^{\mathrm{T}}=x(yA^{\mathrm{T}})^{\mathrm{T}}=(x,yA^{\mathrm{T}}).$$

于是 $\mathscr{A}^*(y)=yA^{\mathrm{T}}$. 所以对任意 $x\in V$，$\mathscr{A}^*(x)=xA^{\mathrm{T}}$. ■

例 7.3.2　设 $\mathbb{R}^{n\times n}$ 是所有 n 阶实方阵集合连同内积 $(X,Y)=\operatorname{Tr}XY^{\mathrm{T}}$ 一起构成的 Euclid 空间，其中 $X,Y\in\mathbb{R}^{n\times n}$. 设 A 是 n 阶实方阵. 定义 $\mathbb{R}^{n\times n}$ 的变换 $\mathscr{A}$ 如下：对任意 $X\in\mathbb{R}^{n\times n}$，令 $\mathscr{A}(X)=XA$. 则 $\mathscr{A}$ 是 $\mathbb{R}^{n\times n}$ 的线性变换. 求 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$.

解　对给定的 $Y\in\mathbb{R}^{n\times n}$，

$$\big(\mathscr{A}(X),Y\big)=(XA,Y)=\operatorname{Tr}(XAY^{\mathrm{T}})=\operatorname{Tr}\big(X(YA^{\mathrm{T}})\big)^{\mathrm{T}}=(X,YA^{\mathrm{T}}),$$

因此 $\mathscr{A}^*(Y)=YA^{\mathrm{T}}$. 所以对任意 $X\in\mathbb{R}^{n\times n}$，$\mathscr{A}^*(X)=XA^{\mathrm{T}}$. ■

例 7.3.3　设 $\mathbb{R}[x]$ 是所有实系数多项式集合连同内积

$$\big(f(x),g(x)\big)=\int_0^1 f(x)g(x)\,\mathrm{d}x$$

构成的 Euclid 空间，其中 $f(x),g(x)\in\mathbb{R}[x]$.

对给定的 $f(x)\in\mathbb{R}[x]$，定义 $\mathbb{R}[x]$ 的变换 $\mathscr{A}_f$ 如下：对任意 $g(x)\in\mathbb{R}[x]$，令

$$\mathscr{A}_f\big(g(x)\big)=f(x)g(x).$$

于是 $\mathscr{A}_f$ 是 $\mathbb{R}[x]$ 的线性变换. 求 $\mathscr{A}_f$ 的伴随变换 $\mathscr{A}_f^*$.

解 设 $h(x)\in\mathbb{R}[x]$,则

$$\begin{aligned}\big(\mathscr{A}_f(g(x)),h(x)\big)&=\int_0^1\big(f(x)g(x)\big)h(x)\,\mathrm{d}x=\int_0^1 g(x)\big(f(x)h(x)\big)\,\mathrm{d}x\\&=\big(g(x),f(x)h(x)\big).\end{aligned}$$

因此 $\mathscr{A}_f^*\big(h(x)\big)=f(x)h(x)$. 即对任意 $g(x)\in\mathbb{R}[x]$,

$$\mathscr{A}_f^*\big(g(x)\big)=f(x)g(x)=\mathscr{A}_f\big(g(x)\big).$$

由 $g(x)$ 的任意性得到,$\mathscr{A}_f^*=\mathscr{A}_f$. ■

伴随变换具有以下性质.

定理 7.3.5 设 $\mathscr{A}$ 与 $\mathscr{B}$ 是 n 维 Euclid 空间 V 的线性变换,$\lambda\in\mathbb{R}$,则

(1) $(\mathscr{A}+\mathscr{B})^*=\mathscr{A}^*+\mathscr{B}^*$; **(2)** $(\lambda\mathscr{A})^*=\lambda\mathscr{A}^*$;

(3) $(\mathscr{A}\mathscr{B})^*=\mathscr{B}^*\mathscr{A}^*$; **(4)** $(\mathscr{A}^*)^*=\mathscr{A}$.

证明 **(1)** 只需证明,对任意 $\beta\in V$,$(\mathscr{A}+\mathscr{B})^*(\beta)=\mathscr{A}^*(\beta)+\mathscr{B}^*(\beta)$. 事实上,由伴随变换的定义,

$$\big((\mathscr{A}+\mathscr{B})(\alpha),\beta\big)=\big(\alpha,(\mathscr{A}+\mathscr{B})^*(\beta)\big).$$

另一方面,

$$\begin{aligned}\big((\mathscr{A}+\mathscr{B})(\alpha),\beta\big)&=\big(\mathscr{A}(\alpha),\beta\big)+\big(\mathscr{B}(\alpha),\beta\big)\\&=\big(\alpha,\mathscr{A}^*(\beta)\big)+\big(\alpha,\mathscr{B}^*(\beta)\big)=\big(\alpha,(\mathscr{A}^*+\mathscr{B}^*)(\beta)\big).\end{aligned}$$

所以 $\big(\alpha,(\mathscr{A}+\mathscr{B})^*(\beta)\big)=\big(\alpha,(\mathscr{A}^*+\mathscr{B}^*)(\beta)\big)$. 由 α 的任意性得到

$$(\mathscr{A}+\mathscr{B})^*(\beta)=(\mathscr{A}^*+\mathscr{B}^*)(\beta).$$

(2) 由伴随变换的定义,对任意 $\alpha,\beta\in V$,

$$\big((\lambda\mathscr{A})(\alpha),\beta\big)=\big(\alpha,(\lambda\mathscr{A})^*(\beta)\big).$$

另一方面,

$$\big((\lambda\mathscr{A})(\alpha),\beta\big)=\big(\lambda\mathscr{A}(\alpha),\beta\big)=\lambda\big(\mathscr{A}(\alpha),\beta\big)=\lambda\big(\alpha,\mathscr{A}^*(\beta)\big)=\big(\alpha,(\lambda\mathscr{A}^*)(\beta)\big).$$

因此 $\big(\alpha,(\lambda\mathscr{A})^*(\beta)\big)=\big(\alpha,(\lambda\mathscr{A}^*)(\beta)\big)$. 由 α 的任意性,

$$(\lambda\mathscr{A})^*(\beta)=\lambda\mathscr{A}^*(\beta),$$

因此 $(\lambda\mathscr{A})^*=\lambda\mathscr{A}^*$.

结论 (3) 与 (4) 的证明方法相同,此处从略. ■

下面的定理给出线性变换与它的伴随变换的方阵表示的联系.

定理 7.3.6 设线性变换 A 在 n 维 Euclid 空间 V 的标准正交基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵为 A,则它的伴随变换 $\mathscr{A}^*$ 在同一组基下的方阵为 A^{T}.

证明 设 $A=(a_{ij})_{n\times n}$,且

$$\mathscr{A}(\xi_1,\xi_2,\ldots,\xi_n)=(\xi_1,\xi_2,\ldots,\xi_n)A.$$

设 $\mathscr{A}^*$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵为 $B=(b_{ij})_{n\times n}$，即

$$\mathscr{A}^*(\xi_1,\xi_2,\ldots,\xi_n)=(\xi_1,\xi_2,\ldots,\xi_n)B.$$

因此

$$\mathscr{A}(\xi_j)=\sum_{k=1}^n a_{kj}\xi_k,\qquad \mathscr{A}^*(\xi_j)=\sum_{k=1}^n b_{kj}\xi_k.$$

由伴随变换的定义，对于 $1\leqslant i,j\leqslant n$，有

$$\big(\mathscr{A}(\xi_i),\xi_j\big)=\big(\xi_i,\mathscr{A}^*(\xi_j)\big),$$

所以

$$\left(\sum_{k=1}^n a_{ki}\xi_k,\ \xi_j\right)=\left(\xi_j,\ \sum_{\ell=1}^n b_{\ell j}\xi_\ell\right),\quad 即\quad \sum_{k=1}^n a_{ki}(\xi_k,\xi_j)=\sum_{\ell=1}^n b_{\ell j}(\xi_j,\xi_\ell).$$

由于 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是标准正交基，因此 $(\xi_i,\xi_j)=\delta_{ij}$. 所以由上式得到

$$\sum_{k=1}^n a_{ki}\delta_{kj}=\sum_{\ell=1}^n b_{\ell j}\delta_{j\ell}.$$

于是 $a_{ji}=b_{ij}$，$1\leqslant i,j\leqslant n$. 这就证明 $B=A^{\mathrm{T}}$. ■

定理 7.3.7　n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 的不变子空间 U 的正交补 $U^\perp$ 是 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$ 的不变子空间.

证明　由伴随变换的定义，对任意 $\alpha\in U$，$\big(\mathscr{A}(\alpha),\beta\big)=\big(\alpha,\mathscr{A}^*(\beta)\big)$. 由于 U 是 $\mathscr{A}$ 的不变子空间，所以 $\mathscr{A}(\alpha)\in U$. 如果 $\beta\in U^\perp$，则 $\big(\mathscr{A}(\alpha),\beta\big)=0$，从而 $\big(\alpha,\mathscr{A}^*(\beta)\big)=0$. 由 α 的任意性得到，$\mathscr{A}^*(\beta)\in U^\perp$. 所以 $U^\perp$ 是 $\mathscr{A}^*$ 的不变子空间. ■

最后给出 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 在 V 的不同标准正交基下的方阵表示之间的联系. 为此先引进定义

定义 7.3.3　设 A 与 B 是 n 阶实方阵，如果存在 n 阶实正交方阵 O，使得 $B=O^{\mathrm{T}}AO$，则称方阵 A 与 B **正交相似**.

容易验证，方阵之间的正交相似关系满足自反性，对称性与传递性. 因此方阵之间的正交相似关系是一种等价关系. 根据这种等价关系，所有 n 阶实方阵的集合 $\mathbb{R}^{n\times n}$ 可以划分为正交相似等价类，即彼此正交相似的方阵归在同一个正交相似等价类，而彼此不正交相似的方阵归在不同的正交相似等价类.

与矩阵的相抵，方阵的相似相类似，关于方阵的正交相似，有两个基本问题，即：在正交相似等价类中如何选取代表元? 如何判定两个方阵是否正交相似? 也即方阵在正交相似下的全系不变量是什么? 这些就是方阵在正交相似下的标准形问题. 本章将讨论某些特殊类型的方阵在正交相似下的标准形.

应当指出，对于正交方阵 O，它的逆方阵 $O^{-1}=O^{\mathrm{T}}$，因此正交相似的方阵一定相似. 反之，相似的方阵并不一定正交相似. 另外，正交相似的方阵一定是相合的，反之则不尽然.

定理 7.3.8 设 A 和 B 是 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 分别在 V 的标准正交基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 与 $\{\eta_1,\eta_2,\dots,\eta_n\}$ 下的方阵，则方阵 A 与 B 正交相似.

证明 由假设

$$\mathscr{A}(\xi_1,\xi_2,\dots,\xi_n)=(\xi_1,\xi_2,\dots,\xi_n)A,\tag{7.3.1}$$

$$\mathscr{A}(\eta_1,\eta_2,\dots,\eta_n)=(\eta_1,\eta_2,\dots,\eta_n)B.\tag{7.3.2}$$

由定理 7.2.5，

$$(\eta_1,\eta_2,\dots,\eta_n)=(\xi_1,\xi_2,\dots,\xi_n)O,\tag{7.3.3}$$

其中 O 是某个 n 阶实正交方阵．因此

$$\mathscr{A}(\eta_1,\eta_2,\dots,\eta_n)=\big(\mathscr{A}(\xi_1,\xi_2,\dots,\xi_n)\big)O.$$

由式 (7.3.1) 得到

$$=(\xi_1,\xi_2,\dots,\xi_n)AO.$$

由式 (7.3.3) 得到

$$=(\eta_1,\eta_2,\dots,\eta_n)O^{\mathrm{T}}AO.\tag{7.3.4}$$

比较式 (7.3.2) 与式 (7.3.4) 得到，$B=O^{\mathrm{T}}AO$．即方阵 A 与 B 正交相似. ∎

习 题 7.3

1. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换．证明，$\operatorname{Tr}A^*A\geqslant 0$，其中等式当且仅当线性变换 $\mathscr{A}$ 为零变换时成立.

2. 设 $\mathscr{A}$ 与 $\mathscr{B}$ 是 n 维 Euclid 空间 V 的线性变换，$\mathscr{A}$ 与 $\mathscr{B}$ 可交换，$\mathscr{A}^*$ 与 $\mathscr{A}$ 可交换．证明，$\mathscr{A}$ 与 $\mathscr{B}^*$ 可交换.

3. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换．证明，$\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$ 的象空间 $\mathscr{A}^*(V)$ 是 $\mathscr{A}$ 的核 $\operatorname{Ker}\mathscr{A}$ 的正交补.

4. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的可逆线性变换．证明 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$ 也是可逆变换，并且

$$(\mathscr{A}^*)^{-1}=(\mathscr{A}^{-1})^*.$$

5. 设 β 与 γ 是 n 维 Euclid 空间 V 的固定向量．证明，由 $\mathscr{A}(\alpha)=(\alpha,\beta)\gamma$ 所定义的变换 $\mathscr{A}$ 是 V 的的线性变换，其中 $\alpha\in V$．并求 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$.

6. 设 $\mathbb{R}_4[x]$ 是所有次数小于 4 的实系数多项式集合连同内积

$$\big(f(x),g(x)\big)=\int_0^1 f(x)g(x)\,\mathrm{d}x$$

构成的 Euclid 空间，其中 $f(x),g(x)\in\mathbb{R}_4[x]$.

设 $\mathscr{D}$ 是 $\mathbb{R}_4[x]$ 的微商变换．求 $\mathscr{D}$ 的伴随变换 $\mathscr{D}^*$.

7. 设 $\mathbb{R}^{n\times n}$ 是所有 n 阶实方阵集合连同内积 $(X,Y)=\operatorname{Tr}XY^{\mathrm{T}}$ 构成的 Euclid 空间，其中 $X,Y\in\mathbb{R}^{n\times n}$．设 P 是固定的 n 阶可逆方阵，由

$$\mathscr{A}_P(X)=P^{-1}XP$$

所定义的变换 $\mathscr{A}_P$ 显然是 $\mathbb{R}^{n\times n}$ 的线性变换，其中 $X\in\mathbb{R}^{n\times n}$．求 $\mathscr{A}_P$ 的伴随变换 $\mathscr{A}_P^*$.

8. 设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的标准正交基，V 的线性变换 $\mathscr{A}$ 的这组基下的方阵 $A=(a_{ij})_{n\times n}$．证明，对任意 $1\leqslant i,j\leqslant n$，有

$$a_{ij}=\big(\mathscr{A}^*(\alpha_i),\alpha_j\big).$$

§7.4　规范变换

本节讨论 n 维 Euclid 空间 V 的一类重要的线性变换.

定义 7.4.1　如果 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 与它的伴随变换 $\mathscr{A}^*$ 可交换,即

$$\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A},$$

则 $\mathscr{A}$ 称为**规范变换**.

根据定理 7.3.6,如果 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 在 V 的一组基下的方阵为 A,则它的伴随变换 $\mathscr{A}^*$ 在同一组基下的方阵为 A^{T},因此可以引进规范方阵的概念如下.

定义 7.4.2　如果 n 阶实方阵 A 与它的转置 A^{T} 可交换,则 $AA^{\mathrm{T}} = A^{\mathrm{T}}A$,则方阵 A 称为**规范方阵**.

关于规范变换,有

定理 7.4.1　设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换,则下述命题等价.

(1) $\mathscr{A}$ 是规范变换;

(2) 对任意 $\alpha \in V$, $\|\mathscr{A}(\alpha)\| = \|\mathscr{A}^*(\alpha)\|$;

(3) $\mathscr{A}$ 在 V 的标准正交基下的方阵为规范方阵.

证明　(1) $\Longrightarrow$ (2)　对任意 $\alpha \in V$,

$$\|\mathscr{A}(\alpha)\|^2 = \big(\mathscr{A}(\alpha), \mathscr{A}(\alpha)\big) = \big(\alpha, \mathscr{A}^*\mathscr{A}(\alpha)\big).$$

因为 $\mathscr{A}$ 为规范变换,所以 $\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A}$,因此

$$\|\mathscr{A}(\alpha)\|^2 = \big(\alpha, \mathscr{A}\mathscr{A}^*(\alpha)\big) = \big(\mathscr{A}\mathscr{A}^*(\alpha), \alpha\big) = \big(\mathscr{A}^*(\alpha), \mathscr{A}^*(\alpha)\big) = \|\mathscr{A}^*(\alpha)\|^2.$$

(2) $\Longrightarrow$ (3)　设 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 是 V 的标准正交基,且

$$\mathscr{A}(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1, \xi_2, \ldots, \xi_n)A,$$

其中 A 是 n 阶实方阵. 由定理 7.3.5,$\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$ 在这组基下的的方阵为 A^{T},即

$$\mathscr{A}^*(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1, \xi_2, \ldots, \xi_n)A^{\mathrm{T}}.$$

记 $A = (a_{ij})_{n\times n}$,则对任意 $1 \leqslant j \leqslant n$,

$$\mathscr{A}(\xi_j) = \sum_{k=1}^{n} a_{kj}\xi_k, \quad \mathscr{A}^*(\xi_j) = \sum_{\ell=1}^{n} a_{j\ell}\xi_\ell.$$

于是

$$\begin{aligned}\big(\mathscr{A}(\xi_i), \mathscr{A}(\xi_j)\big) &= \left(\sum_{k=1}^{n} a_{ki}\xi_k, \sum_{\ell=1}^{n} a_{\ell j}\xi_\ell\right) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki}a_{\ell j}(\xi_k, \xi_\ell) \\ &= \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki}a_{\ell j}\delta_{k\ell} = \sum_{k=1}^{n} a_{ki}a_{kj}.\end{aligned}$$

记 $A^{\mathrm{T}}A=B=(b_{ij})_{n\times n}$. 上式表明 $b_{ij}=\big(\mathscr{A}(\xi_i),\mathscr{A}(\xi_j)\big)$. 同理,

$$\big(\mathscr{A}^*(\xi_i),\mathscr{A}^*(\xi_j)\big)=\sum_{k=1}^{n}a_{ik}a_{jk}.$$

记 $AA^{\mathrm{T}}=C=(c_{ij})_{n\times n}$. 上式表明, $c_{ij}=\big(\mathscr{A}^*(\xi_i),\mathscr{A}^*(\xi_j)\big)$. 由于

$$\begin{aligned}\big(\mathscr{A}(\xi_i+\xi_j),\mathscr{A}(\xi_i+\xi_j)\big)&=\big(\mathscr{A}(\xi_i)+\mathscr{A}(\xi_j),\mathscr{A}(\xi_i)+\mathscr{A}(\xi_j)\big)\\&=\big(\mathscr{A}(\xi_i),\mathscr{A}(\xi_i)\big)+2\big(\mathscr{A}(\xi_i),\mathscr{A}(\xi_j)\big)+\big(\mathscr{A}(\xi_j),\mathscr{A}(\xi_j)\big),\end{aligned}$$

而由条件 (2),

$$\begin{aligned}\big(\mathscr{A}(\xi_i+\xi_j),\mathscr{A}(\xi_i+\xi_j)\big)&=\big(\mathscr{A}^*(\xi_i+\xi_j),\mathscr{A}^*(\xi_i+\xi_j)\big)\\&=\big(\mathscr{A}^*(\xi_i)+\mathscr{A}^*(\xi_j),\mathscr{A}^*(\xi_i)+\mathscr{A}^*(\xi_j)\big)\\&=\big(\mathscr{A}^*(\xi_i),\mathscr{A}^*(\xi_i)\big)+2\big(\mathscr{A}^*(\xi_i),\mathscr{A}^*(\xi_j)\big)\\&\quad+\big(\mathscr{A}^*(\xi_j),\mathscr{A}^*(\xi_j)\big),\end{aligned}$$

所以

$$\big(\mathscr{A}(\xi_i),\mathscr{A}(\xi_j)\big)=\big(\mathscr{A}^*(\xi_i),\mathscr{A}^*(\xi_j)\big).$$

因此 $b_{ij}=c_{ij}, 1\leqslant i,j\leqslant n$, 即 $A^{\mathrm{T}}A=AA^{\mathrm{T}}$.

(3) $\Longrightarrow$ (1) 由定理 7.3.5,

$$\begin{aligned}\mathscr{A}(\xi_1,\xi_2,\ldots,\xi_n)&=(\xi_1,\xi_2,\ldots,\xi_n)A,\\\mathscr{A}^*(\xi_1,\xi_2,\ldots,\xi_n)&=(\xi_1,\xi_2,\ldots,\xi_n)A^{\mathrm{T}},\end{aligned}$$

其中 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的标准正交基. 因此,

$$\begin{aligned}\mathscr{A}^*\mathscr{A}(\xi_1,\xi_2,\ldots,\xi_n)&=(\xi_1,\xi_2,\ldots,\xi_n)A^{\mathrm{T}}A,\\\mathscr{A}\mathscr{A}^*(\xi_1,\xi_2,\ldots,\xi_n)&=(\xi_1,\xi_2,\ldots,\xi_n)AA^{\mathrm{T}}.\end{aligned}$$

由于方阵 A 是规范的, 所以 $A^{\mathrm{T}}A=AA^{\mathrm{T}}$. 因此 $\mathscr{A}^*\mathscr{A}=\mathscr{A}\mathscr{A}^*$. ■

定理 7.4.2 规范变换 $\mathscr{A}$ 的像空间 $\operatorname{Im}\mathscr{A}$ 的正交补 $(\operatorname{Im}\mathscr{A})^{\perp}=\operatorname{Ker}\mathscr{A}$.

证明 设 $\beta\in(\operatorname{Im}\mathscr{A})^{\perp}$. 则对任意 $\alpha\in V$, $\big(\mathscr{A}(\alpha),\beta\big)=0$. 因此 $\big(\alpha,\mathscr{A}^*(\beta)\big)=0$. 由 α 的任意性得到, $\mathscr{A}^*(\beta)=0$. 由定理 7.4.1, $\mathscr{A}(\beta)=0$, 即 $\beta\in\operatorname{Ker}\mathscr{A}$. 所以 $(\operatorname{Im}\mathscr{A})^{\perp}\subseteq\operatorname{Ker}\mathscr{A}$.

反之, 设 $\beta\in\operatorname{Ker}\mathscr{A}$, 则 $\mathscr{A}(\beta)=0$. 由定理 7.4.1, $\mathscr{A}^*(\beta)=0$. 因此对任意 $\alpha\in V$, $\big(\alpha,\mathscr{A}^*(\beta)\big)=0$. 即得 $\big(\mathscr{A}(\alpha),\beta\big)=0$. 由 α 的任意性, $\beta\in(\operatorname{Im}\mathscr{A})^{\perp}$. 所以 $\operatorname{Ker}\mathscr{A}\subseteq(\operatorname{Im}\mathscr{A})^{\perp}$. ■

注 定理 7.4.2 的逆命题不成立. 请读者自己举例说明之.

定理 7.4.3 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的规范变换, $f(\lambda)$ 与 $g(\lambda)$ 是实系数多项式, 则

(1) $f(\mathscr{A})$ 是规范变换;

(2) 设 $f(\lambda)$ 与 $g(\lambda)$ 互素, 且 $f(\mathscr{A})(\alpha)=0, g(\mathscr{A})(\beta)=0, \alpha,\beta\in V$, 则 $\alpha\perp\beta$.

证明 **(1)** 设多项式 $f(\lambda)=\sum a_j\lambda^j, a_0,a_1,\ldots,a_k\in\mathbb{R}$. 则

$$f(\mathscr{A}) = \sum_{j=0}^{k} a_j \mathscr{A}^j.$$

由定理 7.3.5,

$$\big(f(\mathscr{A})\big)^* = \sum_{j=0}^{k} a_j (\mathscr{A}^*)^j = f(\mathscr{A}^*).$$

因为 $\mathscr{A}$ 是规范变换,所以 $\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A}$. 因此

$$f(\mathscr{A}^*)\mathscr{A} = \sum_{j=0}^{k} a_j (\mathscr{A}^*)^j \mathscr{A} = \sum_{j=0}^{k} a_j \mathscr{A} (\mathscr{A}^*)^j = \mathscr{A} f(\mathscr{A}^*).$$

从而 $f(\mathscr{A}^*)\mathscr{A}^\ell = \mathscr{A}^\ell f(\mathscr{A}^*)$,其中 ℓ 是非负整数. 由此得到,

$$f(\mathscr{A}^*)f(\mathscr{A}) = f(\mathscr{A})f(\mathscr{A}^*).$$

这就证明,$f(\mathscr{A})$ 是规范变换.

(2) 因为 $f(\lambda)$ 与 $g(\lambda)$ 互素,所以存在 $u(\lambda), v(\lambda) \in \mathbb{R}[\lambda]$,使得

$$u(\lambda)f(\lambda) + v(\lambda)g(\lambda) = 1.$$

因此 $u(\mathscr{A})f(\mathscr{A}) + v(\mathscr{A})g(\mathscr{A}) = \mathscr{I}$. 由于 $f(\mathscr{A})(\alpha) = 0, g(\mathscr{A})(\beta) = 0$,所以

$$\alpha = \mathscr{I}(\alpha) = v(\mathscr{A})g(\mathscr{A})(\alpha), \qquad \beta = \mathscr{I}(\beta) = u(\mathscr{A})f(\mathscr{A})(\beta).$$

于是

$$\begin{aligned}(\alpha, \beta) &= \big(v(\mathscr{A})g(\mathscr{A})(\alpha),\ u(\mathscr{A})f(\mathscr{A})(\beta)\big) \\ &= \big(v(\mathscr{A})(\alpha),\ u(\mathscr{A})f(\mathscr{A})g(\mathscr{A})^*(\beta)\big).\end{aligned}$$

由结论 (1),$g(\mathscr{A})$ 是规范的. 由定理 7.4.1,$g(\mathscr{A})^*(\beta) = 0$. 因此 $\alpha \perp \beta$. ■

定理 7.4.4　设 $d(\lambda) = (\lambda - a)^2 + b^2$ 是 n 维 Euclid 空间 V 的规范变换 $\mathscr{A}$ 的最小多项式,其中 a, b 为实数,且 $b > 0$. 则 $n = 2k$,k 为正整数. 并且存在 $\mathscr{A}$ 的 2 维不变子空间 $V_1, V_2, \ldots, V_k$,使得

(1) 当 $i \neq j$ 时,$V_i \perp V_j$,即对任意 $\alpha \in V_i$,$\beta \in V_j$,α 与 β 正交;

(2) $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$;

(3) 对于 $j = 1, 2, \ldots, k$,存在 V_j 的标准正交基 $\{\beta_{2j-1}, \beta_{2j}\}$,使得

$$\mathscr{A}(\beta_{2j-1}, \beta_{2j}) = (\beta_{2j-1}, \beta_{2j})\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

证明　如果 n 是奇数,则 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 是奇次实系数多项式,因此 $\mathscr{A}$ 具有实特征值 λ_0. 于是 λ_0 是 $d(\lambda)$ 的根. 但 $d(\lambda) = (\lambda - a)^2 + b^2$ 只有复根 $a \pm ib$,$b \neq 0$,矛盾. 所以 $n = 2k$. 现在对 k 用归纳法证明其余结论成立.

当 $k = 1$ 时,$n = 2$. 记

$$\mathscr{B} = \frac{1}{b}(\mathscr{A} - a\mathscr{I}).$$

因为 $\mathscr{A}$ 是规范的,所以由定理 7.4.3,$\mathscr{B}$ 也是规范的. 由于 $d(\mathscr{A}) = (\mathscr{A} - a\mathscr{I})^2 + b^2\mathscr{I} = 0$,故 $\mathscr{B}^2 + \mathscr{I} = 0$. 因此 V 中存在单位向量 β_1,使得 $\mathscr{B}(\beta_1) \neq 0$. 记 $\beta_2 = -\mathscr{B}(\beta_1)$. 因为 $\mathscr{B}^2 + \mathscr{I} = 0$,所以 $\mathscr{B}(\beta_2) = -\mathscr{B}^2(\beta_1) = \mathscr{I}(\beta_1) = \beta_1$. 于是

$$\mathscr{A}(\beta_1,\beta_2)=(\beta_1,\beta_2)\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

下面证明,$\{\beta_1,\beta_2\}$ 是 V 的标准正交基.

因为 $\mathscr{B}(\beta_1)+\beta_2=0$,$\mathscr{B}(\beta_2)-\beta_1=0$,所以

$$\begin{aligned}
&\|\mathscr{B}(\beta_1)+\beta_2\|^2+\|\mathscr{B}(\beta_2)-\beta_1\|^2\\
=&\|\mathscr{B}(\beta_1)\|^2+2\big(\mathscr{B}(\beta_1),\beta_2\big)+\|\beta_2\|^2+\|\mathscr{B}(\beta_2)\|^2-2\big(\mathscr{B}(\beta_2),\beta_1\big)+\|\beta_1\|^2
\end{aligned}$$

由于 $\mathscr{B}$ 是规范的,所以对任意 α,$\|\mathscr{B}(\alpha)\|=\|\mathscr{B}^*(\alpha)\|$. 因此

$$\begin{aligned}
&=\|\mathscr{B}^*(\beta_2)\|^2+2\big(\mathscr{B}(\beta_1),\beta_2\big)+\|\beta_1\|^2+\|\mathscr{B}^*(\beta_1)\|^2-2\big(\mathscr{B}(\beta_2),\beta_1\big)+\|\beta_2\|^2\\
&=\|\mathscr{B}^*(\beta_2)\|^2+2\big(\beta_1,\mathscr{B}^*(\beta_2)\big)+\|\beta_1\|^2+\|\mathscr{B}^*(\beta_1)\|^2-2\big(\beta_2,\mathscr{B}^*(\beta_1)\big)+\|\beta_2\|^2\\
&=\|\mathscr{B}^*(\beta_2)+\beta_1\|^2+\|\mathscr{B}^*(\beta_1)-\beta_2\|^2\\
&=0.
\end{aligned}$$

所以 $\mathscr{B}^*(\beta_1)=\beta_2$,$\mathscr{B}^*(\beta_2)=-\beta_1$. 于是

$$\begin{aligned}
(\beta_1,\beta_2)&=\big(-\mathscr{B}^*(\beta_2),\beta_2\big)=-\big(\beta_2,\mathscr{B}(\beta_2)\big)\\
&=-(\beta_2,\beta_1)=-(\beta_1,\beta_2).
\end{aligned}$$

所以 $(\beta_1,\beta_2)=0$,即 $\beta_1\perp\beta_2$. 又

$$\|\beta_2\|^2=(\beta_2,\beta_2)=-\big(\mathscr{B}(\beta_1),\beta_2\big)=-\big(\beta_1,\mathscr{B}^*(\beta_2)\big)=(\beta_1,\beta_1)=\|\beta_1\|^2.$$

由于 β_1 是单位向量,因此 β_2 也是单位向量. 这证明,$\{\beta_1,\beta_2\}$ 是 V 的标准正交基. 所以结论对 $k=1$ 成立.

现在设结论对 $k-1$ 成立. 下面证明结论对 k 成立. 仍记

$$\mathscr{B}=\frac{1}{b}(\mathscr{A}-b\mathscr{I}),$$

则 $\mathscr{B}^2+\mathscr{I}=0$. 于是存在单位向量 $\beta_1\in V$,使得 $\mathscr{B}(\beta_1)\neq0$. 记 $\beta_2=-\mathscr{B}(\beta_1)$. 重复上段证明可知,$\beta_1,\beta_2$ 是正交的单位向量,而且

$$\mathscr{A}(\beta_1,\beta_2)=(\beta_1,\beta_2)\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

V 中由 β_1 与 β_2 生成的子空间记为 V_1. 显然,$\{\beta_1,\beta_2\}$ 是 V_1 的标准正交基,而且 V_1 是 $\mathscr{A}$ 的不变子空间. 由定理 7.2.7,$V=V_1\oplus V_1^{\perp}$.

设 $V_1^{\perp}$ 的标准正交基为 $\{\xi_3,\xi_4,\ldots,\xi_{2k}\}$. 则 $\{\beta_1,\beta_2,\xi_3,\xi_4,\ldots,\xi_{2k}\}$ 是 V 的标准正交基,而且

$$\mathscr{A}(\beta_1,\beta_2,\xi_3,\xi_4,\ldots,\xi_{2k})=(\beta_1,\beta_2,\xi_3,\xi_4,\ldots,\xi_{2k})\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

其中

$$A_{11}=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

由定理 7.3.6,

$$\mathscr{A}^*(\beta_1,\beta_2,\xi_3,\xi_4,\dots,\xi_{2k})=(\beta_1,\beta_2,\xi_3,\xi_4,\dots,\xi_{2k})\begin{pmatrix}A_{11}^{\mathrm{T}} & 0\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}}\end{pmatrix},$$

因此

$$\mathscr{A}\mathscr{A}^*(\beta_1,\beta_2,\xi_3,\xi_4,\dots,\xi_{2k})=(\beta_1,\beta_2,\xi_3,\xi_4,\dots,\xi_{2k})\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix}\begin{pmatrix}A_{11}^{\mathrm{T}} & 0\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}}\end{pmatrix},$$

$$\mathscr{A}^*\mathscr{A}(\beta_1,\beta_2,\xi_3,\xi_4,\dots,\xi_{2k})=(\beta_1,\beta_2,\xi_3,\xi_4,\dots,\xi_{2k})\begin{pmatrix}A_{11}^{\mathrm{T}} & 0\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix}.$$

因为 $\mathscr{A}$ 是规范的，所以

$$\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix}\begin{pmatrix}A_{11}^{\mathrm{T}} & 0\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}A_{11}^{\mathrm{T}} & 0\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix}.$$

由此得到

$$A_{11}A_{11}^{\mathrm{T}}+A_{12}A_{12}^{\mathrm{T}}=A_{11}^{\mathrm{T}}A_{11}.$$

取两端方阵的迹，得到 $A_{12}=0$. 于是

$$\mathscr{A}(\xi_3,\xi_4,\dots,\xi_{2k})=(\xi_3,\xi_4,\dots,\xi_{2k})A_{22}.$$

这表明，$V_1^{\perp}$ 是 $\mathscr{A}$ 的不变子空间. 考虑线性变换 $\mathscr{A}$ 在 $V_1^{\perp}$ 的限制 $\mathscr{A}\big|_{V_1^{\perp}}$.

显然 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 是 $\mathscr{A}\big|_{V_1^{\perp}}$ 的化零多项式，因此 $\mathscr{A}\big|_{V_1^{\perp}}$ 的最小多项式 $\widetilde{d}(\lambda)$ 整除 $d(\lambda)$. 由于 $d(\lambda)$ 是实数域上不可约的多项式，因此 $\widetilde{d}(\lambda)=d(\lambda)$. 于是由归纳假设，存在 $\mathscr{A}\big|_{V_1^{\perp}}$ 的 2 维不变子空间 $V_2,V_3,\dots,V_k$，使得

$$V_1^{\perp}=V_2\oplus V_3\oplus\cdots\oplus V_k,$$

其中 $i\neq j$ 时 $V_i\perp V_j$，而且 V_j 具有标准正交基 $\{\beta_{2j-1},\beta_{2j}\}$，使得

$$\mathscr{A}\big|_{V_1^{\perp}}(\beta_{2j-1},\beta_{2j})=(\beta_{2j-1},\beta_{2j})\begin{pmatrix}a & b\\ -b & a\end{pmatrix},$$

其中 $j=2,3,\dots,k$. 由于 $\mathscr{A}\big|_{V_1^{\perp}}$ 是 $\mathscr{A}$ 在 $V_1^{\perp}$ 上的限制，所以 $V_2,V_2,\dots,V_k$ 是 $\mathscr{A}$ 的不变子空间，而且

$$\mathscr{A}(\beta_{2j-1},\beta_{2j})=(\beta_{2j-1},\beta_{2j})\begin{pmatrix}a & b\\ -b & a\end{pmatrix},$$

其中 $j=1,2,\dots,k$. 于是 $V_1,V_2,\dots,V_k$ 即是所需的 $\mathscr{A}$ 的 2 维不变子空间. ■

定理 7.4.5　设规范变换 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 的全部不可约因子是

$$(\lambda-a_1)^2+b_1^2,\quad (\lambda-a_2)^2+b_2^2,\quad \dots\dots,\quad (\lambda-a_s)^2+b_s^2,$$
$$\lambda-\lambda_{2s+1},\quad \lambda-\lambda_{2s+2},\quad \dots\dots,\quad \lambda-\lambda_{2s+t},$$

其中 $a_1,a_2,\dots,a_s,\lambda_{2s+1},\lambda_{2s+2},\dots,\lambda_{2s+t}$ 是实数，而 $b_1,b_2,\dots,b_s$ 是正数. 则

(1) 每个不可约因子都只能在 $d(\lambda)$ 中出现一次，即

$$d(\lambda)=\left((\lambda-a_1)^2+b_1^2\right)\left((\lambda-a_2)^2+b_2^2\right)\cdots\left((\lambda-a_s)^2+b_s^2\right)(\lambda-\lambda_{2s+1})\cdots(\lambda-\lambda_{2s+t});$$

(2) V 中存在两两正交的 $\mathscr{A}$ 的不变子空间 $W_1,W_2,\dots,W_s,W_{2s+1}\cdots,W_{2s+t}$，使得

$\mathscr{A}$ 在 W_j 上的限制 $\mathscr{A}\big|_{W_j}$ 的最小多项式 $d_j(\lambda)$ 为

$$d_j(\lambda)=\begin{cases}(\lambda-a_j)^2+b_j^2, & 1\leqslant j\leqslant s,\\ \lambda-\lambda_j, & 2s+1\leqslant j\leqslant 2s+t;\end{cases}$$

(3) $V=W_1\oplus W_2\oplus\cdots\oplus W_s\oplus W_{2s+1}\oplus\cdots\oplus W_{2s+t}$.

证明 设

$$d(\lambda)=\left((\lambda-a_1)^2+b_1^2\right)^{e_1}\cdots\left((\lambda-a_s)^2+b_s^2\right)^{e_s}(\lambda-\lambda_{2s+1})^{e_{2s+1}}\cdots(\lambda-\lambda_{2s+t})^{e_{2s+t}},$$

其中 $e_1,\ldots,e_t,e_{2s+1},\ldots,e_{2s+t}$ 是正整数.

设某个 $e_j\geqslant 2$,且记相应的不可约因子为 $p_j(\lambda)$. 记

$$g_j(\lambda)=\frac{d(\lambda)}{p_j^2(\lambda)},$$

即 $d(\lambda)=p_j^2(\lambda)g_j(\lambda)$. 因为 $d(\lambda)$ 是 $\mathscr{A}$ 的最小多项式,所以 $d(\mathscr{A})=p_j^2(\mathscr{A})g_j(\mathscr{A})=0$,即对任意 $\alpha\in V$,

$$p_j^2(\mathscr{A})g_j(\mathscr{A})(\alpha)=0.$$

即 $p_j(\mathscr{A})g_j(\mathscr{A})(\alpha)\in\operatorname{Ker}p_i(\mathscr{A})$,同时 $p_j(\mathscr{A})g_j(\mathscr{A})(\alpha)\in\operatorname{Im}p_j(\mathscr{A})$.

因为 $\mathscr{A}$ 是规范的,所以 $p_j(\mathscr{A})$ 也是规范的. 由定理 7.4.2,

$$\operatorname{Ker}p_i(\mathscr{A})=\left(\operatorname{Im}p_j(\mathscr{A})\right)^{\perp}.$$

所以

$$p_j(\mathscr{A})g_j(\mathscr{A})(\alpha)\in\operatorname{Im}p_j(\mathscr{A})\cap\left(\operatorname{Im}p_j(\mathscr{A})\right)^{\perp}=0.$$

即对任意 $\alpha\in V$,

$$p_j(\mathscr{A})g_j(\mathscr{A})(\alpha)=0.$$

这表明 $p_j(\lambda)g_j(\lambda)$ 是 $\mathscr{A}$ 的化零多项式,其次数显然小于 $d(\lambda)$ 的次数,矛盾. 所以每个不可约因子 $p_j(\lambda)$ 只能在 $d(\lambda)$ 中出现一次,即结论 (1) 成立.

现在记

$$f_j(\lambda)=\frac{d(\lambda)}{p_j(\lambda)},$$

其中当 $1\leqslant j\leqslant s$ 时,$p_j(\lambda)=(\lambda-a_j)^2+b_j^2$; 当 $2s+1\leqslant j\leqslant 2s+t$ 时,$p(\lambda)=\lambda-\lambda_j$.

设 $W_j=\operatorname{Ker}p_j(\mathscr{A})$,且取 $\beta\in W_j$,则 $p_j(\mathscr{A})(\beta)=0$. 因此

$$p_j(\mathscr{A})\big(\mathscr{A}(\beta)\big)=\mathscr{A}\big(p_j(\mathscr{A})(\beta)\big)=0.$$

所以 $\mathscr{A}(\beta)\in W_j$. 即 W_j 是 $\mathscr{A}$ 的不变子空间. 记 $\mathscr{A}$ 在 W_j 上的限制为 $\mathscr{A}_j$. 由于

$$p_j(\mathscr{A}_j)(\beta)=p_j(\mathscr{A})(\beta)=0,\quad \beta\in W_j,$$

所以 $p_j(\lambda)$ 是 $\mathscr{A}_j$ 的化零多项式,因此 $\mathscr{A}_j$ 的最小多项式整除 $p_j(\lambda)$. 但 $p_j(\lambda)$ 是实数域上不可约首一多项式,所以 $p_j(\lambda)$ 即是 $\mathscr{A}_j$ 的最小多项式.

显然当 $i\neq j$ 时,$p_i(\lambda)$ 与 $p_j(\lambda)$ 互素. 并且对任意 $\beta_i\in W_i$,$\beta_j\in W_j$,$p_i(\mathscr{A})(\beta_i)=0$,$p_j(\mathscr{A})(\beta_j)=0$. 由定理 7.4.3,$\beta_i\perp\beta_j$. 因此 $W_1,W_2,\ldots,W_s,W_{2s+1},\ldots,W_{2s+t}$ 是两两正交的子空间.

由于多项式 $f_j(\lambda)$ 互素，所以存在实系数多项式 $\{u_j(\lambda)\}$，使得 $\sum_j u_j(\lambda)f_j(\lambda) = 1$. 即 $\sum_j u_j(\mathscr{A})f_j(\mathscr{A}) = \mathscr{I}$. 于是对任意 $\alpha \in V$，

$$\alpha = \mathscr{I}(\alpha) = \sum_j u_j(\mathscr{A})f_j(\mathscr{A})(\alpha).$$

由于

$$\begin{aligned} p_j(\mathscr{A})\big(u_j(\mathscr{A})f_j(\mathscr{A})\big)(\alpha) &= u_j(\mathscr{A})p_j(\mathscr{A})f_j(\mathscr{A})(\alpha) \\ &= u_j(\mathscr{A})d(\mathscr{A})(\alpha) = 0, \end{aligned}$$

所以 $u_j(\mathscr{A})f_j(\mathscr{A})(\alpha) \in W_j$. 因此

$$V = W_1 + W_2 + \cdots + W_s + W_{2s+1} + \cdots + W_{2s+t}.$$

因为诸 W_j 两两正交，所以

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_s \oplus W_{2s+1} \oplus \cdots \oplus W_{2s+t}.$$

这就证明了结论 (2) 与结论 (3) 成立. ■

下面是关于规范变换的一个主要定理.

定理 7.4.6 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的规范变换，则存在 2 维不变子空间 V_1, $V_2, \ldots, V_s$ 和 1 维不变子空间 $V_{2s+1}, V_{2s+2}, \ldots, V_n$，它们两两正交，且

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_s \oplus V_{2s+1} \oplus \cdots \oplus V_n.$$

并且存在 V_j 的标准正交基 $\{\beta_{2j-1}, \beta_{2j}\}$, $j = 1, 2, \ldots, s$，以及 V_k 的单位向量 β_k, $k = 2s+1, 2s+2, \ldots, n$，使得 $\mathscr{A}$ 在 V 的标准正交基 $\beta_1, \beta_2, \ldots, \beta_n$ 下的方阵是如下的准对角形：

$$\operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \ldots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{2s+1}, \lambda_{2s+2}, \ldots, \lambda_n\right), \tag{7.4.1}$$

其中 $b_1, b_2, \ldots, b_s$ 是正数，$a_1, a_2, \ldots, a_s, \lambda_{2s+1}, \lambda_{2s+2}, \ldots, \lambda_n$ 是实数，并且 $a_1 \pm \mathrm{i}b_1, a_2 \pm \mathrm{i}b_2, \ldots, a_s \pm \mathrm{i}b_s, \lambda_{2s+1}, \lambda_{2s+2}, \ldots, \lambda_n$ 是 $\mathscr{A}$ 的全部特征值.

证明 根据定理 7.4.5，可设 $\mathscr{A}$ 的最小多项式 $d(\lambda)$ 为

$$d(\lambda) = \big((\lambda - \widetilde{a}_1)^2 + \widetilde{b}_1^2\big)\big((\lambda - \widetilde{a}_2)^2 + \widetilde{b}_2^2\big)\cdots\big((\lambda - \widetilde{a}_\ell)^2 + \widetilde{b}_\ell^2\big)(\lambda - \widetilde{\lambda}_{2\ell+1})\cdots(\lambda - \widetilde{\lambda}_{2\ell+t}),$$

而且

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_\ell \oplus W_{2\ell+1} \oplus \cdots \oplus W_{2\ell+t}.$$

其中 W_j 是 $\mathscr{A}$ 的不变子空间，显然 $\mathscr{A}|_{W_j}$ 仍是规范的.

当 $1 \leqslant j \leqslant \ell$ 时，$\mathscr{A}$ 在 W_j 上的限制 $\mathscr{A}|_{W_j}$ 的最小多项式是 $(\lambda - \widetilde{a}_j)^2 + \widetilde{b}_j^2$. 因此由定理 7.4.4，

$$W_j = V_{j1} \oplus V_{j2} \oplus \cdots \oplus V_{jk_j},$$

其中 $V_{j1}, V_{j2}, \ldots, V_{jk_j}$ 是两两正交的 $\mathscr{A}|_{W_j}$ 的不变子空间，而且存在 V_{ji} 的标准正交基

$$\{\beta_{2i-1}^{(j)}, \beta_{2i}^{(j)}\}.$$

此时，$\mathscr{A}|_{W_j}$ 在 W_j 的标准正交基

$$\{\beta_1^{(j)},\beta_2^{(j)},\ldots,\beta_{2k_j-1}^{(j)},\beta_{2k_j}^{(j)}\}$$

下的方阵是

$$\operatorname{diag}\Big(\underbrace{\begin{pmatrix}\widetilde{a}_j & \widetilde{b}_j\\ -\widetilde{b}_j & \widetilde{a}_j\end{pmatrix},\ldots,\begin{pmatrix}\widetilde{a}_j & \widetilde{b}_j\\ -\widetilde{b}_j & \widetilde{a}_j\end{pmatrix}}_{k_j\text{ 个}}\Big).$$

当 $2\ell+1\leqslant j\leqslant 2\ell+t$ 时,$\mathscr{A}$ 在 W_j 上的限制 $\mathscr{A}|_{W_j}$ 的最小多项式为 $\lambda-\widetilde{\lambda}_j$. 因此 $\mathscr{A}|_{W_j}=\widetilde{\lambda}_j\mathscr{I}_j$,其中 $\mathscr{I}_j$ 是 W_j 的单位变换. 所以 $\mathscr{A}|_{W_j}$ 的标准正交基

$$\{\beta_1^{(j)},\beta_2^{(j)},\ldots,\beta_{k_j}^{(j)}\}$$

下的方阵是纯量方阵 $\widetilde{\lambda}_j\mathscr{I}_{(k_j)}$,并且 W_j 可以表为基向量 $\beta_i^{(j)}$ 生成的一维子空间 V_{ji} 的直和,即

$$W_j=V_{j1}\oplus V_{j2}\oplus\cdots\oplus V_{jk_j}.$$

显然,子空间 V_{ji} 是 $\mathscr{A}$ 的不变子空间,$i=1,2,\ldots,k_j$,$j=1,2,\ldots,\ell,2\ell+1,\ldots,2\ell+t$,而且

$$V=\bigoplus_{j=1}^{\ell}\bigoplus_{i=1}^{k_j}V_{ji}\oplus\bigoplus_{j=2\ell+1}^{2\ell+t}\bigoplus_{i=1}^{k_j}V_{ji}.$$

同时,

$$\{\beta_1^{(1)},\ldots,\beta_{2k_1}^{(1)},\ldots,\beta_1^{(\ell)},\ldots,\beta_{k_\ell}^{(\ell)},\beta_1^{(2\ell+1)},\ldots,\beta_{k_{2\ell+1}}^{(2\ell+1)},\ldots,\beta_1^{(2\ell+t)},\ldots,\beta_{k_{2\ell+t}}^{(2\ell+t)}\}$$

是 V 的标准正交基,而且规范变换 $\mathscr{A}$ 在这组基下的方阵即是方阵 (7.4.1). ■

定理 7.4.6 的矩阵形式在下面定理中给出.

定理 7.4.7 设 $a_1\pm\mathrm{i}b_1,a_2\pm\mathrm{i}b_2,\ldots,a_s\pm\mathrm{i}b_s,\lambda_{2s+1},\lambda_{2s+2},\ldots,\lambda_n$ 是 n 阶实规范方阵 A 的全部特征值,其中 $a_1,a_2,\ldots,a_s,\lambda_{2s+1},\lambda_{2s+2},\ldots,\lambda_n$ 是实数,$b_1,b_2,\ldots,b_s$ 是正数. 则方阵 A 正交相似于如下的准对角形:

$$\operatorname{diag}\left(\begin{pmatrix}a_1 & b_1\\ -b_1 & a_1\end{pmatrix},\ldots,\begin{pmatrix}a_s & b_s\\ -b_s & a_s\end{pmatrix},\lambda_{2s+1},\lambda_{2s+2},\ldots,\lambda_n\right),$$

并且规范方阵的特征值是规范方阵在正交相似下的全系不变量.

证明 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的标准正交基. 则由

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A$$

便确定 V 的一个线性变换 $\mathscr{A}$. 由于方阵 A 是规范的,所以由定理 7.4.1,$\mathscr{A}$ 是规范变换. 由定理 7.4.5,$\mathscr{A}$ 的最小多项式 $d(\lambda)$ 为

$$d(\lambda)=\big((\lambda-\widetilde{a}_1)^2+\widetilde{b}_1^2\big)\cdots\big((\lambda-\widetilde{a}_\ell)^2+\widetilde{b}_\ell^2\big)(\lambda-\widetilde{\lambda}_{2\ell+1})\cdots(\lambda-\widetilde{\lambda}_{2\ell+t}).$$

即

$$\begin{aligned}d(\lambda)=&(\lambda-\widetilde{a}_1+\mathrm{i}\widetilde{b}_1)(\lambda-\widetilde{a}_1-\mathrm{i}\widetilde{b}_1)\cdots(\lambda-\widetilde{a}_\ell+\mathrm{i}\widetilde{b}_\ell)(\lambda-\widetilde{a}_\ell-\mathrm{i}\widetilde{b}_\ell)\\&\times(\lambda-\widetilde{\lambda}_{2\ell+1})\cdots(\lambda-\widetilde{\lambda}_{2\ell+t}).\end{aligned}$$

其中一次因子两两不同. 由于 $\mathscr{A}$ 的特征值一定是 $d(\lambda)$ 的根,因此 $\widetilde{a}_1\pm\mathrm{i}\widetilde{b}_1,\widetilde{a}_2\pm$

$\mathrm{i}\widetilde{b}_2,\ldots,\widetilde{a}_\ell \pm \mathrm{i}\widetilde{b}_\ell,\widetilde{\lambda}_{2\ell+1},\ldots,\widetilde{\lambda}_{2\ell+t}$ 是 $\mathscr{A}$ 的全部不同的特征值. 所以可设

$$\begin{aligned} \widetilde{a}_j \pm \mathrm{i}\widetilde{b}_j &= a_j \pm \mathrm{i}b_j, && j=1,2,\ldots,\ell,\\ \widetilde{\lambda}_{2\ell+j} &= \lambda_{2s+j}, && j=1,2,\ldots,t. \end{aligned}$$

即

$$d(\lambda)=\big((\lambda+a_1)^2+\mathrm{i}b_1^2\big)\cdots\big((\lambda+a_\ell)^2+\mathrm{i}b_\ell^2\big)(\lambda-\lambda_{2s+1})\cdots(\lambda-\lambda_{2s+t}).$$

于是由定理 7.4.6,存在 V 的标准正交基 $\{\beta_1,\beta_2,\ldots,\beta_n\}$,使得 $\mathscr{A}$ 在这组基下的方阵即为 (7.4.1). 由于线性变换 $\mathscr{A}$ 在不同的标准正交基下的方阵是正交相似的,因此规范方阵 $\mathscr{A}$ 相似于方阵 (7.4.1).

显然,相似方阵的特征值相同,因此正交相似的规范方阵的特征值相同. 反之设规范方阵 A 与 B 的特征值相同,则方阵 A 与 B 都正交相似于同一个形如 (7.4.1) 的准对角方阵 C. 所以 A 与 B 正交相似.

这就证明,规范方阵的特征值是规范方阵在正交相似下的全系不变量. ■

例 7.4.1 设 n 阶实规范方阵 A 与 n 阶实方阵 B 可交换. 证明方阵 A 与 B^{T} 可交换.

证法 1 欲证 $AB^{\mathrm{T}}=B^{\mathrm{T}}A$,即 $AB^{\mathrm{T}}-B^{\mathrm{T}}A=0$.

由例 7.1.3 可知,所有 n 阶实方阵的集合 $\mathbb{R}^{n\times n}$ 连同内积 $(X,Y)=\operatorname{Tr}XY^{\mathrm{T}}$ 构成一个 Euclid 空间,其中 $X,Y\in\mathbb{R}^{n\times n}$.

于是方阵 $X\in\mathbb{R}^{n\times n}$ 的范数 $\|X\|^2=\operatorname{Tr}XX^{\mathrm{T}}$. 由内积的恒正性可知,方阵 X 为零方阵的必要且充分条件是 $\|X\|^2=\operatorname{Tr}XX^{\mathrm{T}}=0$. 因此欲证的结论即化为

$$\operatorname{Tr}(AB^{\mathrm{T}}-B^{\mathrm{T}}A)(AB^{\mathrm{T}}-B^{\mathrm{T}}A)^{\mathrm{T}}=0.$$

由于

$$\begin{aligned} &\operatorname{Tr}(AB^{\mathrm{T}}-B^{\mathrm{T}}A)(AB^{\mathrm{T}}-B^{\mathrm{T}}A)^{\mathrm{T}}=\operatorname{Tr}(AB^{\mathrm{T}}-B^{\mathrm{T}}A)(BA^{\mathrm{T}}-A^{\mathrm{T}}B)\\ =\ &\operatorname{Tr}(AB^{\mathrm{T}}BA^{\mathrm{T}}-AB^{\mathrm{T}}A^{\mathrm{T}}B-B^{\mathrm{T}}ABA^{\mathrm{T}}+B^{\mathrm{T}}AA^{\mathrm{T}}B),\end{aligned}$$

因为 $\operatorname{Tr}X$ 是 $\mathbb{R}^{n\times n}$ 的线性函数,所以

$$=\operatorname{Tr}AB^{\mathrm{T}}BA^{\mathrm{T}}-\operatorname{Tr}AB^{\mathrm{T}}A^{\mathrm{T}}B-\operatorname{Tr}B^{\mathrm{T}}ABA^{\mathrm{T}}+\operatorname{Tr}B^{\mathrm{T}}AA^{\mathrm{T}}B$$

因为 $\operatorname{Tr}XY=\operatorname{Tr}YX$,所以

$$=\operatorname{Tr}A^{\mathrm{T}}AB^{\mathrm{T}}B-\operatorname{Tr}BAB^{\mathrm{T}}A^{\mathrm{T}}-\operatorname{Tr}B^{\mathrm{T}}ABA^{\mathrm{T}}+\operatorname{Tr}AA^{\mathrm{T}}BB^{\mathrm{T}}$$

由于 $AB=BA$,所以

$$=\operatorname{Tr}A^{\mathrm{T}}AB^{\mathrm{T}}B-\operatorname{Tr}ABB^{\mathrm{T}}A^{\mathrm{T}}-\operatorname{Tr}B^{\mathrm{T}}BAA^{\mathrm{T}}+\operatorname{Tr}BB^{\mathrm{T}}AA^{\mathrm{T}}$$

因为方阵 A 是规范的,所以 $AA^{\mathrm{T}}=A^{\mathrm{T}}A$,因此

$$=\operatorname{Tr}AA^{\mathrm{T}}B^{\mathrm{T}}B-\operatorname{Tr}ABB^{\mathrm{T}}A^{\mathrm{T}}-\operatorname{Tr}B^{\mathrm{T}}BAA^{\mathrm{T}}+\operatorname{Tr}BB^{\mathrm{T}}A^{\mathrm{T}}A$$

再由 $\operatorname{Tr}XY=\operatorname{Tr}YX$ 得到,

$$\begin{aligned} &=\operatorname{Tr}B^{\mathrm{T}}BAA^{\mathrm{T}}-\operatorname{Tr}BB^{\mathrm{T}}A^{\mathrm{T}}A-\operatorname{Tr}B^{\mathrm{T}}BAA^{\mathrm{T}}+\operatorname{Tr}BB^{\mathrm{T}}A^{\mathrm{T}}A\\ &=0.\end{aligned}$$

因此 $AB^{\mathrm{T}} = B^{\mathrm{T}}A$,即方阵 A 与 B^{T} 可交换 ■

证法 2 设方阵 A 的全部特征值为 $a_1 \pm \mathrm{i}b_1, a_2 \pm \mathrm{i}b_2, \dots, a_s \pm \mathrm{i}b_s, \lambda_{2s+1}, \lambda_{2s+2}, \dots, \lambda_{2s+t}$,其中 $a_1, a_2, \dots, a_s, \lambda_{2s+1}, \dots, \lambda_{2s+t}$ 是实数,$b_1, b_2, \dots, b_s$ 是正数,而且相应的代数重数为 $e_1, e_2, \dots, e_s, e_{2s+1}, \dots, e_{2s+t}$.

由定理 7.4.7,存在 n 阶实正交方阵 O,使得

$$O^{\mathrm{T}}AO = \mathrm{diag}\big(D_1, D_2, \dots, D_s, D_{2s+1}, \dots, D_{2s+t}\big) = D.$$

其中当 $1 \leqslant j \leqslant s$ 时,

$$D_j = \mathrm{diag}\left(\underbrace{\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, \dots, \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}}_{e_j\text{ 个}}\right);$$

当 $2s+1 \leqslant j \leqslant 2s+t$ 时,$D_j = \lambda_j I_{(e_j)}$.

由 $AB = BA$ 得到,

$$(O^{\mathrm{T}}AO)(O^{\mathrm{T}}BO) = (O^{\mathrm{T}}BO)(O^{\mathrm{T}}AO).$$

记 $\widetilde{B} = O^{\mathrm{T}}BO$,则 $D\widetilde{B} = \widetilde{B}D$. 将方阵 $\widetilde{B}$ 按方阵 D 的分块方式分块,并记 $\widetilde{B} = (B_{ij})$. 于是由 $\widetilde{B}D = D\widetilde{B}$ 得到,

$$B_{ij}D_j = D_iB_{ij}.$$

当 $i \neq j$ 时,方阵 D_i 和 D_j 没有公共的特征值,因此由例 6.6.8,$B_{ij} = 0$. 于是

$$\widetilde{B} = \mathrm{diag}\big(B_1, B_2, \dots, B_s, B_{2s+1}, \dots, B_{2s+t}\big) = D,$$

并且 $B_jD_j = D_jB_j$. 当 $1 \leqslant j \leqslant s$ 时,

$$B_j\,\mathrm{diag}\left(\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, \dots, \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}\right) = \mathrm{diag}\left(\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, \dots, \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}\right)B_j.$$

将 B_j 按方阵 D_j 的分块方式分块为 $B_j = (C_{k\ell})$,其中 $C_{k\ell}$ 是 2 阶方阵. 由上式得到

$$C_{k\ell}\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}C_{k\ell}.$$

记

$$C_{k\ell} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

则

$$\begin{aligned} aa_j - bb_j &= aa_j + cb_j, & ab_j + ba_j &= ba_j + db_j, \\ ca_j - db_j &= -ab_j + ca_j, & cb_j + da_j &= -bb_j + da_j. \end{aligned}$$

因此,$c = -b$,$d = a$,即

$$C_{k\ell}^{\mathrm{T}}\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}C_{k\ell}^{\mathrm{T}},$$

其中 $1 \leqslant k, \ell \leqslant e_j$,即对于 $1 \leqslant j \leqslant s$,有 $B_j^{\mathrm{T}}D_j = D_jB_j^{\mathrm{T}}$.

当 $2s+1 \leqslant j \leqslant 2s+t$ 时,$D_j = \lambda_j I_{(e_j)}$,因此 $B_j^{\mathrm{T}}D_j = D_jB_j^{\mathrm{T}}$. 所以 $\widetilde{B}^{\mathrm{T}}D = D\widetilde{B}^{\mathrm{T}}$. 即

$$O^{\mathrm{T}}B^{\mathrm{T}}O\cdot O^{\mathrm{T}}AO=O^{\mathrm{T}}AO\cdot O^{\mathrm{T}}B^{\mathrm{T}}O.$$

因此 $B^{\mathrm{T}}A=AB^{\mathrm{T}}$. ■

习 题 7.4

1. 举例说明,方阵 A 本身不是规范的,但它的平方 A^2 却是规范的.

2. 证明,规范方阵 A 与 B 正交相似的充分必要条件是,方阵 A 与 B 相似.

3. 方阵 A 与 $A^{\mathrm{T}}A$ 可交换,方阵 A 是否一定是规范的?

4. 证明,一组两两可交换的规范方阵可以同时正交相似于准对角形. 即设 I 是下标集合,规范方阵集合 $\{A_\nu \mid \nu\in I\}$ 满足: 对任意 $\mu,\nu\in I$,$A_\nu A_\mu=A_\mu A_\nu$,则存在正交方阵 O,使得方阵 $O^{\mathrm{T}}A_\nu O$ 为定理 7.4.6 中准对角形 (7.4.1),其中 $\nu\in I$.

5. 证明,n 阶实方阵 A 为规范的当且仅当存在实系数多项式 $f(\lambda)$,使得 $\mathscr{A}^*=f(\mathscr{A})$.

6. 设 $\alpha=\beta+\mathrm{i}\gamma$ 是实规范方阵 A 的属于特征值 λ 的特征向量,β 与 γ 是实向量. 证明,

(1) α 是 A^{T} 属于特征值 $\bar{\lambda}$ 的特征向量; (2) 当 $\lambda\notin\mathbb{R}$ 时,β 与 γ 正交且范数相等.

§7.5 正交变换

本节讨论 n 维 Euclid 空间 V 的一类重要的线性变换,即正交变换. 它是规范变换的特殊情形.

定义 7.5.1 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换. 如果对任意 $\alpha\in V$,

$$\|\mathscr{A}(\alpha)\|=\|\alpha\|,$$

则 $\mathscr{A}$ 称为**正交变换**. 简单地说,保持向量范数不变的线性变换称为正交变换.

显然,n 维 Euclid 空间 V 的单位变换 $\mathscr{I}$ 是正交变换.

设 $U^\perp$ 是 n 维 Euclid 空间 V 的子空间 U 的正交补,则 $V=U\oplus U^\perp$. 即对任意 $\alpha\in V$ 都可以表为 $\alpha=\xi+\eta$,其中 $\xi\in U$,$\eta\in U^\perp$. 定义 V 的变换 $\mathscr{A}$ 为:

$$\mathscr{A}(\alpha)=\xi-\eta.$$

于是 $\mathscr{A}$ 是 V 的线性变换. 由于

$$\begin{aligned}\big(\mathscr{A}(\alpha),\mathscr{A}(\alpha)\big)&=(\xi-\eta,\xi-\eta)=(\xi,\xi)-2(\xi,\eta)+(\eta,\eta)=(\xi,\xi)+(\eta,\eta);\\(\alpha,\alpha)&=(\xi+\eta,\xi+\eta)=(\xi,\xi)+2(\xi,\eta)+(\eta,\eta)=(\xi,\xi)+(\eta,\eta),\end{aligned}$$

所以 $\|\mathscr{A}(\alpha)\|=\|\alpha\|$. 因此 $\mathscr{A}$ 是正交变换. 这个变换称为 V 关于子空间 U 的**反射**.

关于正交变换,有

定理 7.5.1 设 $\mathscr{A}$ 是 n 维 Euclid 空间的线性变换. 则下述命题等价:

(1) $\mathscr{A}$ 是正交变换;

(2) $\mathscr{A}$ 是保内积的,即对任意 $\alpha,\beta\in V$,$\big(\mathscr{A}(\alpha),\mathscr{A}(\beta)\big)=(\alpha,\beta)$;

(3) $\mathscr{A}$ 把 V 的标准正交基变为标准正交基,即设 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的标准正交基,则 $\{\mathscr{A}(\xi_1),\mathscr{A}(\xi_2),\ldots,\mathscr{A}(\xi_n)\}$ 也是 V 的标准正交基;

(4) $\mathscr{A}$ 在 V 的标准正交基下的方阵是正交方阵；

(5) $\mathscr{A}$ 是规范变换，而且 $\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A} = \mathscr{I}$.

证明 **(1) $\Longrightarrow$ (2)** 设 $\alpha, \beta \in V$. 因为 $\mathscr{A}$ 是正交变换，所以

$$\|\mathscr{A}(\alpha+\beta)\| = \|\alpha+\beta\|,$$

即

$$\big(\mathscr{A}(\alpha+\beta), \mathscr{A}(\alpha+\beta)\big) = (\alpha+\beta, \alpha+\beta).$$

于是，

$$\begin{aligned}\big(\mathscr{A}(\alpha+\beta), \mathscr{A}(\alpha+\beta)\big) &= \big(\mathscr{A}(\alpha)+\mathscr{A}(\beta), \mathscr{A}(\alpha)+\mathscr{A}(\beta)\big)\\ &= \big(\mathscr{A}(\alpha), \mathscr{A}(\alpha)\big) + 2\big(\mathscr{A}(\alpha), \mathscr{A}(\beta)\big) + \big(\mathscr{A}(\beta), \mathscr{A}(\beta)\big)\\ &= (\alpha+\beta, \alpha+\beta) = (\alpha, \alpha) + 2(\alpha, \beta) + (\beta, \beta).\end{aligned}$$

因为 $\mathscr{A}$ 是正交变换，所以 $\big(\mathscr{A}(\alpha), \mathscr{A}(\alpha)\big) = (\alpha, \alpha)$，$\big(\mathscr{A}(\beta), \mathscr{A}(\beta)\big) = (\beta, \beta)$. 因此

$$\big(\mathscr{A}(\alpha), \mathscr{A}(\beta)\big) = (\alpha, \beta),$$

即 $\mathscr{A}$ 保内积.

(2) $\Longrightarrow$ (3) 设 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 是 V 的标准正交基，则 $(\xi_i, \xi_j) = \delta_{ij}$. 因为 $\mathscr{A}$ 保内积，所以 $\big(\mathscr{A}(\xi_i), \mathscr{A}(\xi_j)\big) = \delta_{ij}$，即 $\{\mathscr{A}(\xi_1), \mathscr{A}(\xi_2), \ldots, \mathscr{A}(\xi_n)\}$ 是 V 的标准正交基.

(3) $\Longrightarrow$ (4) 设 $\mathscr{A}$ 在 V 的标准正交基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 下的方阵为 A，即

$$\mathscr{A}(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1, \xi_2, \ldots, \xi_n)A,$$

其中 $A = (a_{ij})_{n\times n}$. 则对任意 $j = 1, 2, \ldots, n$，有

$$\mathscr{A}(\xi_j) = \sum_{k=1}^{n} a_{kj}\xi_k.$$

所以

$$\begin{aligned}\big(\mathscr{A}(\xi_i), \mathscr{A}(\xi_j)\big) &= \Big(\sum_{k=1}^{n} a_{ki}\xi_k, \sum_{\ell=1}^{n} a_{\ell j}\xi_\ell\Big) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki}a_{\ell j}(\xi_k, \xi_\ell)\\ &= \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki}a_{\ell j}\delta_{k\ell} = \sum_{k=1}^{n} a_{ki}a_{kj}.\end{aligned}$$

因为 $\{\mathscr{A}(\xi_1), \mathscr{A}(\xi_2), \ldots, \mathscr{A}(\xi_n)\}$ 也是标准正交基，所以 $\big(\mathscr{A}(\xi_i), \mathscr{A}(\xi_j)\big) = \delta_{ij}$. 因此

$$\sum_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij}.$$

这表明，$A^{\mathrm{T}}A = I_{(n)}$. 从而 $A^{\mathrm{T}}A = I_{(n)} = AA^{\mathrm{T}}$，即 A 是正交方阵.

(4) $\Longrightarrow$ (5) 设 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 是 V 的标准正交基，且

$$\mathscr{A}(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1, \xi_2, \ldots, \xi_n)A,$$

其中 A 是 n 阶实正交方阵. 则

$$\mathscr{A}^*(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1, \xi_2, \ldots, \xi_n)A^{\mathrm{T}},$$

其中 A^* 是 $\mathscr{A}$ 的伴随变换. 于是

$$\mathscr{A}\mathscr{A}^*(\xi_1,\xi_2,\ldots,\xi_n)=(\xi_1,\xi_2,\ldots,\xi_n)AA^{\mathrm{T}}=(\xi_1,\xi_2,\ldots,\xi_n)I_{(n)},$$
$$\mathscr{A}^*\mathscr{A}(\xi_1,\xi_2,\ldots,\xi_n)=(\xi_1,\xi_2,\ldots,\xi_n)A^{\mathrm{T}}A=(\xi_1,\xi_2,\ldots,\xi_n)I_{(n)}.$$

因此 $\mathscr{A}^*\mathscr{A}=\mathscr{A}\mathscr{A}^*=\mathscr{I}$.

(5) $\Longrightarrow$ (1)　因为 $\mathscr{A}^*\mathscr{A}=\mathscr{A}\mathscr{A}^*=\mathscr{I}$,所以对任意 $\alpha\in V$,

$$\big(\mathscr{A}(\alpha),\mathscr{A}(\alpha)\big)=\big(\alpha,\mathscr{A}^*\mathscr{A}(\alpha)\big)=(\alpha,\mathscr{I}(\alpha))=(\alpha,\alpha),$$

即 $\|\mathscr{A}(\alpha)\|=\|\alpha\|$. ■

由定理 7.5.1 直接得到,

定理 7.5.2　**(1)** 设线性变换 $\mathscr{A}$ 是正交变换,则 $\mathscr{A}$ 是可逆变换,而且它的逆变换 $\mathscr{A}^{-1}=\mathscr{A}^*$ 也是正交变换;

(2) 设 $\mathscr{A}$ 和 $\mathscr{B}$ 是正交变换,则乘积 $\mathscr{A}\mathscr{B}$ 也是正交变换.

证明　(1) 由定理 7.5.1,对正交变换 $\mathscr{A}$,有 $\mathscr{A}\mathscr{A}^*=\mathscr{I}=\mathscr{A}^*\mathscr{A}$,所以 $\mathscr{A}$ 是可逆的,且 $\mathscr{A}^{-1}=\mathscr{A}^*$. 由于

$$\mathscr{A}^*(\mathscr{A}^*)^*=\mathscr{A}^*\mathscr{A}=\mathscr{I},\quad (\mathscr{A}^*)^*\mathscr{A}^*=\mathscr{A}\mathscr{A}^*=\mathscr{I},$$

所以 $\mathscr{A}^*$ 是正交变换.

(2) 因为 $\mathscr{A}$ 与 $\mathscr{B}$ 是正交变换,所以 $\mathscr{A}\mathscr{A}^*=\mathscr{A}^*\mathscr{A}=\mathscr{I}$,$\mathscr{B}\mathscr{B}^*=\mathscr{B}^*\mathscr{B}=\mathscr{I}$. 因此

$$(\mathscr{A}\mathscr{B})^*\mathscr{A}\mathscr{B}=\mathscr{B}^*(\mathscr{A}^*\mathscr{A})\mathscr{B}=\mathscr{B}^*\mathscr{B}=\mathscr{I}.$$

同理 $(\mathscr{A}\mathscr{B})(\mathscr{A}\mathscr{B})^*=\mathscr{I}$. 即 $\mathscr{A}\mathscr{B}$ 是正交变换. ■

n 维 Euclid 空间 V 的所有正交变换的集合记为 $\mathrm{O}_n(\mathbb{R})$. 在集合 $\mathrm{O}_n(\mathbb{R})$ 中可以按照线性变换的乘法定义两个正交变换的乘积. 由定理 7.5.2,正交变换的乘积仍是正交变换,也就是说,集合 $\mathrm{O}_n(\mathbb{R})$ 在正交变换的乘法下是封闭的. 由于线性变换的乘法满足结合律,所以正交变换的乘法当然也满足结合律; 由于单位变换是正交变换,所以 $\mathrm{O}_n(\mathbb{R})$ 具有单位元 $\mathscr{I}$; 由定理 7.5.2 可知,如果变换 $\mathscr{A}\in\mathrm{O}_n(\mathbb{R})$,则 $\mathscr{A}^{-1}\in\mathrm{O}_n(\mathbb{R})$.

因为 $\mathrm{O}_n(\mathbb{R})$ 对正交变换的乘法封闭,满足结合律,具有单位元,并且每个变换都具有逆变换,所以通常把集合 $\mathrm{O}_n(\mathbb{R})$ 称为 **n 阶实正交群**. 而研究 n 维 Euclid 空间 V 中的几何对象(如超二次曲面)在 n 阶实正交群 $\mathrm{O}_n(\mathbb{R})$ 下的几何不变性质的学科即称为 **Euclid 几何学**.

研究 n 阶实正交群的结构是所谓典型群的一个重要内容. 由兴趣的读者可参阅我国著名数学家华罗庚和万哲先所著的专著《典型群》(上海科学技术出版社,1963 年版).

由定理 7.5.1 可知,正交变换是特殊类型的规范变换. 所以关于规范变换的结论都可以移到正交变换. 这里不再赘述.

定理 7.5.3　正交变换 $\mathscr{A}$ 的特征值 λ_0 的模为 1.

证明 设 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 是 n 维 Euclid 空间 V 的标准正交基. 由定理 7.5.1, $\mathscr{A}$ 在这组基下的方阵 A 是正交方阵. 显然 λ_0 是 A 的特征值. 由例 5.9.1, A 的特征值 λ_0 的模为 1. ■

定理 7.5.4 设 n 维 Euclid 空间 V 的正交变换 $\mathscr{A}$ 的全部特征值为

$$\mathrm{e}^{\mathrm{i}\theta_1},\mathrm{e}^{-\mathrm{i}\theta_1},\dots,\mathrm{e}^{\mathrm{i}\theta_s},\mathrm{e}^{-\mathrm{i}\theta_s},\underbrace{1,\dots,1}_{t\text{个}},\underbrace{-1,\dots,-1}_{n-2s-t\text{个}},$$

其中 $0<\theta_1\leqslant\cdots\leqslant\theta_s<\pi$. 则存在 V 的标准正交基 $\{\xi_1,\xi_2,\dots,\xi_n\}$, 使得 $\mathscr{A}$ 在这组基下的方阵为如下的准对角形:

$$\mathrm{diag}\left(\begin{pmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{pmatrix},\dots,\begin{pmatrix}\cos\theta_s & \sin\theta_s\\ -\sin\theta_s & \cos\theta_s\end{pmatrix},\underbrace{1,\dots,1}_{t\text{个}},\underbrace{-1,\dots,-1}_{n-2s-t\text{个}}\right). \tag{7.5.1}$$

证明 由定理 7.5.1, $\mathscr{A}$ 是规范变换. 记

$$\mathrm{e}^{\mathrm{i}\theta_j}=\cos\theta_j+\mathrm{i}\sin\theta_j.$$

则规范变换 $\mathscr{A}$ 的特征值为

$$\cos\theta_1\pm\mathrm{i}\sin\theta_1,\dots,\cos\theta_s\pm\mathrm{i}\sin\theta_s,\underbrace{1,\dots,1}_{t\text{个}},\underbrace{-1,\dots,-1}_{n-2s-t\text{个}},$$

其中 $\sin\theta_j\neq0$. 由定理 7.4.6, 存在 V 的标准正交基, 使得 $\mathscr{A}$ 在这组基下的方阵即为 (7.5.1). ■

定理 7.5.4 的矩阵形式是

定理 7.5.5 设 n 阶实正交方阵 O 的全部特征值是

$$\mathrm{e}^{\mathrm{i}\theta_1},\mathrm{e}^{-\mathrm{i}\theta_1},\dots,\mathrm{e}^{\mathrm{i}\theta_s},\mathrm{e}^{-\mathrm{i}\theta_s},\underbrace{1,\dots,1}_{t\text{个}},\underbrace{-1,\dots,-1}_{n-2s-t\text{个}},$$

其中 $0<\theta_1\leqslant\cdots\leqslant\theta_s<\pi$. 则方阵 O 正交相似于如下的准对角形:

$$\mathrm{diag}\left(\begin{pmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{pmatrix},\dots,\begin{pmatrix}\cos\theta_s & \sin\theta_s\\ -\sin\theta_s & \cos\theta_s\end{pmatrix},\underbrace{1,\dots,1}_{t\text{个}},\underbrace{-1,\dots,-1}_{n-2s-t\text{个}}\right).$$

并且正交方阵的特征值是正交方阵在正交相似下的全系不变量.

证明 设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的标准正交基. 则由

$$\mathscr{A}(\alpha_1,\alpha_2,\dots,\alpha_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)O$$

便确定 V 的一个线性变换 $\mathscr{A}$.

由定理 7.5.1, $\mathscr{A}$ 是正交变换. 由定理 7.5.4, 存在 V 的标准正交基 $\{\beta_1,\beta_2,\dots,\beta_n\}$, 使得 $\mathscr{A}$ 在这组基下的方阵即为 (7.5.1). 由于同一个线性变换 $\mathscr{A}$ 在不同的标准正交基下的方阵是正交相似的, 所以方阵 A 正交相似于准对角形 (7.5.1).

由于正交相似的方阵一定相似, 而相似的方阵的特征值相同, 所以正交相似的方阵具有相同的特征值. 反之, 设正交方阵 A 与 B 的特征值都是

$$e^{i\theta_1}, e^{-i\theta_1}, \dots, e^{i\theta_s}, e^{-i\theta_s}, \underbrace{1, \dots, 1}_{t\text{个}}, \underbrace{-1, \dots, -1}_{n-2s-t\text{个}},$$

其中 $0 < \theta_1 \leqslant \cdots \leqslant \theta_s < \pi$. 则由上段证明，方阵 A 与 B 都正交相似于准对角方阵 (7.5.1). 因此由正交相似的对称性与传递性得知，方阵 A 与 B 正交相似. 这就证明，正交方阵的特征值是正交方阵在正交相似下的全系不变量. ■

例 7.5.1 设 O 是 n 阶实正交方阵，且 $\operatorname{rank}(O - I_{(n)}) = 1$. 证明方阵 O 正交相似于对角方阵

$$\operatorname{diag}(-1, \underbrace{1, \dots, 1}_{n-1\text{个}}).$$

证明 设方阵 O 的全部特征值是

$$e^{i\theta_1}, e^{-i\theta_1}, \dots, e^{i\theta_s}, e^{-i\theta_s}, \underbrace{1, \dots, 1}_{t\text{个}}, \underbrace{-1, \dots, -1}_{n-2s-t\text{个}},$$

其中 $0 < \theta_1 \leqslant \cdots \leqslant \theta_s < \pi$. 则由定理 7.5.5，存在实正交方阵 O_1，使得 $O_1^{\mathrm{T}} O O_1 = D$，这里 D 是准对角方阵 (7.5.1). 因此

$$O_1^{\mathrm{T}}(O - I_{(n)})O_1 = \operatorname{diag}\left(\begin{pmatrix} \cos\theta_1 - 1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 - 1 \end{pmatrix}, \dots, \begin{pmatrix} \cos\theta_s - 1 & \sin\theta_s \\ -\sin\theta_s & \cos\theta_s - 1 \end{pmatrix},\right.$$
$$\left.\underbrace{0, \dots, 0}_{t\text{个}}, \underbrace{-2, \dots, -2}_{n-2s-t\text{个}}\right).$$

显然方阵的秩在正交相似下是不变的，因此

$$\operatorname{rank} O_1^{\mathrm{T}}(O - I_{(n)})O_1 = 1.$$

所以方阵 $O_1^{\mathrm{T}}(O - I_{(n)})O_1$ 的每个 2 阶子式都是 0，即

$$\det\begin{pmatrix} \cos\theta_j - 1 & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j - 1 \end{pmatrix} = 2(1 - \cos\theta_j) = 0.$$

因此 $\cos\theta_j = 1$，即 $\theta_j = 0$. 但是这不可能，所以方阵 O 的特征值只能是 ± 1. 于是

$$O_1^{\mathrm{T}}(O - I_{(n)})O_1 = \operatorname{diag}(\underbrace{0, \dots, 0}_{t\text{个}}, \underbrace{-2, \dots, -2}_{n-t\text{个}}).$$

由于 $\operatorname{rank} O_1^{\mathrm{T}}(O - I_{(n)})O_1 = 1$，所以 $t = n - 1$. 因此

$$O_1^{\mathrm{T}} O O_1 = \operatorname{diag}(\underbrace{1, \dots, 1}_{n-1\text{个}}, -1).$$

再取初等置换方阵

$$P_{1n} = I_{(n)} - E_{11} - E_{nn} + E_{1n} + E_{n1},$$

其中 E_{ij} 是 (i, j) 位置上的元素为 1、其它元素为 0 的 n 阶方阵. 显然初等置换方阵 P_{1n} 是正交方阵，而且

$$(O_1 P_{1n})^{\mathrm{T}} O (O_1 P_{1n}) = P_{1n}^{\mathrm{T}} O_1^{\mathrm{T}} O O_1 P_{1n} = \operatorname{diag}(-1, \underbrace{1, \dots, 1}_{n-1\text{个}}),$$

其中方阵 O_1P_{1n} 是正交的. ■

例 7.5.2 任意一个 n 阶实正交方阵都可以表为两个 n 阶实对称方阵的乘积.

证明 此例是例 6.6.2 的特例. 这里用正交方阵在正交相似下的标准形重新给出证明.

由定理 7.5.5,存在 n 阶实正交方阵 O_1,使得

$$O = O_1\,\mathrm{diag}\left(\begin{pmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{pmatrix},\dots,\begin{pmatrix}\cos\theta_s & \sin\theta_s\\ -\sin\theta_s & \cos\theta_s\end{pmatrix},\underbrace{1,\dots,1}_{t\,个},\underbrace{-1,\dots,-1}_{n-2s-t\,个}\right)O_1^{\mathrm{T}}.$$

注意到,

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}-\sin\theta & \cos\theta\\ \cos\theta & \sin\theta\end{pmatrix}.$$

记

$$S_1 = O_1\left(\underbrace{\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\dots,\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}}_{s\,个},\underbrace{1,\dots,1}_{n-2s\,个}\right)O_1^{\mathrm{T}},$$

$$S_2 = O_1\,\mathrm{diag}\left(\begin{pmatrix}-\sin\theta_1 & \cos\theta_1\\ \cos\theta_1 & \sin\theta_1\end{pmatrix},\dots,\begin{pmatrix}-\sin\theta_s & \cos\theta_s\\ \cos\theta_s & \sin\theta_s\end{pmatrix},\underbrace{1,\dots,1}_{t\,个},\underbrace{-1,\dots,-1}_{n-2s-t\,个}\right)O_1^{\mathrm{T}}.$$

显然方阵 S_1 与 S_2 都是实对称的,而且 $O = S_1S_2$. ■

习 题 7.5

1. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的保内积变换,即对任意 $\alpha,\beta\in V$,有

$$\big(\mathscr{A}(\alpha),\mathscr{A}(\beta)\big) = (\alpha,\beta).$$

证明,保内积变换 $\mathscr{A}$ 是线性变换,因而 $\mathscr{A}$ 是正交变换. 举例说明,保向量范数的变换不一定是线性变换.

2. 设 α 和 β 是 n 维 Euclid 空间 V 的向量,且 $\|\alpha\| = \|\beta\|$. 证明,存在正交变换 $\mathscr{A}$,使得

$$\mathscr{A}(\alpha) = \beta.$$

3. 设 α_1,α_2 与 β_1,β_2 是 n 维 Euclid 空间 V 的两对向量,且 $\|\alpha_i\| = \|\beta_i\|$, $i = 1,2$,并且向量 α_1 与 α_2 的夹角等于 β_1 与 β_2 的夹角. 证明,存在正交变换 $\mathscr{A}$,使得 $\mathscr{A}(\alpha_i) = \beta_i$, $i = 1,2$.

4. 设 $\alpha_1,\alpha_2,\dots,\alpha_k$ 与 $\beta_1,\beta_2,\dots,\beta_k$ 是 n 维 Euclid 空间 V 的两组向量. 证明,存在适合 $\mathscr{A}(\alpha_j) = \beta_j$, $j = 1,2,\dots,k$ 的正交变换 $\mathscr{A}$ 的充分必要条件是这两组向量的 Gram 方阵相等.

5. 求正交变换 O,使得 $B = O^{\mathrm{T}}AO$ 是正交方阵 A 在正交相似下的标准形.

(1) $A = \dfrac{1}{3}\begin{pmatrix}2 & -1 & 2\\ 2 & 2 & -1\\ -1 & 2 & 2\end{pmatrix}$; (2) $A = \dfrac{1}{2}\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix}$.

6. 设 $O = (a_{ij})$ 是 3 阶实正交方阵,且 $\det O = 1$. 证明,

$$(1-\mathrm{Tr}\,O)^2 + \sum_{1\leqslant i<j\leqslant 3}(a_{ij}-a_{ji})^2 = 4.$$

此结论可否推广?

7. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换. 如果存在 $\mathscr{A}$ 的不变子空间 U,使得对任意 $\alpha \in U$,均有 $\|\mathscr{A}(\alpha)\| = \|\alpha\|$; 对任意 $\alpha \in U^{\perp}$,均有 $\mathscr{A}(\alpha) = 0$,则 $\mathscr{A}$ 称为部分正交的. 证明,

(1) 部分正交变换的伴随变换仍是部分正交的;

(2) 部分正交变换的特征值的绝对值不超过 1.

8. 设 n 阶正交方阵 O 的特征值不等于 -1. 证明方阵 $I_{(n)} + O$ 可逆,方阵

$$K = \left(I_{(n)} - O\right)\left(I_{(n)} + O\right)^{-1}$$

是斜对称方阵,且

$$O = \left(I_{(n)} - K\right)\left(I_{(n)} + K\right)^{-1}.$$

§7.6 自伴变换与斜自伴变换

除正交变换外,还有两类重要的规范变换,即自伴变换与斜自伴变换.

定义 7.6.1 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换. 如果 $\mathscr{A}$ 与它的伴随变换 $\mathscr{A}^*$ 是同一个变换,即 $\mathscr{A} = \mathscr{A}^*$,则 $\mathscr{A}$ 称为**自伴变换**; 如果 $\mathscr{A}$ 满足 $\mathscr{A}^* = -\mathscr{A}$,则 $\mathscr{A}$ 称为**斜自伴变换**.

由伴随变换的定义可以看出,线性变换 $\mathscr{A}$ 是自伴变换的充分必要条件为,对任意 $\alpha, \beta \in V$,均有

$$\left(\mathscr{A}(\alpha), \beta\right) = \left(\alpha, \mathscr{A}(\beta)\right).$$

而线性变换 $\mathscr{A}$ 是斜自伴变换的充分必要条件为,对任意 $\alpha, \beta \in V$,均有

$$\left(\mathscr{A}(\alpha), \beta\right) = -\left(\alpha, \mathscr{A}(\beta)\right).$$

由定义 7.6.1 可以看出,自伴变换与斜自伴变换都是规范变换. 当然,除正交变换、自伴变换以及斜自伴变换外,还有其它的规范变换. 下面先讨论自伴变换.

定理 7.6.1 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 是自伴变换的充分必要条件为,$\mathscr{A}$ 在 V 的标准正交基下的方阵是对称方阵.

证明 设线性变换 $\mathscr{A}$ 在 V 的标准正交基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵是 A,则 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$ 在这组基下的方阵是 A^{T}. 于是 $\mathscr{A}^* = \mathscr{A}$ 等价于 $A^{\mathrm{T}} = A$. ■

定理 7.6.1 表明,如果在 n 维 Euclid 空间 V 中取定一组标准正交基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$,$V$ 的自伴变换 $\mathscr{A}$ 便和它在这组基下的方阵相对应. 这一对应是 V 的所有自伴变换集合到所有 n 阶实对称方阵集合上的一个双射. 于是自伴变换即是是对称方阵的一种几何解释.

由于自伴变换是规范变换,因此关于规范变换的结论可以移到自伴变换上. 当然,由于自伴变换是特殊类型的规范变换,所以相应的结论也带有某种特殊性.

由 §5.9 可知,实对称方阵的特征值都是实数. 所以自伴变换的特征值也都是实数. 于是由定理 7.4.6 得到

定理 7.6.2 设实数 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 n 维 Euclid 空间 V 的自伴变换 $\mathscr{A}$ 的全部

特征值,其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 则存在 V 的一组标准正交基,使得 $\mathscr{A}$ 在这组基下的方阵是如下的对角形:

$$\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n). \tag{7.6.1}$$

证明 因为线性变换 $\mathscr{A}$ 是自伴的,因此 $\mathscr{A}$ 是规范的. 又因为自伴变换 $\mathscr{A}$ 的特征值 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 全为实数,所以由定理 7.4.6 即得结论. ■

定理 7.6.2 的矩阵形式是

定理 7.6.3 设实数 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 n 阶实对称方阵 A 的全部特征值,其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 则方阵 A 正交相似于如下的对角形:

$$\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

而且实对称方阵的特征值是实对称方阵在正交相似下的全系不变量.

证明 设 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 n 维 Euclid 空间 V 的一组标准正交基,则由

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$$

便确定 V 的一个线性变换 $\mathscr{A}$.

由于 A 是实对称方阵,所以由定理 7.6.1,$\mathscr{A}$ 是自伴变换,且 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 $\mathscr{A}$ 的全部特征值. 由定理 7.6.2,存在 V 的标准正交基 $\{\beta_1, \beta_2, \ldots, \beta_n\}$,使得 $\mathscr{A}$ 在这组基下的方阵即为对角形 (7.6.1). 由于同一个线性变换 $\mathscr{A}$ 在不同标准正交基下的方阵是正交相似的,因此方阵 A 正交相似于对角形 (7.6.1).

显然正交相似的方阵具有相同的特征值,所以正交相似的实对称方阵也具有相同的特征值. 因此实对称方阵的特征值是实对称方阵的正交相似不变量. 反之,设实对称方阵 A 与 B 的所有特征值都是 $\lambda_1, \lambda_2, \ldots, \lambda_n$,则由上段证明,方阵 A 正交相似于对角形 (7.6.1),而对角形 (7.6.1) 正交相似于方阵 B. 由正交相似的传递性,实对称方阵 A 与 B 正交相似. 于是实对称方阵的特征值是实对称方阵在正交相似下的全系不变量. ■

定理 7.6.3 中关于全系不变量的结论也可由实对称方阵是规范方阵,而规范方阵的特征值是规范方阵在正交相似下的全系不变量的结论直接得到.

定理 7.6.3 完全解决了实对称方阵在正交相似下的标准形理论的两个基本问题. 于是定理 7.6.3 中的对角形 (7.6.1) 称为实对称方阵在正交相似下的标准形.

定理 7.6.2 也可以用其它方法证明. 这里介绍一种重要的方法,即所谓的分析方法. 这种方法深刻地揭示自伴变换的特征值的几何意义,应予重视. 当然涉及到的分析知识囿于篇幅这里不能一一叙述. 但相信学过多元微积分的读者是能够理解的. 先引进定义.

定义 7.6.2 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的自伴变换,α 是 V 的非零向量. 记

$$R(\alpha) = \frac{(\mathscr{A}(\alpha), \alpha)}{(\alpha, \alpha)},$$

它称为自伴变换 $\mathscr{A}$ 在 V 上的 **Rayleigh 商**.

显然自伴变换 $\mathscr{A}$ 在 V 上的 Rayleigh 商 $R(\alpha)$ 是定义在 V 的所有非零向量集合 V^* 上的实函数. 而且对任意非零实数 a, $R(a\alpha)=R(\alpha)$. 于是将 V 中所有单位向量的集合

$$S=\left\{\alpha\in V \mid \|\alpha\|=1\right\}$$

称为 V 的 n 维单位球面(或单位超球面),则 $\mathscr{A}$ 在 V 上的 Rayleigh 商 $R(\alpha)$ 的值域等于它在 n 维单位球面 S 上的限制的值域.

设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的一组标准正交基. 下面给出自伴变换 $\mathscr{A}$ 在 V 上的 Rayleigh 商 $R(\alpha)$ 在这组基下的表达式.

记 $\alpha\in V^*$ 在这组基下的坐标为 $x=(x_1,x_2,\ldots,x_n)$,即

$$\alpha=x_1\alpha+x_2\alpha_2+\cdots+x_n\alpha_n.$$

自伴变换 $\mathscr{A}$ 在这组基下的方阵为 $A=(a_{ij})_{n\times n}$,即对任意 $j=1,2,\ldots,n$,有

$$\mathscr{A}(\alpha_j)=\sum_{k=1}^{n}a_{kj}\alpha_k.$$

于是内积 $(\alpha,\alpha)=xx^{\mathrm{T}}$, $\big(\mathscr{A}(\alpha),\alpha\big)=xAx^{\mathrm{T}}$. 因此

$$R(\alpha)=\frac{xAx^{\mathrm{T}}}{xx^{\mathrm{T}}},\quad 0\neq x\in\mathbb{R}^n.$$

定理 7.6.4 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的自伴变换,则

(1) $\mathscr{A}$ 在 V 上的 Rayleigh 商 $R(\alpha)$ 具有最小值 λ_n,并且 λ_n 是 Rayleigh 商 $R(\alpha)$ 限定在 n 维单位球面 S 上的最小值,即

$$\lambda_n=\min_{\substack{\alpha\in V\\ \|\alpha\|=1}}R(\alpha);$$

(2) λ_n 是自伴变换 $\mathscr{A}$ 的最小特征值,而且使 Rayleigh 商 $R(\alpha)$ 达到最小值 λ_n 的单位向量 α_n 即是属于特征值 λ_n 的特征向量.

证明 和通常 3 维单位球面一样,n 维单位球面 S 是一个有界闭集. 将自伴变换 $\mathscr{A}$ 在 V 上 Rayleigh 商 $R(\alpha)$ 限定在 n 维单位球面 S 上,则由 $R(\alpha)$ 在标准正交基下的表达式可以看出,$R(\alpha)$ 是定义在 S 上的连续函数. 因此它具有最小值 λ_n.

由于定义在 V 上的 Rayleigh 商 $R(\alpha)$ 的值域与定义在 S 上的值域相同,因此 λ_n 也是定义在 V 上的 Rayleigh 商 $R(\alpha)$ 的最小值. 于是结论 (1) 成立.

设单位向量 α_n 满足

$$R(\alpha_n)=\min_{\substack{\alpha\in V\\ \|\alpha\|=1}}R(\alpha)=\lambda_n.$$

下面证明,α_n即是 $\mathscr{A}$ 的属于特征值 λ_n 的特征向量.

任取 $\beta\in V$,设 t 是实变量. 则

$$f(t)=R(\alpha_n+t\beta)$$

是 t 的实函数,而且 $f(t)$ 在 $t=0$ 处取到最小值 $f(0)$. 对 t 求微商得到,

$$f'(t)=\frac{\mathrm{d}}{\mathrm{d}t}R(\alpha_n+t\beta)=\frac{\mathrm{d}}{\mathrm{d}t}\frac{\big(\mathscr{A}(\alpha_n+t\beta),\alpha_n+t\beta\big)}{(\alpha_n+t\beta,\alpha_n+t\beta)}$$

$$= \frac{\big(\mathscr{A}(\beta),\alpha_n\big)+\big(\mathscr{A}(\alpha_n),\beta\big)+2t\big(\mathscr{A}(\beta),\beta\big)}{(\alpha_n+t\beta,\alpha_n+t\beta)} - f(t)\frac{2(\alpha_n,\beta)+2t(\beta,\beta)}{(\alpha_n+t\beta,\alpha_n+t\beta)},$$

因为 $f(0)=R(\alpha_n)=\lambda_n$,所以

$$f'(0)=\big(\mathscr{A}(\beta),\alpha_n\big)+\big(\mathscr{A}(\alpha_n),\beta\big)-2\lambda_n(\alpha_n,\beta)=0.$$

由于 $\mathscr{A}$ 是自伴的,所以 $\big(\mathscr{A}(\beta),\alpha_n\big)=\big(\beta,\mathscr{A}(\alpha_n)\big)$. 因此

$$f'(0)=2\Big(\big(\beta,\mathscr{A}(\alpha_n)\big)-\lambda_n(\beta,\alpha_n)\Big)=2\big(\beta,\ \mathscr{A}(\alpha_n)-\lambda_n\alpha_n\big)=0.$$

由 β 的任意性得到,

$$\mathscr{A}(\alpha_n)=\lambda_n\alpha_n.$$

因此结论 (2) 成立. ■

定理 7.6.4 即是著名的 Rayleigh 定理. 它可以推广为

定理 7.6.5 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的自伴变换. 则存在 V 的标准正交基

$$\{\alpha_1,\alpha_2,\ldots,\alpha_n\},$$

使得

(1) 设 V_j 是 V 中由向量 $\alpha_{j+1},\alpha_{j+2},\ldots,\alpha_n$ 生成的子空间,则 V_j 与它的的正交补 $V_j^\perp$ 都是自伴变换 $\mathscr{A}$ 的不变子空间,$j=1,2,\ldots,n$,其中约定 $V_n=0$;

(2) 将自伴变换 $\mathscr{A}$ 的 Rayleigh 商 $R(\alpha)$ 限定在 $V_j^\perp$ 上,则定义在 $V_j^\perp$ 上的 Rayleigh 商 $R(\alpha)$ 具有最小值 λ_j,而且 λ_j 即是定义在 $V_j^\perp\cap S$ 上的 Rayleigh 商 $R(\alpha)$ 的最小值,即对于任意 $j=1,2,\ldots,n$,有

$$\lambda_j=\min_{\substack{\alpha\in V_j^\perp\\ \|\alpha\|=1}}R(\alpha);$$

(3) λ_j 是自伴变换 $\mathscr{A}$ 的特征值,而且使得定义在 $V_j^\perp$ 上 Rayleigh 商 $R(\alpha)$ 取到最小值 λ_j 的单位向量 α_j 是自伴变换 $\mathscr{A}$ 的属于特征值 λ_j 的单位特征向量,即

$$R(\alpha_j)=\lambda_j=\min_{\substack{\alpha\in V_j^\perp\\ \|\alpha\|=1}}R(\alpha);$$

(4) 自伴变换 $\mathscr{A}$ 在标准正交基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的方阵是如下对角形

$$\mathrm{diag}\big(\lambda_1,\lambda_2,\ldots,\lambda_n\big),$$

其中 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$.

证明 对空间的维数用归纳法. 当 $n=1$ 时定理显然成立. 现在设定理对 $n-1$ 维 Euclid 空间成立. 下面证明定理对 n 维 Euclid 空间 V 成立.

由定理 7.6.4,将自伴变换 $\mathscr{A}$ 的 Rayleigh 商 $R(\alpha)$ 视为定义在 $V=0^\perp=V_n^\perp$ 的实函数,其最小值即是定义在 $V_n^\perp$ 与 n 维单位球面的交上的 $R(\alpha)$ 的最小值. 而且 λ_n 是 $\mathscr{A}$ 的特征值; 而使得定义在 $V_n^\perp$ 上的 $R(\alpha)$ 取到最小值 λ_n 的单位向量 α_n 是 $\mathscr{A}$ 的属于 λ_n 的特征向量.

由 α_n 生成的子空间记为 V_{n-1},它显然是 $\mathscr{A}$ 的不变子空间. 因为自伴变换 $\mathscr{A}$

是规范的，因此 V_{n-1} 的正交补 $V_{n-1}^{\perp}$ 也是 $\mathscr{A}$ 的不变子空间，且 $V_{n-1}^{\perp}$ 是 $n-1$ 维的.

自伴变换 $\mathscr{A}$ 在 $V_{n-1}^{\perp}$ 上的限制 $\mathscr{A}|_{V_{n-1}^{\perp}}$ 仍是自伴的. 记 $\widetilde{\mathscr{A}}=\mathscr{A}|_{V_{n-1}^{\perp}}$. 显然，$\widetilde{\mathscr{A}}$ 在 $V_{n-1}^{\perp}$ 上的 Rayleigh 商正好是自伴变换 $\mathscr{A}$ 在 V 上的 Rayleigh 商 $R(\alpha)$ 在 $V_{n-1}^{\perp}$ 上的限制. 于是由归纳假设，$V_{n-1}^{\perp}$ 中存在标准正交基 $\{\alpha_1,\alpha_2,\ldots,\alpha_{n-1}\}$，使得 $V_{n-1}^{\perp}$ 中由向量 $\alpha_{j+1},\ldots,\alpha_{n-1}$ 生成的子空间 $\widetilde{V}_j$ 与它的正交补 $\widetilde{V}_j^{\perp}$ 都是 $\widetilde{\mathscr{A}}$ 的不变子空间. 而且 $\widetilde{\mathscr{A}}$ 的 Rayleigh 商在 $\widetilde{V}_j^{\perp}$ 上的限制具有最小值 λ_j，此值恰好是 $\widetilde{\mathscr{A}}$ 的 Rayleigh 商在 $V_j^{\perp}$ 与 $V_{n-1}^{\perp}$ 中 $n-1$ 维单位球面 $\widetilde{S}$ 的交上的限制的最小值，而 $\widetilde{V}_j^{\perp}$ 中使得 $\widetilde{\mathscr{A}}$ 的 Rayleigh 商在 $\widetilde{V}_j^{\perp}$ 上的限制达到最小值 λ_j 的单位向量 α_j 是 $\widetilde{\mathscr{A}}$ 的属于特征值 λ_j 的单位特征向量.

因为 $\alpha_1,\alpha_2,\ldots,\alpha_{n-1}\in V_{n-1}^{\perp}$，$\alpha_n\in V_{n-1}$，所以 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 V 的标准正交基. 而 $\widetilde{V}_j^{\perp}$ 是 V 中由向量 $\alpha_1,\alpha_2,\ldots,\alpha_j$ 生成的子空间，因此 $\widetilde{V}_j^{\perp}$ 即是 V 中由向量 $\alpha_{j+1},\ldots,\alpha_n$ 生成的子空间 V_j 的正交补 $V_j^{\perp}$，并且 $V_j=\widetilde{V}_j\oplus V_{n-1}$.

由于 $\widetilde{\mathscr{A}}$ 是 $\mathscr{A}$ 在 $V_{n-1}^{\perp}$ 上的限制，$\widetilde{V}_j$ 与 $\widetilde{V}_j^{\perp}$ 都是 $\widetilde{\mathscr{A}}$ 的不变子空间，所以 V_j 与 $V_j^{\perp}$ 也都是 $\mathscr{A}$ 的不变子空间. 因此结论 (1) 对 n 维 Euclid 空间成立.

由于 $\widetilde{V}_j^{\perp}=V_j^{\perp}$，并且 $\widetilde{\mathscr{A}}$ 的 Rayleigh 商在 $\widetilde{V}_j^{\perp}$ 上的限制恰好是 $\mathscr{A}$ 的 Rayleigh 商 $R(\alpha)$ 在 $V_j^{\perp}$ 上的限制，因此结论 (2) 与 (3) 对 V 成立.

最后，由于 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 $\mathscr{A}$ 的依次属于特征值 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 的单位特征向量构成的标准正交基，所以

$$\mathscr{A}(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n),$$

即结论 (4) 也成立. ■

与定理 7.6.4 和定理 7.6.5 相对称的结论是下面的定理 7.6.6 和定理 7.6.7.

定理 7.6.6　设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的自伴变换，则

(1) $\mathscr{A}$ 在 V 上的 Rayleigh 商 $R(\alpha)$ 具有最大值 λ_1，并且 λ_1 是 Rayleigh 商 $R(\alpha)$ 限定在 n 维单位球面 S 上的最大值；

(2) λ_1 是自伴变换 $\mathscr{A}$ 的最大特征值，而且使 Rayleigh 商 $R(\alpha)$ 达到最大值 λ_1 的单位向量 α_1 即是属于特征值 λ_1 的特征向量.

定理 7.6.7　设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的自伴变换. 则存在 V 的标准正交基

$$\{\alpha_1,\alpha_2,\ldots,\alpha_n\},$$

使得

(1) 设 V_j 是 V 中由向量 $\alpha_1,\alpha_2\cdots,\alpha_{j-1}$ 生成的子空间，则 V_j 与它的的正交补 $V_j^{\perp}$ 都是自伴变换 $\mathscr{A}$ 的不变子空间，$j=1,2,\ldots,n$，其中约定 $V_1=0$；

(2) 将自伴变换 $\mathscr{A}$ 的 Rayleigh 商 $R(\alpha)$ 限定在 $V_j^{\perp}$ 上，则定义在 $V_j^{\perp}$ 上的 Rayleigh 商 $R(\alpha)$ 具有最大值 λ_j，而且 λ_j 即是定义在 $V_j^{\perp}\cap S$ 上的 Rayleigh 商 $R(\alpha)$ 的最大值，即对于任意 $j=1,2,\ldots,n$，有

$$\lambda_j = \max_{\substack{\alpha\in V_j^{\perp}\\ \|\alpha\|=1}} R(\alpha);$$

(3) λ_j 是自伴变换 $\mathscr{A}$ 的特征值,而且使得定义在 $V_j^{\perp}$ 上 Rayleigh 商 $R(\alpha)$ 取到最大值 λ_j 的单位向量 α_j 是自伴变换 $\mathscr{A}$ 的属于特征值 λ_j 的单位特征向量,即

$$R(\alpha_j) = \lambda_j = \max_{\substack{\alpha\in V_j^{\perp}\\ \|\alpha\|=1}} R(\alpha);$$

(4) 自伴变换 $\mathscr{A}$ 在标准正交基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵是如下对角形

$$\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$.

定理 7.6.6 与定理 7.6.7 的证明和定理 7.6.4 与定理 7.6.5 的证明完全相仿,不再赘述. 由定理 7.6.5 或定理 7.6.7 即得定理 7.6.2,从而给出了定理 7.6.2 的另一证明.

现在转到斜自伴变换.

定理 7.6.8 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 为斜自伴的充分必要条件是,$\mathscr{A}$ 在 V 的标准正交基下的方阵为斜对称方阵.

证明 设线性变换 $\mathscr{A}$ 在 V 的标准正交基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵是 A,则 $\mathscr{A}$ 的伴随变换 $\mathscr{A}^*$ 在这组基下的方阵是 A^{T}. 于是 $\mathscr{A}^* = -\mathscr{A}$ 等价于 $A^{\mathrm{T}} = -A$. ■

定理 7.6.8 表明,如果在 n 维 Euclid 空间 V 中取定一组标准正交基后,斜自伴变换便对应于它在这组基下的方阵——斜自伴方阵. 这一对应是 V 的所有斜自伴变换集合到所有 n 阶实斜对称方阵集合上的一个双射. 因此斜自伴变换即是斜对称方阵的一种几何解释.

由 §5.9 可知,斜对称方阵的非零特征值都是纯虚数,所以斜自伴变换的非零特征值也都是纯虚数. 于是有

定理 7.6.9 设 $\pm \mathrm{i}b_1, \pm \mathrm{i}b_2, \ldots, \pm \mathrm{i}b_s$ 是 n 维 Euclid 空间 V 的自斜伴变换 $\mathscr{A}$ 的全部非零特征值,其中 $2s \leqslant n$,并且 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_s$,则存在 V 的标准正交基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$,使得 $\mathscr{A}$ 在这组基下的方阵为如下的准对角形:

$$\operatorname{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, \underbrace{0, \ldots, 0}_{n-2s\text{ 个}}\right). \tag{7.6.2}$$

证明 因为 $\mathscr{A}$ 是斜自伴变换,因此 $\mathscr{A}$ 是规范变换. 又斜自伴变换 $\mathscr{A}$ 的全部特征值为

$$\pm \mathrm{i}b_1, \pm \mathrm{i}b_2, \ldots, \pm \mathrm{i}b_s, \underbrace{0, \ldots, 0}_{n-2s\text{ 个}},$$

所以由定理 7.4.6 即得定理 7.6.9. ■

定理 7.6.9 的矩阵形式是

定理 7.6.10 设 $\pm \mathrm{i}b_1, \pm \mathrm{i}b_2, \ldots, \pm \mathrm{i}b_s$ 是 n 阶实斜对称方阵 A 的全部非零特征

值，其中 $2s \leqslant n$，并且 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_s$，则 A 正交相似于如下的准对角形：

$$\operatorname{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, \underbrace{0, \ldots, 0}_{n-2s\text{ 个}}\right).$$

而且实斜对称方阵的特征值是实斜对称方阵在正交相似下的全系不变量.

证明　设 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 n 维 Euclid 空间 V 的一组标准正交基，则由

$$\mathscr{A}(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$$

便确定 V 的一个线性变换 $\mathscr{A}$.

由于 A 是实斜对称方阵，所以由定理 7.6.8，$\mathscr{A}$ 是斜自伴变换，由于方阵 A 的所有非零特征值 $\pm ib_1, \pm ib_2, \ldots, \pm ib_s$ 显然是线性变换 $\mathscr{A}$ 的所有非零特征值，由定理 7.6.9，存在 V 的标准正交基 $\{\beta_1, \beta_2, \ldots, \beta_n\}$，使得 $\mathscr{A}$ 在这组基下的方阵即为对角形 (7.6.2). 由于同一个线性变换 $\mathscr{A}$ 在不同标准正交基下的方阵是正交相似的，因此方阵 A 正交相似于对角形 (7.6.2).

由于相似的方阵具有相同的特征值，所以正交相似的方阵的特征值相同. 因此实斜对称方阵的特征值是实斜对称方阵的正交相似不变量. 反之，由于斜对称方阵是规范的，因此由定理 7.4.7，特征值相同的实斜对称方阵相似. 于是实斜对称方阵的特征值是实斜对称方阵在正交相似下的全系不变量. ■

定理 7.6.10 完全解决了实斜对称方阵在正交相似下的标准形问题. 于是定理 7.6.10 中的准对角形 (7.6.2) 称为实斜对称方阵在正交相似下的标准形.

习　题　7.6

1. 证明，n 维 Euclid 空间 V 的自伴变换 $\mathscr{A}$ 与 $\mathscr{B}$ 的乘积 $\mathscr{A}\mathscr{B}$ 仍是自伴的充分必要条件是，$\mathscr{A}$ 与 $\mathscr{B}$ 可交换.

2. 设 $\mathbb{R}[x]$ 是所有实系数多项式集合连同内积

$$\big(f(x), g(x)\big) = \int_0^1 f(x)g(x)\,\mathrm{d}x$$

一起构成的 Euclid 空间，其中 $f(x), g(x) \in \mathbb{R}[x]$. 回答下面的问题：

(1) $\mathbb{R}[x]$ 的线性变换 $\mathscr{A}$ 定义如下：令 $\mathscr{A}\big(f(x)\big) = xf(x)$，其中 $f(x) \in \mathbb{R}[x]$. 线性变换 $\mathscr{A}$ 是否是自伴的?

(2) Euclid 空间 $\mathbb{R}[x]$ 的微商变换 $\mathscr{D}$ 是否是自伴的?

3. 设 $\mathscr{A}$ 与 $\mathscr{B}$ 是 n 维 Euclid 空间 V 的线性变换，$\mathscr{A}$ 与 $\mathscr{A}\mathscr{B}$ 是自伴的，且 $\operatorname{Ker}\mathscr{A} \subseteq \operatorname{Ker}\mathscr{B}$. 证明，存在 V 的自伴变换 $\mathscr{C}$，使得 $\mathscr{C}\mathscr{A} = \mathscr{B}$.

4. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的斜自伴变换. 证明对任意 $\alpha \in V$，$\big(\mathscr{A}(\alpha), \alpha\big) = 0$. 反之是否成立?

5. 设 n 维 Euclid 空间 V 的线性变换 $\mathscr{A}$ 满足 $\mathscr{A}^2 = \mathscr{A}$，且对任意 $\alpha \in V$，$\|\mathscr{A}(\alpha)\| \leqslant \|\alpha\|$. 证明 $\mathscr{A}$ 是自伴的.

6. 设 n 维 Euclid 空间 V 的自伴变换满足 $\mathscr{A}^k = \mathscr{I}$，$k$ 为正整数. 证明 $\mathscr{A}^2 = \mathscr{I}$.

7. 设 $\mathscr{A}$ 是 2 维 Euclid 空间 V 的斜自伴变换. 证明，对任意 $\alpha \in V$，均有

$$\big(\mathscr{A}(\alpha), \mathscr{A}(\beta)\big) = (\det\mathscr{A})(\alpha, \beta).$$

8. 设 $\mathscr{A}$ 与 $\mathscr{B}$ 是 3 维 Euclid 空间 V 的斜自伴变换，$\mathscr{A}\neq 0$，且 $\operatorname{Ker}\mathscr{A}=\operatorname{Ker}\mathscr{B}$. 证明存在实数 λ，使得 $\mathscr{B}=\lambda\mathscr{A}$.

9. 设 λ_1 是 n 阶实对称方阵 $S=(s_{ij})$ 的最大特征值. 证明，

$$\lambda_1\geqslant\frac{1}{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n}s_{k\ell}.$$

10. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的自伴变换，V_0 是 $\mathscr{A}$ 的不变子空间. 证明存在向量 $\alpha_0\in V_0$，使得

$$R(\alpha_0)=\min\left\{R(\alpha)\;\middle|\;\alpha\in V_0^*\right\}.$$

并且 $R(\alpha_0)$ 是自伴变换 $\mathscr{A}$ 的特征值，α_0 是属于特征值 $R(\alpha_0)$ 的特征向量.

11. (Fischer) 设 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 是 n 维 Euclid 空间 V 的自伴变换 $\mathscr{A}$ 的特征值，且 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$. 设 V_k 是 $\mathscr{A}$ 的不变子空间. 证明，对于 $k=1,2,\ldots,n$，有

$$\lambda_{n+1-k}=\min_{V_k}\max\left\{R(\alpha)\;\middle|\;\alpha\in V_k^*\right\}.$$

12. 设所有 n 维实行向量集合连同标准内积构成的 Euclid 空间记为 $\mathbb{R}^n$，A 是 n 阶实对称方阵. 定义方阵 A 的 Rayleigh 商 $R(x)$ 为

$$R(x)=\frac{xAx^{\mathrm{T}}}{xx^{\mathrm{T}}},$$

其中 x 是 $\mathbb{R}^n$ 中非零行向量. 将关于自伴变换的定理 7.6.4 与定理 7.6.5 移到实对称方阵上.

13. 设 $\lambda_1(A)\geqslant\lambda_2(A)\geqslant\cdots\geqslant\lambda_n(A)$ 是 n 阶实对称方阵 A 的所有特征值. 证明，对于任意 $k\times n$ 阶矩阵 X，$k=1,2,\ldots,n$，有

$$\sup_X\frac{\lambda_k(XAX^{\mathrm{T}})}{\lambda_1(XX^{\mathrm{T}})}=\lambda_k(A),\qquad \inf_X\frac{\lambda_1(XAX^{\mathrm{T}})}{\lambda_k(XX^{\mathrm{T}})}=\lambda_{n+1-k}(A).$$

§7.7 正定对称方阵与矩阵的奇异值分解

在 §7.6 讨论自伴变换的特征值的分析性质时，使用了自伴变换 $\mathscr{A}$ 的 Rayleigh 商

$$R(\alpha)=\frac{\big(\mathscr{A}(\alpha),\alpha\big)}{(\alpha,\alpha)},$$

其中重要的是函数

$$f(\alpha)=\big(\mathscr{A}(\alpha),\alpha\big),$$

它称为**自伴变换 $\mathscr{A}$ 的型**.

在 n 维 Euclid 空间 V 中选取一组标准正交基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$，向量 α 在这组基下的坐标记为 $(x_1,x_2,\ldots,x_n)$，自伴变换 $\mathscr{A}$ 在这组基下的方阵记为 A. 显然 A 是对称方阵. 则自伴变换 $\mathscr{A}$ 的型 $f(\alpha)$ 在这组基下便具有如下形式：

$$f(\alpha)=\big(\mathscr{A}(\alpha),\alpha\big)=xAx^{\mathrm{T}},\ x\in\mathbb{R}^n. \tag{7.7.1}$$

由定理 7.6.2 容易得到

定理 7.7.1 设 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 是 n 维 Euclid 空间 V 的自伴变换 $\mathscr{A}$ 的全部特征值. 则存在 V 的标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$，使得自伴变换 $\mathscr{A}$ 的型 $f(\alpha)=\big(\mathscr{A}(\alpha),\alpha\big)$ 在这组基下具有如下形式：

$$f(\alpha)=\big(\mathscr{A}(\alpha),\alpha\big)=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2, \tag{7.7.2}$$

其中 $y=(y_1,y_2,\ldots,y_n)$ 是向量 $\alpha\in V$ 在这组基下的坐标.

证明 由定理 7.6.2,存在 V 的标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$,

$$\mathscr{A}(\xi_j)=\lambda_j\xi_j.$$

由于向量 $\alpha\in V$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标为 $y=(y_1,y_2,\ldots,y_n)$,所以

$$\begin{aligned}f(\alpha)&=\big(\mathscr{A}(\alpha),\alpha\big)=\Big(\mathscr{A}\Big(\sum_{j=1}^n y_j\xi_j\Big),\ \sum_{\ell=1}^n y_\ell\xi_\ell\Big)=\Big(\sum_{j=1}^n\lambda_j y_j\xi_j,\ \sum_{\ell=1}^n y_\ell\xi_\ell\Big)\\&=\sum_{j=1}^n\sum_{\ell=1}^n\lambda_j y_j y_\ell(\xi_j,\xi_\ell)=\sum_{j=1}^n\sum_{\ell=1}^n\lambda_j y_j y_\ell\delta_{j\ell}=\sum_{j=1}^n\lambda_j y_j^2.\end{aligned}$$

■

根据自伴变换的型的选取情况,可以把自伴变换分成若干类型.

定义 7.7.1 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的自伴变换. 如果对任意非零向量 $\alpha\in V$,均有 $\big(\mathscr{A}(\alpha),\alpha\big)>0$ (或 $\big(\mathscr{A}(\alpha),\alpha\big)\geqslant 0$,或 $\big(\mathscr{A}(\alpha),\alpha\big)<0$,或 $\big(\mathscr{A}(\alpha),\alpha\big)\leqslant 0$),则自伴变换 $\mathscr{A}$ 称为**正定的**(或**半正定**,或**负定**,或**半负定的**).

由定义可以看出,正定(负定)自伴变换一定是半正定(半负定)的. 而且自伴变换 $\mathscr{A}$ 为正定(半正定)的当且仅当 $-\mathscr{A}$ 为负定(半负定)的. 于是下面得到的关于正定(半正定)自伴变换的结论都可以自然地移到负定(半负定)自伴变换.

由于自伴变换 $\mathscr{A}$ 在 V 的标准正交基下的方阵是对称方阵,所以关于实对称方阵也有相仿的定义.

定义 7.7.2 设 A 是 n 阶实对称方阵. 如果对任意非零向量 $x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$,均有 $xSx^{\mathrm{T}}>0$ (或 $xSx^{\mathrm{T}}\geqslant 0$,或 $xSx^{\mathrm{T}}<0$,或 $xSx^{\mathrm{T}}\leqslant 0$),则对称方阵 S 称为**正定的**(或**半正定**,或**负定**,或**半负定的**).

定理 7.7.2 设自伴变换 $\mathscr{A}$ 在 n 维 Euclid 空间 V 的标准正交基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的方阵为 A. 则 $\mathscr{A}$ 正定(或半正定)的充分必要条件是,对称方阵 A 正定(或半正定).

证明 设自伴变换 $\mathscr{A}$ 是正定(或半正定)的,那么对任意非零向量 $\alpha\in V$,

$$\big(\mathscr{A}(\alpha),\alpha\big)>0\quad(\text{或 }\big(\mathscr{A}(\alpha),\alpha\big)\geqslant 0).$$

设 $\alpha\in V$ 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的坐标为 $x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$,则

$$\big(\mathscr{A}(\alpha),\alpha\big)=x\mathscr{A}x^{\mathrm{T}}>0,\quad(\text{或}\geqslant 0).$$

因此对称方阵 A 是正定(半正定)的. 反之亦然. ■

引理 7.7.1 设 S 是 n 阶正定对称方阵,O 是任意 n 阶正交方阵. 则方阵 $O^{\mathrm{T}}SO$ 是正定对称方阵. 简单地说,对称方阵的正定性在正交相似下不变.

证明 设 $x=(x_1,x_2,\ldots,x_n)$ 是 $\mathbb{R}^n$ 中任意非零向量. 因为方阵 O 是正交的,所以方阵 O^{T} 可逆. 因此向量 $y=xO^{\mathrm{T}}\in\mathbb{R}^n$ 是非零的. 由于 $(O^{\mathrm{T}}SO)^{\mathrm{T}}=O^{\mathrm{T}}SO$ 是对称的. 而且由于 S 是正定的. 所以

$$x(O^{\mathrm{T}}SO)x^{\mathrm{T}} = ySy^{\mathrm{T}} > 0.$$

因此方阵 $O^{\mathrm{T}}SO$ 是正定对称方阵. ■

注 同理可知,对称方阵的半正定性,负定性以及半负定性都是正交相似下不变的.

下面是关于正定对称方阵的重要定理.

定理 7.7.3 设 $S = (s_{ij})$ 是 n 阶实对称方阵,则下述命题等价:

(1) 方阵 S 是正定的;

(2) 方阵 S 的每个特征值都是正的;

(3) 存在正定对称方阵 S_1,使得 $S = S_1^2$;

(4) 存在可逆方阵 P,使得 $S = P^{\mathrm{T}}P$;

(5) 方阵 S 的每个主子式都是正的;

(6) 方阵 S 的顺序主子式都是正的,这里所谓方阵 S 的顺序主子式是指如下的 n 个主子式: 对于 $k = 1, 2, \ldots, n$,

$$S\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix} = \begin{vmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ s_{21} & s_{22} & \cdots & s_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{k1} & s_{k2} & \cdots & s_{kk} \end{vmatrix};$$

(7) 对每个 $k = 1, 2, \ldots, n$,方阵 S 的所有 k 阶主子式之和都是正的.

证明 (1) $\Longrightarrow$ (2) 由定理 7.6.3,存在 n 阶实正交方阵 O,使得

$$O^{\mathrm{T}}SO = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 S 的所有特征值. 由引理 7.7.1,方阵 $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ 是正定的. 因此对非零向量 $\varepsilon_j = (0, \ldots, 0, \underset{j}{1}, 0 \cdots, 0) \in \mathbb{R}^n$, $j = 1, 2, \ldots, n$,有

$$\varepsilon_j \,\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\varepsilon_j^{\mathrm{T}} = \lambda_j > 0.$$

(2) $\Longrightarrow$ (3) 由定理 7.6.3,存在 n 阶实正交方阵 O,使得

$$O^{\mathrm{T}}SO = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 S 的所有特征值. 因为 $\lambda_j > 0$,所以 $\sqrt{\lambda_j} > 0$, $j = 1, 2, \ldots, n$. 于是

$$S = \left(O^{\mathrm{T}} \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})O\right)\left(O^{\mathrm{T}} \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})O\right).$$

记 $S_1 = O^{\mathrm{T}} \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})O$,则 $S = S_1^2$.

显然方阵 S_1 是对称的. 设 $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,则

$$x \,\mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})x^{\mathrm{T}} = \sqrt{\lambda_1}\, x_1^2 + \sqrt{\lambda_2}\, x_2^2 + \cdots + \sqrt{\lambda_n}\, x_n^2,$$

所以当 x 非零时, $x \,\mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})x^{\mathrm{T}} > 0$. 因此方阵 $\mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$ 是正定的. 由引理 7.7.1,方阵 S_1 是正定的.

(3) $\Longrightarrow$ (4) 因为方阵 $S = S_1^2 = S_1^{\mathrm{T}}S_1$,且 S_1 是正定的,所以由 (2),方阵 S_1 的特

征值都是正的. 即方阵 S_1 可逆. 于是取 $P=S_1$, 即得 $S=P^{\mathrm{T}}P$.

(4) $\Longrightarrow$ (5) 因为方阵 $S=P^{\mathrm{T}}P$, P 可逆, 故由 Binet-Cauchy 公式,

$$S\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} = \sum_{1\leqslant j_1<j_2<\cdots<j_k\leqslant n} P^{\mathrm{T}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} P\begin{pmatrix} j_1 & j_2 & \cdots & j_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}.$$

显然

$$P^{\mathrm{T}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} = P\begin{pmatrix} j_1 & j_2 & \cdots & j_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}.$$

因此

$$S\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} = \sum_{1\leqslant j_1<j_2<\cdots<j_k\leqslant n} \left(P\begin{pmatrix} j_1 & j_2 & \cdots & j_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}\right)^2.$$

如果 $S\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}=0$, 则对每个 $1\leqslant j_1<j_2<\cdots<j_k\leqslant n$, $P\begin{pmatrix} j_1 & j_2 & \cdots & j_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}=0$. 由 Laplace 展开定理

$$\det P = \sum_{1\leqslant j_1<j_2<\cdots<j_k\leqslant n} (-1)^{i_1+\cdots+i_k+j_1+\cdots+j_k} P\begin{pmatrix} j_1 & j_2 & \cdots & j_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} P\begin{pmatrix} j_{k+1} & j_{k+2} & \cdots & j_n \\ i_{k+1} & i_{k+2} & \cdots & i_n \end{pmatrix} = 0.$$

这与方阵 P 可逆的假设相矛盾. 所以对每个 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$, $k=1,2,\ldots,n$,

$$S\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} > 0.$$

(5) $\Longrightarrow$ (6) 显然.

(6) $\Longrightarrow$ (1) 对方阵 S 的阶数 n 用归纳法. 当 $n=1$ 时结论显然成立. 假设结论对 $n-1$ 阶方阵 S 成立. 下面证明结论对 n 阶方阵 S 成立.

把 n 阶方阵 S 分块为

$$S=\begin{pmatrix} S_1 & y^{\mathrm{T}} \\ y & s_{nn} \end{pmatrix},$$

其中 S_1 是 $n-1$ 阶对称方阵, 而 $y\in\mathbb{R}^{n-1}$. 显然, 方阵 S_1 的顺序主子式即是方阵 S 的前 $n-1$ 个顺序主子式, 所以方阵 S_1 的顺序主子式都是正的. 由归纳假设, 方阵 S_1 是正定的. 由 (2), 方阵 S_1 可逆. 于是

$$S=\begin{pmatrix} I_{(n-1)} & 0 \\ yS_1^{-1} & 1 \end{pmatrix}\begin{pmatrix} S_1 & 0 \\ 0 & s_{nn}-yS_1^{-1}y^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} I_{(n-1)} & S_1^{-1}y^{\mathrm{T}} \\ 0 & 1 \end{pmatrix}.$$

设 x 是 $\mathbb{R}^n$ 中非零向量. 记

$$x\begin{pmatrix} I_{(n-1)} & 0 \\ yS_1^{-1} & 1 \end{pmatrix} = (z, z_n),$$

其中 $z\in\mathbb{R}^{n-1}$. 则

$$xSx^{\mathrm{T}} = zS_1z^{\mathrm{T}} + (s_{nn}-yS_1^{-1}y^{\mathrm{T}})z_n^2.$$

由假设, $\det S=(\det S_1)(s_{nn}-yS_1^{-1}y^{\mathrm{T}})>0$, $\det S_1>0$, 故 $s_{nn}-yS_1^{-1}y^{\mathrm{T}}>0$. 因此对任意非零 $x\in\mathbb{R}^n$,

$$xSx^{\mathrm{T}} = zS_1z^{\mathrm{T}} + (s_{nn}-yS_1^{-1}y^{\mathrm{T}})z_n^2 > 0.$$

至此证明了 (1)—(6) 的等价性.

(5) $\Longrightarrow$ (7) 显然.

(7) $\Longrightarrow$ (2) 记方阵 S 的所有 k 阶主子式之和为 s_k, $k=1,2,\ldots,n$. 由 §5.7 可

知，方阵 S 的特征多项式为 $\varphi(\lambda)$ 为

$$\varphi(\lambda)=\lambda^n-s_1\lambda^{n-1}+s_2\lambda^{n-2}-\cdots+(-1)^{n-1}s_{n-1}\lambda+(-1)^n s_n.$$

因为 $s_k>0$，所以多项式 $\varphi(\lambda)$ 的系数是正负相间的，因此 $\varphi(\lambda)$ 不可能具有负根. 由于 $s_n>0$，所以多项式 $\varphi(\lambda)$ 的根不能是零. 这就证明，方阵 S 的特征值都是正的. ■

关于半正定对称方阵也有相应的结论.

定理 7.7.4 设 S 是 n 阶对称方阵，则下述命题等价：

(1) 方阵 S 是半正定的；

(2) 方阵 S 的每个特征值都是非负的；

(3) 存在 n 阶半正定对称方阵 S_1，$\operatorname{rank} S_1=\operatorname{rank} S$，使得 $S=S_1^2$；

(4) 存在 n 阶方阵 P，$\operatorname{rank} P=\operatorname{rank} S$，使得 $S=P^{\mathrm T}P$；

(5) 方阵 S 的每个主子式都是非负的；

(6) 方阵 S 的所有 k 阶主子式之和都是非负的，$k=1,2,\ldots,n$.

证明 **(1) $\Longrightarrow$ (2)** 设 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 是 S 的所有特征值，且 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$. 则由定理 7.6.3，方阵 S 正交相似于对角形 $\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$. 由引理 7.7.1 的注，方阵 $\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$ 是半正定的. 因此对非零向量 $\varepsilon_j=(0,\ldots,0,\underset{j}{1},0,\ldots,0)\in\mathbb{R}^n$，有

$$\varepsilon_j\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)\varepsilon_j^{\mathrm T}=\lambda_j\geqslant 0.$$

(2) $\Longrightarrow$ (3) 记号同上. 由定理 7.6.3，存在 n 阶实正交方阵 O，使得

$$O^{\mathrm T}SO=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n),$$

其中 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n\geqslant 0$. 所以 $\sqrt{\lambda_1}\geqslant\sqrt{\lambda_2}\geqslant\cdots\geqslant\sqrt{\lambda_n}\geqslant 0$，并且

$$S=\left(O^{\mathrm T}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O\right)^2.$$

记 $S_1=O^{\mathrm T}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O$. 显然方阵 S_1 是对称的.

设 $x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$，并且 $\operatorname{rank} S_1$ 等于非负实数 $\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n}$ 中非零的数的个数，即等于方阵 S 的非零特征值的个数，也即 $\operatorname{rank} S_1=\operatorname{rank} S$.

设 $x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$，则

$$x\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})x^{\mathrm T}=\sqrt{\lambda_1}\,x_1^2+\sqrt{\lambda_2}\,x_2^2+\cdots+\sqrt{\lambda_n}\,x_n^2\geqslant 0,$$

因此 $\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})$ 半正定. 由引理 7.7.1 的注，方阵 S_1 是半正定的.

(3) $\Longrightarrow$ (4) 只需取 $P=S_1$ 即得 (4) 成立.

(4) $\Longrightarrow$ (5) 记取自 S 的第 $i_1,i_2,\ldots,i_k$ 行，第 $i_1,i_2,\ldots,i_k$ 列 $(1\leqslant i_1<i_2<\cdots<i_k\leqslant n)$ 交叉位置上的元素构成的 k 阶主子式为

$$S\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}.$$

则由 $S=P^{\mathrm T}P$ 得到，

$$S\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}=\sum_{1\leqslant j_1<j_2<\cdots<j_k\leqslant n}P^{\mathrm T}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}P\begin{pmatrix} j_1 & j_2 & \cdots & j_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}.$$

显然
$$P^{\mathrm{T}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} = P\begin{pmatrix} j_1 & j_2 & \cdots & j_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}.$$
因此
$$S\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} = \sum_{1\leqslant j_1<j_2<\cdots<j_k\leqslant n} \left(P\begin{pmatrix} j_1 & j_2 & \cdots & j_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}\right)^2 \geqslant 0.$$

(5) $\Longrightarrow$ (6) 显然.

(6) $\Longrightarrow$ (1) 记方阵 S 的所有 k 阶主子式之和为 s_k. 由 §5.7 可知,方阵 S 的特征多项式为 $\varphi(\lambda)$ 为
$$\varphi(\lambda) = \lambda^n - s_1\lambda^{n-1} + s_2\lambda^{n-2} - \cdots + (-1)^{n-1}s_{n-1}\lambda + (-1)^n s_n.$$
设 $\lambda_0 = -\mu_0$, μ_0 是正数,则
$$\varphi(\lambda_0) = \varphi(-\mu_0) = (-1)^n \sum_{k=0}^{n} s_k\mu_0^{n-k},$$
其中约定 $s_0 = 1$. 由于 $s_k \geqslant 0, k = 1, 2, \ldots, n$,因此
$$\sum_{k=0}^{n} s_k\mu_0^{n-k} = \mu_0^n + s_1\mu_0^{n-1} + \cdots + s_n \geqslant \mu_0^n > 0,$$
所以负数 λ_0 不可能是多项式 $\varphi(\lambda)$ 的根,即方阵 S 的特征值都是非负的,设为 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 由定理 7.6.3,方阵 S 正交相似于对角形 $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

设 $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,则
$$x\,\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)x^{\mathrm{T}} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2 \geqslant 0.$$
所以 $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ 是半正定的. 由引理 7.7.1 的注,方阵 S 是半正定的. ■

应当指出,定理 7.7.4 中使得半正定对称方阵 $S = S_1^2$ 的半正定方阵 S_1 是唯一的. 即有

定理 7.7.5 设 S 与 S_1 都是 n 阶半正定对称方阵,且 $S = S_1^2$,则方阵 S_1 是唯一的. 而且与方阵 S 可交换的 n 阶方阵 A 也和方阵 S_1 可交换.

证明 设方阵 S 的所有特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$. 由定理 7.7.4 可设 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 设 S_1 与 S_2 是半正定对称方阵,且
$$S = S_1^2 = S_2^2,$$
并且 $\mu_1, \mu_2, \ldots, \mu_n$ 与 $\nu_1, \nu_2, \ldots, \nu_n$ 分别是方阵 S_1 与 S_2 的所有特征值. 由定理 7.7.4,可设 $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n \geqslant 0$ 且 $\nu_1 \geqslant \nu_2 \geqslant \cdots \geqslant \nu_n \geqslant 0$. 由定理 7.6.3,存在正交方阵 O_1 与 O_2,使得
$$S_1 = O_1^{\mathrm{T}}\,\mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_n)O_1, \quad S_2 = O_2^{\mathrm{T}}\,\mathrm{diag}(\nu_1, \nu_2, \ldots, \nu_n)O_2.$$
于是
$$S = O_1^{\mathrm{T}}\,\mathrm{diag}(\mu_1^2, \mu_2^2, \ldots, \mu_n^2)O_1 = O_2^{\mathrm{T}}\,\mathrm{diag}(\nu_1^2, \nu_2^2, \ldots, \nu_n^2)O_2.$$
这表明,$\mu_1^2, \mu_2^2, \ldots, \mu_n^2$ 是方阵 S 的所有特征值,并且 $\mu_1^2 \geqslant \mu_2^2 \geqslant \cdots \geqslant \mu_n^2 \geqslant 0$. 因此 $\mu_j^2 = \lambda_j$. 同理,$\nu_j^2 = \lambda_j$. 于是

$$O_1^{\mathrm{T}}\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)O_1=O_2^{\mathrm{T}}\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)O_2.$$

即

$$O_2O_1^{\mathrm{T}}\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)O_2O_1^{\mathrm{T}}.$$

记 $O_2O_1^{\mathrm{T}}=P=(p_{ij})_{n\times n}$. 比较上式两端方阵的 (i,j) 元素，得到 $p_{ij}\lambda_j=\lambda_i p_{ij}$.

当 $\lambda_i=\lambda_j$ 时，显然 $p_{ij}\sqrt{\lambda_j}=\sqrt{\lambda_i}\,p_{ij}$; 当 $\lambda_i\neq\lambda_j$ 时，由 $p_{ij}\lambda_j=\lambda_i p_{ij}$ 可知 $p_{ij}=0$，因此 $p_{ij}\sqrt{\lambda_j}=\sqrt{\lambda_i}\,p_{ij}$ 仍成立. 所以对任意 $1\leqslant i,j\leqslant n$，均有

$$p_{ij}\sqrt{\lambda_j}=\sqrt{\lambda_i}\,p_{ij}.$$

写成矩阵形式，即得

$$O_2O_1^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})=\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O_2O_1^{\mathrm{T}}.$$

即

$$O_1^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O_1=O_2^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O_2.$$

由此得到，$S_1=S_2$. 这就证明，适合 $S=S_1^2$ 的半正定对称方阵 S_1 是唯一的.

由定理 7.6.3，可设

$$S=O^{\mathrm{T}}\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)O,$$

其中 O 是某个正交方阵. 由 $AS=SA$ 得到，

$$OAO^{\mathrm{T}}\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)OAO^{\mathrm{T}}.$$

由上段唯一性证明得到，

$$OAO^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})=\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})OAO^{\mathrm{T}}.$$

即

$$AO^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O=O^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})OA.$$

显然

$$S=\left(O^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O\right)^2.$$

由唯一性，

$$S_1=O^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_n})O.$$

于是 $AS_1=S_1A$. ■

对半正定对称方阵 S，使得 $S=S_1^2$ 的唯一半正定对称方阵 S_1 称为 **方阵 S 的平方根**，记为 $S^{\frac{1}{2}}$，或 $\sqrt{S}$.

对称方阵在正交相似下的标准形和关于半正定对称方阵的定理 7.7.2 与定理 7.7.4 可以导出实矩阵在正交相抵下的标准形.

定义 7.7.3 设 A 与 B 是 $m\times n$ 实矩阵. 如果存在 m 阶与 n 阶实正交方阵 O_1 与 O_2，使得 $B=O_2AO_2$，则矩阵 A 与 B 称为**正交相抵**的.

显然，如果 $m\times n$ 实矩阵 A 与 B 正交相抵，则 A 与 B 相抵，但反之并不尽然.

$m\times n$ 实矩阵之间的正交相抵关系显然满足自反性、对称性与传递性. 因此它是 $m\times n$ 实矩阵之间的一种等价关系. 于是所有 $m\times n$ 实矩阵的集合 $\mathbb{R}^{m\times n}$ 在正交相抵等价关系下分成正交相抵等价类: 彼此正交相抵的 $m\times n$ 实矩阵归在同一个

正交相抵等价类,而彼此不相抵的 $m\times n$ 实矩阵划归在不同的正交相抵等价类. 关于 $m\times n$ 实矩阵集合 $\mathbb{R}^{m\times n}$ 在正交相抵下分类的两个基本问题是:

(1) 确定 $m\times n$ 实矩阵在正交相抵下的标准形;

(2) 确定 $m\times n$ 实矩阵在正交相抵下的全系不变量.

为讨论正交相抵的需要,引进下面的定义.

定义 7.7.4 设 A 是 $m\times n$ 实矩阵,则 n 阶方阵 $A^{\mathrm{T}}A$ 的非零特征值的算术平方根称为矩阵 A 的**奇异值**.

显然,$(A^{\mathrm{T}}A)^{\mathrm{T}}=A^{\mathrm{T}}A$,所以方阵 $A^{\mathrm{T}}A$ 是对称的. 设 $x\in\mathbb{R}^n$,并记 $y=xA^{\mathrm{T}}$,则 $y\in\mathbb{R}^n$,并且

$$xA^{\mathrm{T}}Ax^{\mathrm{T}}=yy^{\mathrm{T}}=\|y\|\geqslant 0.$$

因此 $A^{\mathrm{T}}A$ 是半正定对称方阵. 由定理 7.7.4,方阵 $A^{\mathrm{T}}A$ 的非零特征值都是正的. 因此如果方阵 $A^{\mathrm{T}}A$ 的全部非零特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_r$,则矩阵 A 的奇异值为 $\sqrt{\lambda_1}$, $\sqrt{\lambda_2},\ldots,\sqrt{\lambda_r}$,并且它们都是正的. 另外,由 §3.3 习题 11 可知,

$$\lambda^m\det\big(\lambda I_{(n)}-A^{\mathrm{T}}A\big)=\lambda^n\det\big(\lambda I_{(m)}-AA^{\mathrm{T}}\big),$$

所以方阵 $A^{\mathrm{T}}A$ 与 AA^{T} 具有相同的非零特征值. 于是矩阵 A 的奇异值也可以定义为方阵 AA^{T} 的非零特征值的算术平方根.

定理 7.7.6 设 $\mu_1,\mu_2,\ldots,\mu_r$ 是 $m\times n$ 实矩阵 A 的全部奇异值,其中 $\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_r>0$,且 $r=\operatorname{rank}A$. 则矩阵 A 正交相抵于如下的准对角形:

$$\operatorname{diag}\big(\operatorname{diag}(\mu_1,\mu_2,\ldots,\mu_r),0\big),\tag{7.7.3}$$

其中 0 是 $(m-r)\times(n-r)$ 零矩阵. 并且矩阵的奇异值是矩阵在正交相似下的全系不变量.

证明 由定理 7.7.4,$A^{\mathrm{T}}A$ 是 n 阶半正定对称方阵,而且它的所有非零特征值都是正的,设为 $\mu_1^2,\mu_2^2,\ldots,\mu_r^2$. 因此存在 n 阶正交方阵 O,使得

$$A^{\mathrm{T}}A=O\operatorname{diag}\big(\mu_1^2,\mu_2^2,\ldots,\mu_r^2,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big)O^{\mathrm{T}}.$$

记 $D=\operatorname{diag}(\mu_1,\mu_2,\ldots,\mu_r)$,$O=(O_1,O_2)$,其中 O_1 是 $n\times r$ 阶子矩阵. 则

$$A^{\mathrm{T}}A=O_1D^2O_1^{\mathrm{T}}.$$

因为方阵 O 是正交的,所以

$$O^{\mathrm{T}}O=\begin{pmatrix}O_1^{\mathrm{T}}\\ O_2^{\mathrm{T}}\end{pmatrix}(O_1,\ O_2)=\begin{pmatrix}O_1^{\mathrm{T}}O_1 & O_1^{\mathrm{T}}O_2\\ O_2^{\mathrm{T}}O_1 & O_2^{\mathrm{T}}O_2\end{pmatrix}=I_{(n)}.$$

因此 $O_1^{\mathrm{T}}O_1=I_{(r)}$,$O_1^{\mathrm{T}}O_2$ 与 $O_2^{\mathrm{T}}O_1$ 分别是 $r\times(n-r)$ 与 $n-r$ 阶零矩阵. 于是由 $A^{\mathrm{T}}A=O_1D^2O_1^{\mathrm{T}}$ 得到

$$(AO_1D^{-1})^{\mathrm{T}}(AO_1D^{-1})=I_{(r)},\qquad (AO_2)^{\mathrm{T}}(AO_2)=0.$$

这表明,$P_1=AO_1D^{-1}$ 是由 r 个两两正交的 m 维实的单位列向量构成的 $m\times r$ 矩

阵,而 AO_2 是 $m\times(n-r)$ 零矩阵. 由于 $m\times r$ 矩阵 P_1 的 r 个列向量可以扩充成 m 维实的列向量空间连同标准内积构成的 Euclid 空间 $\mathbb{R}^m$ 的标准正交基,所以存在 $m\times(m-r)$ 矩阵 P_2,使得 $P=(P_1,P_2)$ 为正交方阵. 因此,

$$P\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix}O^{\mathrm{T}}=\begin{pmatrix}P_1, & P_2\end{pmatrix}\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}O_1^{\mathrm{T}}\\ O_2^{\mathrm{T}}\end{pmatrix}=P_1DO_1^{\mathrm{T}}=AO_1O_1^{\mathrm{T}}.$$

由于方阵 O 是正交的,所以

$$OO^{\mathrm{T}}=\begin{pmatrix}O_1, & O_2\end{pmatrix}\begin{pmatrix}O_1^{\mathrm{T}}\\ O_2^{\mathrm{T}}\end{pmatrix}=O_1O_1^{\mathrm{T}}+O_2O_2^{\mathrm{T}}=I_{(n)}.$$

因此由 $AO_2=0$ 得到

$$P\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix}O^{\mathrm{T}}=AO_1O_1^{\mathrm{T}}=A\left(I_{(n)}-O_2O_2^{\mathrm{T}}\right)=A.$$

这就证明,$m\times n$ 矩阵 A 正交相抵于准对角形 (7.7.3).

设 $m\times n$ 矩阵 A 与 B 正交相抵,即存在 m 阶与 n 阶正交方阵 O_1 与 O_2,使得 $B=O_1AO_2$,则 $B^{\mathrm{T}}B=O_2^{\mathrm{T}}A^{\mathrm{T}}AO_2$. 因此方阵 $B^{\mathrm{T}}B$ 与 $A^{\mathrm{T}}A$ 具有相同的非零特征值. 所以矩阵 A 与 B 具有相同的奇异值,也就是说,矩阵的奇异值是矩阵在正交相抵下的不变量.

反之,设矩阵 A 与 B 具有相同的奇异值 $\mu_1,\mu_2,\ldots,\mu_r$,且 $\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_r>0$. 则由上段证明,矩阵 A 正交相抵于准对角方阵 (7.7.3),而准对角方阵 (7.7.3) 正交相抵于矩阵 B. 由正交相抵的传递性,矩阵 A 与矩阵 B 正交相抵. ■

定理 7.7.6 完全解决了矩阵在正交相抵下的标准形问题,其中准对角方阵 (7.7.3) 称为矩阵在正交相抵下的标准形. 将矩阵 A 表为

$$A=O_1\,\mathrm{diag}\big(\mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_r),0\big)O_2,$$

其中 O_1 与 O_2 是 m 阶与 n 阶正交方阵,称为矩阵 A 的**奇异值分解**.

矩阵的奇异值分解具有重要的应用.

定理 7.7.7(矩阵的极分解) 任意 n 阶实方阵 A 都可以分解为一个半正定对称方阵 S(或 S_1)与一个实正交方阵 O 的乘积,即 $A=SO$(或者 $A=OS_1$),而且其中半正定对称方阵 S(或者 S_1)是由方阵 A 唯一确定的.

证明 设 $\mu_1,\mu_2,\ldots,\mu_r$ 是方阵 A 的所有奇异值,则由定理 7.7.6,存在 n 阶正交方阵 O_1 与 O_2,使得

$$A=O_1\,\mathrm{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big)O_2. \tag{7.7.4}$$

显然

$$A=\left(O_1\,\mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0)O_1^{\mathrm{T}}\right)O_1O_2.$$

记$S=O_1\,\mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0)O_1^{\mathrm{T}}$,$O=O_1O_2$. 显然方阵 S 是对称的,而方阵 O 是正交的. 由定理 7.7.4,方阵 S 是半正定的. 于是 $A=SO$.

设 $A=SO=\widetilde{S}\widetilde{O}$,其中 $\widetilde{S}$ 是半正定对称矩阵,$\widetilde{O}$ 是正交方阵. 则 $AA^{\mathrm{T}}=S^2=\widetilde{S}^2$. 由于方阵 AA^{T} 是半正定的,所以由定理 7.7.5,$\widetilde{S}=S$. 于是,使得 $A=SO$ 成立的半正定对称方阵 S 是唯一的. 由式 (7.7.4),

$$A=O_1O_2\left(O_2^{\mathrm{T}}\operatorname{diag}(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0)O_2\right).$$

记 $O_1O_2=O$,$S_1=O_2^{\mathrm{T}}\operatorname{diag}(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0)O_2$,则方阵 O 是正交的,方阵 S_1 是半正定对称方阵,并且 $A=OS_1$(注意,在 $A=SO$ 与 $A=OS_1$ 中,方阵 O 允许取成同一个正交方阵). 由 $A^{\mathrm{T}}A=S_1^2$ 即得方阵 S_1 的唯一性. ■

应当指出,当方阵 A 可逆时,使得 $A=SO$ 成立的正交方阵 O 也是唯一的; 而当方阵 A 不可逆时,使得 $A=SO$ 成立的正交方阵 O 不再唯一.

例 7.7.1　设 A 与 B 分别是 $m\times n$ 与 $n\times p$ 实矩阵. 证明,

$$\operatorname{Tr}(AB)(AB)^{\mathrm{T}}\leqslant\operatorname{Tr}(AA^{\mathrm{T}})\max\left\{\lambda(BB^{\mathrm{T}})\right\},$$

其中 $\max\left\{\lambda(BB^{\mathrm{T}})\right\}$ 表示方阵 BB^{T} 的最大特征值.

证明　设 $n\times p$ 矩阵 B 的所有奇异值为 $\mu_1,\mu_2,\ldots,\mu_r$,且 $\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_r>0$. 则由定理 7.7.6,存在 n 阶与 p 阶实正交方阵 O_1 与 O_2,使得

$$B=O_1\operatorname{diag}\left(\operatorname{diag}(\mu_1,\mu_2,\ldots,\mu_r),0\right)O_2.$$

其中 0 是一个 $(n-r)\times(p-r)$ 的零矩阵. 因此

$$BB^{\mathrm{T}}=O_1\operatorname{diag}\big(\mu_1^2,\mu_2^2,\ldots,\mu_r^2,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big)O_1^{\mathrm{T}}.$$

所以

$$(AB)(AB)^{\mathrm{T}}=AO_1\operatorname{diag}\big(\mu_1^2,\mu_2^2,\ldots,\mu_r^2,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big)(AO_1)^{\mathrm{T}}.$$

记 $AO_1=(a_{ij})_{m\times n}$,则

$$\operatorname{Tr}(AB)(AB)^{\mathrm{T}}=\sum_{i=1}^{m}\sum_{j=1}^{r}\mu_j^2a_{ij}^2\leqslant\mu_1^2\sum_{i=1}^{m}\sum_{j=1}^{r}a_{ij}^2\leqslant\mu_1^2\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}^2.$$

由奇异值的定义,$\mu_1^2=\max\left\{\lambda(BB^{\mathrm{T}})\right\}$,而

$$\operatorname{Tr}(AA^{\mathrm{T}})=\operatorname{Tr}(AO_1)(AO_1)^{\mathrm{T}}=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}^2,$$

因此

$$\operatorname{Tr}(AB)(AB)^{\mathrm{T}}\leqslant\max\left\{\lambda(BB^{\mathrm{T}})\right\}\operatorname{Tr}(AA^{\mathrm{T}}).$$ ■

习　题　7.7

为方便起见,引用如下记号: $A>0$ 表示 A 是正定对称方阵; $A\geqslant 0$ 表示 A 是半正定对称方阵; $A>B$ 表示对称方阵 A 与 B 之差 $A-B>0$; $A\geqslant B$ 表示对称方阵 A 与 B 之差 $A-B\geqslant 0$.

1. 设 A 与 B 是 n 阶方阵,$A\geqslant 0$,$B\geqslant 0$,且 A^2 与 B^2 正交相似. 证明方阵 A 与 B 正交相似.

2. 设 S 是 n 阶对称方阵. 证明存在唯一的 n 阶对称方阵 S_1,使得 $S=S_1^3$. 方阵 S_1 称为方阵 S 的立方根,记为 $\sqrt[3]{S}$.

3. 设 $A>0, B>0$. 证明 AB 的所有特征值都是正的.

4. 设 $S>0$. 证明存在可逆上三角阵 P,使得 $S=P^{\mathrm{T}}P$.

5. 设 n 阶方阵 $S\geqslant 0$,且 $\operatorname{rank} S=1$. 证明存在非零行向量 $x\in\mathbb{R}^n$,使得 $S=x^{\mathrm{T}}x$.

6. 设 n 阶对称方阵 S 的前 $n-1$ 个顺序主子式大于零,并且 $\det S\geqslant 0$. 证明 $S\geqslant 0$.

7. 设 $A\geqslant 0, B\geqslant 0$. 证明 $\det(A+B)\geqslant \det A$.

8. 设 $S\geqslant 0$. 证明
$$\det S\leqslant S\begin{pmatrix}1&2&\cdots&k\\1&2&\cdots&k\end{pmatrix}S\begin{pmatrix}k+1&k+2&\cdots&n\\k+1&k+2&\cdots&n\end{pmatrix},$$
其中 $k=1,2,\ldots,n$.

9. 设 $S=(a_{ij})_{n\times n}>0$. 证明
$$\det S\leqslant a_{11}a_{22}\cdots a_{nn},$$
并且等式当且仅当方阵 S 为对角方阵时成立.

10. (Hadamard 不等式) 设 $A=(a_{ij})$ 是 n 阶实方阵. 证明
$$\det A\leqslant \prod_{j=1}^{n}\Big(\sum_{i=1}^{n}a_{ij}^2\Big)^{\frac{1}{2}},$$
其中等式当且仅当方阵 A 的 n 个列向量两两正交时成立.

11. 设 S_1 与 S_2 是 n 阶对称方阵,$S_1\geqslant 0$,且 $\det(S_1+\mathrm{i}S_2)=0$,其中 $\mathrm{i}^2=-1$. 证明,存在非零实行向量 $x\in\mathbb{R}^n$,使得 $x(S_1+\mathrm{i}S_2)=0$.

12. 设 $S>0$. 证明,对任意实行向量 x 与 y,
$$(xSy^{\mathrm{T}})^2\leqslant(xSx^{\mathrm{T}})(ySy^{\mathrm{T}}),$$
其中等式当且仅当向量 x 与 y 线性相关时成立.

13. 设 n 阶实方阵 A 的极分解唯一. 证明方阵 A 可逆.

14. 证明,n 阶实方阵 A 规范的充分必要条件是,方阵 A 具有极分解 $A=OS=SO$,其中 $S\geqslant 0$,O 为正交方阵.

15. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 的线性变换. 证明存在 V 的部分正交变换 $\mathscr{B}$ 和半正定自伴变换 $\mathscr{C}$,其中 $\operatorname{Ker}\mathscr{B}=\operatorname{Ker}\mathscr{C}$,使得 $\mathscr{A}=\mathscr{B}\mathscr{C}$,并且变换 $\mathscr{B}$ 与 $\mathscr{C}$ 是唯一的; 证明线性变换 $\mathscr{A}$ 规范的充分必要条件是,变换 $\mathscr{B}$ 与 $\mathscr{C}$ 可交换.

16. 证明,任意 n 阶实方阵都可以分解为三个 n 阶实对称方阵的乘积.

17. 设 A 与 B 是 $m\times n$ 实矩阵. 证明,$AA^{\mathrm{T}}=BB^{\mathrm{T}}$ 的充分必要条件是 $A=BO$,其中 O 是某个 n 阶正交方阵.

18. 证明,实方阵 A 的所有奇异值恰是所有非零特征值的充分必要条件是 $A\geqslant 0$.

19. 设 $\sigma_1(A)\geqslant\sigma_2(A)\geqslant\cdots\geqslant\sigma_r(A)$ 是 n 阶实对称方阵 A 的所有奇异值. 证明,对于任意 $k\times n$ 阶矩阵 X,$k=1,2,\ldots,n$,有
$$\sup_X\frac{\sigma_k(XA)}{\sigma_1(X)}=\sigma_k(A),\qquad \inf_X\frac{\sigma_1(XA)}{\sigma_k(X)}=\sigma_{n+1-k}(A).$$

§7.8　方阵的正交相似

本节将用矩阵方法讨论方阵在正交相似下的标准形. 下面是关于方阵的正交相似的一个重要定理.

定理 7.8.1　设 $a_1 \pm \mathrm{i}b_1, a_2 \pm \mathrm{i}b_2, \dots, a_s \pm \mathrm{i}b_s, \lambda_{2s+1}, \lambda_{2s+2}, \dots, \lambda_n$ 是 n 阶实方阵 A 的所有特征值,其中 $\mathrm{i}^2 = -1, a_1, a_2, \dots, a_s, b_1, b_2, \dots, b_s, \lambda_{2s+1}, \lambda_{2s+2}, \dots, \lambda_n$ 都是实数,且 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_s > 0$. 则方阵 A 正交相似于如下的准下三角形:

$$\begin{pmatrix} A_1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ A_{21} & A_2 & \ddots & & & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & & & & \vdots \\ A_{s1} & A_{s2} & \cdots & A_s & \ddots & & & \vdots \\ \alpha_{2s+1,1} & \alpha_{2s+1,2} & \cdots & \alpha_{2s+1,s} & \lambda_{2s+1} & \ddots & & \vdots \\ \alpha_{2s+2,1} & \alpha_{2s+2,2} & \cdots & \alpha_{2s+2,s} & \beta_{2s+2,2s+1} & \lambda_{2s+2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & 0 \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{ns} & \beta_{n,2s+1} & \cdots & \beta_{n,n-1} & \lambda_n \end{pmatrix}. \tag{7.8.1}$$

其中 $A_{k\ell}$ 是 2 阶子方阵,$\alpha_{k\ell}$ 是 1×2 矩阵,$\beta_{k\ell}$是实数,并且对任意 $j = 1, 2, \dots, s$,存在 2 阶可逆方阵 P_j,使得

$$A_j = P_j^{-1} \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} P_j.$$

证明　对方阵的阶数 n 用归纳法. 当 $n = 1$ 时定理显然成立. 设 $n = 2$. 如果 2 阶方阵 A 的特征值 λ_1 与 λ_2 都是实数,则存在属于特征值 λ_2 的单位特征(列)向量 $\alpha_2 \in \mathbb{R}^2$,使得 $A\alpha_2 = \lambda_2\alpha_2$.

把单位列向量 α_2 扩充成 $\mathbb{R}^2$ 的标准正交基 $\{\alpha_1, \alpha_2\}$,则 $O = (\alpha_1, \alpha_2)$ 是 2 阶正交方阵,

$$AO = A(\alpha_1, \alpha_2) = (\alpha_1, \alpha_2)\begin{pmatrix} a & 0 \\ c & \lambda_2 \end{pmatrix} = O\begin{pmatrix} a & 0 \\ c & \lambda_2 \end{pmatrix}.$$

显然 $a = \lambda_1$,且

$$A = O\begin{pmatrix} \lambda_1 & 0 \\ c & \lambda_2 \end{pmatrix}O^{\mathrm{T}}.$$

因此当 A 的特征值都是实数时定理成立.

如果 A 的特征值都不是实数,则可设 A 的特征值为 $a \pm \mathrm{i}b$. 由于 A 的特征值都不相同,所以 A 相似于对角形 $\mathrm{diag}(a + \mathrm{i}b, a - \mathrm{i}b)$. 显然方阵

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

的特征值也是 $a \pm \mathrm{i}b$,所以方阵 A 与之相似. 由 §6.7,方阵 A 与之实相似. 这就证

明，当 $n=2$ 时定理成立. 现在假设定理对阶数小于 n 的方阵成立. 下面证明定理对 n 成立.

情形 1 n 阶方阵 A 具有实特征值 λ_n. 此时可设 α_n 是属于特征值 λ_n 的单位特征(列)向量，即 $A\alpha_n=\lambda_n\alpha_n$. 把 α_n 扩充成 n 维实列向量空间 $\mathbb{R}^n$ 的标准正交基 $\big(\alpha_1,\alpha_2,\ldots,\alpha_n\big)$. 则 $O=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ 是 n 阶正交方阵，并且

$$AO=\big(A\alpha_1,A\alpha_2,\ldots,A\alpha_n\big)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\begin{pmatrix}A_1 & 0\\ \beta & \lambda_n\end{pmatrix}=O\begin{pmatrix}A_1 & 0\\ \beta & \lambda_n\end{pmatrix},$$

其中 A_1 是 $n-1$ 阶实方阵，β 是 $n-1$ 维行向量. 因此

$$A=O\begin{pmatrix}A_1 & 0\\ \beta & \lambda_n\end{pmatrix}O^{\mathrm{T}}.$$

由归纳假设，存在 $n-1$ 阶正交方阵 O_1，使得 $A_1=O_1B_1O_1^{\mathrm{T}}$，其中 B_1 是形如 (7.8.1) 的准下三角形. 于是

$$A=O\begin{pmatrix}O_1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}B_1 & 0\\ \beta O_1 & \lambda_n\end{pmatrix}\begin{pmatrix}O_1^{\mathrm{T}} & 0\\ 0 & 1\end{pmatrix}O^{\mathrm{T}}.$$

记

$$O_2=O\begin{pmatrix}O_1 & 0\\ 0 & 1\end{pmatrix}.$$

显然 O_2 是正交方阵，且

$$A=O_2\begin{pmatrix}B_1 & 0\\ \beta O_1 & \lambda_n\end{pmatrix}O_2^{\mathrm{T}}.$$

这表明，定理对情形 1 成立.

情形 2 n 阶方阵 A 不具有实特征值 λ_n. 此时 $n=2s$. 设 α_s 是属于特征值 $a_s+\mathrm{i}b_s$ 的特征(列)向量，即 $A\alpha_s=(a_s+\mathrm{i}b_s)\alpha_s$，$\alpha_s\in\mathbb{C}^s$. 记 $\alpha_s=\beta_s+\mathrm{i}\gamma_s$，其中 $\beta_s,\gamma_s\in\mathbb{R}^n$. 则由 $A\alpha_s=(a_s+\mathrm{i}b_s)\alpha_s$ 得到，

$$A(\beta_s,\gamma_s)=(\beta_s,\gamma_s)\begin{pmatrix}a_s & b_s\\ -b_s & a_s\end{pmatrix}.$$

如果 $\lambda\beta_s+\mu\gamma_s=0$，其中 λ,μ 是实数，则

$$A(\beta_s,\gamma_s)\begin{pmatrix}\lambda\\ \mu\end{pmatrix}=(\beta_s,\gamma_s)\begin{pmatrix}a_s & b_s\\ -b_s & a_s\end{pmatrix}\begin{pmatrix}\lambda\\ \mu\end{pmatrix}=0.$$

即得

$$a_s(\lambda\beta_s+\mu\gamma_s)+b_s(\mu\beta_s-\lambda\gamma_s)=0.$$

因为 $\lambda\beta_s+\mu\gamma_s=0$，所以 $b_s(\mu\beta_s-\lambda\gamma_s)=0$. 由于 $b_s>0$，因此

$$\begin{cases}\lambda\beta_s+\mu\gamma_s=0,\\ \mu\beta_s-\lambda\gamma_s=0.\end{cases}$$

由此得到，$(\lambda^2+\mu^2)\beta_s=0$，$(\lambda^2+\mu^2)\gamma_s=0$. 由于 $\alpha\in\mathbb{C}^n$ 是非零的，因此 $\beta_s,\gamma_s\in\mathbb{R}^n$

至少有一个是非零的，所以 $\lambda^2+\mu^2=0$，即 $\lambda=\mu=0$. 这表明，列向量 β_s,γ_s 线性无关. 于是 $\mathbb{R}^n$ 中由向量 β_s 与 γ_s 生成的子空间 V_s 是 2 维的.

设 $\{\xi_{2s-1},\xi_{2s}\}$ 是 V_s 的标准正交基. 则

$$(\xi_{2s-1},\xi_{2s})=(\beta_s,\gamma_s)P_s,$$

其中过渡矩阵 P_s 是 2 阶可逆方阵，并且

$$A(\xi_{2s-1},\xi_{2s})=(\xi_{2s-1},\xi_{2s})P_s^{-1}\begin{pmatrix}a_s & b_s\\ -b_s & a_s\end{pmatrix}P_s.$$

记

$$A_s=P_s^{-1}\begin{pmatrix}a_s & b_s\\ -b_s & a_s\end{pmatrix}P_s,$$

并将 $\{\xi_{2s-1},\xi_{2s}\}$ 扩充成 $\mathbb{R}^n$ 的标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_{2s-1},\xi_{2s}\}$. 则

$$A(\xi_1,\xi_2,\ldots,\xi_{2s-1},\xi_{2s})=(\xi_1,\xi_2,\ldots,\xi_{2s-1},\xi_{2s})\begin{pmatrix}\widetilde{A} & 0\\ B & A_s\end{pmatrix},$$

其中 $\widetilde{A}$ 是 $2s-2$ 阶方阵. 由于 $\{\xi_1,\xi_2,\ldots,\xi_{2s-1},\xi_{2s}\}$ 是 $\mathbb{R}^n$ 的标准正交基，所以 n 阶方阵 $O=(\xi_1,\xi_2,\ldots,\xi_{2s-1},\xi_{2s})$ 是正交的. 由上式，

$$A=O\begin{pmatrix}\widetilde{A} & 0\\ B & A_s\end{pmatrix}O^{\mathrm{T}}.$$

对 $2s-2$ 阶方阵 $\widetilde{A}$ 用归纳假设，则

$$\widetilde{A}=O_1\begin{pmatrix}A_1 & & 0\\ \vdots & \ddots & \\ A_{n-1,1} & \cdots & A_{n-1}\end{pmatrix}O_1^{\mathrm{T}},$$

其中 O_1 是 $2s-2$ 阶正交方阵，而 $A_1,\ldots,A_{s-1}$ 是 2 阶方阵. 于是

$$A=O\begin{pmatrix}O_1 & 0\\ 0 & I_{(2)}\end{pmatrix}\begin{pmatrix}\begin{pmatrix}A_1 & & 0\\ \vdots & \ddots & \\ A_{n-1,1} & \cdots & A_{n-1}\end{pmatrix} & 0\\ BO_1 & A_s\end{pmatrix}\begin{pmatrix}O_1^{\mathrm{T}} & 0\\ 0 & I_{(2)}\end{pmatrix}O^{\mathrm{T}}.$$

记

$$O_2=O\begin{pmatrix}O_1 & 0\\ 0 & I_{(2)}\end{pmatrix},$$

则

$$A=O_2\begin{pmatrix}\begin{pmatrix}A_1 & & 0\\ \vdots & \ddots & \\ A_{n-1,1} & \cdots & A_{n-1}\end{pmatrix} & 0\\ BO_1 & A_s\end{pmatrix}O_2^{\mathrm{T}}.$$

于是定理对情形 2 成立. 这就完成了证明. ■

尽管定理 7.8.1 没有解决方阵在正交相似下的标准形问题,但它却有许多重要的应用.

定理 7.8.2(Schur 定理) 设 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 是 n 阶实方阵 A 的全部特征值. 则

$$\operatorname{Tr} AA^{\mathrm{T}} \geqslant \sum_{j=1}^{n}|\lambda_j|^2, \tag{7.8.2}$$

并且等式当且仅当方阵 A 为规范方阵时成立.

证明 因为方阵 A 的迹 $\operatorname{Tr} A$ 是方阵在相似下的不变量,所以对任意 n 阶正交方阵 O,

$$\operatorname{Tr} AA^{\mathrm{T}} = \operatorname{Tr}(O^{\mathrm{T}}AO)(O^{\mathrm{T}}AO).$$

由定理 7.8.1,存在 n 阶正交方阵 O,使得 $O^{\mathrm{T}}AO$ 为准下三角形 (7.8.1),其中 $\lambda_{2j-1},\lambda_{2j}$ 为 $a_j \pm \mathrm{i}b_j$. 因此

$$\operatorname{Tr} AA^{\mathrm{T}} = \sum_{j=1}^{s}\operatorname{Tr} A_jA_j^{\mathrm{T}} + \sum_{j=2s+1}^{n}\lambda_j^2 + \sigma,$$

其中 σ 是准下三角形 $O^{\mathrm{T}}AO$ 的非对角块上元素的平方和. 所以

$$\operatorname{Tr} AA^{\mathrm{T}} \geqslant \sum_{j=1}^{s}\operatorname{Tr} A_jA_j^{\mathrm{T}} + \sum_{j=2s+1}^{n}\lambda_j^2.$$

由于每个 A_j 是 2 阶的,并且特征值 λ_{2j-1} 与 λ_{2j} 为 $a_j \pm \mathrm{i}b_j, b_j > 0$,所以

$$\det A_j = (a_j + \mathrm{i}b_j)(a_j - \mathrm{i}b_j) = \frac{1}{2}\left(|\lambda_{2j-1}|^2 + |\lambda_{2j}|^2\right).$$

记

$$A_j = \begin{pmatrix} c_j & d_j \\ f_j & g_j \end{pmatrix},$$

则

$$\det A_j = c_jg_j - d_jf_j \leqslant \frac{1}{2}\left(c_j^2 + d_j^2 + f_j^2 + g_j^2\right) = \frac{1}{2}\operatorname{Tr} AA^{\mathrm{T}}.$$

因此,

$$\operatorname{Tr} A_jA_j^{\mathrm{T}} \geqslant |\lambda_{2s-1}|^2 + |\lambda_{2j}|^2,$$

其中等式当且仅当 $2(c_jg_j - d_jf_j) = c_j^2 + d_j^2 + f_j^2 + g_j^2$,即 $c_j = g_j, f_j = -d_j$,也即 A_j 为规范方阵时成立. 所以

$$\operatorname{Tr} AA^{\mathrm{T}} \geqslant \sum_{j=1}^{n}|\lambda_j|^2,$$

其中当且仅当 $O^{\mathrm{T}}AO$ 为准对角方阵而且每个 2 阶对角块为规范方阵,即 $O^{\mathrm{T}}AO$ 为规范方阵时等式成立. 而 $O^{\mathrm{T}}AO$ 为规范方阵显然当且仅当 A 为规范方阵. ■

定理 7.8.3 设 $a_1 \pm \mathrm{i}b_1, a_2 \pm \mathrm{i}b_2, \ldots, a_s \pm \mathrm{i}b_s, \lambda_{2s+1}, \lambda_{2s+2}, \ldots, \lambda_n$ 是 n 阶规范方阵 A 的全部特征值,其中 $\mathrm{i}^2 = -1, a_1, a_2, \ldots, a_s, b_1, b_2, \ldots, b_s, \lambda_{2s+1}, \lambda_{2s+2}, \ldots, \lambda_n$ 都是实数,且 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_s > 0$. 则方阵 A 正交相似于如下的准对角形:

$$\operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \ldots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{2s+1}, \lambda_{2s+2}, \ldots, \lambda_n\right), \tag{7.8.3}$$

并且规范方阵的特征值是规范方阵在正交相似下的全系不变量.

证明　因为方阵是规范的,所以 Schur 定理中式 (7.8.2) 取等式. 由 Schur 定理的证明可知,式 (7.8.2) 取等式时,准下三角形 $O^{\mathrm{T}}AO$ 为准对角方阵,而且每个 2 阶对角块 A_j 应具有形式

$$A_j=\begin{pmatrix} c_j & d_j \\ -d_j & c_j \end{pmatrix}.$$

由于对角块 A_j 的特征值为 $a_j\pm \mathrm{i}b_j$,所以 $c_j=a_j$,且 $d_j=\pm b_j$,可设 $d_j=b_j$. 于是 $O^{\mathrm{T}}AO$ 记为准对角阵 (7.8.3).

至于规范方阵的特征值是规范方阵在正交相似下的全系不变量,定理 7.4.7 已经证明,这里不再赘述. ■

利用定理 7.8.3 就可以得对称方阵、斜对称方阵与正交方阵在正交相似下的标准形.

例 7.8.1　设 A 与 B 都是 n 阶规范方阵,且 AB 也是规范方阵. 证明 BA 是规范方阵.

证明　记方阵 A 的所有特征值为 $\lambda_1(A),\lambda_2(A),\dots,\lambda_n(A)$. 由 Schur 定理,只需证明,

$$\operatorname{Tr}(BA)(BA)^{\mathrm{T}}=\sum_{j=1}^{n}\left|\lambda_j(BA)\right|^2.$$

事实上,　$\operatorname{Tr}(BA)(BA)^{\mathrm{T}}=\operatorname{Tr}BAA^{\mathrm{T}}B^{\mathrm{T}}$

由于对任意同阶方阵 C 与 D,$\operatorname{Tr}CD=\operatorname{Tr}DC$,所以

$$=\operatorname{Tr}AA^{\mathrm{T}}B^{\mathrm{T}}B$$

因为方阵 A 与 B 是规范的,即 $AA^{\mathrm{T}}=A^{\mathrm{T}}A$,$BB^{\mathrm{T}}=B^{\mathrm{T}}B$,所以

$$=\operatorname{Tr}A^{\mathrm{T}}ABB^{\mathrm{T}}=\operatorname{Tr}ABB^{\mathrm{T}}A^{\mathrm{T}}=\operatorname{Tr}(AB)(AB)^{\mathrm{T}}$$

因为方阵 AB 是规范的,所以由 Schur 定理,

$$=\sum_{j=1}^{n}\left|\lambda_j(AB)\right|^2$$

由于方阵 AB 与 BA 的特征多项式相同,所以方阵 AB 与 BA 的特征值相同,因此

$$=\sum_{j=1}^{n}\left|\lambda_j(BA)\right|^2.$$

于是由 Schur 定理,方阵 BA 也是规范的. ■

习　题　7.8

1. 设 λ 是 n 阶实方阵 $A=(a_{ij})$ 的特征值. 记 $a=\max\left\{|a_{ij}|\;\middle|\;1\leqslant i,j\leqslant n\right\}$. 利用 Schur 定理证明,

$$|\lambda|\leqslant na.$$

2. 设 $\lambda_1(A),\lambda_2(A),\dots,\lambda_n(A)$ 是 n 阶实方阵 $A=(a_{ij})$ 的特征值. 证明 Schur 不等式:

$$\sum_{j=1}^{n}\left(\operatorname{Re}\lambda_j(A)\right)^2\leqslant\sum_{1\leqslant i,j\leqslant n}\left|\frac{a_{ij}+a_{ji}}{2}\right|^2;$$

$$\sum_{j=1}^{n}\big(\mathrm{Im}\,\lambda_j(A)\big)^2 \leqslant \sum_{1\leqslant i,j\leqslant n}\left|\frac{a_{ij}-a_{ji}}{2}\right|^2.$$

§7.9 一些例子

例 7.9.1 设 A 与 B 是 n 阶实对称方阵,且 $A>0$. 证明方阵 A 与 B 可同时相合于对角形,即存在 n 阶可逆方阵 P,使得 $P^{\mathrm{T}}AP$ 与 $P^{\mathrm{T}}BP$ 都是对角方阵.

证明 因为 $A>0$,所以存在 n 阶可逆方阵 Q,使得 $A=Q^{\mathrm{T}}Q$. 记 $R=(Q^{\mathrm{T}})^{-1}$,则 $RAR^{\mathrm{T}}=I_{(n)}$. 方阵 RBR^{T} 显然是对称的. 由定理 7.6.3,存在 n 阶正交方阵 O,使得 $ORBR^{\mathrm{T}}O^{\mathrm{T}}$ 是对角方阵. 显然 $ORAR^{\mathrm{T}}O^{\mathrm{T}}=I_{(n)}$. 记 $P=R^{\mathrm{T}}O^{\mathrm{T}}$. 方阵 P 显然可逆,并且 $R^{\mathrm{T}}AP$ 与 $P^{\mathrm{T}}BP$ 都是对角方阵. ■

例 7.9.2 设 S 是 n 阶实对称方阵. 证明 $\mathrm{rank}\,S=n$ 的充分必要条件是,存在 n 阶方阵 A,使得 $SA+A^{\mathrm{T}}S$ 为正定对称方阵.

证明 设 $\mathrm{rank}\,S=n$,则方阵 S 的行列式不为零,所以方阵 S 的每个特征值都不为零. 不妨设方阵 S 的全部特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_p,\lambda_{p+1},\ldots,\lambda_n$,其中 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p>0>\lambda_{p+1}\geqslant\cdots\geqslant\lambda_n$.

于是存在 n 阶正交方阵 O,使得

$$S=O\,\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_p,\lambda_{p+1},\ldots,\lambda_n)O^{\mathrm{T}}.$$

记

$$Q=\mathrm{diag}\big(\sqrt{\lambda_1},\sqrt{\lambda_2},\ldots,\sqrt{\lambda_p},\sqrt{-\lambda_{p+1}},\ldots,\sqrt{-\lambda_n}\big).$$

方阵 Q 显然是可逆的. 记 $P=OQ$. 方阵 P 也是可逆的,并且

$$S=P\,\mathrm{diag}\big(I_{(p)},\ -I_{(n-p)}\big)P^{\mathrm{T}}.$$

令

$$A=\frac{1}{2}(P^{\mathrm{T}})^{-1}\big(I_{(p)},\ -I_{(n-p)}\big)P^{\mathrm{T}}.$$

则方阵 A 可逆,且 $SA+A^{\mathrm{T}}S=PP^{\mathrm{T}}$,即方阵 $SA+A^{\mathrm{T}}S>0$.

反之设存在 n 阶方阵 A,使得 $SA+A^{\mathrm{T}}S>0$. 则由定理 7.7.3,存在 n 阶可逆方阵 P,使得

$$(P^{\mathrm{T}})^{-1}(SA+A^{\mathrm{T}}S)P^{-1}=I_{(n)}.$$

设 $\mathrm{rank}\,S=r<n$,则 $\mathrm{rank}(P^{\mathrm{T}})^{-1}SP^{-1}=r$,且方阵 $(P^{\mathrm{T}})^{-1}SP^{-1}$ 是对称的. 因此存在 n 阶正交方阵 O,使得

$$O^{\mathrm{T}}(P^{\mathrm{T}})^{-1}SP^{-1}O=\mathrm{diag}(D_r,\ 0),$$

其中 $D_r=\mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_r)$. 记 $Q=P^{\mathrm{T}}O$. 则

$$S=Q\,\mathrm{diag}(D_r,\ 0)Q^{\mathrm{T}},\quad SA+A^{\mathrm{T}}S=QQ^{\mathrm{T}}.$$

因此

$$Q\,\mathrm{diag}(D_r,\ 0)Q^{\mathrm{T}}A + A^{\mathrm{T}}Q\,\mathrm{diag}(D_r,\ 0)Q^{\mathrm{T}} = QQ^{\mathrm{T}}.$$

因为方阵 Q 可逆，所以

$$\mathrm{diag}(D_r,\ 0)Q^{\mathrm{T}}A(O^{\mathrm{T}})^{-1} + Q^{-1}A^{\mathrm{T}}Q\,\mathrm{diag}(D_r,\ 0) = I_{(n)}. \tag{7.9.1}$$

将方阵 $Q^{\mathrm{T}}A(O^{\mathrm{T}})^{-1}$ 按方阵 $\mathrm{diag}(D_r,\ 0)$ 的分块方式分块，即记

$$Q^{\mathrm{T}}A(O^{\mathrm{T}})^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

其中 B_{11} 是 r 阶子矩阵. 于是由 (7.9.1) 得到

$$\begin{pmatrix} D_rB_{11} + B_{11}^{\mathrm{T}}D_r & D_rB_{12} \\ B_{12}^{\mathrm{T}}D_r & 0 \end{pmatrix} = I_{(n)}.$$

比较两端方阵的对角元，便得出矛盾. 因此 $\mathrm{rank}\,S = n$. ■

例 7.9.3　设 P, Q 与 R 是 n 阶实方阵，$P > 0, Q > 0$. 证明，方阵 $P - R^{\mathrm{T}}Q^{-1}R > 0$ 的充分必要条件是，方阵 $Q - RP^{-1}R^{\mathrm{T}} > 0$.

证明　因为方阵 P 与 Q 地位是对称的，所以只需证明，如果 $P - R^{\mathrm{T}}Q^{-1}R > 0$，则 $Q - RP^{-1}R^{\mathrm{T}} > 0$. 由于 $P > 0$，故存在 n 阶可逆方阵 T，使得 $P = T^{\mathrm{T}}T$. 因此，

$$P - R^{\mathrm{T}}Q^{-1}R = T^{\mathrm{T}}\big(I_{(n)} - (T^{\mathrm{T}})^{-1}R^{\mathrm{T}}Q^{-1}RT^{-1}\big)T.$$

记

$$A = (T^{\mathrm{T}})^{-1}R^{\mathrm{T}}, \qquad B = Q^{-1}RT^{-1}.$$

因为 $P - R^{\mathrm{T}}Q^{-1}R > 0$，所以方阵 $P - R^{\mathrm{T}}Q^{-1}R$ 可逆. 又方阵 T 可逆，所以方阵 $I_{(n)} - AB$ 可逆，因此方阵 $I_{(n)} - BA$ 也可逆，并且它的逆方阵为

$$\begin{aligned}
\big(I_{(n)} - BA\big)^{-1} &= I_{(n)} + B\big(I_{(n)} - AB\big)^{-1}A \\
&= I_{(n)} + Q^{-1}RT^{-1}\big(I_{(n)} - (T^{\mathrm{T}})^{-1}R^{\mathrm{T}}Q^{-1}RT^{-1}\big)^{-1}(T^{\mathrm{T}})^{-1}R^{\mathrm{T}} \\
&= Q^{-1}\big(Q + R(P - R^{\mathrm{T}}Q^{-1}R)^{-1}R^{\mathrm{T}}\big).
\end{aligned}$$

另一方面，

$$I_{(n)} - BA = I_{(n)} - Q^{-1}RT^{-1}(T^{\mathrm{T}})^{-1}R^{\mathrm{T}} = Q^{-1}(Q - RP^{-1}R^{\mathrm{T}}).$$

所以，

$$Q^{-1}RP^{-1}R^{\mathrm{T}} = Q\big((I_{(n)} - BA)^{-1}\big)^{-1} = Q\big(Q + R(P - R^{\mathrm{T}}Q^{-1}R)^{-1}R^{\mathrm{T}}\big)^{-1}Q. \tag{7.9.2}$$

因为 $P - R^{\mathrm{T}}Q^{-1}R > 0$，故它的逆方阵 $(P - R^{\mathrm{T}}Q^{-1}R)^{-1} > 0$. 因此

$$R(P - R^{\mathrm{T}}Q^{-1}R)^{-1}R^{\mathrm{T}} \geqslant 0.$$

容易证明，如果 $S_1 > 0, S_2 \geqslant 0$，则 $S_1 + S_2 > 0$. 因此方阵

$$Q + R(P - R^{\mathrm{T}}Q^{-1}R)^{-1}R^{\mathrm{T}} > 0.$$

由于方阵 $Q > 0$，所以方阵 Q 可逆，并且 $Q^{\mathrm{T}} = Q$. 于是由 (7.9.2)，

$$Q^{-1}RP^{-1}R^{\mathrm{T}} > 0.$$

■

例 7.9.4 设 $S=(a_{ij})_{n\times n}\geqslant 0$，且 S 的每个行和都为零，即对于 $i=1,\ldots,n$，

$$R_i=\sum_{j=1}^{n}a_{ij}=0.$$

证明，

$$2\max\left\{\sqrt{a_{ii}}\,\middle|\,1\leqslant i\leqslant n\right\}\leqslant\sum_{i=1}^{n}\sqrt{a_{ii}}.$$

证明 因为 $S\geqslant 0$，所以由定理 7.7.4，存在 n 阶方阵 P，使得 $S=P^{\mathrm{T}}P$. 将方阵 P 按列分块，即记 $P=(\alpha_1,\alpha_2,\ldots,\alpha_n)$，其中 $\alpha_j\in\mathbb{R}^n$. 于是

$$\begin{aligned}S=P^{\mathrm{T}}P&=\left(\alpha_1,\alpha_2,\ldots,\alpha_n\right)^{\mathrm{T}}\left(\alpha_1,\alpha_2,\ldots,\alpha_n\right)\\&=\begin{pmatrix}(\alpha_1,\alpha_1)&(\alpha_1,\alpha_2)&\cdots&(\alpha_1,\alpha_n)\\(\alpha_2,\alpha_1)&(\alpha_2,\alpha_2)&\cdots&(\alpha_2,\alpha_n)\\\vdots&\vdots&\ddots&\vdots\\(\alpha_n,\alpha_1)&(\alpha_n,\alpha_2)&\cdots&(\alpha_n,\alpha_n)\end{pmatrix},\end{aligned}$$

其中 $(\alpha_i,\alpha_j)=\alpha_i^{\mathrm{T}}\alpha_j$ 是 $\mathbb{R}^n$ 中向量 α_i 与 α_j 的标准内积. 由此得到 $a_{ij}=(\alpha_i,\alpha_j)$.

因为方阵 S 的每个行和都为零，所以

$$\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}(\alpha_i,\alpha_j)=\left(\alpha_i,\sum_{j=1}^{n}\alpha_j\right)=0.$$

因此

$$\sum_{i=1}^{n}\left(\alpha_i,\sum_{j=1}^{n}\alpha_j\right)=\left(\sum_{i=1}^{n}\alpha_i,\sum_{j=1}^{n}\alpha_j\right)=0.$$

于是

$$\sum_{i=1}^{n}\alpha_i=0,\quad\text{即}\quad\alpha_i=-\sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}}\alpha_j.$$

所以

$$\|\alpha_i\|\leqslant\sum_{\substack{1\leqslant j\leqslant n\\ j\neq i}}\|\alpha_j\|,\quad\text{即}\quad 2\|\alpha_i\|\leqslant\sum_{j=1}^{n}\|\alpha_j\|.$$

记 $\|\alpha_{i_0}\|=\max\left\{\|\alpha_j\|\,\middle|\,j=1,2,\ldots,n\right\}$. 则

$$2\|\alpha_{i_0}\|\leqslant\sum_{j=1}^{n}\|\alpha_j\|.$$

但是 $\|\alpha_j\|=\sqrt{(\alpha_j,\alpha_j)}=\sqrt{a_{jj}}$. 所以上式化为

$$2\max\left\{\sqrt{a_{ii}}\,\middle|\,1\leqslant i\leqslant n\right\}\leqslant\sum_{i=1}^{n}\sqrt{a_{ii}}.$$ ∎

例 7.9.5 设 A 是 n 阶实方阵. 证明

$$\operatorname{Tr}A\leqslant\operatorname{Tr}(AA^{\mathrm{T}})^{\frac{1}{2}},$$

其中等式当且仅当 $A\geqslant 0$ 时成立.

证明 设 $\mu_1,\mu_2,\ldots,\mu_r$ 是方阵 A 的全部奇异值. 则由定理 7.7.6，存在 n 阶正交方阵 O_1 与 O_2，使得

$$\begin{aligned}A &= O_1\,\mathrm{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big)O_2\\ &= O_1\Big(\mathrm{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big)O_2O_1\Big)O_1^{\mathrm{T}}.\end{aligned}$$

由于方阵 A 的迹 $\mathrm{Tr}\,A$ 是方阵在相似下的不变量,所以

$$\mathrm{Tr}\,A = \mathrm{Tr}\Big(\mathrm{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big)O_2O_1\Big).$$

记 $O = O_2O_1 = (b_{ij})$. 显然方阵 O 是正交的. 于是

$$\mathrm{Tr}\,A = \sum_{i=1}^{r}\mu_i b_{ii}.$$

由于方阵 O 是正交的,因此

$$|b_{ii}| \leqslant \Big(\sum_{j=1}^{n} b_{jj}^2\Big)^{\frac{1}{2}} = 1.$$

所以

$$\mathrm{Tr}\,A \leqslant \sum_{i=1}^{r}\mu_i.$$

另一方面,

$$AA^{\mathrm{T}} = O_1\,\mathrm{diag}\big(\mu_1^2,\mu_2^2,\ldots,\mu_r^2,0,\ldots,0\big)O_1^{\mathrm{T}}.$$

所以

$$(AA^{\mathrm{T}})^{\frac{1}{2}} = O_1\,\mathrm{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)O_1^{\mathrm{T}}.$$

因此

$$\mathrm{Tr}(AA^{\mathrm{T}})^{\frac{1}{2}} = \sum_{i=1}^{r}\mu_i \geqslant \mathrm{Tr}\,A.$$

如果 $A \geqslant 0$,则 $A^2 = AA^{\mathrm{T}}$,即 $A = (AA^{\mathrm{T}})^{\frac{1}{2}}$. 因此 $\mathrm{Tr}\,A = \mathrm{Tr}(AA^{\mathrm{T}})^{\frac{1}{2}}$,即等式成立.

反之设等式成立,则依上面的记号,

$$\sum_{i=1}^{r}\mu_i b_{ii} = \sum_{i=1}^{r}\mu_i,\quad \text{即}\quad \sum_{i=1}^{r}\mu_i(1-b_{ii}) = 0.$$

因为 $\mu_i > 0$,且 $1-b_{ii} \geqslant 0$,所以 $b_{ii} = 1, i = 1,2,\ldots,r$. 因此

$$O = O_2O_1 = \begin{pmatrix} I_{(r)} & 0\\ 0 & O_3\end{pmatrix},$$

其中 O_3 是 $n-r$ 阶正交方阵. 于是

$$\begin{aligned}A &= O_1\,\mathrm{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)\begin{pmatrix} I_{(r)} & 0\\ 0 & O_3\end{pmatrix}O_1^{\mathrm{T}}\\ &= O_1\,\mathrm{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)O_1^{\mathrm{T}}.\end{aligned}$$

即 $A \geqslant 0$. ■

例 7.9.6　设 S 是 n 阶实对称方阵,$\mathbb{R}^n$ 是 n 维实行向量空间. 记

$$V_0 = \big\{x \in \mathbb{R}^n \;\big|\; xSx^{\mathrm{T}} = 0\big\}.$$

证明,V_0 为 $\mathbb{R}^n$ 的子空间的充分必要条件是,$S \geqslant 0$,或者 $-S \geqslant 0$.

证明 设 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 是方阵 S 的全部非零特征值，则存在 n 阶正交方阵 O，使得

$$OSO^{\mathrm{T}}=\operatorname{diag}\big(\lambda_1,\lambda_2,\ldots,\lambda_r,\underbrace{0,\ldots,0}_{n-r\text{ 个}}\big).$$

记方阵 O 的 n 个行向量为 $\alpha_1,\alpha_2,\ldots,\alpha_n$，则 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 $\mathbb{R}^n$ 的标准正交基，并且

$$\alpha_jS=\lambda_j\alpha_j,\qquad j=1,2,\ldots,r;$$
$$\alpha_kS=0,\qquad k=r+1,r+2,\ldots,n.$$

设 $\alpha\in\mathbb{R}^n$，则 $\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n$，其中 $a_j\in\mathbb{R}$. 于是

$$\alpha S\alpha^{\mathrm{T}}=\lambda_1a_1^2+\lambda_2a_2^2+\cdots+\lambda_ra_r^2.\tag{7.9.3}$$

必要性 设 V_0 是 $\mathbb{R}^n$ 的子空间. 取 $\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n\in V_0$，由式 (7.9.3)，

$$\alpha S\alpha^{\mathrm{T}}=\lambda_1a_1^2+\lambda_2a_2^2+\cdots+\lambda_ra_r^2.$$

如果某个 $\alpha_j\neq0$，则取

$$\beta=-a_1\alpha_1-a_2\alpha_2-\cdots-a_{j-1}\alpha_{j-1}+a_j\alpha_j-a_{j+1}\alpha_{j+1}-\cdots-a_n\alpha_n.$$

于是

$$\beta S\beta^{\mathrm{T}}=\lambda_1a_1^2+\lambda_2a_2^2+\cdots+\lambda_ra_r^2=0.$$

所以 $\beta\in V_0$. 由于 V_0 是线性空间，所以 $\alpha+\beta=2a_j\alpha_j\in V_0$. 于是

$$(\alpha+\beta)S(\alpha+\beta)^{\mathrm{T}}=4\lambda_ja_j^2.$$

因此 $\lambda_j=0$. 这和 $\lambda_j\neq0$ 的假设相矛盾. 所以 $\alpha=a_{r+1}\alpha_{r+a}+\cdots+a_n\alpha_n$.

反之，由式 (7.9.3)，如果 $\alpha=a_{r+1}\alpha_{r+a}+\cdots+a_n\alpha_n$，则 $\alpha\in V_0$.

设方阵 S 的非零特征值不全同号，即设 $\lambda_i>0$，而 $\lambda_j=-\mu_j<0$，$1\leqslant i\leqslant j\leqslant r$. 取

$$\beta=\frac{1}{\sqrt{\lambda_i}}\alpha_i+\frac{1}{\sqrt{\mu_j}}\alpha_j,$$

则由式 (7.9.3)，

$$\beta S\beta^{\mathrm{T}}=\lambda_i\Big(\frac{1}{\sqrt{\lambda_i}}\Big)^2+\lambda_j\Big(\frac{1}{\sqrt{\mu_j}}\Big)^2=1-1=0,$$

即 $\beta\in V_0$. 另一方面，由上段证明，$\alpha\in V_0$ 当且仅当 $\alpha=a_{r+1}\alpha_{r+a}+\cdots+a_n\alpha_n$. 因此 $\beta\notin V_0$. 矛盾. 因此方阵 S 的所有非零特征值都同号. 这就证明，$S\geqslant0$，或者 $-S\geqslant0$.

充分性 由于 $S\geqslant0$，或者 $-S\geqslant0$，所以方阵 S 的全部特征值都同号. 由式 (7.9.3) 可知，$\alpha\in V_0$ 当且仅当 $\alpha=a_{r+1}\alpha_{r+1}+\cdots+a_n\alpha_n$. 这表明，$V_0$ 是 $\mathbb{R}^n$ 中由 $\alpha_{r+1},\ldots,\alpha_n$ 生成的子空间. ■

例 7.9.7 设方阵 $A=(a_{ij})_{m\times m}$，$B=(b_{ij})_{n\times n}$，mn 阶方阵

$$A\otimes B=\big(a_{ij}B\big)$$

称为方阵 A 与 B 的**张量积**或 **Kronecker 乘积**. 证明，当 $A\geqslant0$ 且 $B\geqslant0$ 时，$A\otimes B\geqslant0$.

证明 设 A,B 分别有正交相似标准形：

$$A=P\operatorname{diag}(\lambda_1,\ldots,\lambda_n)P^{-1},\quad B=Q\operatorname{diag}(\mu_1,\ldots,\mu_n)Q^{-1},$$

其中 P,Q 是正交方阵. 则 $A\otimes B$ 有正交相似标准形:

$$A\otimes B=(P\otimes Q)\big(\operatorname{diag}(\lambda_1,\ldots,\lambda_n)\otimes\operatorname{diag}(\mu_1,\ldots,\mu_n)\big)(P\otimes Q)^{-1},$$

其中 $P\otimes Q$ 仍是正交方阵. $A\otimes B$ 的所有特征值都是非负的,故 $A\otimes B\geqslant 0$. ■

例 7.9.8　设 $A=(a_{ij})_{n\times n}$ 与 $B=(b_{ij})_{n\times n}$ 是 n 阶方阵. 记

$$A\circ B=\begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n}\\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1}b_{n1} & a_{n2}b_{n2} & \cdots & a_{nn}b_{nn}\end{pmatrix}.$$

方阵 $A\circ B$ 称为方阵 A 与 B 的 **Hadamard 乘积**.

证明,当 $A\geqslant 0$ 与 $B\geqslant 0$ 时,Hadamard 乘积 $A\circ B\geqslant 0$.

证明　容易验证,方阵 $A\circ B$ 是对称的. 下面证明,方阵 $A\circ B$ 是半正定的.

事实上,因为 $A\geqslant 0$,所以存在 n 阶方阵 $P=(p_{ij})$,使得 $A=P^{\mathrm T}P$. 因此,

$$a_{ij}=\sum_{k=1}^n p_{ki}p_{kj}.$$

记 $A\circ B=(c_{ij})$,则

$$c_{ij}=a_{ij}b_{ij}=\sum_{k=1}^n p_{ki}p_{kj}b_{ij}.$$

设 $x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$,则

$$xA\circ Bx^{\mathrm T}=\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n b_{ij}x_ip_{ki}x_jp_{kj}=\sum_{k=1}^n\sum_{i=1}^n\sum_{j=1}^n b_{ij}x_ip_{ki}x_jp_{kj}.$$

因为

$$\begin{aligned} d_k&=\sum_{i=1}^n\sum_{j=1}^n b_{ij}x_ip_{ki}x_jp_{kj}\\ &=(x_1p_{k1},x_2p_{k2},\ldots,x_np_{kn})B(x_1p_{k1},x_2p_{k2},\ldots,x_np_{kn})^{\mathrm T},\end{aligned}$$

且 $B\geqslant 0$,所以 $d_k\geqslant 0,k=1,2,\ldots,n$. 因此 $xA\circ Bx^{\mathrm T}\geqslant 0$. 这就证明,$A\circ B\geqslant 0$. ■

例 7.9.9　设 A 是 n 阶三对角方阵,即设

$$A=\begin{pmatrix} a_1 & b_2 & & & & \\ c_2 & a_2 & b_3 & & & \\ & c_3 & a_3 & b_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & c_{n-1} & a_{n-1} & b_n\\ & & & & c_n & a_n\end{pmatrix},$$

其中 $b_jc_j>0$. 证明方阵 A 的全部特征值都是实数,并且互不相等.

证明　先考虑方阵 A 为对称的情形,即 $b_j=c_j,j=2,3,\ldots,n$. 此时方阵 A 的所有特征值都是实数. 下面证明,方阵 A 的特征值两两不同.

由于实对称方阵 A 正交相似于对角形,所以 A 相似于对角形. 因此方阵的最小多项式 $d(\lambda)$ 没有重根.

另一方面 A 的特征方阵 $\lambda I_{(n)}-A$ 的 $n-1$ 阶子式

$$\left(\lambda I_{(n)}-A\right)\left(\begin{smallmatrix}1 & 2 & \cdots & n-1\\ 2 & 3 & \cdots & n\end{smallmatrix}\right)=(-1)^{n-1}b_2b_3\cdots b_n\neq 0.$$

所以方阵 A 的 $n-1$ 阶行列式因子 $D_{n-1}(\lambda)=1$. 于是方阵 A 的不变因子为 $d_1(\lambda)=\cdots=d_{n-1}(\lambda)=1, d_n(\lambda)=D_n(\lambda)$. 由于方阵 A 的第 n 个不变因子 $d_n(\lambda)$ 即是方阵 A 的最小多项式 $d(\lambda)$. 所以 $d(\lambda)=D_n(\lambda)$.

这表明,$d(\lambda)$ 即是方阵 A 的特征多项式 $\varphi(\lambda)$,也即方阵 A 的所有特征值都是最小多项式 $d(\lambda)$ 的根. 而 $d(\lambda)$ 没有重根,所以方阵 A 的特征值两两不同.

现在考虑一般情形. 由于 $b_jc_j>0, j=2,3,\ldots,n$. 所以

$$\frac{b_2b_3\cdots b_j}{c_2c_3\cdots c_j}=\frac{(b_2c_2)(b_3c_3)\cdots(b_jc_j)}{c_2^2c_3^2\cdots c_j^2}>0.$$

因此

$$\sqrt{\frac{b_2b_3\cdots b_j}{c_2c_3\cdots c_j}}>0.$$

记

$$P=\operatorname{diag}\left(1,\sqrt{\frac{b_2}{c_2}},\ldots,\sqrt{\frac{b_2b_3\cdots b_j}{c_2c_3\cdots c_j}}\right),$$

则

$$PAP^{-1}=\begin{pmatrix} a_1 & \sqrt{b_2c_2} & & & & \\ \sqrt{b_2c_2} & a_2 & \sqrt{b_3c_3} & & & \\ & \sqrt{c_3b_3} & a_3 & \sqrt{b_4c_4} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \sqrt{b_{n-1}c_{n-1}} & a_{n-1} & \sqrt{b_nc_n} \\ & & & & \sqrt{b_nc_n} & a_n \end{pmatrix}.$$

记 $PAP^{-1}=S$. 则由上段证明,方阵 S 的所有特征值都是实数,并且两两不同. 由于方阵的特征值是相似不变量,所以方阵 A 的特征值都是实数,并且两两不同. ■

习 题 7.9

1. 设 $A>0$,且方阵 B 与 A 的 k 次幂可交换,k 为正整数. 证明,方阵 B 与 A 可交换.

2. 设 $A>0$,且 B 为对称方阵. 证明多项式 $\det(\lambda A-B)$ 的根都是实数.

3. 设 n 阶对称方阵 $A>0$,B 是 $n\times m$ 列满秩矩阵. 求 $n+m$ 阶对称方阵

$$C=\begin{pmatrix} A & B\\ B^{\mathrm T} & 0\end{pmatrix}$$

的逆方阵. 其中 0 是 m 阶零方阵.

4. 设 $A>0$,$B>0$,且方阵 A 与 B 可交换. 证明 $AB>0$.

5. 设 n 阶实方阵 A 的顺序主子式都不为零．证明，存在对角元全为 1 的 n 阶下三角方阵 P 与 Q，使得

$$A = P\operatorname{diag}(d_1, d_2, \ldots, d_n)Q^{\mathrm{T}},$$

其中对于任意 $k = 1, 2, \ldots, n$，

$$d_k = \frac{A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \cdots & k-1 \\ 1 & 2 & \cdots & k-1 \end{pmatrix}},$$

并约定 $A\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 1$.

6. 设 n 阶对称方阵 $A > 0$．证明，对任意行向量 $x, y \in \mathbb{R}^n$，

$$(xAx^{\mathrm{T}}) + (yA^{-1}y^{\mathrm{T}}) \geqslant 2xy^{\mathrm{T}}.$$

等式成立的充分必要条件是什么？

7. 证明，对任意非零行向量 $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$，均有

$$f(x) = \sum_{j=1}^{n} x_j^2 - \sum_{j=1}^{n-1} x_j x_{j+1} > 0.$$

求 $f(x)$ 的最小值，以及 $f(x)$ 在条件 $x_n = 1$ 下的最小值.

8. 设 n 阶对称方阵 $A > 0$．证明，在 n 维实行向量集合 $\mathbb{R}^n$ 连同标准内积构成的 Euclid 空间中，由不等式 $xAx^{\mathrm{T}} \leqslant 1$ 所定义的区域是有界的，并且它的体积 V 为

$$V = \int_{xAx^{\mathrm{T}} \leqslant 1} \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2} + 1\right)} (\det A)^{\frac{1}{2}},$$

其中 $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

9. 设 $A > 0, B > 0$．证明，方阵 A 与 B 的 Hadamard 乘积 $A \circ B > 0$.

10. 设 $A \leqslant B, C > 0$，且方阵 C 与 A 和 B 都可交换．证明 $AC \leqslant BC$.

11. 设 $0 \leqslant A \leqslant B$．证明 $\det B \geqslant \det A$.

12. 设 $0 < A \leqslant B$．证明 $B^{-1} \leqslant A^{-1}$.

13. 设 $0 \leqslant A \leqslant B$．证明 $\sqrt{A} \leqslant \sqrt{B}$.

14. 设 A 是对称方阵．记

$$|A| = \sqrt{A^2}, \quad A_+ = \frac{1}{2}(|A| + A), \quad A_- = \frac{1}{2}(|A| - A).$$

证明：

(1) $|A|$ 是满足 $A \leqslant |A|, -A \leqslant |A|$ 且与 A 可交换的最小的对称方阵，这里所谓“最小”是指，如果方阵 B 满足 $A \leqslant B, -A \leqslant B$，且与 A 可交换，则 $|A| \leqslant B$；

(2) A_+ 是满足 $A \leqslant A_+$ 且与 A 可交换的最小的半正定对称方阵；

(3) A_- 是满足 $A \leqslant A_-$ 且与 A 可交换的最小的半正定对称方阵；

(4) 设 A 与 B 是可交换的对称方阵，则存在适合 $A \leqslant C, B \leqslant C$ 且与 A 和 B 都可交换的最小对称方阵.

15. 证明，两个 n 阶半正定对称方阵 S_1 与 S_2 可以同时相合于对角形，即存在 n 阶可逆方阵 P，使得 $P^{\mathrm{T}}S_1P$ 与 $P^{\mathrm{T}}S_2P$ 都是对角方阵．(提示：方阵 $S = S_1 + S_2$ 是半正定的)

16. 正定对称方阵的概念可以推广[①]．设 A 是 n 阶实方阵(不必是对称的)，如果对任意非零行向量 $x \in \mathbb{R}^n$，$xAx^{\mathrm{T}} > 0$，则方阵 A 称为正定的．记 $A = S + K$，其中

$$S = \frac{1}{2}(A + A^{\mathrm{T}}), \qquad K = \frac{1}{2}(A - A^{\mathrm{T}}),$$

它们分别成为方阵 A 的对称部分与斜对称部分．证明：

(1) 方阵 A 正定的充分必要条件是，它的对称部分 S 是正定的；

① C. R. Johnson, *Positive Definite Matrices*, *Amer. Math. Monthly*, 1970; 77: 259-264.

(2) 设 $f_A(\lambda)=\det(\lambda S-K)$，则当 A 正定时，$f_A(\lambda)$ 的非零的根是纯虚数；

(3) 设 A 正定，并且 $f_A(\lambda)$ 的所有非零的根为 $\pm\mathrm{i}a_1,\pm\mathrm{i}a_2,\dots,\pm\mathrm{i}a_s$，$a_1\geqslant a_2\geqslant\cdots\geqslant a_s>0$，则 A 相合于如下的准对角形：

$$\operatorname{diag}\left(\begin{pmatrix}1 & a_1\\ -a_1 & 1\end{pmatrix},\dots,\begin{pmatrix}1 & a_s\\ -a_s & 1\end{pmatrix},\underbrace{1,\dots,1}_{n-2s\text{ 个}}\right).$$

并且 $f_A(\lambda)$ 的根是正定方阵在相合下的全系不变量.

17. 设 μ 是实数，C 是 n 阶实方阵，且 $A=\mu I_{(n)}+\mathrm{i}C$ 是 n 阶复正交方阵，即

$$AA^{\mathrm{T}}=I_{(n)}=A^{\mathrm{T}}A,$$

其中 $\mathrm{i}^2=-1$. 证明，方阵 C 是斜对称的，并且

(1) 当 $\operatorname{rank}C<n$ 时，$A=\pm I_{(n)}$；

(2) 当 $\operatorname{rank}C=n$ 时，存在 n 阶实正交方阵 O，使得

$$O^{\mathrm{T}}AO=\operatorname{diag}\Bigg(\underbrace{\begin{pmatrix}\mu & \mathrm{i}\sqrt{\mu^2-1}\\ -\mathrm{i}\sqrt{\mu^2-1} & \mu\end{pmatrix},\dots,\begin{pmatrix}\mu & \mathrm{i}\sqrt{\mu^2-1}\\ -\mathrm{i}\sqrt{\mu^2-1} & \mu\end{pmatrix}}_{\frac{n}{2}\text{ 个}}\Bigg),$$

18. 设 $A=B+\mathrm{i}C$ 是 n 阶复正交方阵，其中 B 与 C 是 n 阶实方阵. 证明，存在 n 阶实正交方阵 O_1 与 O_2，使得

$$O_1AO_2=\operatorname{diag}\left(\underbrace{1,\dots,1}_{n-2t\text{ 个}},\begin{pmatrix}\mu_1 & \mathrm{i}\sqrt{\mu_1^2-1}\\ -\mathrm{i}\sqrt{\mu_1^2-1} & \mu_1\end{pmatrix},\dots,\begin{pmatrix}\mu_t & \mathrm{i}\sqrt{\mu_t^2-1}\\ -\mathrm{i}\sqrt{\mu_t^2-1} & \mu_t\end{pmatrix}\right),$$

其中 $\operatorname{rank}C=2t$，而 $\mu_1,\dots,\mu_t$ 是方阵 B 的所有大于 1 的奇异值，且 1 是 B 的 $n-2t$ 重奇异值.

19. 设 $\lambda_1,\lambda_2,\dots,\lambda_n$ 是 n 阶半正定对称方阵 $A=(a_{ij})$ 的所有特征值，$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{n-1}\geqslant\lambda_n\geqslant 0$. 并且设方阵 A 的每个列和都是零. 证明，

$$\lambda_{n-1}\leqslant\frac{n}{n-1}\min\{a_{jj}\mid 1\leqslant j\leqslant n\}.$$

（提示：对称方阵 $A-\lambda_{n-1}\left(I_{(n)}-n^{-1}J\right)>0$，其中 J 是每个元素都为 1 的 n 阶方阵）

20. 设复方阵 A_1 和 A_2 适合 $A_2=O_1A_1O_2$，其中 O_1 与 O_2 是实正交方阵，则称复方阵 A_1 与 A_2 正交相抵. 证明，复正交方阵的实部的奇异值是复正交方阵在正交相抵下的全系不变量.

21. 设 A 是 n 阶正定对称方阵，$x\in\mathbb{R}^n$. 证明，

$$J_n=\int_{-\infty}^{\infty}\!\!\cdots\!\int_{-\infty}^{\infty}\mathrm{e}^{-(x,Ax)}\,\mathrm{d}x=\frac{\pi^{\frac{n}{2}}}{(\det A)^{\frac{1}{2}}},$$

其中 $(x,Ax)=xAx^{\mathrm{T}}$.

22. 设 A 与 B 是 n 阶实对称方阵，且方阵 A 是正定的. 证明，

$$\int_{-\infty}^{\infty}\!\!\cdots\!\int_{-\infty}^{\infty}\mathrm{e}^{-(x,(A+\mathrm{i}B)x)}\,\mathrm{d}x=\frac{\pi^{\frac{n}{2}}}{\left(\det(A+\mathrm{i}B)\right)^{\frac{1}{2}}},$$

其中 $\mathrm{i}^2=-1$，且 $\left(x,(A+\mathrm{i}B)x\right)=xAx^{\mathrm{T}}+\mathrm{i}xBx^{\mathrm{T}}$.

23. 设 A 与 B 为 n 阶实对称方阵，且方阵 A 是正定的. 证明，$|\det(A+\mathrm{i}B)|\geqslant\det A$，其中 $\mathrm{i}^2=-1$.

24. 设 n 阶对称方阵 A 是正定的. 去掉方阵 A 的第 i 行第 i 列的子矩阵记为 A_i. 记 $Q(x)=xAx^{\mathrm{T}}$，$x=(x_1,x_2,\dots,x_n)\in\mathbb{R}^n$. 证明，$Q(x)$ 在 $x_i=1$ 条件下的最小值是

$$\frac{\det A}{\det A_i}.$$

25. 设 A 与 B 是 n 阶正定对称方阵. A_i 与 B_i 分别是去掉方阵 A 与 B 的第 i 行第 i 列的子矩阵. 证明，

$$\frac{\det(A+B)}{\det(A_i+B_i)}\geqslant\frac{\det A}{\det A_i}+\frac{\det B}{\det B_i}.$$

26. 设 A 是 n 阶正定对称方阵. 证明,

$$\min\left\{\frac{1}{n}\operatorname{Tr}AB \;\middle|\; B\text{ 为 }n\text{ 阶正定方阵且 }\det B=1\right\}=(\det A)^{\frac{1}{n}}.$$

27. 设 A 与 B 是 n 阶正定对称方阵. 证明,

$$\big(\det(A+B)\big)^{\frac{1}{n}}\geqslant(\det A)^{\frac{1}{n}}+(\det B)^{\frac{1}{n}}.$$

28. 设 $\lambda_1(A),\lambda_2(A),\ldots,\lambda_n(A)$ 是 n 阶对称方阵 A 的特征值,并且 $\lambda_1(A)\geqslant\lambda_2(A)\geqslant\cdots\lambda_n(A)$. 证明,

(1) 设 A 与 B 是 n 阶对称方阵,实数 a 满足 $0\leqslant a\leqslant 1$,则

$$\lambda_1\big(aA+(1-a)B\big)\leqslant a\lambda_1(A)+(1-a)\lambda_1(B),$$
$$\lambda_n\big(aA+(1-a)B\big)\geqslant a\lambda_n(A)+(1-a)\lambda_n(B);$$

(2) 当 B 半正定时,

$$\lambda_1(A+B)\geqslant\lambda_1(A),\qquad \lambda_n(A+B)\geqslant\lambda_n(A).$$

§7.10　Euclid 空间的同构

和线性空间一样,需要考虑 Euclid 空间是否同构的问题. 由于一个 Euclid 空间是一个线性空间连同标准内积构成的,所以自然有如下的定义.

定义 7.10.1　设 σ 是 Euclid 空间 V 到 Euclid 空间 W 的一个映射. 如果映射 σ 是线性空间 V 到线性空间 W 上的同构映射,而且映射 σ 是保内积的,即对任意 $\alpha,\beta\in V$,均有

$$\big(\sigma(\alpha),\sigma(\beta)\big)=(\alpha,\beta),$$

其中 $\big(\sigma(\alpha),\sigma(\beta)\big)$ 是 W 中向量 $\sigma(\alpha)$ 与 $\sigma(\beta)$ 的内积,(α,β) 是 V 中向量 α 与 β 的内积,则 σ 称为 Euclid 空间 V 到 Euclid 空间 W 上的**同构映射**,而 Euclid 空间 V 与 W 称为**同构**的.

关于 Euclid 空间 V 到 W 上的同构映射 σ,有

定理 7.10.1　Euclid 空间 V 到 W 上的映射 σ 为同构映射的充分必要条件是,映射 σ 是线性空间 V 到 W 的同构映射,而且映射 σ 是保向量范数的,即对任意 $\alpha\in V$,均有

$$\|\sigma(\alpha)\|=\|\alpha\|.$$

证明　设 σ 是 Euclid 空间 V 到 W 上的同构映射. 由定义 7.10.1,映射 σ 是线性空间 V 到 W 上的同构映射,并且对任意 $\alpha,\beta\in V$,均有 $\big(\sigma(\alpha),\sigma(\beta)\big)=(\alpha,\beta)$. 因此对任意 $\alpha\in V$,均有 $\big(\sigma(\alpha),\sigma(\alpha)\big)=(\alpha,\alpha)$. 即 $\|\sigma(\alpha)\|=\|\alpha\|$. 所以映射 σ 是保向量范数的.

反之设 σ 是 V 到 W 上的同构映射,而且保向量范数. 因此对任意 $\alpha,\beta\in V$,有

$$\|\sigma(\alpha)\|=\|\alpha\|,\qquad\|\sigma(\beta)\|=\|\beta\|,\qquad\|\sigma(\alpha+\beta)\|=\|\alpha+\beta\|.$$

所以
$$\begin{aligned}\|\sigma(\alpha+\beta)\|^2 &= \big(\sigma(\alpha+\beta),\sigma(\alpha+\beta)\big)\\ &= \big(\sigma(\alpha),\sigma(\alpha)\big)+2\big(\sigma(\alpha),\sigma(\beta)\big)+\big(\sigma(\beta),\sigma(\beta)\big)\\ &= \|\sigma(\alpha)\|^2+2\big(\sigma(\alpha),\sigma(\beta)\big)+\|\sigma(\beta)\|^2\\ &= \|\alpha\|^2+2\big(\sigma(\alpha),\sigma(\beta)\big)+\|\beta\|^2\\ &= \|\alpha+\beta\|^2+2\Big(\big(\sigma(\alpha),\sigma(\beta)\big)-(\alpha,\beta)\Big).\end{aligned}$$

即对任意 $\alpha,\beta\in V$,
$$\big(\sigma(\alpha),\sigma(\beta)\big)=(\alpha,\beta).$$

即映射 σ 是保内积的,从而 σ 是 Euclid 空间 V 到 W 上的同构映射. ■

定理 7.10.2 设 σ 是 Euclid 空间 V 到 W 上的同构映射. 则映射 σ 是可逆的,并且它的逆映射 σ^{-1} 是 Euclid 空间 W 到 V 上的同构映射.

证明 由于 Euclid 空间 V 到 W 上的同构映射 σ 是线性空间 V 到 W 上的同构映射,所以 σ 是可逆的,并且它的逆映射 σ^{-1} 是线性空间 W 到 V 上的同构映射. 现在证明,逆映射 σ^{-1} 是保向量范数的.

事实上,对任意 $\beta\in W$,$\beta=\sigma\big(\sigma^{-1}(\beta)\big)$. 由于映射 σ 保向量范数,所以
$$\|\beta\|=\|\sigma\big(\sigma^{-1}(\beta)\big)\|=\|\sigma^{-1}(\beta)\|.$$

因此映射 σ^{-1} 保向量范数.

由定理 7.10.1,映射 σ^{-1} 是 Euclid 空间 W 到 V 上的同构映射. ■

定理 7.10.3 设 σ_1 与 σ_2 分别是 Euclid 空间 V 到 W 与 W 到 U 上的同构映射,则 $\sigma_2\sigma_1$ 是 Euclid 空间 V 到 U 上的同构映射.

证明 显然,σ_1 与 σ_2 分别是线性空间 V 到 W 与 W 到 U 上的同构映射,则 $\sigma_2\sigma_1$ 是线性空间 V 到 U 上的同构映射. 现在证明,映射 $\sigma_2\sigma_1$ 保向量范数.

设 $\alpha\in V$,由于 σ_1 保向量范数,所以 $\|\alpha\|=\|\sigma_1(\alpha)\|$. 因为 $\sigma_1(\alpha)\in W$,且 σ_2 保向量范数,因此 $\|\sigma_1(\alpha)\|=\|\sigma_2\sigma_1(\alpha)\|$. 所以对任意 $\alpha\in V$,$\|\sigma_2\sigma_1(\alpha)\|=\|\alpha\|$,即映射 $\sigma_2\sigma_1$ 保向量范数.

由定理 7.10.1,映射 $\sigma_2\sigma_1$ 是 Euclid 空间 V 到 U 上的同构映射. ■

所有有限维 Euclid 空间的集合记为 $\mathcal{E}$. 显然对任意 $V,W\in\mathcal{E}$,要么 V 与 W 同构,要么 V 与 W 不同构. 于是 Euclid 空间之间的同构关系是集合 $\mathcal{E}$ 中元素之间的一种关系.

容易验证,$\mathcal{E}$ 中元素间的同构关系满足自反性,对称性与传递性. 也即 Euclid 空间 V 与自身同构; 如果 Euclid 空间 V 与 W 同构,则 Euclid 空间 W 与 V 同构; 如果 Euclid 空间 V 与 W 同构,且 Euclid 空间 W 与 U 同构,则 Euclid 空间 V 与 U 同构. 于是 Euclid 空间之间的同构关系是集合 $\mathcal{E}$ 中元素间的一种等价关系. 集合 $\mathcal{E}$ 便按同构等价关系分成同构等价类: 彼此同构的 Euclid 空间归在同一个同构等价类,彼此不同的 Euclid 空间归在不同的同构等价类.

集合 $\mathcal{E}$ 按同构等价关系分类的两个基本问题是:

(1) 同构等价类的代表元是什么?

(2) 两个 Euclid 空间属于同一个等价类的判准是什么? 也即两个 Euclid 空间同构的充分必要条件是什么?

我们知道,所有 n 维实行向量集合 $\mathbb{R}^n$ 连同标准内积构成一个 n 维 Euclid 空间. 所谓 $\mathbb{R}^n$ 的标准内积是指,对 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)\in\mathbb{R}^n$,$x$ 与 y 的标准内积 (x,y) 为

$$(x,y)=x_1y_1+x_2y_2+\cdots+x_ny_n.$$

定理 7.10.4　任意 n 维 Euclid 空间 V 都和 n 维 Euclid 空间 $\mathbb{R}^n$ 同构.

证明　设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维 Euclid 空间 V 的一组标准正交基,则 $\alpha\in V$ 可以唯一地表为

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n,$$

其中 $x_1,x_2,\ldots,x_n\in\mathbb{R}$. 定义映射 $\sigma\colon V\to\mathbb{R}^n$ 如下: 设 $\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\in V$,则令

$$\sigma(\alpha)=x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n.$$

容易验证,映射 σ 是线性空间 V 到 $\mathbb{R}^n$ 上的同构映射. 现在证明,映射 σ 保向量范数. 设 $\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\in V$,则

$$\begin{aligned}(\alpha,\alpha)&=\left(\sum_{i=1}^{n}x_i\alpha_i,\ \sum_{i=j}^{n}x_j\alpha_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j(\alpha_i,\alpha_j)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j\delta_{ij}=\sum_{i=1}^{n}x_i^2=(x,x)=\big(\sigma(\alpha),\sigma(\alpha)\big).\end{aligned}$$

因此 $\|\sigma(\alpha)\|=\|\alpha\|$. 这就证明,映射 σ 保向量范数. 所以映射 σ 是 Euclid 空间 V 到 $\mathbb{R}^n$ 上的同构映射. ■

定理 7.10.5　有限维 Euclid 空间 V 与 W 同构当且仅当 $\dim V=\dim W$.

证明　设有限维 Euclid 空间 V 与 W 同构,则由定义 7.10.1,作为线性空间,V 与 W 同构. 因此 $\dim V=\dim W$.

反之,设 $\dim V=\dim W=n$,则由定理 7.10.4,Euclid 空间 V 同构于 Euclid 空间 $\mathbb{R}^n$,而 $\mathbb{R}^n$ 同构于 W. 由同构关系的传递性,Euclid 空间 V 与 W 同构. ■

定理 7.10.4 与定理 7.10.5 完全解决了所有有限维 Euclid 空间集合 $\mathcal{E}$ 在同构等价关系下分类的两个基本问题. 定理 7.10.5 表明,在 Euclid 空间的一个同构等价类中,所有 Euclid 空间的维数都相同. 定理 7.10.4 表明,如果 Euclid 空间的同构等价类中 Euclid 空间的维数为 n,则可取 Euclid 空间 $\mathbb{R}^n$ 作为这个同构等价类的代表元. 也即在同构意义下,n 维 Euclid 空间 V 可以视为 n 维实的行向量空间 $\mathbb{R}^n$.

第八章

酉 空 间

- 将上一章的概念与方法推广到复线性空间，便在复线性空间上引出相应的结果.
- §8.1 在复线性空间中引进了向量的内积，从而得到了酉空间的概念，并相应产生了复方阵在酉相似下的分类问题.
- §8.2 解决了规范方阵、酉方阵、Hermite 方阵以及斜 Hermite 方阵这些具有重要几何意义的方阵在酉相似下的分类问题.
- §8.3 进一步讨论了正定 Hermite 方阵，并通过奇异值的引进解决了方阵在酉相似下的分类问题.
- 最后用一些例子说明这些结论的应用.
- 由于酉空间是 Euclid 空间的推广，而且在处理有关酉空间的问题采用的方法也和处理 Euclid 空间中相应问题的方法大致相同，所有在提到酉空间有关结论时大部分只叙不证，一提而过.

§8.1 酉空间的定义

先给出复线性空间 V 上内积的定义.

定义 8.1.1 设 (α,β) 是复线性空间 V 上的二元复值函数. 如果 (α,β) 满足

(1) Hermite 对称性 对任意 $\alpha,\beta\in V$，$\overline{(\beta,\alpha)}=(\alpha,\beta)$，这里 $\overline{\mu}$ 表示复数 μ 的共轭复数；

(2) 恒正性 对任意非零向量 $\alpha\in V$，$(\alpha,\alpha)>0$；

(3) 共轭双线性性 对任意 $\alpha,\alpha_1,\alpha_2,\beta,\beta_1,\beta_2\in V$ 和任意 $\lambda_1,\lambda_2,\mu_1,\mu_2\in\mathbb{C}$，

$$(\lambda_1\alpha_1+\lambda_2\alpha_2,\beta)=\lambda_1(\alpha_1,\beta)+\lambda_2(\alpha_2,\beta);\tag{8.1.1}$$

$$(\alpha,\mu_1\beta_1+\mu_2\beta_2)=\overline{\mu}_1(\alpha,\beta_1)+\overline{\mu}_2(\alpha,\beta_2),\tag{8.1.2}$$

则二元复值函数 (α,β) 称为复线性空间 V 的一个**内积**.

应当指出，在复线性空间 V 的内积 (α,β) 的定义中，内积 (α,β) 的 Hermite 对称性与式 (8.1.1) 蕴涵式 (8.1.2). 为了叙述方便，这里还是把它列入定义中.

容易验证，复线性空间 V 的内积 (α,β) 具有下列性质.

性质 8.1.1 对任意 $\alpha,\beta\in V$，$(\alpha,0)=0=(\beta,0)$.

性质 8.1.2 对任意 $\alpha_1,\ldots,\alpha_p,\beta_1,\ldots,\beta_q\in V$ 和 $\lambda_1,\ldots,\lambda_p,\mu_1,\ldots,\mu_1\in\mathbb{C}$，

$$\Big(\sum_{i=1}^{p}\lambda_i\alpha_i,\ \sum_{j=1}^{q}\mu_j\beta_j\Big)=\sum_{i=1}^{p}\sum_{j=1}^{q}\lambda_i\overline{\mu}_j(\alpha_i,\beta_j).$$

性质 8.1.3（Cauchy-Schwartz 不等式） 对任意 $\alpha, \beta \in V$,

$$|(\alpha, \beta)|^2 \leqslant (\alpha, \alpha)(\beta, \beta),$$

并且等式当且仅当向量 α 与 β 线性相关时成立.

现在给出复线性空间 V 的内积 (α, β) 的方阵表示. 设 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 是 n 维复线性空间 V 的一组基. V 中向量 α 与 β 可以唯一地表为

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n,$$
$$\beta = y_1\alpha_1 + x_2\alpha_2 + \cdots + y_n\alpha_n.$$

则向量 α 与 β 的内积为

$$(\alpha, \beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y}_j(\alpha_i, \alpha_j). \tag{8.1.3}$$

记

$$G = \begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \cdots & (\alpha_1, \alpha_n) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \cdots & (\alpha_2, \alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_n, \alpha_1) & (\alpha_n, \alpha_2) & \cdots & (\alpha_n, \alpha_n) \end{pmatrix},$$

则 n 阶方阵 G 称为内积 (α, β) 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的 **Gram 方阵**.

记 $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$, 则式 (8.1.3) 可以写成

$$(\alpha, \beta) = xGy^*,$$

其中 y^* 是 $1 \times n$ 矩阵 y 的共轭转置.

内积 (α, β) 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的 Gram 方阵 G 具有下面的性质.

(1) Gram 方阵 G 是 Hermite 方阵, 即 $G^* = G$;

(2) 对任意非零行向量 $x \in \mathbb{C}^n$, $xGx^* > 0$.

设 H 是 n 阶 Hermite 方阵. 如果对任意非零行向量 $x \in \mathbb{C}^n$, $xHx^* > 0$, 则 H 称为**正定 Hermite 方阵**. Gram 方阵的上述性质表明, Gram 方阵 G 是一个正定 Hermite 方阵.

反之, 设 G 是一个 n 阶正定 Hermite 方阵. 在 n 维复线性空间 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下, 向量 $\alpha, \beta \in V$ 的坐标分别记为 $x, y \in \mathbb{C}^n$. 定义 V 上的二元复值函数 (α, β) 为

$$(\alpha, \beta) = xGy^*.$$

则容易验证, V 上二元复值函数 (α, β) 满足 Hermite 对称性, 恒正性以及共轭双线性性. 因而二元复值函数 (α, β) 是复线性空间 V 的一个内积.

这表明, 在 n 维复线性空间 V 中取定一组基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 之后, V 的内积 (α, β) 便和它在这组基下的 Gram 方阵建立了对应. 这一对应是 V 上所有内积的集合到所有 n 阶 Hermite 方阵集合上的一一对应.

现在讨论 n 维复线性空间 V 的一个内积 (α, β) 在不同基下的 Gram 方阵之间的关系.

设内积 (α,β) 在 V 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 下的 Gram 方阵分别为 G_1 与 G_2，由基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 到基 $\{\beta_1,\beta_2,\dots,\beta_n\}$ 的过渡矩阵为 $P=(p_{ij})$，即

$$(\beta_1,\beta_2,\dots,\beta_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)P.$$

因此，

$$\beta_j=\sum_{k=1}^n p_{kj}\alpha_k.$$

所以

$$(\beta_i,\beta_j)=\left(\sum_{k=1}^n p_{ki}\alpha_k,\ \sum_{\ell=1}^n p_{\ell j}\alpha_\ell\right)=\sum_{k=1}^n\sum_{\ell=1}^n p_{ki}\overline{p}_{\ell j}(\alpha_k,\alpha_\ell).$$

由于 $G_1=\big((\alpha_i,\alpha_j)\big)_{n\times n}$，$G_2=\big((\beta_i,\beta_j)\big)_{n\times n}$，所以上式即为

$$G_2=PG_1P^*.$$

一般地说，设 H_1 与 H_2 是 n 阶 Hermite 方阵．如果存在 n 阶可逆复方阵 P，使得

$$H_2=PH_1P^*,$$

则方阵 H_1 与 H_2 称为**复相合的**.

上面的讨论表明，同一个内积 (α,β) 在不同基下的 Gram 方阵是复相合的．

方阵的复相合关系是方阵间的一种重要的等价关系．以后将详加讨论．

定义 8.1.2 复线性空间 V 连同取定的一个内积 (α,β) 一起称为**酉空间**，仍记为 V.

由于定义酉空间 V 的内积 (α,β) 是恒正的，因此定义 V 中向量 α 的**范数** $\|\alpha\|$ 为

$$\|\alpha\|=\sqrt{(\alpha,\alpha)}.$$

范数为 1 的向量称为**单位向量**.

定义 8.1.3 设 V 是酉空间，(α,β) 是定义酉空间 V 的内积．设 $\alpha,\beta\in V$．如果 $(\alpha,\beta)=0$，则称向量 α 与 β 是**正交的**，记作 $\alpha\perp\beta$.

由性质 8.1.1，酉空间中零向量和每个向量都正交．和 Euclid 空间相仿，有

定理 8.1.1 酉空间 V 中任意 k 个两两正交的非零向量 $\alpha_1,\alpha_2,\dots,\alpha_k$ 是线性无关的.

定理 8.1.2 设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 n 维酉空间 V 的一组基，则 V 中存在一组两两正交的非零向量 $\beta_1,\beta_2,\dots,\beta_n$，使得对每个 k，$\{\beta_1,\beta_2,\dots,\beta_k\}$ 是 V 中由向量 $\alpha_1,\alpha_2,\dots,\alpha_k$ 生成的子空间 V_k 的一组基.

证明 令

$$\begin{aligned}\beta_1&=\alpha_1,\\ \beta_2&=\alpha_2+\lambda_{21}\beta_1,\\ &\cdots\cdots\\ \beta_k&=\alpha_k+\lambda_{k,k-1}\beta_{k-1}+\cdots+\lambda_{k1}\beta_1,\end{aligned}$$

其中 λ_{ij} 是待定常数,$1 \leqslant j < i \leqslant k$. 对 $i > j$,令

$$(\beta_i, \beta_j) = \left(\alpha_i + \sum_{\ell=1}^{i-1} \lambda_{i\ell}\beta_\ell,\ \beta_j\right) = 0.$$

由此得到,对 $j = 1, \ldots, i-1$,

$$\lambda_{ij} = -\frac{(\alpha_i, \beta_j)}{(\beta_j, \beta_j)},$$

即设

$$\begin{aligned}
\beta_1 &= \alpha_1, \\
\beta_2 &= \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1, \\
&\cdots\cdots\cdots \\
\beta_k &= \alpha_k - \frac{(\alpha_k, \beta_{k-1})}{(\beta_{k-1}, \beta_{k-1})}\beta_{k-1} - \cdots - \frac{(\alpha_k, \beta_1)}{(\beta_1, \beta_1)}\beta_1.
\end{aligned}$$

这就说明向量 $\beta_1, \beta_2, \ldots, \beta_k$ 属于由 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 生成的子空间 V_k,且两两正交. 由定理 8.1.1,$\{\beta_1, \beta_2, \ldots, \beta_k\}$ 是 V_k 的一组基. ■

定理 8.1.2 中给出的由 V 的基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 得到两两正交的向量组 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 称为对向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 施行 **Gram-Schmidt 正交化**. 只要将 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 的每个向量 β_j 单位化,即令

$$\xi_j = \frac{\beta_j}{\|\beta_j\|},$$

则 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 即是 V 中由两两正交的单位向量构成的一组基.

一般地说,n 维酉空间 V 中由 n 个两两正交的单位向量构成的基称为 V 的一组标准正交基. 于是得到

定理 8.1.3　n 维酉空间具有标准正交基.

由定理 8.1.2 还可以得到

定理 8.1.4　n 维酉空间 V 中任意一组两两正交的单位向量组 $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ 都可以扩成 V 的一组标准正交基.

下面的定理给出 n 维酉空间 V 中两组标准正交基之间的关系.

定理 8.1.5　设 n 维酉空间 V 中由标准正交基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 到标准正交基 $\{\eta_1, \eta_2, \ldots, \eta_n\}$ 的过渡方阵为 U,即

$$(\eta_1, \eta_2, \ldots, \eta_n) = (\xi_1, \xi_2, \ldots, \xi_n)U.$$

则 U 是酉方阵,即 U 满足

$$UU^* = I_{(n)} = U^*U,$$

其中 U^* 表示方阵 U 的共轭转置.

反之有

定理 8.1.6　设 U 是 n 阶酉方阵,$\{\xi_1, \xi_2, \ldots, \xi_n\}$ 是 n 维酉空间 V 的标准正交基. 则由

$$(\eta_1, \eta_2, \ldots, \eta_n) = (\xi_1, \xi_2, \ldots, \xi_n)U.$$

所确定的向量组 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 V 的标准正交基.

酉方阵是一类重要的方阵. 容易证明,酉方阵 U 是可逆的,并且它的逆方阵 $U^{-1}=U^*$; 同时酉方阵 U 的逆方阵 U^{-1} 仍是酉方阵. 另外,两个酉方阵 U_1 与 U_2 的乘积 U_1U_2 仍是酉方阵. 最后,单位方阵显然是酉方阵.

定理 8.1.5 表明,在 n 维空间 V 中取定一组标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 之后,V 的标准正交基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 便对应于过渡方阵 U. 由于 U 是酉方阵,所以 V 的所有标准正交基的集合便和所有 n 阶酉方阵集合 $\mathrm{U}_n(\mathbb{C})$ 建立了对应. 定理 8.1.6 表明,这一对应是 V 的所有标准正交基集合到 $\mathrm{U}_n(\mathbb{C})$ 上的一一对应.

定理 8.1.2 与定理 8.1.3 可以写成矩阵形式.

定理 8.1.7 任意 n 阶可逆复方阵 A 都可以表为一个酉方阵 U 与一个对角元全为正数的上三角方阵 T 的乘积,即 $A=UT$,并且表法唯一.

设 $\mathbb{C}^n$ 是所有 n 维复行向量集合. 在行向量的加法以及复数与行向量的乘法下,$\mathbb{C}^n$ 是复数域 $\mathbb{C}$ 上的线性空间.

设 $\alpha=(x_1,x_2,\ldots,x_n)$ 与 $\beta=(y_1,y_2,\ldots,y_n)\in\mathbb{C}^n$,定义 $\mathbb{C}^n$ 上二元复值函数 (α,β) 为

$$(\alpha,\beta)=x_1\overline{y}_1+x_2\overline{y}_2+\cdots+x_n\overline{y}_n=\alpha\beta^*.$$

容易验证,二元复值函数 (α,β) 是 $\mathbb{C}^n$ 的一个内积,它称为 $\mathbb{C}^n$ 的**标准内积**. 复向量空间 $\mathbb{C}^n$ 连同标准内积一起构成一个酉空间,仍记为 $\mathbb{C}^n$.

设 U 是一个酉方阵,将酉方阵 U 按行分块,即记 $U=(\alpha_1,\alpha_2,\ldots,\alpha_n)^{\mathrm{T}}$.

显然,$\alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{C}^n$. 由于 U 是酉方阵,所以 $UU^*=I_{(n)}$. 因此,

$$(\alpha_i,\alpha_j)=\alpha_i\alpha_j^*=\delta_{ij}.$$

这表明,$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维酉空间 $\mathbb{C}^n$ 的标准正交基.

反之,如果行向量组 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 $\mathbb{C}^n$ 的标准正交基,显然 $U=(\alpha_1,\alpha_2,\ldots,\alpha_n)^{\mathrm{T}}$ 是一个酉方阵.

因此,复方阵 U 为酉方阵的充分必要条件是,U 的 n 个行向量构成 $\mathbb{C}^n$ 的一组标准正交基. 同样,对于方阵的 n 个列向量,也有类似的结论.

定义 8.1.4 设 W 是酉空间 V 的子空间,$\beta\in V$. 如果 β 与 W 中任意向量 α 都正交,则称向量 β 和子空间 W **正交**,V 中所有与子空间 W 正交的向量的集合称为子空间 W 的**正交补**,记为 $W^{\perp}$.

容易看出,

$$W^{\perp}=\{\beta\in V\mid(\alpha,\beta)=0,\alpha\in W\}.$$

而且 $W^{\perp}$ 是 V 的一个子空间.

定理 8.1.8 对于 n 维酉空间 V 的子空间 W,有 $V=W\oplus W^{\perp}$.

最后转到酉空间的同构.

定义 8.1.5 设 V 和 W 是酉空间. 如果存在复线性空间 V 到 W 上的一个映

射 σ，使得对任意 $\alpha,\beta\in V$，均有

$$\big(\sigma(\alpha),\sigma(\beta)\big)=(\alpha,\beta),$$

其中 (α,β) 是定义酉空间 V 的内积，而 $\big(\sigma(\alpha),\sigma(\beta)\big)$ 是酉空间 W 中向量 $\sigma(\alpha)$ 与 $\sigma(\beta)$ 的内积，即映射 σ 是保内积的，则称 σ 是 V 到 W 上的**同构映射**，而酉空间 V 与 W 称为**同构**的.

容易证明

定理 8.1.9　任意 n 维酉空间 V 都同构于 n 维复行向量空间连同标准内积构成的酉空间 $\mathbb{C}^n$. 有限维酉空间 U 与 W 同构的充分必要条件是，$\dim U=\dim W$.

习　题　8.1

1. 设 a,b,c,d 是复数，$\mathbb{C}^2$ 是 2 维复行向量空间，定义 $\mathbb{C}^2$ 上的二元复值函数 $f(\alpha,\beta)$ 为

$$f(\alpha,\beta)=ax_1\overline{y}_1+bx_2\overline{y}_1+cx_1\overline{y}_2+dx_2\overline{y}_2,$$

其中 $\alpha=(x_1,x_2)$ 与 $\beta=(y_1,y_2)\in\mathbb{C}^2$. 试确定 a,b,c 和 d，使得 $f(\alpha,\beta)$ 是 $\mathbb{C}^2$ 的内积.

2. 证明，酉空间 V 中向量 α 和 β 正交的充分必要条件是，对任意一对复数 a 与 b，

$$\|a\alpha+b\beta\|^2=\|a\alpha\|^2+\|b\beta\|^2.$$

3. 设 V 是 n 维复线性空间. 如果映射 $\sigma: V\to V$ 满足：对任意 $\alpha,\beta\in V$ 和任意复数 λ，

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),\quad \sigma(\lambda\alpha)=\overline{\lambda}\sigma(\alpha),\quad \sigma^2(\alpha)=\alpha,$$

则 σ 称为**共轭映射**. V 中适合 $\sigma(\alpha)=\alpha$ 的向量 α 称为关于共轭映射 σ 的实向量. 记

$$R_\sigma(V)=\big\{\alpha\in V\ \big|\ \sigma(\alpha)=\alpha\big\}.$$

证明

(1) R_σ 是 n 维线性空间；

(2) 每个 $\alpha\in V$ 都可以唯一地表为 $\alpha=\beta+\mathrm{j}\gamma$，其中 $\beta,\gamma\in R_\sigma(V)$，而 j 是复数；

(3) 设 (α_1,α_2) 是 V 的内积. 将内积 (α_1,α_2) 的定义域取为 $R_\sigma(V)$，则 (α_1,α_2) 是 $R_\sigma(V)$ 的内积；　(本题似有误)

(4) 设 (β_1,β_2) 是 $R_\sigma(V)$ 的内积，则

$$(\alpha_1,\alpha_2)=(\beta_1,\beta_2)+(\gamma_1,\gamma_2)+\mathrm{j}\big((\beta_1,\gamma_2)-(\beta_2,\gamma_1)\big)$$

是 V 的内积，其中 $\alpha_1=\beta_1+\mathrm{j}\gamma_1$，$\alpha_2=\beta_2+\mathrm{j}\gamma_2$，且 $\beta_1,\beta_2,\gamma_1,\gamma_2\in R_\sigma(V)$.

4. 证明，任意二阶酉方阵 U 都可以分解为

$$U=\begin{pmatrix}\mathrm{e}^{\mathrm{i}\theta_1} & 0\\ 0 & \mathrm{e}^{\mathrm{i}\theta_2}\end{pmatrix}\begin{pmatrix}\cos\tau & \sin\tau\\ -\sin\tau & \cos\tau\end{pmatrix}\begin{pmatrix}\mathrm{e}^{\mathrm{i}\theta_3} & 0\\ 0 & \mathrm{e}^{\mathrm{i}\theta_4}\end{pmatrix},$$

其中 $\theta_1,\theta_2,\theta_3,\theta_4$ 和 τ 都是实数.

5. 设 n 阶复方阵 A 为 $A=B+\mathrm{i}C$，其中 B 与 C 是实方阵，且 $\mathrm{i}^2=-1$. 证明，方阵 A 为酉方阵的充分必要条件是，方阵 $B^{\mathrm{T}}C$ 是对称的，且 $B^{\mathrm{T}}B+C^{\mathrm{T}}C=I_{(n)}$.

6. 所有 n 阶复方阵构成的复线性空间记为 $\mathbb{C}^{n\times n}$，取内积为 $(A,B)=\mathrm{Tr}\,AB^*$，$A,B\in\mathbb{C}^{n\times n}$. 求 $\mathbb{C}^{n\times n}$ 中所有对角方阵构成的子空间 W 的正交补.

§8.2 复方阵的酉相似

Euclid 空间的线性函数概念可以推广到酉空间.

定义 8.2.1 设 $f(\alpha)$ 是酉空间 V 上的复值函数. 如果对任意 $\alpha_1, \alpha_2 \in V$ 和复数 λ_1, λ_2,

$$f(\lambda_1\alpha_1 + \lambda_2\alpha_2) = \lambda_1 f(\alpha_1) + \lambda_2 f(\alpha_2),$$

则 $f(\alpha)$ 称为 V 的**线性函数**.

n 维酉空间 V 的所有线性函数的集合记为 V^*. 集合 V^* 在通常函数的加法以及复数与函数的乘法下成为一个复线性空间,它称为酉空间 V 的**对偶空间**.

设 (α, β) 是 n 维酉空间 V 的内积. 对给定 $\beta \in V$,

$$f_\beta(\alpha) = (\alpha, \beta)$$

是 V 的一个线性函数. 定义映射 $\sigma: V \to V^*$ 如下: 设 $\beta \in V$,则令

$$\sigma(\beta) = f_\beta.$$

可以验证,映射 σ 是酉空间 V 到线性空间 V^* 的(线性空间)同构映射. 因此 V 与它的对偶空间 V^* 同构.

利用映射 σ,可以证明,如果 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 是 V 的基,则 $\{f_{\beta_1}, f_{\beta_2}, \ldots, f_{\beta_n}\}$ 是 V^* 的一组基,它称为 $\{\beta_1, \beta_2, \ldots, \beta_n\}$ 的**对偶基**.

设 $\mathscr{A}$ 是 n 维酉空间 V 的线性变换,(α, β) 是 V 的内积. 可以证明,对给定的向量 $\beta \in V$,存在唯一的向量 $\widetilde{\beta} \in V$,使得

$$\big(\mathscr{A}(\alpha), \beta\big) = (\alpha, \widetilde{\beta}).$$

定义映射 $\widetilde{\mathscr{A}}: V \to V$ 如下: 设 $\beta \in V$,则令 $\mathscr{A}^*(\beta) = \widetilde{\beta}$.

可以验证,映射 $\mathscr{A}^*$ 是 V 的一个线性变换,它称为线性变换 $\mathscr{A}$ 的**伴随变换**. 伴随变换具有如下性质:

(1) $(\mathscr{A} + \mathscr{B})^* = \mathscr{A}^* + \mathscr{B}^*$; **(2)** $(\lambda\mathscr{A})^* = \overline{\lambda}\mathscr{A}^*$;

(3) $(\mathscr{A}\mathscr{B})^* = \mathscr{B}^*\mathscr{A}^*$; **(4)** $(\mathscr{A}^*)^* = \mathscr{A}$;

(5) 设 V 的线性变换 $\mathscr{A}$ 在 V 的标准正交基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的方阵为 A,则它的伴随变换 $\mathscr{A}^*$ 在同一组基下的方阵为 $\mathscr{A}^*$,即方阵 A 的共轭转置;

(6) n 维酉空间 V 的线性变换 $\mathscr{A}$ 的不变子空间 W 的正交补 $W^\perp$ 是伴随变换 $\mathscr{A}^*$ 的不变子空间.

现在考虑酉空间 V 的一个线性变换在不同的标准正交基下的方阵的表示的关系. 先引进

定义 8.2.2 设 A 与 B 是 n 阶复方阵. 如果存在 n 阶酉方阵 U,使得 $B = U^*AU$,则方阵 A 与 B 称为**酉相似**的.

设 n 维酉空间 V 的线性变换 $\mathscr{A}$ 在 V 的标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 与标准正交基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 下的方阵分别为 A 与 B,即

$$\mathscr{A}(\xi_1,\xi_2,\ldots,\xi_n)=(\xi_1,\xi_2,\ldots,\xi_n)A,$$
$$\mathscr{A}(\eta_1,\eta_2,\ldots,\eta_n)=(\eta_1,\eta_2,\ldots,\eta_n)B.$$

并且设由标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 到标准正交基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 的过渡方阵为 U,即

$$(\eta_1,\eta_2,\ldots,\eta_n)=(\xi_1,\xi_2,\ldots,\xi_n)U.$$

由 §8.1 可知,方阵 U 是酉方阵. 于是,由上式,

$$\begin{aligned}\mathscr{A}(\eta_1,\eta_2,\ldots,\eta_n)&=\mathscr{A}\big((\xi_1,\xi_2,\ldots,\xi_n)U\big)=\big(\mathscr{A}(\xi_1,\xi_2,\ldots,\xi_n)\big)U\\&=(\xi_1,\xi_2,\ldots,\xi_n)AU=(\eta_1,\eta_2,\ldots,\eta_n)U^*AU.\end{aligned}$$

因此 $B=U^*AU$. 所以酉空间 V 中线性变换 $\mathscr{A}$ 在不同的标准正交基下的方阵是酉相似的.

容易验证,方阵之间的酉相似关系满足自反性,对称性与传递性. 因此它是所有 n 阶复方阵集合 $\mathrm{M}_n(\mathbb{C})$ 中元素间的一种等价关系.

于是集合 $\mathrm{M}_n(\mathbb{C})$ 便按酉相似关系分成酉相似等价类: 彼此酉相似的方阵归在同一个酉相似等价类,彼此不酉相似的方阵划归在不同的酉相似等价类.

所有 n 阶复方阵集合 $\mathrm{M}_n(\mathbb{C})$ 在酉相似下分类的两个基本问题是:

(1) 在酉相似等价类中如何选取代表元,即复方阵在酉相似下的标准形是什么?

(2) 如何判定两个复方阵是否酉相似,也即复方阵在酉相似下的全系不变量是什么?

应当指出,如果 n 阶复方阵 A 与 B 酉相似,即 $B=U^*AU$,其中 U 是酉方阵,则 $U^{-1}=U^*$,故 $B=U^{-1}AU$,即方阵 A 与 B 相似. 反之则不然.

关于酉相似,有

定理 8.2.1 设 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 是 n 阶复方阵 A 的全部特征值. 则存在 n 阶酉方阵 U,使得

$$A=U^*\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\ *&&&\lambda_n\end{pmatrix}U.\tag{8.2.1}$$

简单地说,复方阵 A 酉相似于下三角形.

证明 对方阵的阶数 n 用归纳法. 当 $n=1$ 时显然结论成立. 假设结论对 $n-1$ 阶方阵成立,下面证明结论对 n 阶方阵成立.

因为 λ_1 是 n 阶方阵 A 的特征值,所以存在属于特征值 λ_1 的特征(行)向量 $\alpha_1\in\mathbb{C}^n$,即

$$A\alpha_1=\lambda_1\alpha_1.$$

将 α_1 扩充成 n 维酉空间 $\mathbb{C}^n$ 的标准正交基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$,则 $U_1=(\alpha_1,\alpha_2,\ldots,\alpha_n)^{\mathrm{T}}$ 是酉方阵. 并且

$$U_1A=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}A=\begin{pmatrix}\lambda_1&0\\ \beta&A_1\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}=\begin{pmatrix}\lambda_1&0\\ \beta&A_1\end{pmatrix}U_1,$$

其中 A_1 是 $n-1$ 阶方阵. 因此

$$A=U_1^*\begin{pmatrix}\lambda_1&0\\ \beta&A_1\end{pmatrix}U_1.$$

由归纳假设,存在 $n-1$ 阶酉方阵 U_2,使得

$$A_1=U_2^*\begin{pmatrix}\mu_2&&\\ &\ddots&\\ *&&\mu_n\end{pmatrix}U_2=:U_2^*A_2U_2.$$

于是

$$A=U_1^*\begin{pmatrix}1&0\\ 0&U_2^*\end{pmatrix}\begin{pmatrix}\lambda_1&0\\ U_2\beta&A_2\end{pmatrix}\begin{pmatrix}1&0\\ 0&U_2\end{pmatrix}U_1.$$

记

$$U=\begin{pmatrix}1&0\\ 0&U_2\end{pmatrix}U_1.$$

则 U 是 n 阶酉方阵,并且

$$A=U^*\begin{pmatrix}\lambda_1&0\\ U_2\beta&A_2\end{pmatrix}U. \tag{8.2.2}$$

由于方阵 A 的特征值是相似不变量,而下三角方阵的对角元即是下三角方阵的特征值,因此 $\mu_1,\mu_2,\ldots,\mu_n$ 是方阵 A 的特征值. 于是式 (8.2.2) 具有式 (8.2.1) 的形式. ■

定义 8.2.3 满足 $A^*A=AA^*$ 的 n 阶复方阵 A 称为**规范变换**. 满足 $K^*=-K$ 的 n 阶复方阵 K 称为**斜 Hermite 方阵**.

显然,酉方阵、Hermite 方阵与斜 Hermite 方阵都是规范方阵.

定理 8.2.2(Schur 定理) n 阶复方阵 A 相似于对角方阵的充分必要条件是,A 为规范方阵.

证明 设方阵 A 酉相似于对角方阵,则存在 n 阶酉方阵 U,使得

$$A=U^*(\lambda_1,\lambda_2,\ldots,\lambda_n)U.$$

于是

$$A^*A=U^*\operatorname{diag}(\overline{\lambda}_1\lambda_1,\overline{\lambda}_2\lambda_2,\ldots,\overline{\lambda}_n\lambda_n)U=AA^*.$$

因此方阵 A 是规范的.

反之,设方阵 A 是规范的,且它的全部特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_n$,则由定理 8.2.1,存在 n 阶酉方阵 U 使得

$$A = U^* \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} U.$$

于是根据 $A^*A = AA^*$ 即可得到

$$\begin{pmatrix} \overline{\lambda}_1 & & * \\ & \ddots & \\ & & \overline{\lambda}_n \end{pmatrix}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}\begin{pmatrix} \overline{\lambda}_1 & & * \\ & \ddots & \\ & & \overline{\lambda}_n \end{pmatrix} \tag{8.2.3}$$

比较上式两端方阵的 (n,n) 未知上的元素得到

$$\overline{\lambda}_n\lambda_n = \lambda_n\overline{\lambda}_n + \sigma_n^2,$$

其中 σ_n^2 是下三角方阵

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} \tag{8.2.4}$$

的第 n 行上非对角元的绝对值之平方和. 因此 $\sigma_n^2 = 0$. 即下三角方阵 (8.2.4) 中第 n 行上非对角元全为零. 再比较式 (8.2.3) 两端方阵的 $(n-1, n-1)$ 位置上的元素, 如此继续, 即知方阵 (8.2.4) 的非对角元全为零. 所以规范方阵 A 酉相似于对角方阵. ■

由(8.2.1) 还可以得到

定理 8.2.3 (Schur 不等式) 设 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 n 阶复方阵 $A = (a_{ij})$ 的全部特征值. 那么

$$\operatorname{Tr} AA^* = \sum_{1\leqslant i,j\leqslant n} |a_{ij}|^2 \geqslant \sum_{i=1}^{n} |\lambda_i|^2.$$

并且等式当且仅当方阵 A 为规范方阵时成立.

证明略. 由(8.2.2) 得到,

定理 8.2.4 n 阶规范方阵 A 酉相似于对角形. 基存在 n 阶酉方阵 U, 使得

$$U^*AU = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是规范方阵 A 的特征值.

并且规范方阵的特征值是规范方阵在酉相似下的全系不变量.

(8.2.4) 完全解决了规范方阵在酉相似下的标准形问题.

由于酉方阵的特征值的绝对值为 1, 因此由(8.2.4) 得到

定理 8.2.5 n 阶酉方阵 U 酉相似于对角形. 即存在 n 阶酉方阵 U_1, 使得

$$U_1^*UU_1 = \operatorname{diag}(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n}),$$

其中 $e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n}$ 是酉方阵 U 的全部特征值.

并且酉方阵的特征值是酉方阵在酉相似下的全系不变量.

由于 Hermite 方阵是规范方阵,而且 Hermite 方阵的特征值都是实数,所以由(8.2.4) 得到

定理 8.2.6 n 阶 Hermite 方阵 H 酉相似于对角形. 即存在 n 阶酉方阵 U,使得

$$U^*HU = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是方阵 H 的全部特征值,$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$.

并且 Hermite 方阵的特征值是 Hermite 方阵在酉相似下的全系不变量.

由于斜 Hermite 方阵是规范方阵,而且斜 Hermite 方阵的非零特征值都是虚数,所以由(8.2.4) 得到

定理 8.2.7 n 阶斜 Hermite 方阵 K 酉相似于对角形,即存在 n 阶方阵 U,使得

$$U^*KU = \mathrm{diag}(\mathrm{i}\lambda_1, \mathrm{i}\lambda_2, \ldots, \mathrm{i}\lambda_r, \underbrace{0, \ldots, 0}_{n-r\text{个}}),$$

其中 $\mathrm{i}\lambda_1, \mathrm{i}\lambda_2, \ldots, \mathrm{i}\lambda_r$ 是方阵 K 的全部非零特征值,且 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r$.

并且斜 Hermite 方阵的特征值是斜 Hermite 方阵在酉相似下的全系不变量.

上述定理都具有相应的几何形式. 和定理 8.2.1 相应的是

定理 8.2.8 设 $\mathscr{A}$ 是 n 维酉空间 V 的线性变换. 则存在 V 的标准正交基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$,使得线性变换在这组基下的方阵是如下的下三角形:

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ * & & & \lambda_n \end{pmatrix},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是线性变换 $\mathscr{A}$ 的全部特征值.

定义 8.2.4 **(1)** 适合 $\mathscr{A}^*\mathscr{A} = \mathscr{A}\mathscr{A}^*$ 的线性变换 $\mathscr{A}$ 称为**规范变换**;

(2) 适合 $\mathscr{U}^*\mathscr{U} = \mathscr{I} = \mathscr{U}\mathscr{U}^*$ 的线性变换 $\mathscr{U}$ 称为**酉变换**;

(3) 适合 $\mathscr{H}^* = \mathscr{H}$ 的线性变换 $\mathscr{H}$ 称为**自拌变换**;

(4) 适合 $\mathscr{K}^* = -\mathscr{K}$ 的线性变换 $\mathscr{K}$ 称为**斜自拌变换**.

可以证明,n 维酉空间 V 的线性变换 $\mathscr{A}$ 为规范变换(或酉变换,或自拌变换,或斜自拌变换)的充分必要条件是,$\mathscr{A}$ 在 V 的标准正交基下的方阵为规范方阵(或酉方阵,或 Hermite 方阵,或斜 Hermite 方阵). 另外,n 维酉空间 V 的线性变换 $\mathscr{A}$ 为酉变换的充分必要条件是,对任意 $\alpha \in V$,

$$\|\mathscr{A}(\alpha)\| = \|\alpha\|.$$

n 维酉空间 V 的所有酉变换的集合记为 $\mathrm{U}_n(\mathbb{C})$.

容易验证,n 维酉空间 V 的单位变换 $\mathscr{I} \in \mathrm{U}_n(\mathbb{C})$; 如果 $\mathscr{U} \in \mathrm{U}_n(\mathbb{C})$,则 $\mathscr{U}^{-1} \in \mathrm{U}_n(\mathbb{C})$; 如果 $\mathscr{U}_1, \mathscr{U}_2 \in \mathrm{U}_n(\mathbb{C})$,则 $\mathscr{U}_1\mathscr{U}_2 \in \mathrm{U}_n(\mathbb{C})$. 由于集合 $\mathrm{U}_n(\mathbb{C})$ 具有这些性质,所以集合 $\mathrm{U}_n(\mathbb{C})$ 也称为 n 维酉空间 V 上的**酉变换群**,或简称**酉群**.

酉群是一类重要的典型群. 这里不拟讨论.

与定理 8.2.2 相应的是

定理 8.2.9　设 $\mathscr{A}$ 是 n 维酉空间 V 的线性变换. 则 $\mathscr{A}$ 为规范变换的充分必要条件是,存在 V 的一组标准正交基,使得 $\mathscr{A}$ 在这组基下的方阵为对角方阵.

定理 8.2.5,定理 8.2.6 与定理 8.2.7 都有相应的几何形式. 请读者自行补出.

最后给出一些例子.

例 8.2.1　证明,n 阶复方阵 A 为规范方阵的充分必要条件是,存在复系数多项式 $f(\lambda)$,使得 $A^*=f(A)$.

证明　充分性是显然的. 下面证明必要性.

设 $\lambda_1,\lambda_2,\ldots,\lambda_t$ 是规范方阵 A 的全部不同特征值,它们的代数重数分别是 n_1, $n_2,\ldots,n_t$. 由定理 8.2.4,存在 n 阶酉方阵 U,使得

$$A=U^*\operatorname{diag}\big(\lambda_1I_{(n_1)},\lambda_2I_{(n_2)},\ldots,\lambda_tI_{(n_t)}\big)U.$$

由于方阵 A(酉)相似于对角形,所以方阵 A 的最小多项式为

$$d(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_t).$$

记 $d(\lambda)=(\lambda-\lambda_j)d_j(\lambda)$, $j=1,2,\ldots,t$. 并记

$$\varphi_j(\lambda)=\frac{\overline{\lambda}_j}{d_j(\lambda_j)}d_j(\lambda).$$

设 $D=\operatorname{diag}\big(\lambda_1I_{(n_1)},\lambda_2I_{(n_2)},\ldots,\lambda_tI_{(n_t)}\big)$. 则对每个 i,

$$\varphi_j\big(\lambda_iI_{(n_i)}\big)=\begin{cases}0, & \text{当 } i\neq j \text{ 时};\\ \overline{\lambda}_jI_{(n_j)}, & \text{当 } i=j \text{ 时}.\end{cases}$$

因此,对每个 j,

$$\begin{aligned}\varphi_j(D)&=\operatorname{diag}\big(\varphi_j(\lambda_1I_{(n_1)}),\varphi_j(\lambda_2I_{(n_2)}),\ldots,\varphi_j(\lambda_tI_{(n_t)})\big)\\&=\operatorname{diag}\big(0,\ldots,0,\overline{\lambda}_jI_{(n_j)},0,\ldots,0\big).\end{aligned}$$

令多项式 $f(\lambda)=\sum\limits_{j=1}^{n}\varphi_j(\lambda)$. 则

$$f(D)=\sum_{j=1}^{n}\varphi_j(D)=\operatorname{diag}\big(\overline{\lambda}_1I_{(n_1)},\overline{\lambda}_2I_{(n_2)},\ldots,\overline{\lambda}_tI_{(n_t)}\big)=\overline{D}.$$

所以

$$f(A)=f(U^*DU)=U^*f(D)U=U^*\overline{D}U=A^*.$$ ■

例 8.2.2　设 n 阶复方阵 A 与酉方阵 U 适合 $UA=AU^{\mathrm{T}}$. 证明,存在 n 阶酉方阵 V,使得 $U=V^2$,并且 $VA=AV^{\mathrm{T}}$.

证明　由定理 8.2.5,存在 n 阶酉方阵 U_1,使得

$$U=U_1^*\operatorname{diag}\big(\mathrm{e}^{\mathrm{i}\theta_1},\mathrm{e}^{\mathrm{i}\theta_2},\ldots,\mathrm{e}^{\mathrm{i}\theta_n}\big)U_1,$$

其中 $\mathrm{e}^{\mathrm{i}\theta_1},\mathrm{e}^{\mathrm{i}\theta_2},\ldots,\mathrm{e}^{\mathrm{i}\theta_n}$ 是 U 的全部特征值. 记

$$V = U_1^* \operatorname{diag}\left(\mathrm{e}^{\mathrm{i}\frac{\theta_1}{2}}, \mathrm{e}^{\mathrm{i}\frac{\theta_2}{2}}, \ldots, \mathrm{e}^{\mathrm{i}\frac{\theta_n}{2}}\right) U_1,$$

则 $V^2 = U$.

由于 $UA = AU^{\mathrm{T}}$,所以

$$\operatorname{diag}\left(\mathrm{e}^{\mathrm{i}\theta_1}, \mathrm{e}^{\mathrm{i}\theta_2}, \ldots, \mathrm{e}^{\mathrm{i}\theta_n}\right) U_1 A U_1^{\mathrm{T}} = U_1 A U_1^{\mathrm{T}} \operatorname{diag}\left(\mathrm{e}^{\mathrm{i}\theta_1}, \mathrm{e}^{\mathrm{i}\theta_2}, \ldots, \mathrm{e}^{\mathrm{i}\theta_n}\right).$$

记 $U_1 A U_1^{\mathrm{T}} = B = (b_{k\ell})$. 比较上式两端方阵的 (k, ℓ) 位置上的元素得到,

$$\mathrm{e}^{\mathrm{i}\theta_k} b_{k\ell} = \mathrm{e}^{\mathrm{i}\theta_\ell} b_{k\ell}.$$

由此得到

$$\mathrm{e}^{\mathrm{i}\frac{\theta_k}{2}} b_{k\ell} = \mathrm{e}^{\mathrm{i}\frac{\theta_\ell}{2}} b_{k\ell}.$$

于是得到

$$\operatorname{diag}\left(\mathrm{e}^{\mathrm{i}\frac{\theta_1}{2}}, \mathrm{e}^{\mathrm{i}\frac{\theta_2}{2}}, \ldots, \mathrm{e}^{\mathrm{i}\frac{\theta_n}{2}}\right) U_1 A U_1^{\mathrm{T}} = U_1 A U_1^{\mathrm{T}} \operatorname{diag}\left(\mathrm{e}^{\mathrm{i}\frac{\theta_1}{2}}, \mathrm{e}^{\mathrm{i}\frac{\theta_2}{2}}, \ldots, \mathrm{e}^{\mathrm{i}\frac{\theta_n}{2}}\right).$$

即得到,$VA = AV^{\mathrm{T}}$. ■

习 题 8.2

1. 设 α 与 β 分别是 n 维酉空间 V 的线性变换 $\mathscr{A}$ 与伴随变换 $\mathscr{A}^*$ 的特征向量. 证明,如果它们所属的特征值不共轭,则它们相互正交.

2. 证明,n 维酉空间 V 的线性变换 $\mathscr{A}$ 为规范的充分必要条件是,$\mathscr{A}$ 的每个特征向量也是它的伴随变换 $\mathscr{A}^*$ 的特征向量.

3. 证明,n 维酉空间 V 的线性变换 $\mathscr{A}$ 为规范的充分必要条件是,$\mathscr{A}$ 的不变子空间也是它的伴随变换 $\mathscr{A}^*$ 的不变子空间.

4. 证明,n 维酉空间 V 的线性变换 $\mathscr{A}$ 为规范的充分必要条件是,$\mathscr{A}$ 的每个不变子空间的正交补是 $\mathscr{A}$ 的不变子空间.

5. 设规范方阵 A 与方阵 B 可交换. 证明,方阵 A 与方阵 B^* 可交换.

6. 设规范方阵 A 与规范方阵 B 可交换. 证明 AB 是规范方阵.

7. 设 A 与 B 是规范方阵. 证明方阵 A 与 B 酉相似的充分必要条件是,方阵 A 与 B 相似.

8. 证明,两两可交换的 n 阶规范方阵集合可以同时酉相似于对角形.

9. 设 n 阶规范方阵 $A = B + \mathrm{i}C, B^* = B, C^* = C$,方阵 A 的任意两个特征值的实部与虚部分别不相等,且 x 是方阵 A, B 与 C 中某个方阵的特征向量. 证明,存在复数 λ,实数 μ 与 ν,使得

$$Ax = \lambda x, \qquad Bx = \mu x, \qquad Cx = \nu x,$$

并且 $\lambda = \mu + \mathrm{i}\nu$.

10. 所有 n 阶复方阵的集合连同内积 $(A, B) = \operatorname{Tr} AB^*$ 构成的酉空间记为 $\mathbb{C}^{n\times n}$,其中 $A, B \in \mathbb{C}^{n\times n}$. 设 $G \in \mathbb{C}^{n\times n}$. 定义 $\mathbb{C}^{n\times n}$ 的线性变换 $\mathscr{T}_G(A) = GA$. 证明,$\mathscr{T}_G$ 为酉变换的充分必要条件是,G 为酉方阵.

11. 设 W 是 n 维酉空间 V 的子空间,则 $V = W \oplus W^\perp$,即对每个 $\alpha \in V$,存在唯一一对向量 β, γ,其中 $\beta \in W, \gamma \in W^\perp$,使得 $\alpha = \beta + \gamma$. 定义 V 的线性变换 $\mathscr{A}$ 为: $\mathscr{A}(\alpha) = \beta - \gamma$. 证明 $\mathscr{A}$ 是酉变换.

12. 设 $\mathscr{A}$ 是 n 维酉空间 V 的自拌变换. 证明,

(1) 对任意 $\alpha \in V$, $\|\alpha + \mathrm{i}\mathscr{A}(\alpha)\| = \|\alpha - \mathrm{i}\mathscr{A}(\alpha)\|$;

(2) $\alpha+\mathrm{i}\mathscr{A}(\alpha)=\beta+\mathrm{i}\mathscr{A}(\beta)$ 的充分必要条件是 $\alpha=\beta$;

(3) $\mathscr{I}-\mathrm{i}\mathscr{A}$ 与 $\mathscr{I}+\mathrm{i}\mathscr{A}$ 都是可逆的;

(4) 变换

$$\mathscr{U}=(\mathscr{I}-\mathrm{i}\mathscr{A})(\mathscr{I}+\mathrm{i}\mathscr{A})^{-1}$$

是酉变换,它称为 $\mathscr{A}$ 的 **Cayley 变换**.

13. 设方阵 S 与 T 分别是实对称与实斜对称方阵,且 $\det(I_{(n)}-T-\mathrm{i}S)\neq 0$. 证明,

$$(I_{(n)}+T+\mathrm{i}S)(I_{(n)}-T-\mathrm{i}S)^{-1}$$

是酉方阵.

14. 设 n 阶复方阵 O 满足 $OO^{\mathrm{T}}=I_{(n)}$, $O^*=O$,则方阵 O 称为正交 Hermite 的. 证明,正交 Hermite 方阵 O 实正交相似于如下的准对角形:

$$\operatorname{diag}\left(\begin{pmatrix} a_1 & \mathrm{i}b_1 \\ -\mathrm{i}b_1 & a_1\end{pmatrix},\dots,\begin{pmatrix} a_s & \mathrm{i}b_s \\ -\mathrm{i}b_s & a_s\end{pmatrix},\underbrace{1,\dots,1}_{t\text{ 个}},\underbrace{-1,\dots,-1}_{n-2s-t\text{ 个}}\right),$$

其中 $a_1,a_2,\dots,a_s$ 和 $b_1,b_2,\dots,b_s$ 都是实数,且 $a_j^2-b_j^2=1$, $j=1,2,\dots,s$.

§8.3 正定 Hermite 方阵与矩阵的奇异值分解

实对称方阵的正定性与半正定性等概念都可以推广.

定义 8.3.1 设 H 是 n 阶 Hermite 方阵. 如果对任意非零复行向量 $\alpha\in\mathbb{C}^n$,依次有

$$\alpha H\alpha^*>0、\alpha H\alpha^*\geqslant 0、\alpha H\alpha^*<0、\alpha H\alpha^*\leqslant 0,$$

则方阵 H 依次称为**正定**、**半正定**、**负定**与**半负定**的,并且依次记为

$$H>0、H\geqslant 0、H<0、H\leqslant 0.$$

与正定实对称方阵相仿,可以证明

定理 8.3.1 设 H 是 n 阶 Hermite 方阵. 则下列命题等价:

(1) 方阵 H 是正定的;

(2) 方阵 H 的每个特征值都是正的;

(3) 存在正定 Hermite 方阵 H_1,使得 $H=H_1^2$;

(4) 存在可逆复方阵 P,使得 $H=P^*P$;

(5) 方阵 H 的每个主子式都是正的;

(6) 方阵 H 的每个顺序主子式都是正的;

(7) 对每个 $k=1,2,\dots,n$,方阵 H 的所有 k 阶主子式之和都是正的.

关于半正定 Hermite 方阵,有

定理 8.3.2 设 H 是 n 阶 Hermite 方阵. 则下列命题等价:

(1) 方阵 H 是半正定的;

(2) 方阵 H 的每个特征值都是非负的;

(3) 存在半正定 Hermite 方阵 H_1, $\operatorname{rank} H_1 = \operatorname{rank} H$, 使得 $H = H_1^2$;

(4) 存在 n 阶复方阵 P, $\operatorname{rank} P = \operatorname{rank} H$, 使得 $H = P^*P$;

(5) 方阵 H 的每个主子式都是非负的;

(6) 对每个 $k = 1, 2, \ldots, n$, 方阵 H 的所有 k 阶主子式之和都是非负的.

定理 8.3.3 设 H 与 H_1 是 n 阶半正定 Hermite 方阵, 且 $H = H_1^2$, 则方阵 H_1 是唯一的. 并且与方阵 H 可交换的方阵 A 也和方阵 H_1 可交换.

实矩阵的正交相抵可以推广到复矩阵.

定义 8.3.2 设 A 与 B 是 $m \times n$ 复矩阵. 如果存在 m 阶与 n 阶酉方阵 U_1 与 U_2, 使得 $B = U_1AU_2$, 则矩阵 A 与 B 称为**酉相抵**的.

定义 8.3.3 设 A 是 $m \times n$ 复矩阵. 则 n 阶半正定 Hermite 方阵 A^*A 的非零特征值的算术平方根称为矩阵 A 的**奇异值**.

矩阵 A 的奇异值总是正数, 而且它也可以定义为 m 阶半正定 Hermite 方阵 AA^* 的非零特征值的平方根.

定理 8.3.4 (矩阵的奇异值分解) 设 $\mu_1, \mu_2, \ldots, \mu_r$ 是 $m \times n$ 阶复矩阵 A 的所有奇异值, $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_r > 0$. 则矩阵 A 酉相抵如下的标准形:

$$\operatorname{diag}\big(\operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_r),\ 0\big),$$

其中 0 是 $(m-r) \times (n-r)$ 零矩阵. 即存在 m 阶与 n 阶酉方阵 U_1 与 U_2, 使得

$$A = U_1 \operatorname{diag}\big(\operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_r),\ 0\big) U_2. \tag{8.3.1}$$

并且矩阵的奇异值是矩阵在酉相抵下的全系不变量.

利用复矩阵在酉相抵下的标准形, 容易证明

定理 8.3.5 (矩阵的极分解) 设 A 是 n 阶复方阵. 则存在 n 阶半正定 Hermite 方阵 H_1 与 H_2, 以及酉方阵 U, 使得

$$A = H_1U = UH_2,$$

其中 H_1 与 H_2 由方阵 A 唯一确定.

利用矩阵在酉相抵下的标准形, 可以给出第 3 章中关于矩阵的广义逆的存在性的一个简捷证明. 重述一下广义逆的定义. 设 A 是 $m \times n$ 矩阵, 则适合矩阵方程组

$$\begin{cases} AXA = A, \\ XAX = X, \\ (AX)^* = AX, \\ (XA)^* = XA, \end{cases} \tag{P}$$

的 $n \times m$ 矩阵 X 称为矩阵 A 的 **Moore-Penrose 广义逆**, 记为 A^+.

定理 8.3.6 任意 $m \times n$ 矩阵 A 的 Moore-Penrose 广义逆 A^+ 存在而且唯一.

证明 设 $\mu_1, \mu_2, \ldots, \mu_r$ 是矩阵 A 的全部奇异值. 由定理 8.3.4, 存在 m 阶与 n

阶酉方阵 U_1 与 U_2,使得

$$A = U_1 \operatorname{diag}\big(\operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_r),\ 0\big)U_2,$$

其中 0 是 $(m-r)\times(n-r)$ 零矩阵. 取

$$X = U_2^* \operatorname{diag}\big(\operatorname{diag}({\mu^{-1}}_1, {\mu^{-1}}_2, \ldots, {\mu^{-1}}_r),\ 0\big)U_1^*,$$

其中 0 是 $(n-r)\times(m-r)$ 零矩阵. 容易证明,$n\times m$ 矩阵 X 是矩阵方程组 (P) 的解. 因此 Moore-Penrose 广义逆 A^+ 存在. 唯一性的证明仍如定理 3.7.2. ■

定理 8.3.6 的证明中给出了矩阵 A 的 Moore-Penrose 广义逆 A^+ 的表达式. 即如果 $m\times n$ 矩阵 A 的奇异值分解为

$$A = U_1 \operatorname{diag}(D,\ 0)U_2,$$

其中 $D=\operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_r)$,且 $\mu_1, \mu_2, \ldots, \mu_r$ 为矩阵的全部奇异值,则

$$A^+ = U_2^* \operatorname{diag}(D^{-1},\ 0)U_1^*.$$

习　题　8.3

1. 设 n 阶复方阵 $A=B+\mathrm{i}C$,其中

$$B=\frac{1}{2}(A+\overline{A}),\qquad C=-\frac{\mathrm{i}}{2}(A-\overline{A}),$$

并且 A 是半正定 Hermite 方阵. 证明 $\operatorname{rank} A\leqslant \operatorname{rank} B$,$\operatorname{rank} C\leqslant \operatorname{rank} B$.

2. 设 H 是 n 阶正定 Hermite 方阵,A 是 $n\times m$ 列满秩矩阵. 求逆方阵

$$\begin{pmatrix} H & A \\ A^* & 0 \end{pmatrix}^{-1}.$$

3. 设 A 是 n 阶复方阵,$\mathbb{C}^n$ 是 n 维复行向量空间. 记

$$K(A)=\big\{x\in\mathbb{C}^n \mid xA=0\big\},$$

它称为方阵 A 的零空间. 设 H_1 与 H_2 是 n 阶 Hermite 方阵,其中 $H_1\geqslant 0$,$\operatorname{rank} H_1=r$,且 $K(H_1)\subseteq K(H_2)$. 证明,存在 $n\times r$ 列满秩矩阵 P 与 r 阶实对角方阵 D,使得 $H_1=PP^*$ 且 $H_2=PDP^*$.

4. 设 H_1 与 H_2 是 n 阶正定 Hermite 方阵. 证明方阵 H_1H_2 的特征值都是正的.

5. 设 λ_1 与 λ_n 是 n 阶正定 Hermite 方阵 H 的最大与最小特征值,α 是任意 n 维非零复行向量. 证明

$$\frac{\lambda_n}{\lambda_1}\leqslant f(\alpha)=\frac{(\alpha\alpha^*)^2}{(\alpha H^{-1}\alpha^*)(\alpha H\alpha^*)}\leqslant 1.$$

6. 设 A 与 B 是 n 阶 Hermite 方阵. 证明,

$$\operatorname{Tr}(AB)^2\leqslant \operatorname{Tr}A^2B^2,$$

并且等式当且仅当 $AB=BA$ 时成立.

7. 设 A 与 B 是 n 阶 Hermite 方阵. 证明,

$$2\operatorname{Tr}AB\leqslant \operatorname{Tr}A^2+\operatorname{Tr}B^2,$$

并且等式当且仅当 $A=B$ 时成立.

8. 证明,复方阵 A 的每个奇异值都是 A 的特征值当且仅当方阵 A 为半正定 Hermite 方阵.

9. 设 A 是规范方阵. 证明 $A^+A=AA^+$.

10. 设 A 是 $m\times n$ 列满秩矩阵. 证明,$A^+=A^*$ 的充分必要条件是 $A^*A=I_{(n)}$.

§8.4 一些例子

例 8.4.1 设 H_1 与 H_2 是 n 阶 Hermite 方阵,且 $H_1>0$. 证明,Hermite 方阵 $H_1+H_2>0$ 的充分必要条件是,方阵 $H_1^{-1}H_2$ 的特征值都大于 -1.

证明 因为 $H_1>0$,所以由定理 8.3.1,存在 n 阶可逆复方阵 P,使得

$$(P^*)^{-1}H_1P^{-1}=I_{(n)}.$$

显然方阵 $(P^*)^{-1}H_2P^{-1}$ 仍是 Hermite 的. 因此存在 n 阶酉方阵 U,使得

$$U^*(P^*)^{-1}H_2P^{-1}U=\mathrm{diag}\big(\mu_1,\mu_2,\dots,\mu_n\big).$$

记 $Q=P^*U$. 显然方阵 Q 可逆,并且

$$H_1=QQ^*,\qquad H_2=Q\,\mathrm{diag}\big(\mu_1,\mu_2,\dots,\mu_n\big)Q^*.$$

于是,

$$H_1+H_2=Q\,\mathrm{diag}\big(1+\mu_1,1+\mu_2,\dots,1+\mu_n\big)Q^*.$$

由于 $H_1+H_2>0$,所以 $\mathrm{diag}\big(1+\mu_1,1+\mu_2,\dots,1+\mu_n\big)>0$. 因此 $1+\mu_i>0$,即 $\mu_i>-1$. 由于

$$H_1^{-1}H_2=(Q^*)^{-1}\,\mathrm{diag}\big(\mu_1,\mu_2,\dots,\mu_n\big)Q^*,$$

所以 $\mu_1,\mu_2,\dots,\mu_n$ 是方阵 $H_1^{-1}H_2$ 的特征值. 于是方阵 $H_1^{-1}H_2$ 的特征值都大于 -1.

反之设方阵 $H_1^{-1}H_2$ 的特征值都大于 -1. 则由上段证明,$\mu_1,\mu_2,\dots,\mu_n$ 都大于 -1. 因此

$$\mathrm{diag}\big(1+\mu_1,1+\mu_2,\dots,1+\mu_n\big)>0.$$

所以方阵 $H_1+H_2>0$. ■

例 8.4.2(樊畿 (Ky Fan) 与 O. Tausky) 设 H_1 与 H_2 是 n 阶 Hermite 方阵,$H_1>0$,且 H_1H_2 是 Hermite 方阵. 证明,$H_1H_2>0$ 的充分必要条件是,Hermite 方阵 H_2 的特征值都是正的.

证明 因为 $H_1>0$,所以由定理 8.3.1,方阵 H_1 的特征值都是正的. 因此逆方阵 H_1^{-1} 的特征值也都是正的. 于是方阵 $H_1^{-1}>0$. 由定理 8.3.1,存在某个 n 阶可逆复方阵 P,使得

$$(P^*)^{-1}H_1^{-1}P^{-1}=I_{(n)}.$$

显然方阵 $(P^*)^{-1}H_2P^{-1}$ 仍是 Hermite 的. 所以存在 n 阶酉方阵 U,使得

$$U^*(P^*)^{-1}H_2P^{-1}U=\mathrm{diag}(\mu_1,\mu_2,\dots,\mu_n).$$

又

$$U^*(P^*)^{-1}H_1^{-1}P^{-1}U=I_{(n)}.$$

记 $Q=U^*P$. 则方阵 Q 可逆,并且

$$H_1^{-1}=Q^*Q,\qquad H_2=Q^*\,\mathrm{diag}(\mu_1,\mu_2,\dots,\mu_n)Q.$$

因此
$$H_1H_2 = Q^{-1}\operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n)Q.$$

由此可知，如果 $H_1H_2 > 0$，则方阵 H_1H_2 的特征值 $\mu_1, \mu_2, \ldots, \mu_n$ 都是正的，所以方阵 $H_2 > 0$．反之设方阵 $H_2 > 0$，则 $\operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n) > 0$．所以方阵 H_1H_2 的特征值 $\mu_1, \mu_2, \ldots, \mu_n$ 都是正的．因此 $H_1H_2 > 0$．■

例 8.4.3（O. Tausky） 设 n 阶正定 Hermite 方阵 $H = A + \mathrm{i}B$，其中 A 与 B 是 n 阶实方阵，且 $\mathrm{i}^2 = -1$．证明 $\det A \geqslant \det H$．并且等式当且仅当 $B = 0$ 时成立．

证明　容易验证，方阵 A 与 B 分别是实对称与实斜对称的，而且因为 $H > 0$，所以
$$A > 0.$$

由定理 7.7.3，存在某个 n 阶可逆实方阵 P，使得
$$(P^{\mathrm{T}})^{-1}AP^{-1} = I_{(n)}.$$

显然方阵 $(P^{\mathrm{T}})^{-1}BP^{-1}$ 仍是实斜对称的．因此由定理 7.6.3，存在 n 阶实正交方阵 O，使得
$$O^{\mathrm{T}}(P^{\mathrm{T}})^{-1}BP^{-1}O = \operatorname{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, \underbrace{0, \ldots, 0}_{n-2s\text{ 个}}\right),$$
其中 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_s$，且 $\operatorname{rank} B = 2s$．又
$$O^{\mathrm{T}}(P^{\mathrm{T}})^{-1}AP^{-1}O = I_{(n)}.$$
记 $Q = P^{\mathrm{T}}O$．显然方阵 Q 可逆，并且 $A = QQ^{\mathrm{T}}$，
$$B = Q\operatorname{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, \underbrace{0, \ldots, 0}_{n-2s\text{ 个}}\right)Q^{\mathrm{T}}.$$
于是
$$H = A + \mathrm{i}B = Q\operatorname{diag}\left(\begin{pmatrix} 1 & \mathrm{i}b_1 \\ -\mathrm{i}b_1 & 1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 & \mathrm{i}b_s \\ -\mathrm{i}b_s & 1 \end{pmatrix}, \underbrace{1, \ldots, 1}_{n-2s\text{ 个}}\right)Q^{\mathrm{T}}.$$
因此
$$\det H = \det Q \det Q^{\mathrm{T}}(1 - b_1^2)(1 - b_2^2)\cdots(1 - b_s^2).$$
由于 $H > 0$，所以方阵 H 的 2 阶子式
$$\det\begin{pmatrix} 1 & \mathrm{i}b_j \\ -\mathrm{i}b_j & 1 \end{pmatrix} = 1 - b_j^2 > 0,$$
即 $|b_j| < 1$．所以
$$\det H \leqslant \det QQ^{\mathrm{T}} = \det A.$$
其中等式当且仅当 $b_1 = b_2 = \cdots = b_s = 0$．即方阵 $B = 0$ 时成立．■

例 8.4.4（华罗庚） 设 A 与 B 是 n 阶复方阵，$I_{(n)} - A^*A > 0$ 且 $I_{(n)} - B^*B > 0$．证明，
$$\left|\det(I_{(n)} - A^*B)\right|^2 \geqslant \det(I_{(n)} - A^*A)\det(I_{(n)} - B^*B).$$

证明 因为

$$\begin{pmatrix} I_{(n)} & 0 \\ -A^* & I_{(n)} \end{pmatrix}\begin{pmatrix} I_{(n)} & B \\ A^* & I_{(n)} \end{pmatrix} = \begin{pmatrix} I_{(n)} & B \\ 0 & I_{(n)} - A^*B \end{pmatrix}.$$

所以

$$\det\begin{pmatrix} I_{(n)} & B \\ A^* & I_{(n)} \end{pmatrix} = \det(I_{(n)} - A^*B).$$

同理

$$\det\begin{pmatrix} I_{(n)} & -A \\ -B^* & I_{(n)} \end{pmatrix} = \det(I_{(n)} - B^*A) = \overline{\det(I_{(n)} - A^*B)}.$$

所以

$$\begin{aligned}\left|\det(I_{(n)} - A^*B)\right|^2 &= \det\begin{pmatrix} I_{(n)} & B \\ A^* & I_{(n)} \end{pmatrix}\begin{pmatrix} I_{(n)} & -A \\ -B^* & I_{(n)} \end{pmatrix} \\ &= \det\begin{pmatrix} I_{(n)} - BB^* & B - A \\ A^* - B^* & I_{(n)} - A^*A \end{pmatrix}.\end{aligned}$$

由于 $I_{(n)} - A^*A > 0$,所以方阵 $I_{(n)} - A^*A$ 可逆. 因此

$$\begin{aligned}&\begin{pmatrix} I_{(n)} & -(B-A)(I_{(n)} - A^*A)^{-1} \\ 0 & I_{(n)} \end{pmatrix}\begin{pmatrix} I_{(n)} - BB^* & B - A \\ A^* - B^* & I_{(n)} - A^*A \end{pmatrix} \\ &= \begin{pmatrix} I_{(n)} - BB^* - (B-A)(I_{(n)} - A^*A)^{-1}(A^* - B^*) & 0 \\ A^* - B^* & I_{(n)} - A^*A \end{pmatrix}.\end{aligned}$$

于是

$$\begin{aligned}\left|\det(I_{(n)} - A^*B)\right|^2 = &\det(I_{(n)} - A^*A) \\ &\times \det\big((I_{(n)} - BB^*) + (B-A)(I_{(n)} - A^*A)^{-1}(B^* - A^*)\big).\end{aligned}$$

由方阵的极分解定理,存在酉方阵 U 与半正定 Hermite 方阵 H,使得 $B = UH$. 因此由假设,

$$I_{(n)} - B^*B = I_{(n)} - H^2 > 0.$$

而

$$I_{(n)} - BB^* = I_{(n)} - UH^2U^* = U(I_{(n)} - H^2)U^*.$$

由于 Hermite 方阵的正定性在酉相合下不变,所以 $I_{(n)} - BB^* > 0$.

另一方面,由于 $I_{(n)} - A^*A > 0$,所以 $\big(I_{(n)} - A^*A\big)^{-1} > 0$. 因此方阵

$$(B-A)(I_{(n)} - A^*A)^{-1}(B-A)^* \geqslant 0.$$

于是

$$\det\big((I_{(n)} - BB^*) + (B-A)(I_{(n)} - A^*A)^{-1}(B-A)^*\big) \geqslant \det(I_{(n)} - BB^*).$$

但 $\det(I_{(n)} - BB^*) = \det(I_{(n)} - B^*B)$. 所以

$$\left|\det(I_{(n)} - A^*B)\right|^2 \geqslant \det(I_{(n)} - A^*A)\det(I_{(n)} - B^*B).$$ ∎

例 8.4.5 设 n 阶 Hermite 方阵 H 的秩为 r. 证明,

$$r \geqslant \frac{(\mathrm{Tr}\, H)^2}{\mathrm{Tr}\, H^2}.$$

证明 因为方阵 H 是 Hermite 的,所以存在 n 阶酉方阵 U,使得

$$UHU^* = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是方阵 H 的全部特征值.

由于 $\mathrm{rank}\, H = r$,所以方阵 $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ 的秩也为 r. 因此可设 $r_{r+1} = \cdots = r_n = 0$,且 $\lambda_1, \lambda_2, \ldots, \lambda_r$ 全不为零. 所以

$$(\mathrm{Tr}\, H)^2 = (\lambda_1 + \lambda_2 + \cdots + \lambda_r)^2.$$

利用 Cauchy 不等式得到

$$(\mathrm{Tr}\, H)^2 \leqslant r \sum_{i=1}^{r} \lambda_i^2.$$

显然方阵 H^2 的所有非零特征值为 $\lambda^2{}_1, \lambda^2{}_2, \ldots, \lambda^2{}_r$. 因此

$$\mathrm{Tr}\, H^2 = \sum_{i=1}^{r} \lambda_i^2.$$

于是,$(\mathrm{Tr}\, H)^2 \leqslant r\, \mathrm{Tr}\, H^2$. ■

例 8.4.6 (Mitchell) 证明方阵 A 相似于对角形的充分必要条件是,存在正定 Hermite 方阵 H,使得方阵 HAH^{-1} 为规范方阵.

证明 设方阵 A 相似于对角形,即存在可逆方阵 P,使得

$$A = P\, \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) P^{-1}.$$

根据方阵的极分解,存在正定 Hermite 方阵 H_1 与酉方阵 U,使得 $P = H_1 U$. 因此

$$H_1^{-1} A H_1 = U\, \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) U^*.$$

取 $H = H_1^{-1}$,显然方阵 H 是正定 Hermite 的,而且方阵 HAH^{-1} 是规范的.

反之,因为方阵 HAH^{-1} 是规范的,所以存在酉方阵 U,使得

$$UHAH^{-1}U^* = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

记 $P = UH$. 显然,方阵 P 可逆,而且

$$PAP^{-1} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

即方阵 A 相似于对角形. ■

例 8.4.7 设 $\mu_1, \mu_2, \ldots, \mu_r$ 是 n 阶复对称方阵 S 的全部奇异值,$\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_r > 0$. 证明,存在 n 阶酉方阵 U,使得

$$USU^{\mathrm{T}} = \mathrm{diag}\big(\mu_1, \mu_2, \ldots, \mu_r, \underbrace{0, \ldots, 0}_{n-r\text{个}}\big).$$

证明 由定理 8.3.4,存在 n 阶酉方阵 U_1 与 U_2,使得

$$S = U_1\, \mathrm{diag}\big(\mu_1, \mu_2, \ldots, \mu_r, \underbrace{0, \ldots, 0}_{n-r\text{个}}\big) U_2.$$

因为方阵 S 是对称的,所以

$$\overline{U}_2U_1\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)=\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)(\overline{U}_2U_1)^{\mathrm{T}}.$$

显然方阵 $\overline{U}_2U_1$ 是酉方阵. 由例 8.2.2,存在 n 阶酉方阵 V,使得 $\overline{U}_2U_1=V^2$,并且

$$V\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)=\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)V^{\mathrm{T}}.$$

于是

$$\begin{aligned}VU_1^*S\overline{U}_1V^{\mathrm{T}}&=V\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)U_2\overline{U}_1V^{\mathrm{T}}\\&=\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big)V^{\mathrm{T}}\overline{V}\cdot\overline{V}V^{\mathrm{T}}\\&=\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,0,\ldots,0\big).\end{aligned}$$

记 $U=VU_1^*$. 显然方阵 U 是酉方阵,并且

$$USU^{\mathrm{T}}=\operatorname{diag}\big(\mu_1,\mu_2,\ldots,\mu_r,\underbrace{0,\ldots,0}_{n-r\text{个}}\big).$$ ■

注 设 A 与 B 是 n 阶复方阵. 如果存在 n 阶酉方阵 U,使得 $B=UAU^{\mathrm{T}}$,则方阵 A 与 B 称为酉相合的.

容易验证,方阵之间的酉相合关系满足自反性,对称性与传递性. 所以方阵之间的酉相合关系式所有 n 阶复方阵集合中元素之间的一种等价关系. 于是所有 n 阶复方阵集合在酉相合等价关系下分成酉相合等价类.

例 8.4.7 表明,复对称方阵酉相合于对角形,而且非零对角元是方阵的奇异值. 由此不难证明,复对称方阵的奇异值是复对称方阵在酉相合下的全系不变量.

习 题 8.4

1. 设 n 阶 Hermite 方阵 $H=(h_{ij})>0$. 证明

$$\det H\leqslant h_{11}h_{22}\cdots h_{nn}.$$

2. 设 $A=(a_{ij})$ 是 n 阶复方阵. 证明

$$\big|\det A\big|^2\leqslant\prod_{i=1}^{n}\Big(\sum_{j=1}^{n}|a_{ij}|^2\Big).$$

3. 设 n 阶 Hermite 方阵 $H>0$. 证明

$$\det H\leqslant H\big(\begin{smallmatrix}1&2&\cdots&r\\1&2&\cdots&r\end{smallmatrix}\big)H\big(\begin{smallmatrix}(r+1)&(r+2)&\cdots&n\\(r+1)&(r+2)&\cdots&n\end{smallmatrix}\big).$$

4. 设 n 阶复方阵 A 的每个元素的模都等于 1. 证明

$$\big|\det A\big|^2\leqslant n^n.$$

5. 设 H_1 与 H_2 是正定 Hermite 方阵,且 H_1-H_2 是正定的. 证明,方阵 $H_2^{-1}-H_1^{-1}$ 是正定的.

6. 设 A 是 $m\times n$ 行满秩复矩阵,B 为 $n\times p$ 复矩阵. 证明

$$\det\big(B^*(I_{(n)}-A^*(AA^*)^{-1}A)B\big)\leqslant\det B^*B.$$

第九章

双线性函数

✍ 本章讨论线性空间上的双线性函数，以及由此产生的方阵在相合下的分类问题.

✍ §9.1 首先给出双线性函数的概念，然后指出，在线性空间的不同基下，同一个双线性函数的矩阵表示是彼此相合的方阵. 于是，在几何上研究双线性函数，就相当于研究方阵在相合下的分类.

✍ 最有意义的双线性函数是对称双线性函数. 它对应于对称方阵.

✍ §9.2 详尽地讨论了对称方阵在相合下的分类，并以此解决了二次型的问题.

✍ §9.3 讨论了斜对称双线性函数，也即斜对称方阵在相合下的分类.

✍ 在复线性空间上，将双线性函数稍加推广，便得到共轭双线性函数的概念.

✍ 由共轭双线性函数在不同基下的矩阵表示引出了方阵在复相合下的分类问题.

✍ §9.4 讨论了具有重要意义的一类特殊的共轭双线性函数——Hermite 共轭双线性函数，也即研究了 Hermite 方阵在复相合下的分类问题，从而解决了 Hermite 型的相应问题.

§9.1 双线性函数

先给出下面的

定义 9.1.1 设 V 是数域 $\mathbb{F}$ 上的 n 维线性空间，$\alpha,\beta\in V$，$f(\alpha,\beta)$ 是定义在 V 上且取值在数域 $\mathbb{F}$ 的二元函数，如果对于任意向量 $\alpha,\beta,\alpha_1,\alpha_2,\beta_1,\beta_2\in V$，以及任意纯量 $\lambda_1,\lambda_2,\mu_1,\mu_2\in\mathbb{F}$，均有

$$f(\lambda_1\alpha_1+\lambda_2\alpha_2,\beta)=\lambda_1 f(\alpha_1,\beta)+\lambda_2 f(\alpha_2,\beta);\tag{9.1.1}$$

$$f(\alpha,\mu_1\beta_1+\mu_2\beta_2)=\mu_1 f(\alpha,\beta_1)+\mu_2 f(\alpha,\beta_2),\tag{9.1.2}$$

则 $f(\alpha,\beta)$ 称为 V 上的**双线性函数**.

例如，设 $L_1(\alpha)$ 与 $L_2(\alpha)$ 是 V 上的线性函数，$\alpha,\beta\in V$. 定义

$$f(\alpha,\beta)=L_1(\alpha)L_2(\beta).$$

容易验证，$f(\alpha,\beta)$ 是 V 上满足条件 (9.1.1) 和条件 (9.1.2) 的二元函数. 所以，$f(\alpha,\beta)$ 是 V 上的双线性函数.

又例如，设 V 是数域 $\mathbb{F}$ 上的所有 $m\times n$ 矩阵构成的 $m\times n$ 维线性空间，$X,\overline{Y}\in V$. 对于给定的 $A\in V$，定义 V 上二元函数 $f_A(X,\overline{Y})$ 为

$$f_A(X,\overline{Y})=\mathrm{Tr}(X^{\mathrm{T}}A\overline{Y}).$$

容易验证,二元函数 $f_A(X,\overline{Y})$ 满足条件 (9.1.1) 与条件 (9.1.2),因此,$f_A(X,\overline{Y})$ 是 V 上的双线性函数.

根据定义 9.1.1,容易证明,数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数 $f(\alpha,\beta)$ 具有如下性质.

命题 9.1.1 设 $f(\alpha,\beta)$ 是 V 上的双线性函数,则对任意 $\alpha,\beta\in V$,

$$f(\alpha,0)=0=f(0,\beta).$$

命题 9.1.2 设 $f(\alpha,\beta)$ 是 V 上的双线性函数,则对任意向量 $\alpha_1,\alpha_2,\dots,\alpha_p,\beta_1,\beta_2,\dots,\beta_q\in V$,以及任意 $\lambda_1,\lambda_2,\dots,\lambda_p,\mu_1,\mu_2,\dots,\mu_q\in\mathbb{F}$,

$$f\Big(\sum_{i=1}^{p}\lambda_i\alpha_i,\ \sum_{j=1}^{q}\mu_j\beta_j\Big)=\sum_{i=1}^{p}\sum_{j=1}^{q}\lambda_i\mu_jf(\alpha_i,\beta_j). \tag{9.1.3}$$

现在给出数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数 $f(\alpha,\beta)$ 在 V 的基下的方阵表示. 设 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 是 V 的一组基,向量 α 与 β 在这组基下的坐标分别是 $x=(x_1,x_2,\dots,x_n)$ 与 $y=(y_1,y_2,\dots,y_n)$,即

$$\alpha=\sum_{i=1}^{n}x_i\xi_i,\quad \beta=\sum_{j=1}^{n}y_j\xi_j.$$

则由式 (9.1.3),

$$f(\alpha,\beta)=f\Big(\sum_{i=1}^{n}x_i\xi_i,\ \sum_{j=1}^{n}y_j\xi_j\Big)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_jf(\xi_i,\xi_j).$$

记 n 阶方阵 A 为

$$A=\big(f(\xi_i,\xi_j)\big)=\begin{pmatrix} f(\xi_1,\xi_1) & f(\xi_1,\xi_2) & \cdots & f(\xi_1,\xi_n)\\ f(\xi_2,\xi_1) & f(\xi_2,\xi_2) & \cdots & f(\xi_2,\xi_n)\\ \vdots & \vdots & \ddots & \vdots\\ f(\xi_n,\xi_1) & f(\xi_n,\xi_2) & \cdots & f(\xi_n,\xi_n)\end{pmatrix}.$$

则

$$f(\alpha,\beta)=xAy^{\mathrm{T}}. \tag{9.1.4}$$

方阵 A 称为**双线性函数 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵.**

容易看出,V 上两个不同的双线性函数 $f(\alpha,\beta)$ 与 $g(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵 $A=\big(f(\xi_i,\xi_j)\big)$ 与 $B=\big(g(\xi_i,\xi_j)\big)$ 是不同的.

数域 $\mathbb{F}$ 上的 n 维线性空间 V 的所有双线性函数集合记为 $L(V,V,\mathbb{F})$.

设 $f,g\in L(V,V,\mathbb{F})$,$\alpha,\beta\in V$,定义

$$(f+g)(\alpha,\beta)=f(\alpha,\beta)+g(\alpha,\beta).$$

容易验证,$(f+g)(\alpha,\beta)$ 是 V 上的双线性函数,即 $f+g\in L(V,V,\mathbb{F})$. 双线性函数 $(f+g)(\alpha,\beta)$ 称为双线性函数 (α,β) 与 $g(\alpha,\beta)$ 的和.

设 $f\in L(V,V,\mathbb{F})$,$\lambda\in\mathbb{F}$,$\alpha,\beta\in V$,定义

$$(\lambda f)(\alpha,\beta)=\lambda f(\alpha,\beta).$$

显然 $(\lambda f)(\alpha,\beta)$ 是 V 上的双线性函数,即 $\lambda f \in L(V,V,\mathbb{F})$. 双线性函数 $(\lambda f)(\alpha,\beta)$ 称为双线性函数 $f(\alpha,\beta)$ 与纯量 λ 的**乘积**.

于是在集合 $L(V,V,\mathbb{F})$ 中引进了双线性函数的加法以及纯量与双线性函数的乘法. 可以验证,集合 $L(V,V,\mathbb{F})$ 在如此的加法与乘法下构成一个线性空间.

定理 9.1.1　数域 $\mathbb{F}$ 上的所有 n 阶方阵 A 构成的线性空间记为 $\mathbb{F}^{n\times n}$. 则线性空间 $L(V,V,\mathbb{F})$ 与 $\mathbb{F}^{n\times n}$ 同构. 从而

$$\dim L(V,V,\mathbb{F}) = (\dim V)^2.$$

证明　设 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 n 维线性空间 V 的一组基,双线性函数 $f\in L(V,V,\mathbb{F})$ 在这组基下的方阵为 $A=\big(f(\xi_i,\xi_j)\big)$. 建立由线性空间 $L(V,V,\mathbb{F})$ 到 $\mathbb{F}^{n\times n}$ 的映射 σ 如下: 对于任意 $f\in(V,V,\mathbb{F})$,令

$$\sigma(f)=A.$$

由于不同的双线性函数 $f,g\in L(V,V,\mathbb{F})$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵 $A=\big(f(\xi_i,\xi_j)\big)$ 与 $B=\big(g(\xi_i,\xi_j)\big)$ 是不同的,因此 $\sigma(f)\neq\sigma(g)$. 这表明映射 σ 是单射.

现在设 $A=(a_{ij})\in\mathbb{F}^{n\times n}$. 定义 V 上的二元函数 $f(\alpha,\beta)$ 为

$$f(\alpha,\beta)=xAy^{\mathrm{T}}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j,$$

其中 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$ 分别是向量 α 与 β 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标. 容易验证,$f(\alpha,\beta)$ 是 V 上的双线性函数,即 $f\in L(V,V,\mathbb{F})$.

由于 ξ_i 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标为

$$\varepsilon_i=(0,\ldots,0,\underset{i}{1},0,\ldots,0),$$

因此

$$f(\xi_i,\xi_j)=a_{ij},$$

其中 $1\leqslant i,j\leqslant n$. 所以 A 是双线性函数 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵. 于是 $\sigma(f)=A$. 这表明映射 σ 是满射.

从而 σ 是线性空间 $L(V,V,\mathbb{F})$ 到 $\mathbb{F}^{n\times n}$ 上的双射.

设 $f,g\in L(V,V,\mathbb{F})$,并且双线性函数 $f(\alpha,\beta)$ 与 $g(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵依次为 $A=\big(f(\xi_i,\xi_j)\big)$ 与 $B=\big(g(\xi_i,\xi_j)\big)$,即 $\sigma(f)=A,\sigma(g)=B$. 则对于 $\lambda,\mu\in\mathbb{F}$,

$$(\lambda f+\mu g)(\xi_i,\xi_j)=(\lambda f)(\xi_i,\xi_j)+(\mu g)(\xi_i,\xi_j)=\lambda f(\xi_i,\xi_j)+\mu g(\xi_i,\xi_j).$$

这表明,$(\lambda f+\mu g)(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵 $\big((\lambda f+\mu g)(\xi_i,\xi_j)\big)$ 满足

$$\big((\lambda f+\mu g)(\xi_i,\xi_j)\big)=\lambda\big(f(\xi_i,\xi_j)\big)+\mu\big(g(\xi_i,\xi_j)\big)=\lambda A+\mu B.$$

于是

$$\begin{aligned}\sigma(\lambda f+\mu g)&=\big((\lambda f+\mu g)(\xi_i,\xi_j)\big)=\lambda A+\mu B\\&=\lambda\sigma(f)+\mu\sigma(g).\end{aligned}$$

这表明,映射 σ 保线性运算. 所以映射 σ 是线性空间 $L(V,V,\mathbb{F})$ 到 $\mathbb{F}^{n\times n}$ 上的同构映射.

由于线性空间 $L(V,V,\mathbb{F})$ 与 $\mathbb{F}^{n\times n}$ 同构,而且线性空间 $\mathbb{F}^{n\times n}$ 是 n^2 维的,所以

$$\dim L(V,V,\mathbb{F})=(\dim V)^2 \qquad \blacksquare$$

现在考虑一个双线性函数在不同基下的方阵表示之间的联系. 为此重述一下方阵相合的概念. 设 A 与 B 是数域 $\mathbb{F}$ 上的 n 阶方阵. 如果存在数域 $\mathbb{F}$ 上的 n 阶可逆方阵 P,使得 $B=P^{\mathrm{T}}AP$,其中 P^{T} 是方阵 P 的转置,则方阵 A 与 B 称为相合的.

定理 9.1.2 设数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 与 $\{\eta_1,\eta_2,\dots,\eta_n\}$ 下的方阵分别是 $A=\big(f(\xi_i,\xi_j)\big)$ 与 $B=\big(f(\eta_i,\eta_j)\big)$,并且

$$(\eta_1,\eta_2,\dots,\eta_n)=(\xi_1,\xi_2,\dots,\xi_n)P, \tag{9.1.5}$$

其中 P 是数域 $\mathbb{F}$ 上的 n 阶可逆方阵. 则

$$B=P^{\mathrm{T}}AP.$$

简单地说,同一个双线性函数在不同基下的方阵是相合的.

证明 记 $P=(p_{ij})$. 则由式 (9.1.5),对于 $k=1,2,\dots,n$,

$$\eta_k=\sum_{\ell=1}^{n}p_{\ell k}\xi_\ell.$$

因此由式 (9.1.3),

$$f(\eta_i,\eta_j)=f\Big(\sum_{k=1}^{n}p_{ki}\xi_k,\ \sum_{\ell=1}^{n}p_{\ell j}\xi_\ell\Big)=\sum_{k=1}^{n}\sum_{\ell=1}^{n}p_{ki}p_{\ell j}f(\xi_k,\xi_\ell).$$

将此式写成矩阵形式,即得到

$$f(\eta_i,\eta_j)=(p_{1i},p_{2i},\dots,p_{ni})\begin{pmatrix} f(\xi_1,\xi_1) & f(\xi_1,\xi_2) & \cdots & f(\xi_1,\xi_n)\\ f(\xi_2,\xi_1) & f(\xi_2,\xi_2) & \cdots & f(\xi_2,\xi_n)\\ \vdots & \vdots & \ddots & \vdots\\ f(\xi_n,\xi_1) & f(\xi_n,\xi_2) & \cdots & f(\xi_n,\xi_n)\end{pmatrix}\begin{pmatrix} p_{1j}\\ p_{2j}\\ \vdots\\ p_{nj}\end{pmatrix},$$

其中 $1\leqslant i,j\leqslant n$. 于是得到

$$B=\big(f(\eta_i,\eta_j)\big)=P^{\mathrm{T}}\big(f(\xi_i,\xi_j)\big)P=P^{\mathrm{T}}AP. \qquad \blacksquare$$

由于方阵的秩是方阵在相合下的不变量,即相合的方阵具有相同的秩,因此定理 9.1.2 表明,双线性函数 $f(\alpha,\beta)$ 在 V 的某组基下的方阵的秩并不依赖于基的选取,而是由双线性函数 $f(\alpha,\beta)$ 自身所确定的.

于是双线性函数 $f(\alpha,\beta)$ 在某组基下的方阵的秩便定义为 $\boldsymbol{f(\alpha,\beta)}$ **的秩**,记为 $\operatorname{rank} f$. 如果双线性函数 $f(\alpha,\beta)$ 的秩为 n,即 $f(\alpha,\beta)$ 在 V 的基下的方阵是可逆的,则 $f(\alpha,\beta)$ 称为**非退化**的. 否则双线性函数 $f(\alpha,\beta)$ 称为**退化**的.

应当指出,在第 7 章中所讨论的 n 维实线性空间 V 的内积 (α,β) 是 V 上的双线性函数. 利用内积 (α,β),可以引进向量的正交性,即如果向量 α 与 β 的内积

$(\alpha,\beta)=0$,则向量 α 与 β 称为正交的. 向量关于内积 $f(\alpha,\beta)$ 的正交性可以推广到一般双线性函数.

定义 9.1.2　设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的线性函数. 如果向量 $\alpha,\beta\in V$ 满足 $f(\alpha,\beta)=0$,则称向量 α 关于双线性函数 f **左正交**于向量 β,并称向量 β 关于 f **右正交**于向量 α,且分别记作 $\alpha\perp_L\beta$,$\beta\perp_R\alpha$.

一般地说,向量关于双线性函数 $f(\alpha,\beta)$ 的正交性并不是对称的,也就是说,向量 α 关于 f 左正交于向量 β 并不意味着向量 β 关于 f 也左正交于向量 α.

向量关于双线性函数的正交性可以推广到子空间情形.

定义 9.1.3　设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数,W 是 V 的子空间,$\alpha\in V$. 如果对任意向量 $\beta\in W$,均有 $f(\alpha,\beta)=0$ (或 $f(\beta,\alpha)=0$),则向量 α 称为关于 f **左(或右)正交**于子空间 W,记为 $\alpha\perp_L W$ (或 $\alpha\perp_R W$).

V 中所有左正交于子空间 W 的向量集合记为 $W^{\perp_L}$,即

$$W^{\perp_L}=\left\{\alpha\in V\ \middle|\ f(\alpha,\beta)=0,\text{对任意 }\beta\in W\right\}.$$

同样,V 中所有右正交于子空间 W 的向量集合记为 $W^{\perp_R}$,即

$$W^{\perp_R}=\left\{\beta\in V\ \middle|\ f(\alpha,\beta)=0,\text{对任意 }\alpha\in W\right\}.$$

特别,$V^{\perp_L}$ 与 $V^{\perp_R}$ 分别称为 V 关于双线性函数 $f(\alpha,\beta)$ 的**左根基**与**右根基**.

容易验证,$W^{\perp_L}$ 与 $W^{\perp_R}$ 都是 V 的子空间,并且有

命题 9.1.3　设 W 与 U 是 V 的子空间,且 $W\subseteq U$. 则有

(1) $U^{\perp_L}\subseteq W^{\perp_L}$,$U^{\perp_R}\subseteq W^{\perp_R}$;　　**(2)** $\left(W^{\perp_L}\right)^{\perp_R}\supseteq W$, $\left(W^{\perp_R}\right)^{\perp_L}\supseteq W$.

请试者根据定义自证之.

下面的定理给出双线性函数的秩与左根基及右根基的联系.

定理 9.1.3　对数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数 $f(\alpha,\beta)$,下述等式成立:

$$\dim V^{\perp_L}=\dim V^{\perp_R}=\dim V-\operatorname{rank}f.$$

证明　设 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵为 $A=\big(f(\xi_i,\xi_j)\big)$,向量 $\alpha,\beta\in V$ 在这组基下的坐标分别为 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$. 则由式 (9.1.4),

$$f(\alpha,\beta)=xAy^{\mathrm{T}}.$$

设 $\alpha\in V^{\perp_L}$,则对任意 $\beta\in V$,$f(\alpha,\beta)=0$. 因此对任意行向量 $y\in\mathbb{F}^n$,$xAy^{\mathrm{T}}=0$. 所以 $xA=0$. 这表明,$V^{\perp_L}$ 中向量 α 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标 x 是齐次方程组 $xA=0$ 的解. 齐次方程组 $xA=0$ 的解空间记为 W_A. 建立左根基 $V^{\perp_L}$ 到 W_A 的映射 σ 如下: 设 $\alpha\in V^{\perp_L}$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标为 x,则令

$$\sigma(\alpha)=x.$$

由于左根基 $V^{\perp_L}$ 中不同的向量在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标是不同的,所以

σ 是单射. 设 $x \in W_A$，则 V 中存在向量 α，使得 α 在基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 下的坐标即为 x. 由于 $x \in W_A$，所以 $xA = 0$. 因此对任意 $y \in \mathbb{F}^n$，$xAy^{\mathrm{T}} = 0$，所以对任意 $\beta \in V$，$f(\alpha, \beta) = 0$. 即 $\alpha \in V^{\perp_L}$. 而且 $\sigma(\alpha) = x$. 这表明，σ 是满射.

从而 σ 是左根基 $V^{\perp_L}$ 到解空间 W_A 的双射. 另外，容易验证，σ 保线性运算.

这就证明了 σ 是左根基 $V^{\perp_L}$ 到解空间 W_A 的同构映射，即 $V^{\perp_L}$ 与 W_A 同构. 而

$$\dim W_A = n - \operatorname{rank} A,$$

所以 $\dim V^{\perp_L} = \dim V - \operatorname{rank} f$.

同理可证，$\dim V^{\perp_R} = \dim V - \operatorname{rank} f$. ■

由定理 9.1.3 直接得到下面的推论.

推论 9.1.1 数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数 $f(\alpha, \beta)$ 非退化的充分必要条件是，空间 V 关于 f 的左根基 $V^{\perp_L}$ 为零子空间.

将上述推论中的左根基 $V^{\perp_L}$ 改为右根基 $V^{\perp_R}$，结论仍成立.

设 $f(\alpha, \beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数. 取定向量 $\alpha_0 \in V$. 则 $f(\alpha_0, \beta)$ 显然是空间 V 上的一个线性函数. 这个线性函数记为 $L(\alpha_0)$，即

$$\big(L(\alpha_0)\big)(\beta) = f(\alpha_0, \beta).$$

V 上的线性函数 $L(\alpha_0)$ 显然属于线性空间 V 的对偶空间 V^*. 建立线性空间 V 到它的对偶空间 V^* 的映射 $\mathscr{A}$ 如下：设 $\alpha \in V$，则令

$$\mathscr{A}(\alpha) = L(\alpha).$$

关于映射 $\mathscr{A}$，有

命题 9.1.4 映射 $\mathscr{A}$ 是线性空间 V 到它的对偶空间 V^* 的线性映射.

证明 设 $\lambda_1, \lambda_2 \in \mathbb{F}$，$\alpha_1, \alpha_2 \in V$，则

$$\mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2) = L(\lambda_1\alpha_1 + \lambda_2\alpha_2).$$

因此对任意 $\beta \in V$，

$$\big(L(\lambda_1\alpha_1 + \lambda_2\alpha_2)\big)(\beta) = f(\lambda_1\alpha_1 + \lambda_2\alpha_2, \beta).$$

由于 $f(\alpha, \beta)$ 是双线性的，所以

$$\big(L(\lambda_1\alpha_1 + \lambda_2\alpha_2)\big)(\beta) = \lambda_1 f(\alpha_1, \beta) + \lambda_2 f(\alpha_2, \beta).$$

由 β 的任意性得到

$$\begin{aligned}\big(L(\lambda_1\alpha_1 + \lambda_2\alpha_2)\big)(\beta) &= \lambda_1\big(L(\alpha_1)\big)(\beta) + \lambda_2\big(L(\alpha_2)\big)(\beta)\\ &= \big(\lambda_1 L(\alpha_1) + \lambda_2 L(\alpha_2)\big)(\beta).\end{aligned}$$

所以
$$\begin{aligned}\mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2) &= L(\lambda_1\alpha_1 + \lambda_2\alpha_2) = \lambda_1 L(\alpha_1) + \lambda_2 L(\alpha_2)\\ &= \lambda_1\mathscr{A}(\alpha_1) + \lambda_2\mathscr{A}(\alpha_2).\end{aligned}$$

因此映射 $\mathscr{A}$ 是线性的. ■

命题 9.1.5　线性空间 V 到对偶空间 V^* 的线性映射 $\mathscr{A}$ 的核 $\operatorname{Ker}\mathscr{A}$ 等于空间 V 关于双线性函数 $f(\alpha,\beta)$ 的左根基 $V^{\perp_L}$.

证明　设 $\alpha\in\operatorname{Ker}\mathscr{A}$，则 $\mathscr{A}(\alpha)=L(\alpha)=0$，即 $L(\alpha)$ 是 V 上的零线性函数. 因此对任意 $\beta\in V$，

$$\big(L(\alpha)\big)(\beta)=f(\alpha,\beta)=0.$$

所以 $\alpha\in V^{\perp_L}$，即 $\operatorname{Ker}\mathscr{A}\subseteq V^{\perp_L}$.

将上述证明过程反推，即得 $V^{\perp_L}\subseteq\operatorname{Ker}\mathscr{A}$. 所以 $\operatorname{Ker}\mathscr{A}=V^{\perp_L}$.　■

利用映射 $\mathscr{A}$，可以给出双线性函数非退化的另一个充分必要条件.

定理 9.1.4　数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数 $f(\alpha,\beta)$ 非退化的充分必要条件是，对于 V 上的任意线性函数 $g(\beta)$，总存在向量 $\alpha\in V$，使得 $g(\beta)=f(\alpha,\beta)$，其中向量 β 遍历 V 中所有的向量.

证明　**必要性**　设双线性函数 $f(\alpha,\beta)$ 是非退化的，则由定理 9.1.3 的推论，线性空间 V 关于 f 的左根基 $V^{\perp_L}=0$.

由命题 9.1.5，线性空间 V 到对偶空间 V^* 的映射 $\mathscr{A}$ 的核 $\operatorname{Ker}\mathscr{A}=0$.

根据线性映射的象空间与核的维数定理可知，$\dim\operatorname{Im}\mathscr{A}=\dim V$，其中 $\operatorname{Im}\mathscr{A}$ 是映射 $\mathscr{A}$ 的象空间. 由于线性空间 V 与对偶空间 V^* 的维数相同，所以

$$\dim\operatorname{Im}\mathscr{A}=\dim V^*.$$

因此 $\operatorname{Im}\mathscr{A}=V^*$. 这表明，映射 $\mathscr{A}$ 是满射. 所以对 V 上的任意函数 $g(\beta)$，即 $g\in V^*$，必有向量 $\alpha\in V$，使得 $\mathscr{A}(\alpha)=L(\alpha)=g$. 因此对任意 $\beta\in V$，

$$g(\beta)=\big(L(\alpha)\big)(\beta)=f(\alpha,\beta).$$

充分性　设对于 V 上的任意线性函数 $g(\beta)$，总存在向量 $\alpha\in V$，使得 $g(\beta)=f(\alpha,\beta)$. 则

$$g(\beta)=f(\alpha,\beta)=\big(L(\alpha)\big)(\beta).$$

所以 $g=L(\alpha)=\mathscr{A}(\alpha)$. 这表明，线性空间 V 到对偶空间 V^* 的映射 $\mathscr{A}$ 是满射.

由线性映射的象空间与核的维数定理可知，$\dim\operatorname{Ker}\mathscr{A}=0$，即 $\operatorname{Ker}\mathscr{A}=0$. 由命题 9.1.5，$V^{\perp_L}=0$. 由定理 9.1.3 的推论，双线性函数 $f(\alpha,\beta)$ 是非退化的.　■

与定理 9.1.4 相对应的是

定理 9.1.5　数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数 $f(\alpha,\beta)$ 非退化的充分必要条件是，对 V 上任意线性函数 $g(\alpha)$，总存在向量 $\beta\in V$，使得 $g(\alpha)=f(\alpha,\beta)$，其中 α 遍历 V 中所有的向量.

证明　取定向量 $\beta\in V$，显然，$f(\alpha,\beta)$ 是 V 上的线性函数，记之为 $\big(R(\beta)\big)(\alpha)$.

定义线性空间 V 到对偶空间 V^* 的映射 $\mathscr{B}$ 如下：设 $\beta\in V$，则令

$$\mathscr{B}(\beta)=R(\beta).$$

和命题 9.1.4 与命题 9.1.5 相仿，可以证明 $\mathscr{B}$ 是由线性空间 V 到对偶空间 V^* 的线性映射，并且 $\operatorname{Ker}\mathscr{B}=V^{\perp_R}$. 再仿照定理 9.1.4 的证明，即可证明定理 9.1.5.　■

前面曾经指出，向量关于双线性函数 $f(\alpha,\beta)$ 的正交性一般是不对称的．这对讨论向量关于双线性函数 $f(\alpha,\beta)$ 的正交性必然带来许多的麻烦．自然要问，对哪些双线性函数 $f(\alpha,\beta)$，向量关于 $f(\alpha,\beta)$ 的正交性是对称的?

为了讨论这个问题，引进下面的定义.

定义 9.1.4 设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数．如果对任意向量 $\alpha,\beta\in V$，均有

$$f(\beta,\alpha)=f(\alpha,\beta),$$

则 $f(\alpha,\beta)$ 称为**对称**的；如果对任意 $\alpha,\beta\in V$，均有

$$f(\beta,\alpha)=-f(\alpha,\beta),$$

则 $f(\alpha,\beta)$ 称为**斜对称**或者**交代**的.

定理 9.1.6 设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数．则向量关于 $f(\alpha,\beta)$ 的正交性是对称的当且仅当双线性函数 $f(\alpha,\beta)$ 是对称的，或者是斜对称的.

证明 设向量关于双线性函数 $f(\alpha,\beta)$ 的正交性是对称的．也就是说，如果向量 α 关于 f 左正交于向量 β，则向量 β 关于 f 也左正交于向量 α，即若 $f(\alpha,\beta)=0$，则 $f(\beta,\alpha)=0$.

任取向量 $\alpha,\beta,\gamma\in V$．定义向量 $\xi=f(\alpha,\beta)\gamma-f(\alpha,\gamma)\beta$．则

$$f(\alpha,\xi)=f(\alpha,\beta)f(\alpha,\gamma)-f(\alpha,\gamma)f(\alpha,\beta)=0.$$

即 $\alpha\perp_L\xi$．由正交的对称性，$\xi\perp_L\alpha$，即对任意向量 $\alpha,\beta,\gamma\in V$，

$$f(\xi,\alpha)=f(\alpha,\beta)f(\gamma,\alpha)-f(\alpha,\gamma)f(\beta,\alpha)=0. \tag{9.1.6}$$

在上式中令 $\beta=\alpha$，则得到

$$f(\alpha,\alpha)\big(f(\gamma,\alpha)-f(\alpha,\gamma)\big)=0. \tag{9.1.7}$$

现在我们要根据式 (9.1.7) 证明，要么对所有 $\alpha,\gamma\in V$，均有 $f(\alpha,\gamma)=f(\gamma,\alpha)$，要么对所有 $\alpha\in V$，均有 $f(\alpha,\alpha)=0$.

事实上，若不然，则存在向量 $\eta,\zeta\in V$，使得 $f(\eta,\zeta)\neq f(\zeta,\eta)$，同时还存在向量 $\varepsilon\in V$，使得 $f(\varepsilon,\varepsilon)\neq 0$．因此由式 (9.1.7)，$f(\eta,\eta)=f(\zeta,\zeta)=0$，并且 $f(\varepsilon,\eta)=f(\eta,\varepsilon)$，$f(\varepsilon,\zeta)=f(\zeta,\varepsilon)$．由于 $f(\eta,\zeta)\neq f(\zeta,\eta)$，所以由式 (9.1.6) 得到，

$$f(\varepsilon,\eta)=f(\eta,\varepsilon)=0,\qquad f(\varepsilon,\zeta)=f(\zeta,\varepsilon)=0.$$

于是，

$$f(\eta,\varepsilon+\zeta)=f(\eta,\zeta)\neq f(\zeta,\eta)=f(\varepsilon+\zeta,\eta).$$

再由式 (9.1.7) 得到，$f(\varepsilon+\zeta,\varepsilon+\zeta)=0$．但是，

$$f(\varepsilon+\zeta,\varepsilon+\zeta)=f(\varepsilon,\varepsilon)+f(\varepsilon,\zeta)+f(\zeta,\varepsilon)+f(\zeta,\zeta)=f(\varepsilon,\varepsilon)\neq 0.$$

这就产生矛盾．所以要么对所有的 $\alpha,\gamma\in V$，$f(\gamma,\alpha)=f(\alpha,\gamma)$，要么对所有 $\alpha\in V$，

$f(\alpha,\alpha)=0$. 前者表明,双线性函数 $f(\alpha,\beta)$ 是对称的. 至于后者,设 $\alpha,\beta\in V$,则

$$\begin{aligned}f(\alpha+\beta,\alpha+\beta)&=f(\alpha,\alpha)+f(\beta,\alpha)+f(\alpha,\beta)+f(\beta,\beta)\\&=f(\alpha,\beta)+f(\beta,\alpha)=0.\end{aligned}$$

因此 $f(\beta,\alpha)=-f(\alpha,\beta)$. 所以后者表明,双线性函数 $f(\alpha,\beta)$ 是斜对称的.

这就证明了必要性. 至于充分性,则是显然的. ∎

对于对称或斜对称双线性函数 $f(\alpha,\beta)$,向量关于 $f(\alpha,\beta)$ 的正交性是对称的. 因此 $f(\alpha,\beta)$ 的左根基 $V^{\perp_L}$ 即是 $f(\alpha,\beta)$ 的右根基 $V^{\perp_R}$. 于是可以定义 $f(\alpha,\beta)$ 的根基 $V^\perp$ 为

$$V^\perp=\left\{\alpha\in V\;\middle|\;f(\alpha,\beta)=0,\text{对任意 }\beta\in V\right\}.$$

定理 9.1.7　任意一个双线性函数都可以唯一地表为一个对称双线性函数与一个斜对称双线性函数之和.

具体地说,设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数,则存在 V 上的对称双线性函数 $g(\alpha,\beta)$ 与斜对称双线性函数 $h(\alpha,\beta)$,使得

$$f(\alpha,\beta)=g(\alpha,\beta)+h(\alpha,\beta).$$

而且 $g(\alpha,\beta)$ 与 $h(\alpha,\beta)$ 是唯一的.

证明　取

$$g(\alpha,\beta)=\frac{1}{2}\big(f(\alpha,\beta)+f(\beta,\alpha)\big),\qquad h(\alpha,\beta)=\frac{1}{2}\big(f(\alpha,\beta)-f(\beta,\alpha)\big).$$

则 $g(\alpha,\beta)$ 与 $h(\alpha,\beta)$ 显然是 V 上的双线性函数. 对任意 $\alpha,\beta\in V$,

$$g(\beta,\alpha)=\frac{1}{2}\big(f(\beta,\alpha)+f(\alpha,\beta)\big)=\frac{1}{2}\big(f(\alpha,\beta)+f(\beta,\alpha)\big)=g(\alpha,\beta),$$

$$h(\beta,\alpha)=\frac{1}{2}\big(f(\beta,\alpha)-f(\alpha,\beta)\big)=-\frac{1}{2}\big(f(\alpha,\beta)-f(\beta,\alpha)\big)=-h(\alpha,\beta).$$

这表明,$g(\alpha,\beta)$ 与 $h(\alpha,\beta)$ 分别是对称与斜对称双线性函数. 显然

$$f(\alpha,\beta)=g(\alpha,\beta)+h(\alpha,\beta).$$

这就证明了双线性函数 $f(\alpha,\beta)$ 可以表为对称双线性函数 $g(\alpha,\beta)$ 与斜对称双线性函数 $h(\alpha,\beta)$ 之和.

设 $g_1(\alpha,\beta)$ 与 $h_1(\alpha,\beta)$ 分别是对称与斜对称双线性函数,并且

$$f(\alpha,\beta)=g_1(\alpha,\beta)+h_1(\alpha,\beta).$$

则

$$f(\beta,\alpha)=g_1(\beta,\alpha)+h_1(\beta,\alpha)=g_1(\alpha,\beta)-h_1(\alpha,\beta).$$

所以

$$g_1(\alpha,\beta)=\frac{1}{2}\big(f(\alpha,\beta)+f(\beta,\alpha)\big)=g(\alpha,\beta),$$

$$h_1(\alpha,\beta)=\frac{1}{2}\big(f(\alpha,\beta)-f(\beta,\alpha)\big)=h(\alpha,\beta).$$

这就证明,将双线性函数 $f(\alpha,\beta)$ 表为对称与斜对称双线性函数之和的表法是唯一的. ∎

数域 $\mathbb{F}$ 上的 n 维线性空间 V 的所有对称双线性函数集合与所有斜对称双线性函数集合分别记为 $S(V,V,\mathbb{F})$ 与 $K(V,V,\mathbb{F})$. 容易验证,它们是线性空间 $L(V,V,\mathbb{F})$ 的子空间. 定理 9.1.7 表明,线性空间 $L(V,V,\mathbb{F})$ 是子空间 $S(V,V,\mathbb{F})$ 与 $K(V,V,\mathbb{F})$ 的直和.

对称与斜对称双线性函数是两类重要的双线性函数. 接下来的两节将进行专门讨论.

习 题 9.1

1. 设 $\mathbb{R}^2$ 是 2 维实的行向量空间,向量 $\alpha=(\alpha_1,\alpha_2)$, $\beta=(y_1,y_2)\in\mathbb{R}^2$. 判断在 $\mathbb{R}^2$ 上,下列函数是否是双线性函数:

(1) $f(\alpha,\beta)=1$; (2) $f(\alpha,\beta)=(x_1-y_1)^2+x_2y_2$;

(3) $f(\alpha,\beta)=(x_1+y_1)^2-(x_2-y_2)^2$; (4) $f(\alpha,\beta)=x_1y_2-x_2y_1$.

2. 所有 2×3 实矩阵构成的实线性空间记为 $\mathbb{R}^{2\times3}$. 设方阵 A 为

$$A=\begin{pmatrix}1&2\\3&4\end{pmatrix}.$$

对于 $X,Y\in\mathbb{R}^{2\times3}$,定义空间 $\mathbb{R}^{2\times3}$ 上的双线性函数 $f(X,\overline{Y})$ 为

$$f(X,\overline{Y})=\mathrm{Tr}(X^{\mathrm{T}}A\overline{Y}).$$

(1) 求双线性函数 $f(X,\overline{Y})$ 在空间 $\mathbb{R}^{2\times3}$ 的基 $\{E_{11},E_{12},E_{13},E_{21},E_{22},E_{23}\}$ 下的方阵,其中 E_{ij} 是第 i 行第 j 列交叉位置上的元素为 1 而其它元素为 0 的 2×3 矩阵, $1\leqslant i\leqslant2$, $1\leqslant j\leqslant3$;

(2) 判断双线性函数 $f(X,\overline{Y})$ 是否是非退化的.

3. 设 $\mathbb{C}^{n\times n}$ 是所有 n 阶复方阵构成的复线性空间, V 是空间 $\mathbb{C}^{n\times n}$ 中所有迹为零的方阵构成的子空间. 对于 $X,\overline{Y}\in\mathbb{C}^{n\times n}$,定义空间 $\mathbb{C}^{n\times n}$ 上双线性函数 $f(X,\overline{Y})$ 为

$$f(X,\overline{Y})=n\,\mathrm{Tr}(X\overline{Y})-(\mathrm{Tr}\,X)(\mathrm{Tr}\,\overline{Y}).$$

证明

(1) $f(X,\overline{Y})$ 是空间 $\mathbb{C}^{n\times n}$ 上退化的双线性函数;

(2) 取双线性函数 $f(X,\overline{Y})$ 的定义域为 V,则 V 上双线性函数 $f(X,\overline{Y})$ 是非退化的.

(3) 设 A 是 n 阶非零的斜 Hermite 方阵,即 $A^*=-A$,这里 A^* 是 A 的共轭转置,则 $f(A,A)\leqslant0$.

4. 设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数. 证明,对空间 V 的任意子空间 V_1 与 V_2 有

$$(V_1+V_2)^{\perp L}=V_1^{\perp L}\cap V_2^{\perp L};\qquad (V_1+V_2)^{\perp R}=V_1^{\perp R}\cap V_2^{\perp R}.$$

如果 $f(\alpha,\beta)$ 非退化,则有

$$(V_1\cap V_2)^{\perp L}=V_1^{\perp L}+V_2^{\perp L};\qquad (V_1\cap V_2)^{\perp R}=V_1^{\perp R}+V_2^{\perp R}.$$

5. 设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数, W 是 V 的子空间,将 $f(\alpha,\beta)$ 的定义域限定在 W 上时, $f(\alpha,\beta)$ 是非退化的. 证明

$$V=W\oplus W^{\perp L}=W\oplus W^{\perp R}.$$

6. 设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的非退化双线性函数, $h(\alpha,\beta)$ 是 V 上的双线性函数. 证明

(1) 存在 V 的唯一线性变换 $\mathscr{A}_h$,使得对任意 $\alpha,\beta\in V$, $h(\alpha,\beta)=f\big(\mathscr{A}_h(\alpha),\beta\big)$;

(2) 双线性函数 $h(\alpha,\beta)$ 非退化的充分必要条件是,线性变换 A_h 可逆;

(3) 存在 V 的唯一可逆线性变换 $\mathscr{B}$,使得对任意 $\alpha,\beta\in V$,$f(\beta,\alpha)=f\big(\mathscr{B}(\alpha),\beta\big)$.

7. 设 $f(\alpha,\beta)$ 是 V 上的非退化双线性函数,$\mathscr{A}$ 是 V 的线性变换,证明,存在 V 的唯一线性变换 $\mathscr{A}^*$,使得对任意 $\alpha,\beta\in V$,$f\big(\mathscr{A}(\alpha),\beta\big)=f\big(\alpha,\mathscr{A}^*(\beta)\big)$.

8. 设 $f(\alpha,\beta)$ 是 V 上的双线性函数. 证明,存在 V 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\ldots,\beta_n\}$,使得

$$\big(f(\alpha_i,\beta_i)\big)=\mathrm{diag}\big(\underbrace{1,1,\ldots,1}_{r\text{个}},\underbrace{0,0,\ldots,0}_{n-r\text{个}}\big),$$

其中 $r=\mathrm{rank}\,f$.

9. 设 $f(\alpha,\beta)$ 是 V 上的双线性函数. 对于给定的向量 $\beta,\gamma\in V$,定义 V 到自身的映射 $\beta\otimes\gamma$ 如下: 设 $\alpha\in V$,则令 $(\beta\otimes\gamma)(\alpha)=f(\alpha,\beta)\gamma$. 显然 $\beta\otimes\gamma$ 是 V 的线性变换. 求线性变换 $\beta\otimes\gamma$ 的迹 $\mathrm{Tr}(\beta\otimes\gamma)$.

10. 设 $f(\alpha,\beta)$ 是 V 上的双线性函数. 证明,$f(\alpha,\beta)$ 可以分解为两个线性函数的乘积的充分必要条件是,$f(\alpha,\beta)$ 的秩为 1.

§9.2　对称双线性函数与二次型

设 V 是数域 $\mathbb{F}$ 上的 n 维线性空间,$S(V,V,\mathbb{F})$ 是 V 上所有对称双线性函数的集合.

定理 9.2.1　集合 $S(V,V,\mathbb{F})$ 与数域 $\mathbb{F}$ 上所有 n 阶对称方阵集合 $S(n,\mathbb{F})$ 之间存在一个一一对应.

证明　设 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的一组基,$f(\alpha,\beta)\in S(V,V,\mathbb{F})$ 在这组基下的方阵为

$$S=\big(f(\xi_i,\xi_j)\big)=(s_{ij}).$$

由于 $f(\alpha,\beta)\in S(V,V,\mathbb{F})$,所以 $f(\xi_j,\xi_i)=f(\xi_i,\xi_j)$,即 $S^{\mathrm{T}}=S$,$S\in S(n,\mathbb{F})$.

建立集合 $S(V,V,\mathbb{F})$ 到 $S(n,\mathbb{F})$ 的映射 σ 如下: 对于 $f(\alpha,\beta)\in S(V,V,\mathbb{F})$,令

$$\sigma(f)=S,$$

其中 S 是双线性函数 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵.

由于不同的双线性函数在同一组基下的方阵是不同的,所以 σ 是单射.

设 $S\in S(n,\mathbb{F})$. 定义 V 上二元函数 $f(\alpha,\beta)$ 为

$$f(\alpha,\beta)=xSy^{\mathrm{T}},$$

其中 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$ 分别是 V 中向量 α 与 β 在基 $(\xi_1,\xi_2,\ldots,\xi_n)$ 下的坐标.

由定理 9.1.1 的证明,$f(\alpha,\beta)$ 是 V 上的双线性函数. 由于

$$f(\beta,\alpha)=ySx^{\mathrm{T}}=(xSy^{\mathrm{T}})^{\mathrm{T}}=xSy^{\mathrm{T}}=f(\alpha,\beta),$$

其中 $\alpha,\beta\in V$,所以 $f(\alpha,\beta)$ 是对称的,即 $f(\alpha,\beta)\in S(V,V,\mathbb{F})$.

显然方阵 S 是双线性函数 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵. 因此由映射

σ 的定义，$\sigma(f)=S$. 所以 σ 是满射. 从而 σ 是双射. ■

定理 9.2.2 设 $f(\alpha,\beta)\in S(V,V,\mathbb{F})$. 则存在 V 的一组基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$，使得 $f(\alpha,\beta)$ 在这组基下的方阵是如下的对角形：

$$\operatorname{diag}\big(a_1,a_2,\ldots,a_r,\underbrace{0,0,\ldots,0}_{n-r\text{个}}\big), \tag{9.2.1}$$

其中 $a_1,a_2,\ldots,a_r\in\mathbb{F}$ 均非零，而 $r=\operatorname{rank} f$.

换句话说，$f(\alpha,\beta)$ 在这组基下的表达式为

$$f(\alpha,\beta)=a_1x_1y_1+a_2x_2y_2+\cdots+a_rx_ry_r,$$

其中 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$ 分别是 V 中向量 α 与 β 在这组基下的坐标.

证明 对空间的维数 n 用归纳法. 当 $n=1$ 时定理显然成立.

假设定理对 $n-1$ 维线性空间成立. 现在证明定理对 n 维空间也成立.

如果对称双线性函数是零函数，即对任意 $\alpha,\beta\in V$，$f(\alpha,\beta)=0$，则定理显然成立. 因此可设 $f(\alpha,\beta)$ 是非零的. 于是存在 $\alpha_0,\beta_0\in V$，使得 $f(\alpha_0,\beta_0)\neq 0$. 由于

$$f(\alpha_0+\beta_0,\alpha_0+\beta_0)-f(\alpha_0-\beta_0,\alpha_0-\beta_0)=4f(\alpha_0,\beta_0)\neq 0,$$

因此存在 $\xi_1\in V$，使得 $f(\xi_1,\xi_1)\neq 0$. 记 $a_1=f(\xi_1,\xi_1)$.

V 中由向量 ξ_1 生成的子空间记为 W. 由于 $f(\alpha,\beta)\in S(V,V,\mathbb{F})$，所以向量关于 f 的正交性是对称的. 因此可以不区分“左正交”与“右正交”，而统称为“正交”. 子空间 W 关于 $f(\alpha,\beta)$ 的正交子空间记为 $W^{\perp}$，即

$$W^{\perp}=\big\{\beta\in V\ \big|\ f(\alpha,\beta)=0,\text{对任意 }\alpha\in W\big\}.$$

现在证明，$V=W\oplus W^{\perp}$. 为此先证明 $W\cap W^{\perp}=0$. 事实上，设 $\alpha\in W\cap W^{\perp}$，则 $\alpha\in W$. 由于 W 是向量 ξ_1 生成的子空间，所以 $\alpha=\lambda_0\xi_1$，$\lambda_0\in\mathbb{F}$. 由于 $\alpha\in W\cap W^{\perp}$，所以

$$f(\alpha,\alpha)=\lambda_0^2f(\xi_1,\xi_1)=0.$$

但 $f(\xi_1,\xi_1)\neq 0$，所以 $\lambda_0=0$. 因此 $\alpha=0$，即 $W\cap W^{\perp}=0$.

其次证明 $V=W+W^{\perp}$. 设 $\alpha\in V$，取 $\beta=\alpha-\lambda\xi_1$，其中 λ 是待定参数. 令 $f(\xi_1,\beta)=0$，则得到

$$f(\xi_1,\alpha)-\lambda f(\xi_1,\xi_1)=0.$$

由于 $f(\xi_1,\xi_1)\neq 0$，所以

$$\lambda=\frac{f(\xi_1,\alpha)}{f(\xi_1,\xi_1)},$$

于是

$$\beta=\alpha-\frac{f(\xi_1,\alpha)}{f(\xi_1,\xi_1)}\xi_1\in W^{\perp}.$$

因此

$$\alpha=\frac{f(\xi_1,\alpha)}{f(\xi_1,\xi_1)}\xi_1+\left(\alpha-\frac{f(\xi_1,\alpha)}{f(\xi_1,\xi_1)}\xi_1\right)\in W+W^{\perp}.$$

这就证明了 $V \subseteq W + W^{\perp}$. 显然 $W + W^{\perp} \subseteq V$，所以 $V = W \oplus W^{\perp}$.

由于 $\dim W = 1$，$\dim V = n$，所以 $\dim W^{\perp} = n-1$. 将双线性函数 $f(\alpha,\beta)$ 的定义域限制在 $W^{\perp}$ 上，则 $f(\alpha,\beta)$ 是 $n-1$ 维空间 $W^{\perp}$ 上的对称双线性函数.

由归纳假设，存在 $W^{\perp}$ 的一组基 $\{\xi_2,\xi_3,\dots,\xi_n\}$，使得 $W^{\perp}$ 上的双线性函数 $f(\alpha,\beta)$ 在这组基下的方阵是如下的对角形：

$$\operatorname{diag}\big(a_2,a_3,\dots,a_s,\underbrace{0,0,\dots,0}_{n-s\text{ 个}}\big),$$

其中 $a_2,\dots,a_s$ 全不为零.

因为 $V = W \oplus W^{\perp}$，所以 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 是 V 的基，而且 $f(\alpha,\beta)$ 在这组基下的方阵即为

$$\operatorname{diag}\big(a_1,a_2,\dots,a_s,\underbrace{0,0,\dots,0}_{n-s\text{ 个}}\big).$$

由 $f(\alpha,\beta)$ 的秩的定义可知，$s = \operatorname{rank} f$. ■

应当指出，在定理 9.2.2 中，由于对称双线性函数 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵为对角方阵 (9.2.1)，所以当 $i \neq j$ 时，$f(\xi_i,\xi_j) = 0$，也即 ξ_i, ξ_j 关于 $f(\alpha,\beta)$ 是正交的.

一般地说，如果 V 的基 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 中任意两个向量 α_i 与 α_j 关于 $f(\alpha,\beta)$ 是正交的，则这组基称为 V 的关于 $f(\alpha,\beta)$ 的正交基.

定理 9.2.2 表明，对于任意一个对称双线性函数 $f(\alpha,\beta)$，在 V 中都存在关于 $f(\alpha,\beta)$ 的正变基 $\{\xi_1,\xi_2,\dots,\xi_n\}$，使得 $f(\alpha,\beta)$ 在这组正交基下的方阵是对角形 (9.2.1)，而且 V 中由基向量 $\xi_{r+1},\dots,\xi_n$ 生成的子空间即是 $f(\alpha,\beta)$ 的根基 $V^{\perp}$.

定理 9.2.2 可以写成矩阵形式如下.

定理 9.2.3　设 $S \in S(n,\mathbb{F})$. 则存在数域 $\mathbb{F}$ 上的 n 阶可逆方阵 P，使得

$$P^{\mathrm{T}}SP = \operatorname{diag}\big(a_1,a_2,\dots,a_r,\underbrace{0,0,\dots,0}_{n-r\text{ 个}}\big),$$

其中 $a_1,a_2,\dots,a_r$ 是数域 $\mathbb{F}$ 中非零的数，而 r 是对称方阵 S 的秩.

简单地说，数域 $\mathbb{F}$ 上的 n 阶对称方阵相合于对角形.

证明　设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的一组基，向量 $\alpha,\beta \in V$ 在这组基下的坐标分别是 $x = (x_1,x_2,\dots,x_n)$ 与 $y = (y_1,y_2,\dots,y_n)$. 定义 V 上的二元函数 $f(\alpha,\beta)$ 如下：

$$f(\alpha,\beta) = xSy^{\mathrm{T}}.$$

由定理 9.2.1 的证明，$f(\alpha,\beta) \in S(V,V,\mathbb{F})$. 由定理 9.2.2，存在 V 的关于 $f(\alpha,\beta)$ 的正交基 $\{\xi_1,\xi_2,\dots,\xi_n\}$，使得 $f(\alpha,\beta)$ 在这组基下的方阵为对角方阵 (9.2.1). 由定理 9.1.2，方阵 S 与对角方阵 (9.2.1) 是相合的. ■

现在转到 n 维实线性空间 V 上的对称双线性函数，即 $\mathbb{F} = \mathbb{R}$ 的情形.

定理 9.2.4　设 $f(\alpha,\beta) \in S(V,V,\mathbb{R})$，则存在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$，使得 $f(\alpha,\beta)$

在这组基下的方阵是如下的对角形：

$$\operatorname{diag}\big(I_{(p)}, -I_{(q)}, 0_{(n-p-q)}\big), \tag{9.2.2}$$

其中 $0_{(n-p-q)}$ 是 $n-p-q$ 阶零方阵，并且 $p+q=r$ 是 $f(\alpha,\beta)$ 的秩.

换句话说，$f(\alpha,\beta)$ 在这组基下的表达式为

$$f(\alpha,\beta) = x_1y_1 + x_2y_2 + \cdots + x_py_p - x_{p+1}y_{p+1} - \cdots - x_{p+q}y_{p+q},$$

其中 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$ 分别是 V 中向量 α 与 β 在这组基下的坐标.

证明 由定理 9.2.2，存在 V 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$，使得 $f(\alpha,\beta)$ 在这组基下的方阵为对角形 (9.2.1). 适当调整基向量 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 的次序，可以假定 $a_1,a_2,\ldots,a_p$ 都是正的，而 $a_{p+1},\ldots,a_{p+q}$ 都是负的，其中 $p+q=r=\operatorname{rank} f$. 记

$$\begin{aligned}
\xi_i &= \frac{1}{\sqrt{a_i}}\alpha_i, && i=1,2,\ldots,p;\\
\xi_j &= \frac{1}{\sqrt{-a_j}}\alpha_j, && j=p+1,\ldots,p+q;\\
\xi_k &= \alpha_k, && k=p+q+1,\ldots,n.
\end{aligned}$$

显然 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的一组基，并且 $f(\alpha,\beta)$ 在这组基下的方阵即为 (9.2.2). ■

现在考察定理 9.2.4 中对角形 $\operatorname{diag}\big(I_{(p)},-I_{(q)},0_{(n-p-q)}\big)$ 的对角元 $+1$ 与 -1 的个数 p 与 q 的几何意义. 为此引进

定义 9.2.1 设 $f(\alpha,\beta)$ 是 n 维实线性空间 V 上的对称双线性函数.

如果对任意非零向量 $\alpha\in V$，均有 $f(\alpha,\alpha)>0$，则 $f(\alpha,\beta)$ 称为**正定的**；

如果对任意向量 $\alpha\in V$，均有 $f(\alpha,\alpha)\geqslant 0$，则 $f(\alpha,\beta)$ 称为**半正定**的.

可以类似定义负定与半负定的对称双线性函数.

按照定理 9.2.4 的记号，V 中由基向量 $\xi_1,\xi_2,\ldots,\xi_p$、$\xi_{p+1},\ldots,\xi_{p+q}$ 与 $\xi_{p+q+1},\ldots,\xi_n$ 生成的子空间依次记为 V^+，V^- 与 $V^\perp$. 显然，$\dim V^+=p$，$\dim V^-=q$，并且

$$V = V^+ \oplus V^- \oplus V^\perp,$$

其中 $V^\perp$ 是 $f(\alpha,\beta)$ 的根基. 对于向量 $\alpha = a_1\xi_1 + a_2\xi_2 + \cdots + a_p\xi_p$，显然

$$f(\alpha,\alpha) = a_1^2 + a_2^2 + \cdots + a_p^2 \geqslant 0,$$

并且等式当且仅当 $\alpha=0$ 时成立. 也就是说，将 $f(\alpha,\beta)$ 限定在 V^+ 上，则 $f(\alpha,\beta)$ 是实线性空间 V^+ 上的正定对称双线性函数. 同理. 将 $f(\alpha,\beta)$ 限定在 V^- 上，则 $f(\alpha,\beta)$ 是实线性空间 V^- 上的负定对称双线性函数.

定理 9.2.5 设 $f(\alpha,\beta)$ 是 n 维实线性空间 V 上的对称双线性函数，则空间 V 可以分解为子空间 V^+，V^- 与 $V^\perp$ 的直和. 即

$$V = V^+ \oplus V^- \oplus V^\perp,$$

使得 $f(\alpha,\beta)$ 在 V^+ 与 V^- 上的限制分别是正定与负定的,而 $V^\perp$ 是 $f(\alpha,\beta)$ 的根基. 如果空间 V 还可以分解为子空间子空间 V_1^+, V_1^- 与 $V_1^\perp$ 的直和,即

$$V = V_1^+ \oplus V_1^- \oplus V_1^\perp,$$

使得 $f(\alpha,\beta)$ 在 V_1^+ 与 V_1^- 上的限制分别是正定与负定的,则

$$\dim V_1^+ = \dim V^+, \qquad \dim V_1^- = \dim V^-.$$

证明 定理的前半结论是定理 9.2.4 的另一种叙述方式. 下面证明后半结论.

设 $\alpha \in V_1^+ \cap (V^- \oplus V^\perp)$. 由于 $\alpha \in V_1^+$,所以 $f(\alpha,\alpha) \geqslant 0$. 由于 $\alpha \in V^- \oplus V^\perp$,所以

$$\alpha = \beta + \gamma,$$

其中 $\beta \in V^-, \gamma \in V^\perp$. 因此

$$f(\alpha,\alpha) = f(\beta,\beta) + f(\beta,\gamma) + f(\gamma,\beta) + f(\gamma,\gamma).$$

因为 $\gamma \in V^\perp$,所以

$$f(\beta,\gamma) = f(\gamma,\beta) = f(\gamma,\gamma) = 0.$$

因此 $f(\alpha,\alpha) = f(\beta,\beta)$. 由于 $\beta \in V^-$,所以 $f(\alpha,\alpha) = f(\beta,\beta) \leqslant 0$. 于是 $f(\alpha,\alpha) = 0$. 由于 $f(\alpha,\beta)$ 在 V_1^+ 上的限制是正定的,因此 $\alpha = 0$. 即 $V_1^+ \cap (V^- \oplus V^\perp) = 0$. 这表明,

$$V_1^+ + (V^- \oplus V^\perp) = V_1^+ \oplus (V^- \oplus V^\perp) \subseteq V.$$

所以

$$\dim V_1^+ + \dim(V^- \oplus V^\perp) \leqslant \dim V.$$

由于 $V = V^+ \oplus V^- \oplus V^\perp$,所以

$$\dim(V^- \oplus V^\perp) = \dim V - \dim V^+.$$

于是 $\dim V_1^+ - \dim V^+ \leqslant 0$. 即 $\dim V_1^+ \leqslant \dim V^+$.

考察 $V^+ \cap (V_1^- \oplus V^\perp)$,即可得到 $\dim V^+ \leqslant \dim V_1^+$. 因此 $\dim V_1^+ = \dim V^+$.

同理可证,$\dim V_1^- = \dim V^-$. ∎

定理 9.2.5 说明,对于 n 维实线性空间 V 上的对称双线性函数 $f(\alpha,\beta)$,空间 V 可以分解为三个部分 V^+, V^- 与 $V^\perp$,使得 $f(\alpha,\beta)$ 在 V^+ 与 V^- 上的限制分别是正定与负定的,而 $V^\perp$ 是 $f(\alpha,\beta)$ 的根基,并且子空间 V^+, V^- 与 $V^\perp$ 的维数与分解的方式无关,它们是由 $f(\alpha,\beta)$ 自身所决定的. 因此 $p = \dim V^+$ 与 $q = \dim V^-$ 分别称为 $f(\alpha,\beta)$ 的**正、负惯性指数**,它们之差 $p - q$ 称为 $f(\alpha,\beta)$ 的**符号差**,记为 $\delta(f)$. 显然,

$$\operatorname{rank} f = p + q, \qquad \delta(f) = 2p - \operatorname{rank} f.$$

于是定理 9.2.5 称为关于对称双线性函数的**惯性定理**.

定理 9.2.4 和定理 9.2.5 可以用来解决实对称方阵在相合下分类的问题.

设 $S \in S(n,\mathbb{R})$,且 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维实线性空间 V 的一组基. 令

$$f(\alpha,\beta) = xSy^{\mathrm{T}},$$

其中 x 与 y 分别是向量 α 与 β 在这组基下的坐标. 则 $f(\alpha,\beta)$ 是 V 上的对称双线性函数. 由定理 9.2.4,存在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵为对角形

$$\mathrm{diag}\big(I_{(p)},-I_{(q)},0_{(n-p-q)}\big).$$

由于对称双线性函数 $f(\alpha,\beta)$ 在不同基下的方阵是相合的,所以方阵 S 相合于对角形 $\mathrm{diag}\big(I_{(p)},-I_{(q)},0_{(n-p-q)}\big)$.

如果方阵 S 还相合于对角形 $\mathrm{diag}\big(I_{(p')},-I_{(q')},0_{(n-p'-q')}\big)$,则存在 V 的基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵即为 $\mathrm{diag}\big(I_{(p')},-I_{(q')},0_{(n-p'-q')}\big)$.

由定理 9.2.5, $p'=p, q'=q$. 也就是说,方阵 S 所相合的对角形 $\mathrm{diag}\big(I_{(p)},-I_{(q)},0_{(n-p-q)}\big)$ 的对角元中 $+1$ 的个数 p 与 -1 的个数 q 是由方阵 S 自身所决定的. 它们也就分别称为方阵 S 的正、负惯性指数. 而正、负惯性指数 p 与 q 之差称为方阵 S 的符号差,记为 $\delta(S)$. 显然 $p+q=\mathrm{rank}\,S$, $\delta(S)=2p-\mathrm{rank}\,S$. 由于

$$\mathrm{diag}\big(I_{(p)},-I_{(q)},0_{(n-p-q)}\big)$$

是由方阵 S 自身所决定的,所以它称为方阵 S 在相合下的标准形.

设方阵 $S_1,S_2\in S(n,\mathbb{R})$ 相合,即存在 n 阶可逆实方阵 P,使得

$$S_2=P^{\mathrm{T}}S_1P.$$

则取 n 维实线性空间 V 的一组基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$,并令 $f(\alpha,\beta)=xS_1y^{\mathrm{T}}$,其中 x 与 y 分别是向量 α 与 β 在这组基下的坐标. 显然 $f(\alpha,\beta)$ 是空间 V 上的对称双线性函数. 记

$$(\eta_1,\eta_2,\ldots,\eta_n)=(\xi_1,\xi_2,\ldots,\xi_n)P.$$

则 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 V 的基,并且 $f(\alpha,\beta)$ 在这组基下的方阵即为 S_2.

由上一段的说明,方阵 S_1 与 S_2 的正、负惯性指数也就是 $f(\alpha,\beta)$ 的正、负惯性指数. 因此 S_1 与 S_2 的正、负惯性指数相同,从而它们的符号差相同. 由于相合的方阵一定是相抵的,所以 S_1 与 S_2 的秩相同. 这表明,实对称方阵的秩和符号差是实对称方阵在相合下的不变量.

反之,设实对称方阵 S_1 与 S_2 的秩与符号差相同,则由 $p+q=\mathrm{rank}\,S$ 与 $p-q=\delta(S)$,即知方阵 S_1 与 S_2 的正、负惯性指数相同,分别记为 p 与 q. 由上段说明,方阵 S_1 与 S_2 都相合于对角形

$$\mathrm{diag}\big(I_{(p)},-I_{(q)},0_{(n-p-q)}\big).$$

由于相合关系满足对称性与传递性,所以方阵 S_1 与 S_2 相合. 这表明,实对称方阵的秩与符号差是实对称方阵在相合下的全系不变量. 综合上述,即得

定理 9.2.6 设 n 阶实对称方阵 S 的正、负惯性指数分别为 p 与 q,则方阵 S 相合于如下的标准形:

$$\mathrm{diag}\big(I_{(p)},-I_{(q)},0_{(n-p-q)}\big).$$

并且实对称方阵的秩与符号差是实对称方阵在相合下的全系不变量.

由于 Euclid 空间是赋予一个给定的内积的实线性空间，所以可以讨论 Euclid 空间上的对称双线性函数.

定义 9.2.2　n 维 Euclid 空间 V 上的对称双线性函数 $f(\alpha,\beta)$ 在 V 的标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵 S 的特征值称为 $f(\alpha,\beta)$ 的**特征值**.

由定理 9.1.2 可以看出，V 上的对称线性函数 $f(\alpha,\beta)$ 在 V 的不同标准正交基下的方阵是相合的. 由于 V 的一组标准正交基到另一组标准正交基的过渡矩阵是正交方阵，所以 $f(\alpha,\beta)$ 在 V 的不同标准正交基下的方阵是正交相似的. 而正交相似的方阵具有相同的特征值. 所以，$f(\alpha,\beta)$ 的特征值的定义与标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 的选取无关. 也就是说，对称双线性函数 $f(\alpha,\beta)$ 的特征值是有确切定义的.

定理 9.2.7　设 $\lambda_1,\lambda_2,\ldots,\lambda_r$ 是 n 维 Euclid 空间 V 的对称双线性函数 $f(\alpha,\beta)$ 的全部非零特征值，$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r$，$r=\operatorname{rank} f$. 则存在 V 的标准正交基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$，使得 $f(\alpha,\beta)$ 在这组基下的方阵为如下的对角形：

$$\operatorname{diag}\big(\lambda_1,\lambda_2,\ldots,\lambda_r,\underbrace{0,0,\ldots,0}_{n-r\text{ 个}}\big).\tag{9.2.3}$$

换句话说，$f(\alpha,\beta)$ 在这组基下的表达式为

$$f(\alpha,\beta)=\lambda_1x_1y_1+\lambda_2x_2y_2+\cdots+\lambda_rx_ry_r,\tag{9.2.4}$$

其中 $x=(x_1,\ldots,x_n)$ 与 $y=(y_1,\ldots,y_n)$ 分别是 V 中向量 α 与 β 在这组基下的坐标.

证明　设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 V 的标准正交基，且 $f(\alpha,\beta)=\widetilde{x}S\widetilde{y}^{\mathrm{T}}$，其中 $\widetilde{x}$ 与 $\widetilde{y}$ 分别是向量 α 与 β 在这组基下的坐标，而 S 是 n 阶实对称方阵. 由定理 7.6.3，存在 n 阶实正交方阵 O，使得

$$O^{\mathrm{T}}SO=\operatorname{diag}\big(\lambda_1,\lambda_2,\ldots,\lambda_r,\underbrace{0,0,\ldots,0}_{n-r\text{ 个}}\big),$$

其中 $\lambda_1,\lambda_2,\ldots,\lambda_r$ 是方阵 S（也即 $f(\alpha,\beta)$）的特征值. 令

$$(\xi_1,\xi_2,\ldots,\xi_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)O,$$

则 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的标准正交基.

设向量 α 与 β 在这组基下的坐标分别为 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$. 则 $\widetilde{x}=xO^{\mathrm{T}}$，$\widetilde{y}=yO^{\mathrm{T}}$. 因此

$$\begin{aligned}f(\alpha,\beta)&=\widetilde{x}S\widetilde{y}^{\mathrm{T}}=xO^{\mathrm{T}}SOy^{\mathrm{T}}=x\Big(\operatorname{diag}\big(\lambda_1,\lambda_2,\ldots,\lambda_r,\underbrace{0,0,\ldots,0}_{n-r\text{ 个}}\big)\Big)y^{\mathrm{T}}\\&=\lambda_1x_1y_1+\lambda_2x_2y_2+\cdots+\lambda_rx_ry_r.\end{aligned}$$

■

对于复线性空间 V 上的对称双线性函数，由定理 9.2.2 得到

定理 9.2.8　设 $f(\alpha,\beta)\in S(V,V,\mathbb{C})$，则存在 n 维复线性空间 V 的一组基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$，使得 $f(\alpha,\beta)$ 在这组基下的方阵为如下的对角形：

$$\operatorname{diag}\big(I_{(r)},\ 0_{(n-r)}\big), \tag{9.2.5}$$

其中 $r=\operatorname{rank} f$. 换句话说,

$$f(\alpha,\beta)=x_1y_1+x_2y_2+\cdots+x_ry_r,$$

其中 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$ 分别是 V 中向量 α 与 β 在这组基下的坐标.

证明 由定理 9.2.2,存在 V 的基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵为对角形 (9.2.1). 于是当 $1\leqslant i\neq j\leqslant n$ 时,$f(\alpha_i,\alpha_j)=0$; 当 $1\leqslant i\leqslant r$ 时,$f(\alpha_i,\alpha_i)=a_i$; 而当 $r+1\leqslant i\leqslant n$ 时,$f(\alpha_i,\alpha_i)=0$. 令

$$\xi_i=\frac{1}{\sqrt{a_i}}\alpha_i,\ \ 1\leqslant i\leqslant r;\qquad \xi_j=\alpha_j,\ \ r+1\leqslant j\leqslant n.$$

显然 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 V 的基. 容易验证,当 $1\leqslant i\neq j\leqslant n$ 时,$f(\xi_i,\xi_j)=0$; 当 $1\leqslant i\leqslant r$ 时,$f(\xi_i,\xi_i)=1$; 而当 $r+1\leqslant i\leqslant n$ 时,$f(\xi_i,\xi_i)=0$. 所以 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵即为 (9.2.5). ■

定理 9.2.8 的矩阵形式是

定理 9.2.9 设 $S\in S(n,\mathbb{C})$. 则存在 n 阶可逆复方阵 P,使得

$$P^{\mathrm{T}}SP=\operatorname{diag}\big(I_{(r)},0_{(n-r)}\big),$$

其中 $r=\operatorname{rank} S$,即 n 阶复对称方阵 S 相合于对角形 (9.2.5). 而且复对称方阵的秩是复对称方阵在相合下的全系不变量.

证明 设 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 是 n 维复线性空间 V 的一组基. 定义 V 上的对称双线性函数 $f(\alpha,\beta)$ 为 $f(\alpha,\beta)=xSy^{\mathrm{T}}$,其中 x 与 y 分别是 V 中向量 α 与 β 在这组基下的坐标.

于是 S 即是 $f(\alpha,\beta)$ 在这组基下的方阵. 由定理 9.2.7,存在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵为对角形 (9.2.5). 由于一个双线性函数在不同基下的方阵是相合的,所以方阵 S 相合于对角形 (9.2.5).

设 n 阶复对称方阵 S_1 与 S_2 相合,即存在 n 阶可逆复方阵 P,使得 $S_2=P^{\mathrm{T}}S_1P$,因此 S_1 与 S_2 相抵,所以方阵 S_1 与 S_2 的秩相等,也就是说,复对称方阵的秩是复对称方阵在相合下的不变量.

反之,设 n 阶复对称方阵 S_1 与 S_2 的秩相等,且为 r,则由上段证明,方阵 S_1 与 S_2 都相合于对角形 (9.2.5). 由于方阵的相合关系满足对称性与传递性,所以方阵 S_1 与 S_2 相合.

这就证明了复对称方阵的秩是复对称方阵在相合下的全系不变量. ■

定理 9.2.9 完全解决了复对称方阵在相合下的标准形问题. 定理 9.2.9 中对角形 (9.2.5) 称为复对称方阵在相合下的标准形.

与对称双线性函数密切关联的是所谓二次型.

定义 9.2.3 设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的对称双线性函数.

则 $Q(\alpha)=f(\alpha,\alpha)$ 称为空间 V 上的**二次型**.

显然二次型 $Q(\alpha)$ 是空间 V 上的一元函数.

由定理 9.2.1 的证明,对称双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵 $S=(a_{ij})$ 是数域 $\mathbb{F}$ 上的 n 阶对称方阵,并且 $f(\alpha,\beta)=xSy^{\mathrm{T}}$,其中 x 与 y 分别是向量 α 与 β 在这组基下的坐标. 因此

$$Q(\alpha)=xSx^{\mathrm{T}}=\sum_{1\leqslant i,j\leqslant n}a_{ij}x_ix_j=\sum_{i=1}^{n}a_{ii}x_i^2+2\sum_{1\leqslant i<j\leqslant n}a_{ij}x_ix_j. \tag{9.2.6}$$

式 (9.2.6) 你为二次型 $Q(\alpha)$ 在基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的表达式,而方阵 S 称为二次型 $Q(\alpha)$ 在这组基下的方阵.

反之,设 $S=(a_{ij})$ 是数域 $\mathbb{F}$ 上的 n 阶对称方阵. 令 $f(\alpha,\beta)=xSy^{\mathrm{T}}$. 则 $f(\alpha,\beta)$ 是空间 V 上的对称双线性函数. 并且 $Q(\alpha)=xSx^{\mathrm{T}}$ 是 V 上的二次型.

有关对称双线性函数 $f(\alpha,\beta)$ 的术语都可以移到二次型 $Q(\alpha)$.

例如,数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的对称双线性函数 $f(\alpha,\beta)$ 的秩称为由 $f(\alpha,\beta)$ 决定的二次型 $Q(\alpha)$ 的秩. n 维实线性空间 V 上的对称双线性函数 $f(\alpha,\beta)$ 的正、负惯性指数与符号差分别称为 $f(\alpha,\beta)$ 决定的二次型 $Q(\alpha)$ 的正、负惯性指数与符号差. 如果 n 维实线性空间 V 上的对称双线性函数 $f(\alpha,\beta)$ 是正定(或者半正定、负定与半负定)的,则由 $f(\alpha,\beta)$ 决定的二次型 $Q(\alpha)$ 称为正定(或者半正定、负定与半负定)的.

由于二次型 $Q(\alpha)$ 是由对称双线性函数 $f(\alpha,\beta)$ 所决定的,所以有关对称双线性函数 $f(\alpha,\beta)$ 的结论都可以移到二次型 $Q(\alpha)$.

定理 9.2.10　设 $Q(\alpha)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的二次型. 则存在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$,使得

$$Q(\alpha)=a_1x_1^2+a_2x_2^2+\cdots+a_rx_r^2, \tag{9.2.7}$$

其中 $x=(x_1,x_2,\dots,x_n)$ 是向量 α 在这组基下的坐标,而 $a_1,a_2,\dots,a_r\in\mathbb{F}$ 均不为零,r 为二次型 $Q(\alpha)$ 的秩.

定理 9.2.11　设 $Q(\alpha)$ 是 n 维实线性空间 V 上的二次型,则存在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$,使得

$$Q(\alpha)=x_1^2+x_2^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+q}^2, \tag{9.2.8}$$

其中 $x=(x_1,x_2,\dots,x_n)$ 是向量 α 在这组基下的坐标,而 p 与 q 分别是二次型 $Q(\alpha)$ 的正、负惯性指数.

定理 9.2.12　设 $Q(\alpha)$ 是 n 维 Euclid 空间 V 上的二次型,则存在 V 的标准正交基 $\{\xi_1,\xi_2,\dots,\xi_n\}$,使得

$$Q(\alpha)=\lambda_1x_1^2+\lambda_2x_2^2+\cdots+\lambda_rx_r^2, \tag{9.2.9}$$

其中 $x=(x_1,x_2,\dots,x_n)$ 是向量 α 在这组基下的坐标,而 $\lambda_1,\lambda_2,\dots,\lambda_r$ 是二次型 $Q(\alpha)$ 的全部非零特征值. 也即决定二次型 $Q(\alpha)$ 的对称双线性函数 $f(\alpha,\beta)$ 的全

部非零特征值，并且 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r$.

应当指出，如果 n 维实线性空间 V 上的二次型 $Q(\alpha)$ 是正定的，并且它在 V 的基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 下的方阵为 S，即 $Q(\alpha) = xSx^{\mathrm{T}}$，其中 $x = (x_1, x_2, \ldots, x_n)$ 是向量 α 在这组基下的坐标. 显然 $x \in \mathbb{R}^n$. 由于 $Q(\alpha)$ 是正定的，所以对任意非零 $\alpha \in V$，$Q(\alpha) > 0$，即对任意非零 $x \in \mathbb{R}^n$，$xSx^{\mathrm{T}} > 0$. 因此方阵 S 是正定的. 反之亦然. 这表明，二次型 $Q(\alpha)$ 为正定的充分必要条件是，它的方阵 S 是正定的. 于是可以利用§7.7 中关于正定对称方阵的结论来判定二次型 $Q(\alpha)$ 的正定性. 当然也可以用定理 9.2.11 来判定.

由定理 9.2.11 直接得到，二次型 $Q(\alpha)$ 半正定的充分必要条件是，它的秩等于符号差. 而 $Q(\alpha)$ 正定的充分必要条件是，它的秩等于符号差，并且秩为 n.

定理 9.2.13 设 $Q(\alpha)$ 是 n 维复线性空间 V 上的二次型，则存在 V 的基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$，使得

$$Q(\alpha) = x_1^2 + x_2^2 + \cdots + x_r^2, \tag{9.2.10}$$

其中 $x = (x_1, x_2, \ldots, x_n)$ 是向量 α 在这组基下的坐标，而 r 是二次型 $Q(\alpha)$ 的秩.

最后介绍二次型化简方法.

一、Euclid 空间的二次型的化简

设 n 维 Euclid 空间 V 的二次型 $Q(\alpha)$ 在 V 的标准正交基 $\{\xi_1, \ldots, \xi_n\}$ 下的表达式为

$$Q(\alpha) = xSx^{\mathrm{T}} = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{1\leqslant i<j\leqslant n} a_{ij}x_ix_j,$$

其中 $x = (x_1, x_2, \ldots, x_n)$ 是向量 α 在基 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 下的坐标，而 $S = (a_{ij})$ 是二次型 $Q(\alpha)$ 在这组基下的方阵.

问题是寻求 V 的一组标准正交基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$，使得

$$Q(\alpha) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2,$$

其中 $y = (y_1, y_2, \ldots, y_n)$ 是向量 α 在基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$ 下的坐标，而 $\lambda_1, \lambda_2, \ldots, \lambda_r$ 是 $Q(\alpha)$ 的全部非零特征值，$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r$.

这就是所谓化二次型 $Q(\alpha)$ 为**主轴形式**问题.

化二次型 $Q(\alpha)$ 为主轴形式的步骤如下.

(1) 利用对称方阵 S 的特征多项式 $\varphi(\lambda) = \det\big(\lambda I_{(n)} - S\big)$，求出方阵 S 的全部不同特征值 $\lambda_1, \lambda_2, \ldots, \lambda_t$，$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_t$，以及它们的代数重数 $e_1, e_2, \ldots, e_t$；

(2) 对每个 i，确定属于 λ_i 的特征子空间，也即确定齐次方程组 $x\big(\lambda I_{(n)} - S\big) = 0$ 的解空间 V_{λ_i}，其中 x 是 n 维实的行向量. 解空间 V_{λ_i} 是 e_i 维的，$i = 1, 2, \ldots, t$；

(3) 解空间 V_{λ_i} 是 n 维实的行向量空间 $\mathbb{R}^n$ 的子空间，而 $\mathbb{R}^n$ 在标准内积下成为一个 Euclid 空间. 求出解空间 V_{λ_i} 的标准正交基，记为 $\{\eta_1^{(i)}, \eta_2^{(i)}, \ldots, \eta_{e_i}^{(i)}\}$；

(4) 将解空间 $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_t}$ 的标准正交基合并，即得 $\mathbb{R}^n$ 的标准正交基

$$\left\{\eta_1^{(1)},\dots,\eta_{e_1}^{(1)},\eta_1^{(2)},\dots,\eta_{e_2}^{(2)},\dots,\eta_1^{(t)},\dots,\eta_{e_t}^{(t)}\right\}.$$

记

$$O=\left(\eta_1^{(1)},\dots,\eta_{e_1}^{(1)},\eta_1^{(2)},\dots,\eta_{e_2}^{(2)},\dots,\eta_1^{(t)},\dots,\eta_{e_t}^{(t)}\right)^{\mathrm{T}}.$$

则 O 是 n 阶实正交方阵,并且

$$OSO^{\mathrm{T}}=\mathrm{diag}\left(\lambda_1 I_{(e_1)},\lambda_2 I_{(e_2)},\dots,\lambda_t I_{(e_t)}\right);$$

(5) 设 $(\xi_1,\xi_2,\dots,\xi_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)O^{\mathrm{T}}$,则 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 是 V 的标准正交基,而且 $Q(\alpha)$ 在这组基下具有主轴形式:

$$Q(\alpha)=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_t y_t^2.$$

二、实线性空间的二次型的化简

设 n 维实线性空间 V 的二次型 $Q(\alpha)$ 在 V 的基 $\{\alpha_1,\dots,\alpha_n\}$ 下的表达式为

$$Q(\alpha)=xSx^{\mathrm{T}}=\sum_{i=1}^{n}a_{ii}x_i^2+2\sum_{1\leqslant i<j\leqslant n}a_{ij}x_ix_j, \tag{9.2.11}$$

其中 $x=(x_1,x_2,\dots,x_n)$ 是向量 α 在这组基下的坐标,而 $S=(a_{ij})$ 是 $Q(\alpha)$ 在此基下的方阵.

问题是寻求 V 的一组基 $\{\xi_1,\xi_2,\dots,\xi_n\}$,使得 $Q(\alpha)$ 在这组基下的表达式为

$$Q(\alpha)=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2, \tag{9.2.12}$$

其中 $y=(y_1,y_2,\dots,y_n)$ 是向量 α 在此基下的坐标,而 p 与 q 分别是 $Q(\alpha)$ 的正、负惯性指数. 这就是所谓化二次型为标准形问题.

如果 $Q(\alpha)$ 是零函数,则无需化简. 因此设 $Q(\alpha)\neq 0$,也即方阵 $S=(a_{ij})\neq 0$.

(1) 设表达式 (9.2.11) 中含有平方项 $a_{11}x_1^2$,即 $a_{11}\neq 0$,则式 (9.2.11) 可以改写为

$$Q(\alpha)=a_{11}x_1^2+2\Big(\sum_{j=2}^{n}a_{1j}x_j\Big)x_1+\sum_{2\leqslant i,j\leqslant n}a_{ij}x_ix_j.$$

将上式配方,即

$$\begin{aligned}Q(\alpha)=a_{11}\left(x_1^2+2\Big(\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\Big)x_1+\Big(\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\Big)^2\right)&\\ -a_{11}\Big(\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\Big)^2+\sum_{2\leqslant i,j\leqslant n}a_{ij}x_ix_j.&\end{aligned} \tag{9.2.13}$$

记

$$Q_1(\beta)=\sum_{2\leqslant i,j\leqslant n}a_{ij}x_ix_j-a_{11}\Big(\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\Big)^2.$$

令

$$\begin{cases}y_1=\sqrt{|a_{11}|}\Big(x_1+\displaystyle\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\Big), & \\ y_i=x_i, & i=2,3,\dots,n.\end{cases}$$

即令

$$(y_1,y_2,\dots,y_n)=(x_1,x_2,\dots,x_n)\begin{pmatrix}1&0&\cdots&0\\ \dfrac{a_{12}}{a_{11}}&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ \dfrac{a_{1n}}{a_{11}}&0&\cdots&1\end{pmatrix}\begin{pmatrix}\sqrt{|a_{11}|}&0&\cdots&0\\ 0&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&1\end{pmatrix},\tag{9.2.14}$$

则式 (9.2.13) 化为

$$Q(\alpha)=\pm y_1^2+Q_1(\beta),\tag{9.2.15}$$

其中 y_1^2 的系数的正负号与 a_{11} 相同. 由 (9.2.14) 可以看出,由 (9.2.11) 化为 (9.2.15) 所作的变换是坐标变换. 相应的基变换是

$$(\xi_1,\xi_2,\dots,\xi_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)\begin{pmatrix}1&-\dfrac{a_{12}}{a_{11}}&\cdots&-\dfrac{a_{1n}}{a_{11}}\\ 0&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&1\end{pmatrix}\begin{pmatrix}\dfrac{1}{\sqrt{|a_{11}|}}&0&\cdots&0\\ 0&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&1\end{pmatrix}.$$

而式 (9.2.15) 是 $Q(\alpha)$ 在基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的表达式. 式 (9.2.15) 中的 $Q_1(\beta)$ 为

$$Q_1(\beta)=\sum_{2\leqslant i,j\leqslant n}a_{1j}y_iy_j-a_{11}\Big(\sum_{j=2}^n\frac{a_{1j}}{a_{11}}y_j\Big)^2.$$

它是 V 中向量 $\xi_1,\xi_2,\dots,\xi_n$ 生成的 $n-1$ 维实线性空间的二次型.

(2) 设式 (9.2.11) 中不含平方项 $a_{11}x_1^2$,但含某个平方项 $a_{ii}x_i^2$, $2\leqslant i\leqslant n$,则令

$$\begin{cases}y_1=x_i,\quad y_i=x_1,\\ y_j=x_j,\quad j\neq1,\quad j\neq i,\quad 2\leqslant j\leqslant n.\end{cases}$$

即令

$$(y_1,y_2,\dots,y_n)=(x_1,x_2,\dots,x_n)P_{1i},\tag{9.2.16}$$

其中 P_{1i} 是对换 n 阶单位方阵 $I_{(n)}$ 的第 1 行与第 i 行所得到的初等置换方阵. 式 (9.2.16) 给出的变换是坐标变换,相应的基变换是

$$(\xi_1,\xi_2,\dots,\xi_n)=(\alpha_1,\alpha_2,\dots,\alpha_n)P_{1i}.$$

易知 $Q(\alpha)$ 在基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的表达式含有平方项 $a_{ii}y_1^2$. 于是就化为 (1).

(3) 设表达式 (9.2.11) 中不含平方项,即 $a_{11},a_{22},\dots,a_{nn}$ 全为零,且 $a_{12}\neq0$. 将式 (9.2.11) 改写为

$$Q(\alpha)=2a_{12}x_1x_2+2\Big(\sum_{j=3}^n(a_{1j}x_1+a_{2j}x_2)x_j\Big)+2\sum_{3\leqslant i<j\leqslant n}a_{ij}x_ix_j.\tag{9.2.17}$$

令

$$\begin{cases}y_1=\dfrac12x_1+\dfrac12x_2,\quad y_2=\dfrac12x_1-\dfrac12x_2,\\ y_i=x_i,\qquad\qquad i=3,4,\dots,n.\end{cases}$$

即令

$$(y_1,y_2,\ldots,y_n)=(x_1,x_2,\ldots,x_n)\begin{pmatrix}\frac{1}{2} & -\frac{1}{2} & 0 & \cdots & 0\\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}. \tag{9.2.18}$$

与坐标变换 (9.2.18) 相应的基变换是

$$(\xi_1,\xi_2,\ldots,\xi_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)\begin{pmatrix}1 & -1 & 0 & \cdots & 0\\ 1 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}.$$

$Q(\alpha)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的表达式为

$$Q(\alpha)=2a_{12}y_1^2-2a_{12}y_2^2+2\sum_{j=3}^{n}(a_{1j}+a_{2j})y_1y_j+2\sum_{j=3}^{n}(a_{1j}-a_{2j})y_2y_j+2\sum_{3\leqslant i<j\leqslant n}a_{ij}y_iy_j.$$

于是化为 (1).

(4) 设表达式 (9.2.11) 不含平方项,即 $a_{11},a_{22},\ldots,a_{nn}$ 全为零,且 $a_{12}=0$,但某个 $a_{ij}\neq 0, i\neq j$. 令

$$(y_1,y_2,\ldots,y_n)=(x_1,x_2,\ldots,x_n)P_{1i}P_{2j},$$

即作基变换

$$(\xi_1,\xi_2,\ldots,\xi_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)P_{1i}P_{2j},$$

也即对换基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 中向量 α_1 与 α_i 以及向量 α_2 与 α_j 的位置,得 V 的一组新基,并记新基为 $\{\xi_1,\xi_2,\ldots,\xi_n\}$. 则 $Q(\alpha)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的表达式不含平方项,但乘积项 y_1y_2 的系数为 $a_{ij}\neq 0$,于是化为 (3).

这表明,式 (9.2.11) 可以经过基变换,使得 $Q(\alpha)$ 在新基下的表达式为 (9.2.15). 然后再对 $n-1$ 维实线性空间上的二次型 $Q_1(\beta)$ 重复上面的过程,即可将式 (9.2.11) 化为式 (9.2.12). 上述化二次型 $Q(\alpha)$ 为标准形的方法实质上是配方法. 这种方法也可以用矩阵形式表述. 设二次型 $Q(\alpha)$ 在基 $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 下的方阵 $S=(a_{ij})\neq 0$. 如果方阵 S 的对角元 $a_{11}=0$,而 $a_{ii}\neq 0, 2\leqslant i\leqslant n$,则

$$S_1=P_{1i}^{\mathrm{T}}SP_{1i}\begin{pmatrix}a_{ii} & *\\ * & *\end{pmatrix}, \tag{9.2.19}$$

其中 P_{1i} 是对换 n 阶单位方阵 $I_{(n)}$ 的第 1 行与第 i 行所得到的初等置换方阵.

对方阵 S 前乘与后乘同一个初等置换方阵,也即对调方阵 S 的第 i 行与第 j 行,然后对调第 i 列与第 j 列,称为对方阵 S 施行一次**同步置换**.

式 (9.2.19) 表明,当 $a_{11}=0$,而 $a_{ii}\neq 0$ 时,方阵 S 可以经过同步置换,使得到的方阵 S_1 的西北角元素不为零,而且方阵 S_1 是对称的. 显然方阵 S 与 S_1 相合,如果方

阵 S 的对角元全为零,则因方阵 $S \neq 0$,故有某个元素 $a_{ij} \neq 0, i \neq j$. 于是

$$S_2 = P_{2j}^{\mathrm{T}}(P_{1i}^{\mathrm{T}} S P_{1i}) P_{2j} = \begin{pmatrix} 0 & a_{ij} & & * \\ a_{ij} & 0 & & \\ & & \ddots & \\ * & & & 0 \end{pmatrix}$$

即对称方阵 S 可以经同步置换化为对称方阵 S_2,而方阵 S_2 的第 1 行与第 2 列交叉位置上的元素不为零. 所以当方阵 S 的对角元为零时,可以设方阵 S 的元素 $a_{12} \neq 0$. 取 n 阶可逆实方阵 P 为

$$P = \operatorname{diag}\left(\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, I_{(n-2)}\right),$$

则

$$S_3 = P^{\mathrm{T}} S P = \begin{pmatrix} \begin{pmatrix} 2a_{12} & 0 \\ 0 & 2a_{12} \end{pmatrix} & * \\ * & * \end{pmatrix}.$$

对称方阵 S 相合于 S_3,而且 S_3 的西北角元素不为零.

上面的讨论表明,如果对称方阵 $S = (a_{ij}) \neq 0$,则不妨设 $a_{11} \neq 0$. 于是仿照式 (9.2.14) 的基变换形式,

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\dfrac{a_{12}}{a_{11}} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\dfrac{a_{1n}}{a_{11}} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 & -\dfrac{a_{12}}{a_{11}} & \cdots & -\dfrac{a_{1n}}{a_{11}} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} - \dfrac{a_{12}a_{12}}{a_{11}} & \cdots & a_{2n} - \dfrac{a_{12}a_{1n}}{a_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{2n} - \dfrac{a_{12}a_{1n}}{a_{11}} & \cdots & a_{nn} - \dfrac{a_{1n}a_{1n}}{a_{11}} \end{pmatrix}. \tag{9.2.20}$$

将上式写成分块矩阵形式. 设

$$S = \begin{pmatrix} a_{11} & \beta \\ \beta^{\mathrm{T}} & S_{22} \end{pmatrix},$$

其中 S_{22} 是 $n-1$ 阶子方阵,β 是 $1 \times (n-1)$ 子矩阵. 则

$$\begin{pmatrix} 1 & 0 \\ -a_{11}^{-1}\beta^{\mathrm{T}} & I_{(n-1)} \end{pmatrix} \begin{pmatrix} a_{11} & \beta \\ \beta^{\mathrm{T}} & S_{22} \end{pmatrix} \begin{pmatrix} 1 & -a_{11}^{-1}\beta \\ 0 & I_{(n-1)} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & S_{22} - a_{11}^{-1}\beta^{\mathrm{T}}\beta \end{pmatrix}. \tag{9.2.21}$$

顺带指出,式 (9.2.21) 是式 (9.2.20) 的分块矩阵形式. 式 (9.2.20) 源于 (1) 对式 (9.2.11) 的二次型 $Q(\alpha)$ 进行配方. 而式 (9.2.21) 是第 3 章中提到的 Schur 公式的特殊情形. 因此可以说 Schur 公式是配方法的一种推广.

由于

$$\begin{pmatrix} \frac{1}{|a_{11}|} & 0 \\ 0 & I_{(n-1)} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} a_{11} & 0 \\ 0 & S_{22} - a_{11}^{-1}\beta^{\mathrm{T}}\beta \end{pmatrix} \begin{pmatrix} \frac{1}{|a_{11}|} & 0 \\ 0 & I_{(n-1)} \end{pmatrix} = \begin{pmatrix} \frac{a_{11}}{|a_{11}|} & 0 \\ 0 & S_{22} - a_{11}^{-1}\beta^{\mathrm{T}}\beta \end{pmatrix},$$

其中 $S_{22} - a_{11}^{-1}\beta^{\mathrm{T}}\beta$ 是 $n-1$ 阶实对称方阵，且 $\frac{a_{11}}{|a_{11}|} = \pm 1$.

于是只要对 $S_{22} - a_{11}^{-1}\beta^{\mathrm{T}}\beta$ 重复上面的做法，即可求得 n 阶可逆实方阵 Q，使得方阵相合于对角形

$$\operatorname{diag}\left(\delta_1, \delta_2, \ldots, \delta_r, 0, \ldots, 0\right),$$

其中 $\delta_i = \pm 1, 1 \leqslant i \leqslant r$，且 $r = \operatorname{rank} S$. 再作适当的同步置换，即可将方阵 S 化为相合下的标准形.

由于正交相似的方阵一定是相合的，因此也可以将实对称方阵 S 按 Euclid 空间中二次型化简的方法先化为正交相似下的标准形，然后再化为相合下的标准形.

至于 n 维复线性空间上的二次型的化简，方法大体同实线性空间.

例 9.2.1 设 3 维 Euclid 空间 V 的二次型 $Q(\alpha)$ 在 V 的标准正交基 $\left\{\alpha_1, \alpha_2, \alpha_3\right\}$ 下的表达式为

$$Q(\alpha) = x_1x_2 + x_2x_3,$$

其中 $x = (x_1, x_2, x_3)$ 是 V 中向量 α 在基 $\left\{\alpha_1, \alpha_2, \alpha_3\right\}$ 下的坐标. 将二次型 $Q(\alpha)$ 化为主轴形式.

解 将二次型 $Q(\alpha)$ 在基 $\left\{\alpha_1, \alpha_2, \alpha_3\right\}$ 下的表达式写成矩阵形式

$$Q(\alpha) = xSx^{\mathrm{T}},$$

其中

$$S = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

则方阵 S 的特征多项式

$$\varphi(\lambda) = \det\left(\lambda I_{(3)} - S\right) = \lambda^3 - \frac{1}{2}\lambda.$$

所以方阵 S 的特征值为

$$\frac{\sqrt{2}}{2}, \quad -\frac{\sqrt{2}}{2}, \quad 0.$$

方阵 S 的属于特征值 $\frac{\sqrt{2}}{2}$ 的单位特征向量 ξ_1 在基 $\left\{\alpha_1, \alpha_2, \alpha_3\right\}$ 下的坐标记为 $x = (x_1, x_2, x_3)$，则 x 满足方程 $x\left(\frac{\sqrt{2}}{2}I_{(3)} - S\right) = 0$，即得齐次方程组：

$$\begin{cases} \frac{\sqrt{2}}{2}x_1 - \frac{1}{2}x_2 = 0, \\ -\frac{1}{2}x_1 + \frac{\sqrt{2}}{2}x_2 - \frac{1}{2}x_3 = 0, \\ -\frac{\sqrt{2}}{2}x_2 + \frac{1}{2}x_3 = 0. \end{cases}$$

解得 $x_1=x_3,x_2=\sqrt{2}x_3$.

由于 ξ_1 是单位向量,所以 $x_1^2+x_2^2+x_3^2=\frac{x_3^2}{4}=1$,因此 $x_3=\pm\frac{1}{2}$. 取 $x_3=\frac{1}{2}$,则

$$x=\left(\frac{1}{2},\frac{\sqrt{2}}{2},\frac{1}{2}\right).$$

同理可以求得分别属于方阵 S 的特征值 $-\frac{\sqrt{2}}{2}$ 与 0 的单位特征向量 ξ_2 与 ξ_3 在基 $\{\alpha_1,\alpha_2,\alpha_3\}$ 下的坐标为

$$\left(\frac{1}{2},-\frac{\sqrt{2}}{2},\frac{1}{2}\right)\quad 与\quad \left(\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}\right).$$

由于属于方阵 S 的不同特征值的特征向量是正交的,所以

$$O=\begin{pmatrix}\frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{\sqrt{2}}{2} & \frac{1}{2}\\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2}\end{pmatrix}$$

是正交方阵. 而且

$$O^{\mathrm{T}}SO=\mathrm{diag}\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0\right),\qquad (\xi_1,\xi_2,\xi_3)=(\alpha_1,\alpha_2,\alpha_3)O.$$

设向量 α 在基 $\{\xi_1,\xi_2,\xi_3\}$ 下的坐标为 $y=(y_1,y_2,y_3)$,则 $x=yO^{\mathrm{T}}$. 因此

$$\begin{aligned}Q(\alpha)&=xSx^{\mathrm{T}}=y\left(O^{\mathrm{T}}SO\right)y^{\mathrm{T}}\\&=y\,\mathrm{diag}\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0\right)y^{\mathrm{T}}=\frac{\sqrt{2}}{2}y_1^2-\frac{\sqrt{2}}{2}y_2^2.\end{aligned}$$

上式即是二次型 $Q(\alpha)$ 在基 $\{\xi_1,\xi_2,\xi_3\}$ 下的表达式,而且是主轴形式. ■

例 9.2.2 把 4 维实线性空间 V 的二次型

$$Q(\alpha)=x_1^2+x_2^2+x_3^2-2x_4^2-2x_1x_2+2x_1x_3-2x_1x_4+2x_2x_3-4x_2x_4 \tag{9.2.22}$$

化为标准形.

解 将二次型 $Q(\alpha)$ 配方,

$$\begin{aligned}Q(\alpha)&=x_1^2+2x_1(-x_2+x_3-x_4)+x_2^2+x_3^2-2x_4^2+2x_2x_3-4x_2x_4\\&=\left(x_1^2+2x_1(-x_2+x_3-x_4)+(-x_2+x_3-x_4)^2\right)-(-x_2+x_3-x_4)^2\\&\quad+x_2^2+x_3^2-2x_4^2+2x_2x_3-4x_2x_4\\&=(x_1-x_2+x_3-x_4)^2-3x_4^2+4x_2x_3-6x_2x_4+2x_3x_4.\end{aligned}$$

作坐标变换

$$\begin{cases}u_1=x_1-x_2+x_3-x_4,\\ u_i=x_i,\quad i=2,3,4.\end{cases}$$

于是

$$Q(\alpha)=u_1^2-3u_4^2+4u_2u_3-6u_2u_4+2u_2u_4.$$

再配方,

$$
\begin{aligned}
Q(\alpha) &= u_1^2 - \left(\frac{1}{3}(3u_2-u_3)^2 + 2(3u_2-u_3)u_4 + 3u_4^2\right) + 3u_2^2 + 2u_2u_3 + \frac{1}{3}u_3^2 \\
&= u_1^2 - \left(\sqrt{3}\,u_2 - \frac{1}{\sqrt{3}}u_3 + \sqrt{3}\,u_4\right)^2 + \left(\sqrt{3}\,u_2 + \frac{1}{\sqrt{3}}u_3\right)^2.
\end{aligned}
$$

令

$$
\begin{cases}
y_1 = u_1, & y_2 = \sqrt{3}\,u_2 + \dfrac{1}{\sqrt{3}}u_3, \\
y_3 = \sqrt{3}\,u_2 - \dfrac{1}{\sqrt{3}}u_3 + \sqrt{3}\,u_4, & y_4 = u_4.
\end{cases}
$$

则

$$
Q(\alpha) = y_1^2 + y_2^2 - y_3^2. \tag{9.2.23}
$$

由形式 (9.2.22) 的二次型 $Q(\alpha)$ 化为标准形 (9.2.23) 的坐标变换是

$$
\begin{cases}
y_1 = x_1 - x_2 + x_3 - x_4, & y_2 = \sqrt{3}\,x_2 + \dfrac{1}{\sqrt{3}}x_3, \\
y_3 = \sqrt{3}\,x_2 - \dfrac{1}{\sqrt{3}}x_3 + \sqrt{3}\,u_4, & y_4 = x_4.
\end{cases}
$$

■

例 9.2.3　把 3 维实线性空间 V 上的二次型

$$
Q(\alpha) = x_1x_2 + x_2x_3
$$

化为标准形.

解　把二次型 $Q(\alpha)$ 写成矩阵形式:

$$
Q(\alpha) = (x_1, x_2, x_3)\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := xSx^{\mathrm{T}}. \tag{9.2.24}
$$

记 $Q = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, 则 $Q^{\mathrm{T}}SQ = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & -1 & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$. 记 $R = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$, 则

$$
R^{\mathrm{T}}(Q^{\mathrm{T}}SQ)R = \mathrm{diag}(1, -1, 0).
$$

记 $P = (QR)^{-1}$, 则

$$
S = P^{\mathrm{T}}\,\mathrm{diag}(1, -1, 0)P,
$$

因此

$$
Q(\alpha) = xP^{\mathrm{T}}\,\mathrm{diag}(1, -1, 0)Px^{\mathrm{T}}.
$$

作坐标变换 $y = xP^{\mathrm{T}}$, 即

$$
\begin{cases}
y_1 = \dfrac{1}{2}x_1 + \dfrac{1}{2}x_2 + \dfrac{1}{2}x_3, & y_2 = -\dfrac{1}{2}x_1 + \dfrac{1}{2}x_2 - \dfrac{1}{2}x_3, \\
y_3 = x_3.
\end{cases}
$$

则二次型 $Q(\alpha)$ 化为标准形

$$
Q(\alpha) = y\,\mathrm{diag}(1, -1, 0)y^{\mathrm{T}} = y_1^2 - y_2^2.
$$

■

例 9.2.4　设 S 与 $S-\alpha^{\mathrm{T}}\alpha$ 是 n 阶可逆实对称方阵,其中 a 是 n 维实的行向量. 证明对称方阵 S 的符号差 $\delta(S)$ 满足

$$\delta(S)=\begin{cases}\delta(S-\alpha^{\mathrm{T}}\alpha)+2, & \alpha S^{-1}\alpha^{\mathrm{T}}>1;\\ \delta(S-\alpha^{\mathrm{T}}\alpha), & \alpha S^{-1}\alpha^{\mathrm{T}}<1.\end{cases}$$

证明　考虑 $n+1$ 阶实对称方阵

$$S_0=\begin{pmatrix}1 & \alpha\\ \alpha^{\mathrm{T}} & S\end{pmatrix}.$$

因为

$$\begin{pmatrix}1 & 0\\ -\alpha^{\mathrm{T}} & I_{(n)}\end{pmatrix}\begin{pmatrix}1 & \alpha\\ \alpha^{\mathrm{T}} & S\end{pmatrix}\begin{pmatrix}1 & -\alpha\\ 0 & I_{(n)}\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & S-\alpha^{\mathrm{T}}\alpha\end{pmatrix},$$

$$\begin{pmatrix}1 & -\alpha S^{-1}\\ 0 & I_{(n)}\end{pmatrix}\begin{pmatrix}1 & \alpha\\ \alpha^{\mathrm{T}} & S\end{pmatrix}\begin{pmatrix}1 & 0\\ -S^{-1}\alpha^{\mathrm{T}} & I_{(n)}\end{pmatrix}=\begin{pmatrix}1-\alpha S\alpha^{\mathrm{T}} & 0\\ 0 & S\end{pmatrix},$$

所以 $n+1$ 阶实对称方阵

$$S_1=\begin{pmatrix}1 & 0\\ 0 & S-\alpha^{\mathrm{T}}\alpha\end{pmatrix}\quad 与\quad S_2=\begin{pmatrix}1-\alpha S\alpha^{\mathrm{T}} & 0\\ 0 & S\end{pmatrix}$$

相合,因此它们的符号差相同. 显然 $\delta(S_1)=1+\delta(S-\alpha^{\mathrm{T}}\alpha)$.

当 $\alpha S^{-1}\alpha^{\mathrm{T}}>1$,即 $1-\alpha S^{-1}\alpha^{\mathrm{T}}<0$ 时,$\delta(S_2)=\delta(S)-1$. 因此,$1+\delta(S-\alpha^{\mathrm{T}}\alpha)=\delta(S)-1$,即 $\delta(S)=\delta(S-\alpha\alpha^{\mathrm{T}})+2$.

当 $\alpha S^{-1}\alpha^{\mathrm{T}}<1$,即 $1-\alpha S^{-1}\alpha^{\mathrm{T}}>0$ 时,$\delta(S_2)=\delta(S)+1$. 因此 $1+\delta(S-\alpha^{\mathrm{T}}\alpha)=\delta(S)+1$,即 $\delta(S)=\delta(S-\alpha^{\mathrm{T}}\alpha)$. ■

习　题　9.2

1. 已知有理数域 $\mathbb{Q}$ 上的 3 阶对称方阵 S 为

$$S=\begin{pmatrix}-2 & 3 & 5\\ 3 & 1 & -1\\ 5 & -1 & 4\end{pmatrix}.$$

求 $\mathbb{Q}$ 上的 3 阶可逆方阵 P,使得 $P^{\mathrm{T}}SP$ 是对角形.

2. 求有理数域 $\mathbb{Q}$ 上的 2 阶可逆方阵 P,使得

$$P^{\mathrm{T}}\operatorname{diag}(5,5)P=\operatorname{diag}(1,1).$$

3. 求下列实对称方阵在相合(通过实方阵)下的标准形.

(1) $\begin{pmatrix}2 & 5 & 8\\ 5 & 3 & 1\\ 8 & 1 & 0\end{pmatrix}$;　　(2) $\begin{pmatrix}0_{(n)} & I_{(n)}\\ I_{(n)} & 0_{(n)}\end{pmatrix}$;

(3) $\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 2 & 2 & 2\\ 1 & 2 & 3 & 3\\ 1 & 2 & 3 & 4\end{pmatrix}$;　　(4) $\begin{pmatrix}0_{(n)} & I_{(n)} & 0\\ I_{(n)} & 0_{(n)} & 0\\ 0 & 0 & 1\end{pmatrix}$;

(5) $\begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & \iddots & 0 \\ 0 & \iddots & \iddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}_{n\times n}$; (6) $\begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \cdots & 0 & 0 & \frac{1}{2} & 1 \end{pmatrix}_{n\times n}$.

4. 把下列实线性空间 V 上的二次型化为标准形.

(1) $2x_1^2+3x_2^2+10x_1x_2+16x_1x_3+2x_2x_3$; (2) $x_1^2+2x_2^2+2x_3^2+4x_4^2+x_1x_2+x_2x_4$;

(3) $\sum\limits_{j=1}^{n-1} x_jx_{j+1}$; (4) $\sum\limits_{1\leqslant i<j\leqslant n} x_ix_j$;

(5) $\sum\limits_{1\leqslant i<j\leqslant n} (-1)^{i+j}x_ix_j$; (6) $\sum\limits_{1\leqslant i<j\leqslant n} |i-j|x_ix_j$.

5. 证明,如果二次型 $Q(\alpha)=0$ 当且仅当 $\alpha=0$,则 $Q(\alpha)$ 或者是正定的,或者是负定的.

6. 设二次型 $Q(\alpha)=xSx^{\mathrm{T}}$,其中方阵 S 的顺序子式 $S\begin{pmatrix} 1 & 2 & \cdots & j \\ 1 & 2 & \cdots & j \end{pmatrix}\neq 0$, $j=1,2,\dots,n$. 证明,二次型 $Q(\alpha)$ 可以化为

$$Q(\alpha)=S\begin{pmatrix}1\\1\end{pmatrix}y_1^2+\frac{S\begin{pmatrix}1&2\\1&2\end{pmatrix}}{S\begin{pmatrix}1\\1\end{pmatrix}}y_2^2+\cdots+\frac{S\begin{pmatrix}1&2&\cdots&n\\1&2&\cdots&n\end{pmatrix}}{S\begin{pmatrix}1&2&\cdots&n-1\\1&2&\cdots&n-1\end{pmatrix}}y_n^2.$$

7. 求下列复对称方阵在相合(通过复方阵)下的标准形,其中 $\mathrm{i}^2=-1$.

(1) $\begin{pmatrix} 1 & 1+\mathrm{i} & 2+\mathrm{i} & \cdots & n+\mathrm{i} \\ 1+\mathrm{i} & 1 & 0 & \cdots & 0 \\ 2+\mathrm{i} & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ n+\mathrm{i} & 0 & \cdots & 0 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} 1 & \frac{\mathrm{i}}{2} & 0 & 0 & \cdots & 0 \\ \frac{\mathrm{i}}{2} & 1 & \frac{\mathrm{i}}{2} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\mathrm{i}}{2} & 1 & \frac{\mathrm{i}}{2} \\ 0 & \cdots & 0 & 0 & \frac{\mathrm{i}}{2} & 1 \end{pmatrix}_{n\times n}$;

(3) $\sum\limits_{j=1}^{n-1} x_jx_{j+1}$; (4) $\sum\limits_{1\leqslant k<\ell\leqslant n} (k+\mathrm{i}\ell)x_kx_\ell$.

8. 设 $f(\alpha,\beta)$ 是 n 维实线性空间 V 上的双线性函数,并且对任意非零向量 $\alpha\in V$, $f(\alpha,\alpha)>0$. 证明,存在 V 的一组基 $(\xi_1,\xi_2,\dots,\xi_n)$,使得 $f(\alpha,\beta)$ 在这组基下的方阵是如下的准对角形:

$$\operatorname{diag}\left(\begin{pmatrix} 1 & a_1 \\ -a_1 & 1 \end{pmatrix},\dots,\begin{pmatrix} 1 & a_s \\ -a_s & 1 \end{pmatrix},\underbrace{1,1,\dots,1}_{n-2s\text{个}}\right),$$

其中 $a_1\geqslant a_2\geqslant\cdots\geqslant a_s\geqslant 0$.

9. 定义所有 n 阶实方阵构成的实线性空间 $\mathbb{R}^{n\times n}$ 上的对称双线性函数 $f(X,Y)$ 为

$$f(X,Y)=\operatorname{Tr}XY^{\mathrm{T}}.$$

其中 $X,Y\in\mathbb{R}^{n\times n}$. 求 $f(X,Y)$ 的正、负惯性指数.

10. 设

$$Q(\alpha)=\sum_{1\leqslant i,j\leqslant n} a_{ij}x_ix_j$$

是 n 维实线性空间 V 的正定二次型. 证明

$$Q(\beta)=\sum_{\substack{1\leqslant i,j\leqslant n\\ i,j\neq k}}\left(a_{ij}-\frac{a_{ik}a_{jk}}{a_{kk}}\right)x_ix_j$$

是关于自变量 $x_1,x_2,\dots,x_{k-1},x_{k+1},\dots,x_n$ 的正定二次型.

11. 设 $f(\alpha,\beta)$ 是 n 维实的行向量集合连同标准内积构成的 Euclid 空间 $\mathbb{R}^n$ 上的双线性函数，$\mathrm{O}_n(\mathbb{R})$ 是 Euclid 空间 $\mathbb{R}^n$ 的所有正交变换的集合. 如果对任意 $\mathscr{A}\in\mathrm{O}_n(\mathbb{R})$，均有

$$f\big(\mathscr{A}(\alpha),\mathscr{A}(\beta)\big)=f(\alpha,\beta),$$

则 $f(\alpha,\beta)$ 称为在 $\mathrm{O}_n(\mathbb{R})$ 下是不变的. 求所有在 $\mathrm{O}_n(\mathbb{R})$ 下不变的双线性函数 $f(\alpha,\beta)$.

12. 将上一习题中实数域 $\mathbb{R}$ 改为复数域 $\mathbb{C}$，即求 n 维复的行向量空间 $\mathbb{C}$ 上的所有在 $\mathrm{O}_n(\mathbb{C})$ 下不变的双线性函数 $f(\alpha,\beta)$，这里 $\mathrm{O}_n(\mathbb{C})$ 是所有 n 阶复正交方阵的集合.

13. 设 $\mathbb{C}^2$ 是所有 2 维复的行向量 $\alpha=(x_1,x_2)$ 构成的复线性空间，$Q(\alpha)=x_1^2-x_2^2$ 是 $\mathbb{C}^2$ 的二次型. 设线性变换 $\mathscr{A}:\mathbb{C}^n\to\mathbb{C}^2$ 满足，对任意 $\alpha\in\mathbb{C}^2$，

$$Q\big(\mathscr{A}(\alpha)\big)=Q(\alpha),$$

则 $\mathscr{A}$ 称为保二次型 $Q(\alpha)$ 的. 证明

(1) 设 $\mathscr{A}$ 在 $\mathbb{C}^2$ 的基 $\{\varepsilon_1,\varepsilon_2\}$ 下的方阵为 $A=(a_{ij})$，其中 $\varepsilon_1=(1,0)$，$\varepsilon_2=(0,1)$，则 $a_{22}=\pm a_{11}$，$a_{21}=\pm a_{12}$，$a_{11}^2-a_{12}^2=1$；

(2) 如果 $\det A=1$，则存在非零复数 c，使得

$$A=\frac{1}{2}\begin{pmatrix} c+\dfrac{1}{c} & c-\dfrac{1}{c}\\ c-\dfrac{1}{c} & c+\dfrac{1}{c}\end{pmatrix};$$

如果 $\det A=-1$，则存在非零复数 c，使得

$$A=\frac{1}{2}\begin{pmatrix} c+\dfrac{1}{c} & c-\dfrac{1}{c}\\ -c+\dfrac{1}{c} & -c-\dfrac{1}{c}\end{pmatrix}.$$

14. 伪 Euclid 空间. 所谓 Euclid 空间是指赋以内积 (α,β) 的实线性空间 V，而内积 (α,β) 是 V 上的正定对称双线性函数. Euclid 空间概念之推广即是伪 Euclid 空间. 其定义如下：n 维实线性空间 V 上的非退化对称双线性函数 (α,β) 称为 V 的一个内积. 实线性空间 V 连同取定的一个内积 (α,β) 称为**伪 Euclid 空间**，内积 (α,β) 的正惯性指数 p 称为伪 Euclid 空间 V 的**指数**. 如果 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 满足

$$(\xi_i,\xi_j)=\varepsilon_i\delta_{ij},$$

其中 $1\leqslant i,j\leqslant n$，当 $1\leqslant i\leqslant p$ 时，$e_i=1$，当 $p+1\leqslant i\leqslant n$ 时，$\varepsilon_i=-1$，则 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 称为伪 Euclid 空间 V 的一组标准正交基. 如果线性变换 $\mathscr{A}:V\to V$ 满足，对任意 $\alpha,\beta\in V$，

$$\big(\mathscr{A}(\alpha),\mathscr{A}(\beta)\big)=(\alpha,\beta),$$

则 $\mathscr{A}$ 称为**伪正交变换**. 证明

(1) 伪正交变换是可逆的，并且它的逆变换仍是伪正交变换；

(2) 伪正交变换的乘积仍是伪正交变换.

§9.3　斜对称双线性函数

本节讨论数域 $\mathbb{F}$ 上的 n 维线性空间 V 的斜对称双线性函数.

容易证明,对于斜对称双线性函数,下述命题成立.

命题 9.3.1　数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数 $f(\alpha,\beta)$ 为斜对称的充分必要条件是,对任意向量 $\alpha\in V$, $f(\alpha,\alpha)=0$.

命题 9.3.2　数域 $\mathbb{F}$ 上的 n 维线性空间 V 的双线性函数 $f(\alpha,\beta)$ 为斜对称的充分必要条件是, $f(\alpha,\beta)$ 在 V 的基下的方阵是斜对称的.

命题 9.3.3　设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的斜对称双线性函数,则 V 中向量关于 $f(\alpha,\beta)$ 的正交性是对称的.

命题 9.3.4　数域 $\mathbb{F}$ 上的 n 维线性空间 V 的所有斜对称双线性函数集合记为 $K(V,V,\mathbb{F})$,数域 $\mathbb{F}$ 上的所有 n 阶斜对称方阵集合记为 $K(n,\mathbb{F})$. 则集合 $K(V,V,\mathbb{F})$ 与 $K(n,\mathbb{F})$ 之间存在双射.

命题 9.3.3 是定理 9.1.6 的直接结论. 其它几个命题的证明留给读者作练习.

定理 9.3.1　设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的斜对称双线性函数,则存在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵是如下的准对角形:

$$\operatorname{diag}\Bigg(\underbrace{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\ldots,\begin{pmatrix}0&1\\-1&0\end{pmatrix}}_{s\text{ 个}},\underbrace{0,0,\ldots,0}_{n-2s\text{ 个}}\Bigg),\tag{9.3.1}$$

其中 $\operatorname{rank} f=2s$. 换句话说,

$$f(\alpha,\beta)=(x_1y_2-x_2y_1)+\cdots+(x_{2s-1}y_{2s}-x_{2s}y_{2s-1}),\tag{9.3.2}$$

其中 $x=(x_1,\ldots,x_n)$ 与 $y=(y_1,\ldots,y_n)$ 分别是 V 中向量 α 与 β 在这组基下的坐标.

证明　对空间 V 的维数 n 用归纳法. 当 $n=1$ 时结论显然成立.

假设结论对维数小于 n 的空间成立. 下面证明结论对 n 维空间 V 成立.

如果 $f(\alpha,\beta)$ 是零函数,即对任意 $\alpha,\beta\in V$, $f(\alpha,\beta)=0$,则结论显然成立. 因此可设存在向量 $\eta,\zeta\in V$,使得 $f(\eta,\zeta)=b\neq 0$.

记 $\xi_1=\eta$, $\xi_2=b^{-1}\zeta$,且

$$f(\xi_1,\xi_1)=f(\xi_2,\xi_2)=0,\qquad f(\xi_1,\xi_2)=1.$$

如果存在数 $a_1,a_2\in\mathbb{F}$,使得 $a_1\xi_1+a_2\xi_2=0$,则

$$f(\xi_1,a_1\xi_1+a_2\xi_2)=a_2f(\xi_1,\xi_2)=0.$$

因此 $a_2=0$. 同理可证 $a_1=0$. 于是向量 ξ_1,ξ_2 线性无关. 所以 V 中由向量 ξ_1 与 ξ_2 生成的子空间 W 是 2 维的.

子空间 W 关于 $f(\alpha,\beta)$ 的正交子空间记为 $W^{\perp}$. 我们将证明, $V=W\oplus W^{\perp}$.

事实上,设 $\alpha\in W\cap W^{\perp}$. 由于 $\alpha\in W$,而子空间 W 是向量 ξ_1 与 ξ_2 生成的,所以

$$\alpha=a_1\xi_1+a_2\xi_2,$$

其中 $a_1,a_2\in\mathbb{F}$. 另一方面,由于 $\alpha\in W^{\perp}$,所以

$$\begin{aligned}f(\xi_1,\alpha)&=a_2f(\xi_1,\xi_2)=0,\\ f(\xi_2,\alpha)&=a_1f(\xi_2,\xi_1)=-a_1f(\xi_2,\xi_1)=0.\end{aligned}$$

因此 $a_1=a_2=0$,即 $\alpha=0$. 所以 $W\cap W^{\perp}=0$. 于是

$$W+W^{\perp}=W\oplus W^{\perp}.$$

其次,设 $\alpha\in V$. 考虑向量 $\beta=\alpha-\lambda_1\xi_1-\lambda_2\xi_2$,其中 λ_1,λ_2 是待定常数. 令

$$\begin{aligned}f(\xi_1,\beta)&=f(\xi_1,\alpha)-\lambda_2f(\xi_1,\xi_2)=0,\\ f(\xi_2,\beta)&=f(\xi_2,\alpha)+\lambda_1f(\xi_1,\xi_2)=0.\end{aligned}$$

则得到 $\lambda_1=-f(\xi_2,\alpha),\lambda_2=f(\xi_1,\alpha)$. 于是

$$\alpha=\big(-f(\xi_2,\alpha)\xi_1+f(\xi_1,\alpha)\xi_2\big)+\big(\alpha+f(\xi_2,\alpha)\xi_1-f(\xi_1,\alpha)\xi_2\big),$$

其中

$$-f(\xi_2,\alpha)\xi_1+f(\xi_1,\alpha)\xi_2\in W,\qquad \alpha+f(\xi_2,\alpha)\xi_1-f(\xi_1,\alpha)\xi_2\in W^{\perp},$$

即 $\alpha\in W+W^{\perp}$. 因此 $V=W+W^{\perp}=W\oplus W^{\perp}$.

把斜对称双线性函数 $f(\alpha,\beta)$ 限定在 $W^{\perp}$ 上,则 $f(\alpha,\beta)$ 是 $n-2$ 维线性空间 $W^{\perp}$ 上的斜对称双线性函数. 由归纳假设,存在 $W^{\perp}$ 的基 $\{\xi_3,\xi_4,\dots,\xi_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵是如下的准对角形:

$$\operatorname{diag}\Bigg(\underbrace{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\dots,\begin{pmatrix}0&1\\-1&0\end{pmatrix}}_{t\text{ 个}},\underbrace{0,0,\dots,0}_{n-2-2t\text{ 个}}\Bigg).$$

由于 $V=W\oplus W^{\perp}$,所以 $\{\xi_1,\xi_2,\xi_3,\dots,\xi_n\}$ 是 V 的基,并且 $f(\alpha,\beta)$ 在此基下的方阵为

$$\operatorname{diag}\Bigg(\underbrace{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\dots,\begin{pmatrix}0&1\\-1&0\end{pmatrix}}_{t+1\text{ 个}},\underbrace{0,0,\dots,0}_{n-2-2t\text{ 个}}\Bigg).$$

显然 $2(t+1)=\operatorname{rank}f$. ■

定理 9.3.1 的矩阵形式在以下定理中给出.

定理 9.3.2 设 K 是数域 $\mathbb{F}$ 上的 n 阶斜对称方阵. 则存在数域 $\mathbb{F}$ 上的 n 阶可逆方阵 P,使得

$$P^{\mathrm{T}}KP=\operatorname{diag}\Bigg(\underbrace{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\dots,\begin{pmatrix}0&1\\-1&0\end{pmatrix}}_{s\text{ 个}},\underbrace{0,0,\dots,0}_{n-2s\text{ 个}}\Bigg),$$

其中 $2s=\operatorname{rank}K$. 换句话说,数域 $\mathbb{F}$ 上的 n 阶斜对称方阵相合(通过数域 $\mathbb{F}$ 上的方

阵）于标准形 (9.3.1)，斜对称方阵的秩是斜对称方阵在相合下的全系不变量.

证明　设 $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的一组基，定义 V 上的二元函数 $f(\alpha,\beta)$ 为

$$f(\alpha,\beta)=xKy^{\mathrm{T}},$$

其中 x 与 y 分别是 V 中向量 α 与 β 在这组基下的坐标.

由于方阵 K 是斜对称的，所以 $f(\alpha,\beta)$ 是 V 上的斜对称双线性函数. 由定理 9.3.1 存在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$，使得 $f(\alpha,\beta)$ 在这组基下的方阵为准对角形 (9.3.1).

由于同一个双线性函数在不同的基下的方阵是相合的，所以方阵 K 相合于准对角方阵 (9.3.1). 显然 $2s=\operatorname{rank}K$.

设斜对称方阵 K_1 与 K_2 相合，则方阵 K_1 与 K_2 显然相抵，所以 $\operatorname{rank}K_1=\operatorname{rank}K_2$. 即斜对称方阵的秩是斜对称方阵在相合下的不变量.

反之，设斜对称方阵 K_1 与 K_2 的秩相等，且其秩都是 $2s$，则它们都相合于准对角形 (9.3.1). 由于方阵的相合关系具有对称性与传递性，所以方阵 K_1 与 K_2 相合. 因此斜对称方阵的秩是斜对称方阵在相合下的全系不变量.　■

定理 9.3.2 完全解决了斜对称方阵在相合下的标准形问题.

习　题　9.3

1. 设 4 阶斜对称方阵 K 为

$$K=\begin{pmatrix}0&2&-1&3\\-2&0&4&-2\\1&-4&0&1\\-3&2&-1&0\end{pmatrix}.$$

求 4 阶有理系数可逆方阵 P，使得

$$P^{\mathrm{T}}KP=\operatorname{diag}\left(\begin{pmatrix}0&1\\-1&0\end{pmatrix},\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right)$$

2. 设 V 是数域 $\mathbb{F}$ 上的 n 维线性空间，$L(V,V,\mathbb{F})$ 是 V 上的所有双线性函数构成的数域 $\mathbb{F}$ 上的线性空间. 对任意 $f\in L(V,V,\mathbb{F})$，记

$$(\mathscr{P}(f))(\alpha,\beta)=\frac{1}{2}f(\alpha,\beta)-\frac{1}{2}f(\beta,\alpha).$$

显然 $\mathscr{P}(f)\in L(V,V,F)$. 定义空间 $L(V,V,\mathbb{F})$ 的变换 $\mathscr{P}$ 如下：对于 $f\in L(V,V,\mathbb{F})$，令 f 在 $\mathscr{P}$ 下的象为 $\mathscr{P}(f)$. 证明

(1) $\mathscr{P}$ 是 $L(V,V,\mathbb{F})$ 的线性变换，并且 $\mathscr{P}^2=\mathscr{P}$；

(2) $\mathscr{P}$ 的秩为 $\frac{1}{2}n(n-1)$；

(3) 设 $\mathscr{B}$ 是 V 的线性变换，对任意 $f\in L(V,V,\mathbb{F})$，记

$$(\widetilde{\mathscr{B}}(f))(\alpha,\beta)=f(\mathscr{B}(\alpha),\mathscr{B}(\beta)).$$

显然，$\widetilde{\mathscr{B}}(f)\in L(V,V,\mathbb{F})$. 定义空间 $L(V,V,\mathbb{F})$ 的变换 $\widetilde{\mathscr{B}}$ 如下：对于 $f\in L(V,V,\mathbb{F})$，令 f 在 $\widetilde{\mathscr{B}}$ 下的象为 $\widetilde{\mathscr{B}}(f)$. 则 $\widetilde{\mathscr{B}}$ 是 $L(V,V,\mathbb{F})$ 的线性变换，而且和 $\mathscr{P}$ 可交换.

3. 设 V 是数域 $\mathbb{F}$ 上的 n 维线性空间，$L_1(\alpha)$ 与 $L_2(\alpha)$ 是 V 上的线性函数. 证明

$$f(\alpha,\beta)=L_1(\alpha)L_2(\beta)-L_1(\beta)L_2(\alpha)$$

是 V 上的斜对称双线性函数，而且当且仅当 $L_1, L_2\in V^*$ 线性相关时，$f(\alpha,\beta)$ 为零函数.

4. 设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的斜对称双线性函数. 证明，$\operatorname{rank} f=2$ 的充分必要条件是，存在线性无关的 $L_1, L_2\in V^*$，使得

$$f(\alpha,\beta)=L_1(\alpha)L_2(\beta)-L_1(\beta)L_2(\alpha).$$

5. 设 $\mathbb{R}^3$ 是 3 维实的行向量空间，$f(\alpha,\beta)$ 是 $\mathbb{R}^3$ 上的斜对称双线性函数. 证明，存在 $L_1, L_2\in(\mathbb{R}^3)^*$，使得

$$f(\alpha,\beta)=L_1(\alpha)L_2(\beta)-L_1(\beta)L_2(\alpha).$$

6. 设 V 是数域 $\mathbb{F}$ 上的 n 维线性空间，$f(\alpha,\beta)$ 与 $g(\alpha,\beta)$ 是 V 上的斜对称双线性函数. 证明 $\operatorname{rank} f=\operatorname{rank} g$ 的充分必要条件是，存在 V 的线性变换 $\mathscr{A}$ 使得对任意 $\alpha,\beta\in V$，均有 $f\big(\mathscr{A}(\alpha),\mathscr{A}(\beta)\big)=g(\alpha,\beta)$.

7. 辛几何. 所谓 Euclid 空间是赋以一个给定的内积 (α,β) 的实线性空间，而内积 (α,β) 是正定对称双线性函数，它当然是非退化的. 如果将内积取成非退化斜对称双线性函数，则引出所谓辛空间. 其定义如下：设 $f(\alpha,\beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的非退化斜对称双线性函数，则 $f(\alpha,\beta)$ 称为 V 的一个**辛内积**. 线性空间 V 连同一个取定的辛内积 $f(\alpha,\beta)$ 称为**辛空间**. 显然辛空间 V 应是偶数维的. 设 $\dim V=n=2k$. 如果辛空间 V 的向量组 $\{\alpha_1,\alpha_2,\ldots,\alpha_k,\beta_1,\beta_2,\ldots,\beta_k\}$ 适合

$$f(\alpha_i,\alpha_j)=0,\qquad f(\beta_i,\beta_j)=0,\qquad f(\alpha_i,\beta_j)=\delta_{ij},$$

其中 $1\leqslant i,j\leqslant k$，$\delta_{ij}$ 是 Kronecker 符号，则

$$\{\alpha_1,\alpha_2,\ldots,\alpha_k,\beta_1,\beta_2,\ldots,\beta_k\}$$

称为 V 的一组**辛基**. 如果线性变换 $\mathscr{A}: V\to V$ 适合 $f\big(\mathscr{A}(\alpha),\mathscr{A}(\beta)\big)=f(\alpha,\beta)$，则 $\mathscr{A}$ 称为**辛变换**. 如果 $\mathscr{A}: V\to V$ 是辛空间 V 的线性变换，$\alpha,\beta\in V$，则由

$$f\big(\mathscr{A}(\lambda),\beta\big)=f\big(\alpha,\widetilde{\mathscr{A}}(\beta)\big)$$

所定义的变换 $\widetilde{\mathscr{A}}: V\to V$ 称为 $\mathscr{A}$ 的**辛伴随变换**. 如果 $\widetilde{\mathscr{A}}=\mathscr{A}$（或者 $\widetilde{\mathscr{A}}=-\mathscr{A}$），则线性变换 $\mathscr{A}$ 称为**辛自伴**（或者**辛斜自伴**）的. 证明

(1) 每一个辛空间 V 都具有辛基；

(2) 设 $h(\alpha,\beta)$ 与 $g(\alpha,\beta)$ 是 V 的辛内积，则存在可逆线性变换 $\mathscr{A}: V\to V$，使得对任意 $\alpha,\beta\in V$，

$$g(\alpha,\beta)=h\big(\mathscr{A}(\alpha),\mathscr{A}(\beta)\big);$$

(3) 对每个线性变换 $\mathscr{A}: V\to V$，均有 $(\widetilde{\mathscr{A}})^{\sim}=\mathscr{A}$，并且都可以唯一地分解为一个辛自伴变换与一个辛斜自伴变换的和；

(4) 辛空间 V 的辛斜自伴变换 $\mathscr{A}$ 在 V 的辛基下的方阵 A 具有如下形式：

$$A=\begin{pmatrix} M & N\\ K & L\end{pmatrix},$$

其中 M, N, K 与 L 都是数域 $\mathbb{F}$ 上的 k 阶方阵，并且 $N=N^*$，$K=K^*$，$L=-M^*$，这里 $\square^*$ 表示方阵的共轭转置.

§9.4　共轭双线性函数与 Hermite 型

本节将推广双线性函数的概念.

定义 9.4.1　设 $f(\alpha,\beta)$ 是 n 维复线性空间 V 上的二元函数. 如果对任意向量 $\alpha,\beta,\alpha_1,\alpha_2,\beta_1,\beta_2\in V$，以及任意复数 $\lambda_1,\lambda_2,\mu_1,\mu_2\in\mathbb{C}$，均有

$$f(\lambda_1\alpha_1+\lambda_2\alpha_2,\beta)=\lambda_1 f(\alpha_1,\beta)+\lambda_2 f(\alpha_2,\beta) \tag{9.4.1}$$

$$f(\alpha,\mu_1\beta_1+\mu_2\beta_2)=\overline{\mu}_1 f(\alpha,\beta_1)+\overline{\mu}_2 f(\alpha,\beta_2) \tag{9.4.2}$$

其中 $\overline{\mu}$ 表示复数 μ 的共轭复数，则二元函数 $f(\alpha,\beta)$ 称为**共轭双线性**的.

容易看出，V 上的共轭双线性函数 $f(\alpha,\beta)$ 具有如下性质.

命题 9.4.1　设 $f(\alpha,\beta)$ 是 V 上的共轭双线性函数，则对任意 $\alpha,\beta\in V$，

$$f(\alpha,0)=0=f(0,\beta).$$

命题 9.4.2　设 $f(\alpha,\beta)$ 是 V 上的共轭双线性函数，则对任意 $\alpha_1,\ldots,\alpha_p,\beta_1,\ldots,\beta_q\in V$，$\lambda_1,\ldots,\lambda_p,\mu_1,\ldots,\mu_q\in\mathbb{C}$，

$$f\Big(\sum_{k=1}^{p}\lambda_k\alpha_k,\ \sum_{\ell=1}^{q}\mu_\ell\beta_\ell\Big)=\sum_{k=1}^{p}\sum_{\ell=1}^{q}\lambda_k\overline{\mu}_\ell f(\alpha_k,\beta_\ell). \tag{9.4.3}$$

现在给出 V 上的共轭双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵表示. 设向量 $\alpha,\beta\in V$ 在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标分别是

$$x=(x_1,x_2,\ldots,x_n)\quad 与\quad y=(y_1,y_2,\ldots,y_n),$$

即

$$\alpha=\sum_{k=1}^{n}x_k\xi_k,\qquad \beta=\sum_{\ell=1}^{n}y_\ell\xi_\ell.$$

则由式 (9.4.3)，

$$f(\alpha,\beta)=f\Big(\sum_{k=1}^{n}x_k\xi_k,\ \sum_{\ell=1}^{n}y_\ell\xi_\ell\Big)=\sum_{1\leqslant k,\ell\leqslant n}x_k\overline{y}_\ell f(\xi_k,\xi_\ell). \tag{9.4.4}$$

记 n 阶方阵 $A=\big(f(\xi_k,\xi_\ell)\big)_{n\times n}$，则上式化为

$$f(\alpha,\beta)=xAy^*, \tag{9.4.5}$$

其中 $y^*=\overline{y}^{\mathrm{T}}$ 是 $y=(y_1,y_2,\ldots,y_n)$ 的共轭转置.

方阵 A 称为共轭双线性函数 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵. 而式 (9.4.4) 称为 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的表达式.

显然，不同的共轭双线性函数在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵是不同的.

反之，设 A 是 n 阶复方阵，则令

$$f(\alpha,\beta)=xAy^*,$$

其中 x 与 y 分别是向量 α,β 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标.

容易验证,$f(\alpha,\beta)=xAy^*$ 是 V 上的共轭双线性函数.

这表明,如果在 V 中取定一组基,并建立 V 上的所有共轭双线性函数的集合到所有 n 阶复方阵集合的映射 σ: 共轭双线性函数在映射 σ 下的象为它在这组基下的方阵,则映射 σ 是双射. 所以,V 上的所有共轭双线性函数集合与所有 n 阶复方阵集合之间存在一一对应.

为了给出 V 上的共轭双线性函数 $f(\alpha,\beta)$ 在不同基下的方阵表示之间的关系,先引进下面的定义.

定义 9.4.2 设 A 与 B 是 n 阶复方阵. 如果存在 n 阶可逆复方阵 P,使得 $B=P^*P$,其中 $P^*=\overline{P}^{\mathrm{T}}$ 是方阵 P 的共轭转置. 则方阵 A 与 B 称为**复相合**的.

容易验证,复方阵间的复相合关系满足自反性,对称性与传递性. 因此复相合关系是所有 n 阶复方阵集合 $\mathbb{C}^{n\times n}$ 中方阵间的一种等价关系. 集合 $\mathbb{C}^{n\times n}$ 便按照复相合等价关系划分为复相合等价类.

定理 9.4.1 设 n 维复线性空间 V 上的共轭双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 与 $\{\eta_1,\eta_2,\dots,\eta_n\}$ 下的方阵分别为 $A=\big(f(\xi_k,\xi_\ell)\big)$ 与 $B=\big(f(\eta_k,\eta_\ell)\big)$,并且

$$(\eta_1,\eta_2,\dots,\eta_n)=(\xi_1,\xi_2,\dots,\xi_n)\overline{P},$$

其中 $\overline{P}$ 是 n 阶可逆复方阵 P 的共轭. 则

$$B=P^*AP.$$

也就是说,同一个共轭双线性函数在不同基下的方阵是复相合的.

证明 与定理 9.1.2 的证明相仿,略. ■

由于复相合的方阵是相抵的,因此它们的秩相等. 所以根据定理 9.4.1,共轭双线性函数 $f(\alpha,\beta)$ 在 V 的基下的方阵的秩定义为 $f(\alpha,\beta)$ 的秩. 记为 $\operatorname{rank} f$.

特别,如果 $\operatorname{rank} f=\dim V$,则 $f(\alpha,\beta)$ 称为非退化的.

在共轭双线性函数中,重要的是 Hermite 共轭双线性函数. 其定义如下.

定义 9.4.3 设 $f(\alpha,\beta)$ 是 V 上的共轭双线性函数. 如果对任意向量 $\alpha,\beta\in V$,有

$$\overline{f(\beta,\alpha)}=f(\alpha,\beta),$$

则 $f(\alpha,\beta)$ 称为 **Hermite** 的.

定理 9.4.2 设 V 上的共轭双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵为 $A=\big(f(\xi_k,\xi_\ell)\big)_{n\times n}$,则 $f(\alpha,\beta)$ 为 Hermite 的充分必要条件是,方阵 A 为 Hermite 方阵,即方阵 A 满足 $A^*=A$.

证明 由于 $f(\alpha,\beta)$ 是 Hermite 的,所以对任意 $1\leqslant k,\ell\leqslant n$,有

$$\overline{f(\xi_\ell,\xi_k)}=f(\xi_k,\xi_\ell).$$

也即 $A^*=A$. 因此方阵 A 是 Hermite 方阵.

反之,设 A 是 Hermite 方阵. 由于 A 是 $f(\alpha,\beta)$ 在基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵,

所以
$$f(\alpha,\beta)=xAy^*,$$
其中 x 与 y 分别是向量 α 与 β 在基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的坐标. 于是
$$\overline{f(\alpha,\beta)}=\overline{xAy^*}=\overline{x}\overline{A}y^{\mathrm{T}}=\left(\overline{x}\overline{A}y^{\mathrm{T}}\right)^{\mathrm{T}}=yA^*x^*.$$
由于 A 是 Hermite 方阵,所以 $A^*=A$. 因此对任意 $\alpha,\beta\in V$,
$$\overline{f(\alpha,\beta)}=yAx^*=f(\beta,\alpha).$$
即 $f(\alpha,\beta)$ 是 Hermite 的. ■

由定理 9.4.2 可以得到,V 上的所有 Hermite 共轭双线性函数集合与所有 Hermite 方阵集合之间存在一一对应.

利用 Hermite 方阵在酉相似下的标准形,可以证明

定理 9.4.3 设 H 是 n 阶 Hermite 方阵. 则存在 n 阶可逆复方阵 P,使得
$$P^*HP=\operatorname{diag}\left(I_{(p)},-I_{(q)},0_{(n-p-q)}\right),\tag{9.4.6}$$
其中 $0_{(n-p-q)}$ 是 $n-p-q$ 阶零方阵,并且 $p+q=r$.

简单地说,Hermite 方阵 H 复相合于对角形 (9.4.6).

证明 设 $\lambda_1,\lambda_2,\ldots,\lambda_r$ 是 Hermite 方阵 H 的全部非零特征值,$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r$. 则存在 n 阶酉方阵 U,使得
$$U^*HU=\operatorname{diag}\left(\lambda_1,\lambda_2,\ldots,\lambda_r,\underbrace{0,0,\ldots,0}_{n-r\text{ 个}}\right).$$
可设 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p>0>\lambda_{p+1}\geqslant\cdots\geqslant\lambda_{p+q}$,$p+q=r$. 记
$$Q=\operatorname{diag}\left(\frac{1}{\sqrt{\lambda_1}},\ldots,\frac{1}{\sqrt{\lambda_p}},\frac{1}{\sqrt{-\lambda_{p+1}}},\ldots,\frac{1}{\sqrt{-\lambda_{p+q}}},\underbrace{1,1,\ldots,1}_{n-r\text{ 个}}\right).$$
则方阵 Q 可逆,并且
$$Q^*U^*HUQ=\operatorname{diag}\left(I_{(p)},-I_{(q)},0_{(n-p-q)}\right).$$
记 $P=UQ$. 显然方阵 P 可逆,并且式 (9.4.6) 成立. ■

由定理 9.4.3 的证明可以看出,式 (9.4.6) 中 p 与 p 分别是 Hermite 方阵 H 的正特征值与负特征值的个数.

由定理 9.4.3 立即得到

定理 9.4.4 设 $f(\alpha,\beta)$ 是 n 维复线性空间 V 上的 Hermite 共轭双线性函数,则存在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵为对角形 (9.4.6):
$$\operatorname{diag}\left(I_{(p)},-I_{(q)},0_{(n-p-q)}\right).$$
换句话说,$f(\alpha,\beta)$ 在这组基下的表达式为
$$f(\alpha,\beta)=x_1\overline{y}_1+\cdots+x_p\overline{y}_p-x_{p+1}\overline{y}_{p+1}-\cdots-x_{p+q}\overline{y}_{p+q},$$
其中 $x=(x_1,x_2,\ldots,x_n)$ 与 $y=(y_1,y_2,\ldots,y_n)$ 分别是 α 与 β 在这组基下的坐标.

现在分析定理 9.4.4 中式 (9.4.6) 的非负整数 p 与 q 的意义.

定义 9.4.4 设 $f(\alpha,\beta)$ 是 n 维复线性空间 V 上的 Hermite 共轭双线性函数.

如果对任意非零向量 $\alpha\in V$, $f(\alpha,\alpha)>0$, 则 $f(\alpha,\beta)$ 称为**正定的**;

如果对任意向量 $\alpha\in V$, $f(\alpha,\alpha)\geqslant 0$, 则 $f(\alpha,\beta)$ 称为**半正定**的.

可以类似地定义负定或半负定的 Hermite 共轭双线性函数.

对于 n 维复线性空间 V 上的 Hermite 共轭双线性函数 $f(\alpha,\beta)$, 记

$$V^{\perp}=\left\{\alpha\in V \,\middle|\, f(\alpha,\beta)=0,\text{对任意 }\beta\in V\right\}.$$

容易验证, $V^{\perp}$ 是 V 的子空间, 它称为 $f(\alpha,\beta)$ 的**根基**.

定理 9.4.4 说明, 对于 V 上的 Hermite 共轭双线性函数 $f(\alpha,\beta)$, 在 V 中存在一组基

$$\{\xi_1,\xi_2,\dots,\xi_n\},$$

使得对任意 k 与 ℓ, $1\leqslant k\neq\ell\leqslant n$, $f(\xi_k,\xi_\ell)=0$. 具有这一性质的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 称为关于 $f(\alpha,\beta)$ 的**正交基**. 于是定理 9.4.4 断言, 关于 $f(\alpha,\beta)$ 的正交基是存在的.

其次, V 中由基向量 $\{\xi_1,\xi_2,\dots,\xi_p\}$ 生成的子空间记为 V^+, 由基向量 $\{\xi_{p+1},\dots,\xi_{p+q}\}$ 生成的子空间记为 V^-, 而由基向量 $\{\xi_{p+q+1},\dots,\xi_n\}$ 生成的子空间记为 W, 则

$$V=V^+\oplus V^-\oplus W.$$

并且对任意非零向量 $\alpha\in V^+$, 均有 $f(\alpha,\alpha)>0$, 即 $f(\alpha,\beta)$ 限定在 V^+ 上是正定的. 同样, $f(\alpha,\beta)$ 在 V^- 上的限制是负定的.

最后, 容易证明, W 是关于 $f(\alpha,\beta)$ 的根基 $V^{\perp}$. 于是有

定理 9.4.5 设 $f(\alpha,\beta)$ 是 n 维复线性空间 V 上的 Hermite 共轭双线性函数. 则 V 可以分解为子空间 V^+, V^- 与 $V^{\perp}$ 的直和, 使得 $f(\alpha,\beta)$ 在 V^+ 上的限制是正定的, 在 V^- 上的限制是负定的, 而 $V^{\perp}$ 是关于 $f(\alpha,\beta)$ 的根基.

如果 V 还可以分解为子空间 V_1^+, V_1^- 与 $V_1^{\perp}$ 的直和, 使得 $f(\alpha,\beta)$ 在 V_1^+ 与 V_1^- 上的限制分别是正定与负定的, 则 $\dim V_1^+=\dim V^+$, $\dim V_1^-=\dim V^-$.

定理 9.4.5 说明, 对于 n 维复线性空间 V 上的 Hermite 共轭双线性函数 $f(\alpha,\beta)$, 可以将它的定义域 V 分成三个部分: V^+, V^-, $V^{\perp}$. 而且子空间 V^+, V^- 与 $V^{\perp}$ 的维数与分解的方式无关, 是由 $f(\alpha,\beta)$ 自身所确定的.

因此, V^+ 与 V^- 的维数分别称为 $f(\alpha,\beta)$ 的**正、负惯性指数**. 而 V^+ 与 V^- 的维数之差, 也即正、负惯性指数之差称为 $f(\alpha,\beta)$ 的**符号差**, 记为 $\delta(f)$.

由定理 9.4.4, 如果 Hermite 共轭双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\dots,\xi_n\}$ 下的方阵为对角形 (9.4.6), 则 $f(\alpha,\beta)$ 的正、负惯性指数即是对角元中 $+1$ 的个数 p 与 -1 的个数 q.

现在考虑 n 阶 Hermite 方阵在复相合下的标准形.

设 H 是 n 阶 Hermite 方阵, $\{\alpha_1,\alpha_2,\dots,\alpha_n\}$ 是 n 维复线性空间 V 的一组基. 令

$$f(\alpha,\beta)=xHy^*,$$

其中 x 与 y 分别是 V 中向量 α 与 β 在这组基下的坐标.

定理 9.4.2 表明,$f(\alpha,\beta)$ 是 V 上的 Hermite 共轭双线性函数. 由定理 9.4.3,方阵 H 复相合于对角形

$$\operatorname{diag}\left(I_{(p)},-I_{(q)},0_{(n-p-q)}\right).$$

如果方阵 H 复相合于另一个对角形 $\operatorname{diag}\left(I_{(p')},-I_{(q')},0_{(n-p'-q')}\right)$,那么根据定理 9.4.5,$p=p'$,$q=q'$. 因此对角方阵 $\operatorname{diag}\left(I_{(p)},-I_{(q)},0_{(n-p-q)}\right)$ 的对角元中 $+1$ 的个数 p 与 -1 的个数 q 是由方阵 H 自身决定的.

于是对角形 (9.4.6) 称为 Hermite 方阵 H 在复相合下的标准形,而 p 与 q 分别称为方阵 H 的正、负惯性指数,正、负惯性指数之差称为方阵 H 的符号差,记为 $\delta(H)$. 显然,

$$p+q=r,\quad \delta(H)=2p-r.$$

定理 9.4.6　n 阶 Hermite 方阵 H 复相合于如下的标准形:

$$\operatorname{diag}\left(I_{(p)},-I_{(q)},\ 0_{(n-p-q)}\right),$$

其中 p 与 q 分别是方阵 H 的正、负惯性指数. 而且 Hermite 方阵 H 的秩与符号差是 Hermite 方阵在复相合下的全系不变量.

证明　只需证明后一结论. 设 Hermite 方阵 H_1 与 H_2 复相合,即设 $H_2=P^*H_1P$,其中 P 是 n 阶可逆方阵. 显然方阵 H_1 与 H_2 相抵,所以方阵 H_1 与 H_2 的秩相等. 下面证明,方阵 H_1 与 H_2 的符号差相等. 事实上,设 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 是 n 维复线性空间 V 的一组基. 定义 V 上的二元函数 $f(\alpha,\beta)$ 为

$$f(\alpha,\beta)=xH_2y^*,$$

其中 x 与 y 分别是向量 α 与 β 在这组基下的坐标.

由定理 9.4.2,$f(\alpha,\beta)$ 是 Hermite 共轭双线性函数. 设

$$(\eta_1,\eta_2,\ldots,\eta_n)=(\xi_1,\xi_2,\ldots,\xi_n)\overline{P}^{-1}.$$

由于方阵 P 可逆,所以 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 是 V 的一组基.

向量 α 在基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 下的坐标记为 $\widetilde{x}$,则 $\widetilde{x}=xP^*$. 因此,

$$f(\alpha,\beta)=xH_2y^*=\widetilde{x}H_1\widetilde{y}^*,$$

其中 $\widetilde{x}$ 与 $\widetilde{y}$ 分别是向量 α 与 β 在基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 下的坐标. 所以方阵 H_1 是 Hermite 共轭双线性函数 $f(\alpha,\beta)$ 在基 $\{\eta_1,\eta_2,\ldots,\eta_n\}$ 下的方阵. 也就是说,复相合的 Hermite 方阵 H_1 与 H_2 决定同一个共轭双线性函数 $f(\alpha,\beta)$.

由定理 9.4.4 与定理 9.4.5,存在 V 的基 $\{\zeta_1,\zeta_2,\ldots,\zeta_n\}$,使得 $f(\alpha,\beta)$ 在这组基下的方阵即为

$$\operatorname{diag}\left(I_{(p)},-I_{(q)},\ 0_{(n-p-q)}\right),$$

其中 p 与 q 分别是 $f(\alpha,\beta)$ 的正、负惯性指数.

由于一个共轭双线性函数在不同基下的方阵是复相合的，所以方阵 H_1 与 H_2 都复相合于对角形 $\mathrm{diag}\big(I_{(p)}, -I_{(q)}, 0_{(n-p-q)}\big)$. 这表明，方阵 H_1 与 H_2 的正、负惯性指数都是 p 与 q. 由于方阵 H_1 与 H_2 的秩相等，因此它们的符号差也相等. 所以 Hermite 方阵的秩与符号差是 Hermite 方阵在复相合下的不变量.

反之，设 Hermite 方阵 H_1 与 H_2 的秩和符号差分别相等，则它们的正、负惯性指数分别相等，且设它们的正、负惯性指数分别是 p 与 q. 于是由定理前一结论，方阵 H_1 与 H_2 都复相合于对角形

$$\mathrm{diag}\big(I_{(p)}, -I_{(q)}, 0_{(n-p-q)}\big).$$

由于复相合关系满足对称性与传递性，所以方阵 H_1 与 H_2 复相合. 这就证明，Hermite 方阵的秩与符号差是 Hermite 方阵在复相合下的全系不变量. ■

最后讨论 Hermite 型.

定义 9.4.5 设 $f(\alpha,\beta)$ 是 n 维复线性空间 V 上的 Hermite 共轭双线性函数. 则 $H(\alpha)=f(\alpha,\alpha)$ 称为 **V 上的 Hermite 型**，其中 $\alpha \in V$.

如果 $f(\alpha,\beta)$ 是正定（或者半正定、负定与半负定）的，则 Hermite 型 $H(\alpha)$ 称为**正定**（或者**半正定**、**负定**与**半负定**）的.

设 Hermite 共轭双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\ldots,\xi_n\}$ 下的方阵为 $H=(h_{k\ell})$，其中 $h_{k\ell}=f(\xi_k,\xi_\ell)$，即

$$f(\alpha,\beta)=xHy^*,$$

其中 x 与 y 分别是向量 α 与 β 在这组基下的坐标. 则

$$H(\alpha)=f(\alpha,\alpha)=xHx^*=\sum_{1\leqslant k,\ell\leqslant n} h_{k\ell}x_k\overline{x}_\ell.$$

上式称为 Hermite 型 $H(\alpha)$ 在这组基下的表达式，方阵 H 称为 $H(\alpha)$ 在这组基下的方阵. 而 $f(\alpha,\alpha)$ 的秩，正、负惯性指数与符号差即称为 $H(\alpha)$ 的秩，正、负惯性指数与符号差.

由定理 9.4.4 立即得到

定理 9.4.7 设 $H(\alpha)$ 是 n 维复线性空间 V 上的 Hermite 型，则存在 V 的基 $\{\xi_1, \xi_2,\ldots,\xi_n\}$，使得

$$H(\alpha)=x_1\overline{x}_1+\cdots+x_p\overline{x}_p-x_{p+1}\overline{x}_{p+1}-\cdots-x_{p+q}\overline{x}_{p+q},$$

其中 p 与 q 分别是 Hermite 型 $H(\alpha)$ 的正、负惯性指数，而 $x=(x_1,x_2,\ldots,x_n)$ 是向量 α 在这组基下的坐标.

与二次型相仿，可以证明

定理 9.4.8 n 维复线性空间 V 上的 Hermite 型 $H(\alpha)$ 为正定（或半正定）的必要与充分条件是，Hermite 型 $H(\alpha)$ 在 V 的基下的方阵为正定（或半正定）Hermite 方阵.

利用 Hermite 方阵在酉相似下的标准形,可以证明

定理 9.4.9 设 $H(\alpha)$ 是 n 维酉空间 V 上的 Hermite 型. 则存在 V 的一组标准正交基 $\{\xi_1, \xi_2, \ldots, \xi_n\}$,使得

$$H(\alpha) = \lambda_1 x_1 \overline{x}_1 + \lambda_2 x_2 \overline{x}_2 + \cdots + \lambda_r x_r \overline{x}_r,$$

其中 r 是 $H(\alpha)$ 的秩,而 $x = (x_1, x_2, \ldots, x_n)$ 是向量 α 在这组基下的坐标.

习 题 9.4

1. 把下列 Hermite 型 $H(\alpha)$ 化为标准形,其中 $\mathrm{i}^2 = -1$:

(1) $H(\alpha) = \sum_{j=1}^{n-1}(x_j\overline{x}_{j+1} + x_{j+1}\overline{x}_j)$; (2) $H(\alpha) = \sum_{1\leqslant k,\ell\leqslant n} |k - \mathrm{i}\ell| x_k \overline{x}_\ell$.

2. 求下列 Hermite 型 $H(\alpha)$ 的秩与符号差.

(1) $H(\alpha) = a\sum_{j=1}^{n} x_j\overline{x}_j + \sum_{1\leqslant k\neq\ell\leqslant n}(x_k\overline{x}_\ell + x_\ell\overline{x}_k)$; (2) $H(\alpha) = \sum_{1\leqslant k\neq\ell\leqslant n} |x_k - x_\ell|^2$.

3. 设 n 维复线性空间 V 上的 Hermite 型为

$$H(\alpha) = \sum_{1\leqslant k,\ell\leqslant n}(ak\ell + k + \ell)x_k\overline{x}_\ell,$$

其中 a 是常数. 证明,$H(\alpha)$ 的秩与符号差和复数 a 无关.

4. 设 r 与 s 分别是 n 维复线性空间 V 上的 Hermite 型 $H(\alpha)$ 的秩与符号差. 证明,存在 V 的子空间 W,使得 $\dim W = \dfrac{1}{2}(r - s)$,并且对任意非零向量 $\alpha \in W$ 均有 $H(\alpha) < 0$.

5. 把下列 Hermite 方阵在复相合下化为标准形.

(1) $$\begin{pmatrix} 1 & 1+\mathrm{i} & 2+\mathrm{i} & \cdots & (n-1)+\mathrm{i} \\ 1-\mathrm{i} & 1 & 0 & \cdots & 0 \\ 2-\mathrm{i} & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ (n-1)-\mathrm{i} & 0 & \cdots & 0 & 1 \end{pmatrix};$$

(2) $$\begin{pmatrix} 1 & \frac{\mathrm{i}}{2} & 0 & 0 & \cdots & 0 \\ -\frac{\mathrm{i}}{2} & 1 & \frac{\mathrm{i}}{2} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{\mathrm{i}}{2} & 1 & \frac{\mathrm{i}}{2} \\ 0 & \cdots & 0 & 0 & -\frac{\mathrm{i}}{2} & 1 \end{pmatrix}_{n\times n}.$$

附 录

線性代數五講

前　言

从现代数学的观点来重新审视与认识数学基础课是数学教育现代化的一个重要途径.

外微分形式使微积分从古典走向现代,而模的理论使线性代数从古典走向现代. 这本小书就是从模的观点来重新审视与认识线性代数. 本书不是大学数学基础课线性代数的教材,读者的对象是已经念过线性代数与近世代数这两门基础课的大学生,希望他们阅读过这本小书后,能在高一个层次上来认识线性代数.

线性代数是研究线性空间(向量空间)、模和其上的线性变换以及与之有关的问题(如线性、双线性、二次函数等)的数学学科. 在本书的第一讲中介绍了向量空间、线性变换以及其它一些基本概念; 在第二讲中讨论了向量空间以及其上的线性泛函与对偶空间,其上的双线性形式、二次型及度量向量空间,正交几何与辛几何的分类,还有大家十分熟悉的内积空间; 第三讲中讨论了向量空间上的线性变换以及与之相关伴随算子及内积空间上的共轭算子; 第四讲与第五讲是本书的主要部分,是用模的观点来重新审视与认识线性代数. 在第四讲中定义了环上的模,尤其着重讨论了主理想整环上的模及其分解定理. 若 $\mathbb{F}$ 是域,则 $\mathbb{F}$ 上的多项式环 $\mathbb{F}[x]$ 是主理想整环; 若 V 是域 $\mathbb{F}$ 上的向量空间,给定 V 上的一个线性变换 τ 后,可以定义数乘,使得 V 可以看作主理想整环 $\mathbb{F}[x]$ 上的模. 从这个观点可以将第四讲中关于主理想整环上的模的理论与定理,“翻译”成为向量空间中的理论与定理,而这些内容正是大家在大学数学基础课线性代数中熟悉的主要内容,但这已是从更高一个层次,从模的观点来重新审视与认识它们了.

我要深切感谢上海交通大学章璞教授、中国科学技术大学李炯生教授、赵林城教授和北京师范大学张英伯教授,他们先后分别认真地阅读了本书的书稿,并提出了十分宝贵的意见,他们花了很多精力,对书稿反复修改使本书增色不少. 我也要深切感谢程艺、叶向东、陈发来与刘太顺教授对本书的关心与支持. 科学出版社杨波、李鹏奇与姚莉丽等同志为本书出版作了很大的努力,使我感激不尽. 我还要感谢余华敏小姐为精心打印本书所付出的辛勤劳动.

这本小书中的错误与不妥之处一定不少,还望读者不吝赐教.

龚　昇

二〇〇四年六月于北京玉海园

一些基本的代数结构

在这一讲中,首先讨论了线性代数的研究对象,然后回顾了群、环、域以及向量空间、线性变换等基本代数结构的定义. 由于今后的讨论大部分都是在主理想整环上进行的,所以着重讨论了主理想整环的一些基本性质.

A1.1 线性代数所研究的对象

什么是线性代数? 它所研究的对象是什么?

要说清楚这点,先得弄清楚什么是代数. 而代数的定义又是随时代的变化而不断的变化,不妨十分简略地回顾一下.

小学里学习的数学叫算术,主要是讨论数字的一些运算,这些内容人们很早就已经知道,并沿用了几千年,直到后来,产生了"数字符号化",才彻底改变了这种状况."数字符号化"就是用符号代替数字. 这件事在我国发生在宋元时代(约公元 13 世纪五六十年代),当时有"天元术"及"四元术". 也就是将未知数记作"天"元,后来将两个、三个及四个未知数记作"天"、"地"、"人"、"物"等四元,也就是相当于现在用 x, y, z, w 来表达四个未知数. 有了这些"元",也就可以解一些代数方程与联立方程组了. 在西方,彻底完成数字符号化是在公元 16 世纪."数字符号化"的产生标志着代数学"史前时期"的结束和代数学的诞生. 它包括了一元二次方程的求解,多元(一般为一元、二元至多四元)一次方程的求解等. 而这些正是目前中学代数课程的内容.

从公元 17 到 18 世纪中期,代数被理解为在代数符号上进行计算的数学,如解三次、四次代数方程,给出了这些方程的解法及根的具体表达式,建立了一些代数恒等式如二项式定理等. 从公元 18 到 19 世纪代数学的首要问题是求代数方程

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0 \qquad \text{(A1.1.1)}$$

的根式解,即推导出方程的系数经加、减、乘、除并开方所构成的公式来表示方程的根. 在已知一次、二次、三次、及四次代数方程的根式解后,不知多少人企图找出五次及更高次代数代数方程的根式解,但都以失败告终. 直到 1770 年,J. Lagrange (1736–1813) 看到了五次及高次方程不可能做到这点. 又过了半个世纪,1824 年 N. Abel (1802–1829) 解决了这个问题,即对于五次和五次以上一般方程求根式解是不可能的. 但什么样的代数方程能根式可解,这是 1880 年由 E. Galois (1811–1832) 彻底解决的. 他证明了: 方程根式可解当且仅当它的 Galois 群可解. Abel 与 Galois

不仅解决了三百年来无法解决的著名难题,更重要的是:他们为了解决这个难题,建立起“域”和“群”的概念,为后来近世代数的产生作了准备.

与此相关的问题是要证明方程 (A1.1.1) 的根的存在性. 即若方程 (A1.1.1) 的系数都是复数,则至少有一个复数根. 这就是著名的代数基本定理. 18 世纪末,C. F. Gauss (1777–1855) 给出了这个定理的证明.

从 19 世纪中叶,代数学最终从方程式论转向代数运算的研究. 代数学及代数运算的一般理论与近代观点于 20 世纪初在 D. Hilbert (1862–1943), E. Steinitz (1871–1928), A. E. Noether (1882–1935), E. Artin (1898–1932) 等人的影响下得以明确.

近世代数的主要内容是集合及这些集合上的代数运算. 集合本身和作为代数运算的载体的集合是不加区分的,故实质上研究的是代数运算本身. 说更仔细一些,考虑非空集合 S 上一个或几个二元运算. 运算作用在集合两个元素之间得到的元素仍在集合中,对集合施行运算要适合一些法则(或称公理),则集合对于运算成一**代数结构**. **研究代数结构的性质是近世代数的内容与任务**. 主要的代数结构有:群、环、体、域、模等,这将在下一节详细定义之.

特别要强调的是:研究一个代数结构,除了要了解它的内部构造和结构性质外,一个重要而基本的方法是研究这个代数结构的表示,或这个代数结构上的模. 例如作为一个群上的模,粗略地说,就是这个群在一个向量空间上的作用,作用的效果如何当然反映出这个群本身的性质;而一个环上的模,粗略地说,就是这个环在一个 Abel 群上的作用. 模本身既可以看成一个代数结构,更重要的,它是一个代数结构在另一个代数结构上的作用. 因此,可以说,**现代代数学的两大主题是结构与表示理论**.

线性代数是研究线性空间(向量空间)、模和其上的线性变换以及与之有关的问题(如线性、双线性、二次函数等)的数学学科. 也就是说,此时,代数结构是线性空间. 仅仅讨论线性空间的结构是不够的,还要考虑线性变换在其上的作用. 从表示论的观点看,带有线性变换的线性空间就是主理想整环上的模. 这就是本书希望用模的观点来考虑线性代数的出发点.

从这里还可以看出,线性代数所研究的是:线性空间;模是线性空间的扩充;作用在线性空间上的线性变换,大致上说,线性变换就是将一个线性空间映到另一个线性空间,且保持线性空间上的运算的映射;定义在线性空间上的线性泛函及其推广双线性形式,而二次型不过是双线性形式的特例. 因此,可以说“线性”是线性代数的灵魂;线性代数只考虑“线性”的问题,而“非线性”的问题就不在讨论之列了.

A1.2 主理想整环

回顾一下一些重要的代数结构的定义.

1. 群 (group) 这是最基本、最重要的代数结构.

群是一非空集合 G,其上有一个二元运算 $*$(通常称为乘法)满足:

(1) 封闭性 对所有 $a, b \in G$,则

$$a * b \in G;$$

(2) 结合律 对所有 $a, b, c \in G$,则

$$(a * b) * c = a * (b * c);$$

(3) 有单位元 G 中存在一元素 e,使得对任意 $a \in G$,都有

$$e * a = a * e = a;$$

(4) 有逆元 对 G 中每一元素 a,存在 G 中的元素 a^{-1},使得

$$a * a^{-1} = a^{-1} * a = e.$$

若群 G 还满足

(5) 交换律 对所有 $a, b \in G$,有

$$a * b = b * a,$$

则称 G 为 **Abel 群**或**交换群**. 通常此时运算 $*$ 称为加法,并用 $+$ 来替代.

若集合 G 只满足 (1) 和 (2),则称 G 为**半群 (semi-group)**.

2. 环 (ring) 环是一个非空集合 R,有两个二元运算,加法(常记作 $+$)及乘法(常用毗连表示)满足:

(1) 加法 Abel 群 R 对加法成一个 Abel 群;

(2) 乘法半群 R 对乘法成一个半群;

(3) 分配律 对所有 $a, b \in R$,有

$$a(b + c) = ab = ac \quad 及 \quad (a + b)c = ac + bc.$$

若环 R 还满足:

(4) 交换律 对所有 $a, b \in R$,有

$$ab = ba,$$

则称 R 为**交换环**. 若环 R 中有元素 e,使得

$$ae = ea,$$

则称 R 为有**单位元 (identity) 的环**,单位元常记作 1.

本书中讨论的大多数环均为有单位元的交换环.

使 $c1=0$ 的最小正整数 c,称为环 R 的**特征 (characteristic)**. 若没有这样的 c,则称 R 的特征为 0.

3. 整环,也叫作整域 (intergral domain) R 为交换环,非零元素 $r \in R$ 称为**零因子 (zero divisor)**,如果存在非零元素 $s \in R$,使得 $rs=0$. 一个无零因子的有单位元的交换环称为**整环**. 与此等价的说法是:

一个满足**消去律 (cancellation law)** 的有单位元的交换环称为整环.

所谓消去律是,若 $x, y, r \in R$ 且 $r \neq 0$,则 $rx = ry$ 蕴涵 $x = y$.

4. 体,也叫除环 (division ring) 或拟域 (skew field) 一个有单位元的环称为**体**,如果所有非零元素全体对乘法作成群. 也就是有除法的有单位元的环.

5. 域 (field) 可交换的体称为域.

6. 主理想整环 (principal ideal domain) 若 R 是环,R 中一个子集 I 称为 R 的一个**理想 (ideal)**,如果

(1) I 对 R 中的加法成 Abel 群; **(2)** 若 $a \in I, r \in R$,则 $ar \in I$ 及 $ra \in I$.

如果将条件 (1) 易以

(1′) 对 I 中任意两个 a, b,则 $a-b \in I$,

易证条件 (1) 和 (2) 与条件 (1′) 和 (2) 等价. 有些书上用条件 (1) 和 (2) 来定义理想.

若 R 为有单位元的交换环,S 为 R 的一个子集,则易知集合

$$\langle s_1, \ldots, s_n \rangle = \left\{ r_1 s_1 + \cdots + r_n s_n \;\middle|\; r_i \in R,\ s_i \in S \right\}$$

为 R 的一个理想,称其为**由 S 生成 (generated) 的理想**. 称由一个元素 a 生成的理想

$$\langle a \rangle = \left\{ ra \;\middle|\; r \in R \right\}$$

为**由 a 生成的主理想 (principal ideal)**.

每一个理想都是主理想的整环称为**主理想整环**.

今后讨论将着重在主理想整环上进行,不妨再对此多说几句. 整数全体 $\mathbb{Z}$ 是整环,且也是主理想整环. 这是因为 $\mathbb{Z}$ 的任一理想 I 都是有 I 中的最小正整数 a 生成的.

若 $\mathbb{F}$ 为域,所有系数均在 $\mathbb{F}$ 上的单变量多项式的集合 $\mathbb{F}[x]$ 是一个有单位元的交换环. 若 $p(x), q(x) \in \mathbb{F}[x]$,且 $p(x)q(x)=0$,则有 $p(x)=0$ 或 $q(x)=0$,故 $\mathbb{F}[x]$ 是一个整环. 不但如此,要证明以下一条十分重要且有用的定理.

定理 A1.2.1 $\mathbb{F}[x]$ 是主理想整环.

证明 设 I 是 $\mathbb{F}[x]$ 的一个理想,$m(x)$ 是 I 中最低次的首一多项式 (monic polynomial),即首项系数为 1 的多项式. 首先看出,在 I 中,这样的多项式是唯一的. 事实上,若还有一个首一的多项式 $n(x)$,且 $\deg n(x) = \deg m(x)$,则

$$b(x)=m(x)-n(x)\in I.$$

将 $b(x)$ 乘以其最高次项系数的逆,得一首一多项式 $b_1(x)$,而 $b_1(x)\in I$,但 $\deg b_1(x)=\deg b(x)<\deg m(x)$. 故 $b_1(x)=0$,因此 $b(x)=0$,即 $m(x)=n(x)$.

现在来证 I 由 $m(x)$ 生成.

因为 I 是理想,$m(x)\in I$,故 $\langle m(x)\rangle\subseteq I$. 往证反方向的包含关系. 若 $p(x)\in I$,则 $p(x)$ 用 $m(x)$ 相除,得到

$$p(x)=m(x)q(x)+r(x),$$

这里 $r(x)=0$ 或 $0\leqslant\deg r(x)<\deg m(x)$. 由于 I 是理想,故

$$r(x)=p(x)-m(x)q(x)\in I.$$

将 $r(x)$ 乘以其最高此项系数的逆,得一首一多项式 $r_1(x)$,而 $r_1(x)\in I$,故

$$0\leqslant\deg r_1(x)=\deg r(x)<\deg m(x).$$

由于 $m(x)$ 的次数的最小性,所以 $r_1(x)=0$,故 $r(x)=0$,即

$$p(x)=q(x)m(x)\in\langle m(x)\rangle.$$

这就证明了 $I\subseteq\langle m(x)\rangle$. 因此,$I=\langle m(x)\rangle$. ■

还可以证明如下命题.

命题 A1.2.1 若 $p_1(x),\ldots,p_n(x)\in\mathbb{F}[x]$,则

$$\langle p_1(x),\ldots,p_n(x)\rangle=\langle\gcd(p_1(x),\ldots,p_n(x))\rangle,$$

这里 $\gcd(p_1(x),\ldots,p_n(x))$ 为 $p_1,\ldots,p_n$ 的最大公因子.

证明 令 $I=\langle p_1(x),\ldots,p_n(x)\rangle$,由定理 A1.2.1 可知,有 I 中唯一的一个最低次的首一多项式 $m(x)$,使得 $I=\langle m(x)\rangle$. 由于 $p_i(x)\in\langle m(x)\rangle$,故有多项式 $a_i(x)\in\mathbb{F}[x]$,使得

$$p_i(x)=a_i(x)m(x).$$

因此,$m(x)\mid p_i(x)$,即 $m(x)$ 是 $p_1(x),\ldots,p_n(x)$ 的公因子.

若有 $q(x)\mid p_i(x)$,则 $p_i(x)\in\langle q(x)\rangle$. 由于 $\langle p_1(x),\ldots,p_n(x)\rangle$ 是包含 $p_1(x),\ldots,p_n(x)$ 的最小理想,故

$$\langle m(x)\rangle=\langle p_1(x),\ldots,p_n(x)\rangle\subseteq\langle q(x)\rangle.$$

因此 $m(x)\in\langle q(x)\rangle$,即 $q(x)\mid m(x)$,故 $m(x)$ 为 $p_1(x),\ldots,p_n(x)$ 的最大公因子. ■

值得注意的是:由两个变数 x 与 y 的多项式的全体组成的多项式环 $R=\mathbb{F}[x,y]$ 是整环,但不再是主理想整环.

下面来证一个有用且重要的主理想整环上的素因子分解定理.

先给出一些定义. 设 R 是整环.

(1) $r,s\in R$,称 r **可除 (divide)** s,记作 $r\mid s$,若存在 $x\in R$,使得 $s=xr$;

(2) $u\in R$ 称为一个**可逆元 (unit)**,若有 $v\in R$,使得 $uv=1$;

(3) 若 $a, b \in R$,称 a, b **相伴 (associate)**,若有 R 中可逆元 u,使得 $a = ub$;

(4) 一个非零非可逆元 $p \in R$ 称为**素元 (prime)**,若 $p \mid ab$ 蕴涵 $p \mid a$ 或 $p \mid b$;

(5) 一个非零可逆元 $p \in R$ 称为**不可约元 (irreducible)**,若 $p = ab$ 蕴涵 a 或 b 是可逆元.

由此可以得到

(1) $u \in R$ 为逆元当且仅当 $\langle u \rangle = R$; (2) r, s 相伴当且仅当 $\langle r \rangle = \langle s \rangle$;

(3) r 可除 s 当且仅当 $\langle s \rangle \subseteq \langle r \rangle$;

(4) r 真除 (properly divide) s 当且仅当 $\langle s \rangle \subsetneqq \langle r \rangle$.

所谓 r 真除 (properly divide) s,是指 $s = xr$,其中 x 不是一个可逆元.

对整数环 $\mathbb{Z}$,一个整数是素元(素数)当且仅当它是不可约元. 但一般来说,这两者是不一致的. 但对于主理想整环,这两者却是一致的.

定理 A1.2.2 若 R 是一个主理想整环,则 R 中的一个元素是素元当且仅当它是不可约元.

证明 若 p 是素元,令 $p = ab$,则 $p \mid ab$,因此 $p \mid a$ 或 $p \mid b$. 若 $p \mid a$,则 $a = xp$,于是

$$p = ab = xpb.$$

由于 R 是整环,故消去律成立,在上式中消去 p 后得到 $1 = xb$,因此 b 是一个可逆元,故 p 不可约. 需要注意的是,在这部分证明中只用到 R 是整环,并未用到 R 是主理想整环,故这部分对 R 是整环也成立. 即在整环中素元一定是不可约元.

下面证明不可约元一定是素元. 先来证明: 若 $r \in R$ 是不可约元,则主理想 $\langle r \rangle$ 是**极大理想 (maximal ideal)**,即 $\langle r \rangle \neq R$,且不存在理想 $\langle a \rangle$ 使得 $\langle r \rangle \subsetneqq \langle a \rangle \subsetneqq R$.

若有 $\langle a \rangle$,使得 $\langle r \rangle \subseteq \langle a \rangle \subseteq R$,则 $r = xa, x \in R$. 由于 r 为不可约元,故 a 或 x 为可逆元. 若 a 为可逆元,则由前述 (1),$\langle a \rangle = R$; 若 x 为可逆元,则由前述 (2),$\langle a \rangle = \langle xa \rangle = \langle r \rangle$. 这得到矛盾,故 $\langle r \rangle$ 为极大理想.

若 r 为不可约元,且 $r \mid ab$,要证 $r \mid a$ 或 $r \mid b$,即 r 是素元. 由前述 (3),$ab \in \langle r \rangle$. 由刚才已证的知道 $\langle r \rangle$ 是极大理想,要证 $a \in \langle r \rangle$ 或 $b \in \langle r \rangle$. 若 $a \notin \langle r \rangle$,由于 $\langle r \rangle$ 是极大理想,故 $\langle a, r \rangle = R$,因此有 $x, y \in R$,使得 $1 = xa + yr$. 将此式两边又乘以 b,得 $b = xab + yrb$,由 $r \mid ab$ 得 $r \mid xab$,又显然有 $r \mid yrb$,因此 $r \mid b$,即 $b \in \langle r \rangle$.

同样可证,若 $b \notin \langle r \rangle$,则有 $a \in \langle r \rangle$. 这就证明了 r 为素元. ■

如果环 R 有理想序列 $I_1, I_2, \ldots$,满足

$$I_i \subseteq I_{i+1}, \quad i = 1, 2, \ldots,$$

则称 $\{I_i\}$ 为一个**理想升链 (ascending chain of ideals)**.

先来证明如下命题.

命题 A1.2.2 若 R 是主理想整环,则任一理想升链 $\{\langle a_i \rangle\}$ 一定是有限的,即存在某个正整数 m,使得

$$\langle a_m \rangle = \langle a_{m+1} \rangle = \langle a_{m+2} \rangle = \cdots.$$

证明 令 $I = \bigcup_i \langle a_i \rangle$,先证明 I 是 R 一个理想.

事实上,对于任意 $b, c \in I$,则 b, c 分别属于某个 $\langle a_j \rangle$ 或 $\langle a_k \rangle$. 不妨假设 $j \leqslant k$,于是 $\langle a_j \rangle \subseteq \langle a_k \rangle$. 因此 $b \in \langle a_k \rangle, b - c \in \langle a_k \rangle$. 对于任意 $d \in R, b \in \langle a_j \rangle$,得 $bd \in \langle a_j \rangle$, $ab \in \langle a_j \rangle$. 因此 $bd \in I, db \in I$. 所以 I 是 R 的一个理想.

由于 R 是主理想整环,故 $I = \langle f \rangle$. 由 I 的定义 f 属于某个 $\langle a_m \rangle$,从而 $I \subseteq \langle a_m \rangle$. 反之,显然 $\langle a_m \rangle \subseteq I$. 所以 $I = \langle a_m \rangle$ 对于任意大于 m 的整数 n 有 $\langle a_m \rangle \subseteq \langle a_n \rangle \subseteq I$,由 $\langle a_m \rangle = I$ 可得 $\langle a_n \rangle = I$. 最后得到

$$I = \langle a_m \rangle = \langle a_{m+1} \rangle = \langle a_{m+2} \rangle = \cdots. \qquad \blacksquare$$

由此可得如下定理.

定理 A1.2.3(主理想整环上素元分解定理) 若 R 是主理想整环,则任一 $r \in R$, $r \neq 0$ 可以写成

$$r = u p_1 p_2 \cdots p_n,$$

这里 u 是可逆元, $p_1, \ldots, p_n$ 是素元. 并且除去排列次序及可逆元 u 外,这样的因子分解是唯一的.

证明 由定理 A1.2.2, R 是主理想整环时,素元与不可约元是一致的,故只要将 $r \in R$ 分解为不可约元的乘积即可.

若 $r \in R$,如果 r 为不可约元,则定理已证. 若不是,则 $r = r_1 r_2$,而 r_1, r_2 都不是可逆元,若 r_1, r_2 都是不可约元,则定理已证. 否则,若 r_2 不是不可约元,则 $r_2 = r_3 r_4$,而 r_3, r_4 都不是可逆元. 这个步骤一直进行下去,则 r 分解为

$$r = r_1 r_2 = r_1(r_3 r_4) = (r_1 r_3)(r_5 r_6) = (r_1 r_3 r_5)(r_7 r_8) = \cdots.$$

每步分解将 r 分解为非可逆元的乘积,但这种分解经过有限步后必然停止. 这是因为

$$r_2 \mid r, \quad r_4 \mid r_2, \quad r_6 \mid r_4, \ldots,$$

故由上述 (3),得到一个上升的理想序列

$$\langle r \rangle \subseteq \langle r_2 \rangle \subseteq \langle r_4 \rangle \subseteq \langle r_6 \rangle \subseteq \cdots. \tag{A1.2.1}$$

由于所有 r_i 均不可逆,故由上述 (4),上式的包含是真包含. 如果这种分解不能停止,则得到一个上升的理想无穷序列. 然而由 命题 A1.2.2,这个升链一定有限,即有 n 使得

$$\langle r_{2n} \rangle = \langle r_{2(n+1)} \rangle = \cdots,$$

这与理想序列 (A1.2.1) 中的包含关系是真包含相矛盾. 于是

$$r = r_1 r_3 \cdots r_{2n-1} r_{2n},$$

这里 r_{2n} 是不可约元. 记 $r_1 r_3 \cdots r_{2n-1} = s$,则 $r = s r_{2n}$. 对 s 重复上面的证明,则可分解性得证.

利用主理想整环上素元与不可约元的一致性,几乎是重复整数环 $\mathbb{Z}$ 的算术基本定理(唯一素数分解定理)的唯一性的证明,可以给出 定理 A1.2.3 的唯一性证明.

此处从略. ■

A1.3 向量空间与线性变换

在前面 A1.1 中已经讲到,线性代数是研究线性空间,即向量空间、模和其上的线性变换以及与之有相关的问题,如线性函数、双线性形式等等的数学学科.

先来定义线性空间. **线性空间 (linear space)**,也称**向量空间 (vector space)**,来源于解析几何中三维空间的推广. 当时向量是定义为有大小、有方向的量,现在给出的定义是抽象的定义,使用的范围当然要广泛得多.

定义 A1.3.1 若 $\mathbb{F}$ 是域,其中元素称为**纯量 (scalar)**. $\mathbb{F}$ 中的一个**向量空间**为一个非空集合 V,它的元素称为**向量 (vector)**,有运算 $+$,对 $(u,v)\in V\times V$,有 $u+v\in V$;以及 $\mathbb{F}$ 与 V 的运算数乘,用毗连表示,对 $(r,u)\in\mathbb{F}\times V$,有 $ru\in V$,且满足以下这些条件:

(1) V 对 $+$ 作成 Abel 群;

(2) $\mathbb{F}$ 对 V 的数乘满足: 对所有的 $r,s\in\mathbb{F}$, $u,v\in V$ 有

分配律

$$r(u+v)=ru+rv,\qquad (r+s)u=ru+su;$$

结合律

$$(rs)u=r(su);$$

及

$$1u=u.$$

这样定义的向量空间当然要比解析几何中定义的向量空间要广泛得多. 例如:

(1) 若 $\mathbb{F}$ 为域,所有将 $\mathbb{F}$ 映到 $\mathbb{F}$ 的函数的全体是一个向量空间;

(2) 所有元素取自域 $\mathbb{F}$ 的 $m\times n$ 矩阵的全体,对矩阵加法与矩阵的数乘成一个向量空间,记作 $\mathcal{M}_{m\times n}(\mathbb{F})$; 若 $m=n$,则记作 $\mathcal{M}_n(\mathbb{F})$.

再来定义线性变换. 粗略地说,线性变换是将一个向量空间映到另一个向量空间,且保持向量空间中的运算的映射.

定义 A1.3.2 若 V 与 W 是 $\mathbb{F}$ 上的两个向量空间,映射 $\tau: V\to W$ 称为一个**线性变换 (linear transformation)**,若对任意 $r,s\in\mathbb{F}$ 及 $u,v\in V$,有

$$\tau(ru+sv)=r\tau(u)+s\tau(v).$$

记从 V 到 W 的线性变换的全体为 $\mathcal{L}(V,W)$.

称线性变换 $\tau: V\to V$ 为 V 上的**线性算子 (linear operator)**. 记 V 上所有的线性算子的全体为 $\mathcal{L}(V)$.

在各种代数结构中,就有一种结构叫作代数,定义如下.

定义 A1.3.3 若 $\mathbb{F}$ 为域，$\mathbb{F}$ 上的一个**代数 (algebra)** $\mathcal{A}$ 为一个非空集合 $\mathcal{A}$，且有两种运算：加法（记作 $+$），乘法（用毗连表示）以及 $\mathbb{F}$ 对 $\mathcal{A}$ 的运算数乘（也用毗连表示）满足以下规律：

(1) 对加法与 $\mathbb{F}$ 对 $\mathcal{A}$ 的数乘，$\mathcal{A}$ 是一个向量空间；

(2) 对加法与乘法，$\mathcal{A}$ 是一个有单位元的环；

(3) 若 $r \in \mathbb{F}$ 及 $a, b \in \mathcal{A}$，有

$$r(ab) = (ra)b = a(rb).$$

也就是说，代数是有向量乘法的向量空间，代数是可以对每个元素进行数乘的环. 也可以说代数既是向量空间又同时是环，是向量空间与环的结合.

若 V 是 $\mathbb{F}$ 上的一个向量空间，对于 $\mathcal{L}(V)$，取 $\mathcal{L}(V)$ 中两个元素的乘法为映射的复合，取 $\mathcal{L}(V)$ 中的恒等映射为 $\mathcal{L}(V)$ 中乘法的单位元，则容易验证：$\mathcal{L}(V)$ 确实是 $\mathbb{F}$ 上的一个代数. 这个代数的元素就是 $\mathcal{L}(V)$ 的元素. 当然，还有其它的代数，如李代数、Clifford 代数等等.

线性变换与向量空间是相辅相成的，是相互依存的. 向量空间是线性变换的载体，没有向量空间，线性变换无用武之地，对它进行研究也就没多大意义. 反之，向量空间本身如果没有线性变换作用于其上，则向量空间是死的，没有多少话可说.

如在 A1.1 中所说的，近世代数的主要内容是集合与这些集合上的代数运算. 集合本身和作为代数运算的载体的集合是不加区分的，故实质上研究的是代数运算本身. 而线性代数实质上是在研究 $\mathcal{L}(V)$. 如上所述，$\mathcal{L}(V)$ 的确是一个代数.

讨论数域 $\mathbb{F}$ 上带有一个线性变换 τ 的线性空间 V，从模的观点就是讨论 $\mathbb{F}[x]$ 上的模. 这就是本书从模的观点来讨论线性代数的出发点.

A1.4 同构、等价、相似与相合

若 S, T 是两个集合，$f: S \to T$ 是从 S 到 T 的一个映射.

称 f 为**单射 (injective)** 或**一对一 (one to one)**，若 $x \neq y$ 蕴涵 $f(x) \neq f(y)$；

称 f 为**满射 (surjective)** 或**映上 (onto)**，若 $f(S) = T$；

称 f 为**双射 (bijective)**，若 f 既是单射又是满射；

称 $f(S) = \{f(s) \mid s \in S\}$ 为 f 的**像 (image)**，记作 $\operatorname{Im} f$.

若 V, W 是 $\mathbb{F}$ 上的两个向量空间，$\tau \in \mathcal{L}(V, W)$，称 $\{s \in S \mid f(s) = 0\}$ 为映射 f 的**核 (kernel)**，记作 $\operatorname{Ker} f$. 则有

(1) τ 是满射当且仅当 $\operatorname{Im} \tau = W$；　　**(2)** τ 是单射当且仅当 $\operatorname{Ker} \tau = 0$.

若线性变换 $\tau \in \mathcal{L}(V, W)$ 是双射，则称 τ 是从 V 到 W 的**同构变换 (isomorphism)**，称线性空间 V 与 W **同构**，记作 $V \approx W$.

同构是线性变换中极为重要的概念,两个向量空间是同构的,则有线性变换,使这两个空间的点一一对应,且保持线性关系不变. 这时我们往往将这两个向量空间视为同一个. 如对向量空间进行分类,就是指在同构意义下的分类.

更一般地,有等价关系. 若 S 是一非空集合,S 上的一个二元关系 $\sim$ 称为 S 上的**等价关系 (equivalence relation)**,若它满足如下三个条件:

(1) 自反性 (reflexivity) 对所有 $a\in S$,有 $a\sim a$;

(2) 对称性 (symmetry) 对所有 $a,b\in S$,有 $a\sim b$ 蕴涵 $b\sim a$;

(3) 传递性 (transitivity) 对所有 $a,b,c\in S$,有 $a\sim b$,$b\sim c$ 蕴涵 $a\sim c$.

若 $a\in S$,集合 $[a]=\{b\in S \mid b\sim a\}$ 称为 a 的**等价类 (equivalent class)**. 若 S 是非空集合,S 的一个**划分 (partition)** 是 S 的一个非空子集的集合 $\{A_1,\ldots,A_n,\ldots\}$ 满足

(1) $A_i\cap A_j=\varnothing$ 对所有 $i\neq j$ 都成立; **(2)** $S=A_1\cup\cdots\cup A_n\cup\cdots$.

此时,这些 A_i 称为 S 的一个**块 (block)**.

显然,若 $\sim$ 是 S 的一个等价关系,则由 $\sim$ 所得到的不同的等价类是 S 划分的块. 反之,若 $\mathcal{P}$ 是 S 的一个划分,定义

$$a\sim b \Longleftrightarrow a,b \text{ 在 } \mathcal{P} \text{ 的同一类中},$$

则 $\sim$ 是 S 上的一个等价关系,它的等价类就是 $\mathcal{P}$ 的块.

于是,S 的等价关系与 S 的划分是一一对应的.

设 $\sim$ 是 S 的等价关系,S 的一个子集 C 称为对 $\sim$ 而言的**标准形式 (canonical form)**,若对每一个 $s\in S$,在 C 中有唯一的一个 c,使得 $c\sim s$.

显然,对于向量空间,同构就是等价关系,在后面几讲中,将讨论在这个等价关系下,向量空间的标准形式.

若 $A,B\in\mathcal{M}_n(\mathbb{F})$,称 A 与 B **等价 (equivalent)**,若存在可逆阵 P,Q,使得

$$A=PBQ;$$

称 A 与 B **相似 (similar)**,若存在可逆阵 P,使得

$$A=PBP^{-1};$$

称 A 与 B **相合 (congruent)**,若存在可逆阵 P,使得

$$A=PBP^{\mathrm{T}},$$

这里 P^{T} 为矩阵 P 的转置.

易见,这三种矩阵的关系都是等价关系. 在一个有限维向量空间上同一个线性算子在不同基下所对应的矩阵之间的关系是相似关系(详见 A3.1). 本书的第五讲中将给出相似关系下的标准形式. 在一个有限维向量空间上同一个双线性形式在不同基下所对应的矩阵之间的关系是相合关系(详见 A2.3). 本书的第三讲中将给出相合关系下的标准形式. 在两个有限维向量空间之间的线性变换,在两个向

量空间的各自取定一组的基下所对应的矩阵之间的关系是等价关系(详见 A3.1).

若 $A \in \mathcal{M}_n(\mathbb{F})$,熟知对 A 有三个初等运算:

(1) 对 A 中的一行(列)乘以非零的 $r \in \mathbb{F}$;

(2) 将 A 中的两行(列)交换;

(3) 将 A 中的一行(列)乘以非零的 $r \in \mathbb{F}$ 加到另一行(列)上.

对 A 进行行(列)的初等运算相当于对 A 左(右)乘以相应的矩阵. 不难证明,任意 $A \in \mathcal{M}_n(\mathbb{F})$,经过行与列的初等运算可以变为

$$N_k = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix},$$

这里 $k \leqslant n$,I_k 是 k 阶单位方阵. 因此,在矩阵等价这个等价关系下,矩阵 $A \in \mathcal{M}_n(\mathbb{F})$ 的标准形式就是 N_k,$k = 0, 1, 2, \ldots, n$.

向量空间

在本书一开始,就明确指出,线性代数是研究线性空间,即向量空间、模和其上的线性变换及其与之相关问题的数学学科. 在 A1.3 中,定义 A1.3.1 以及定义 A1.3.2 分别给出了向量空间与线性变换的定义. 在这一讲中,将详细讨论向量空间.

A2.1 基与矩阵表示

关于向量空间有以下这些常规、常用的定义.

1. S 是数域 $\mathbb{F}$ 上的向量空间 V 的子集,如果将 V 的加法与 $\mathbb{F}$ 对 V 的数乘限制在 S 上,S 也成为一个向量空间,则称 S 为 V 的**子空间**.

2. 若 $V_1,\dots,V_n$ 是域 $\mathbb{F}$ 上的向量空间,令

$$V=\left\{(v_1,\dots,v_n)\ \middle|\ v_i\in V_i,\ i=1,\dots,n\right\},$$

且在其上定义加法

$$(u_1,\dots,u_n)+(v_1,\dots,v_n)=(u_1+v_1,\dots,u_n+v_n),$$

$\mathbb{F}$ 对 V 的数乘为

$$r(v_1,\dots,v_n)=(rv_1,\dots,rv_n),$$

这里 $r\in\mathbb{F}$,则 V 成为一个向量空间,称为向量空间 $V_1,\dots,V_n$ 的**直和 (direct sum)**,记作

$$V=V_1\oplus\cdots\oplus V_n.$$

若 S 是向量空间 V 的一个子空间,且有子空间 T,使得 $V=S\oplus T$,则称 T 为 S 的补 **(complement)**,记作 S^c. 可证 V 的任一子空间一定有补.

3. 向量空间 V 中的一个非空子集 S 称为**线性无关 (linearly independent)**,如果由

$$r_1v_1+\cdots+r_nv_n=0$$

可导出 $r_1=\cdots=r_n=0$,这里 $v_i\in S$,$r_i\in\mathbb{F}$.

V 中一个子集如果不是线性无关,则称其为**线性相关 (linearly dependent)**.

4. 向量空间 V 的一个集合 T 称为**生成 (span)** V,如果 V 中的每个向量可以写成 T 中的一些向量的线性组合,即对每个 $v\in V$,可以写成

$$v=r_1u_1+\cdots+r_mu_m,$$

这里 $r_i \in \mathbb{F}, u_i \in T$.

向量空间 V 中由子集 S 所有元素的线性组合的全体组成 V 中的一个子空间，记作

$$\langle S\rangle = \operatorname{span} S = \left\{r_1v_1 + \cdots + r_nv_n \;\middle|\; r_i \in \mathbb{F},\ v_i \in S,\ n = 1, 2, \ldots\right\}.$$

5. 向量空间 V 的一个线性无关且生成 V 的子集，称为 V 的一组**基 (basis)**. 向量空间 V 的基的基数 (cardinality) 称为 V 的**维数 (dimension)**，记作 $\dim V$. 当基为有限集时，这就是基中元素的个数.

这样定义的基是否存在? 这样定义的维数是否合理?

命题 A2.1.1 除了零空间 $\{0\}$ 之外，任意向量空间一定存在一组基.

证明 设 V 是非零向量空间，V 中线性无关的子集的全体记作 $\mathcal{A}$. 任取一个非零向量组成的集合就是一个线性无关子集，故 $\mathcal{A}$ 非空.

在 $\mathcal{A}$ 中可按集合的包含关系"$\subseteq$"定义一个偏序，若 $I_1 \subseteq I_2 \subseteq \cdots$ 是 V 中线性无关子集的一条链，则 $U = \bigcup_i I_i$ 仍为一个线性无关子集，故任一条链必有上界. 因此，由 Zorn 引理①，$\mathcal{A}$ 中必有极大元，即 V 有极大线性无关组 (maximal linearly independent set) S，即 S 是线性无关的，但任意真包含有 S 的集合一定不是线性无关的，于是 S 一定生成 V. 否则必有向量 $v \in V - S$，它不是 S 中的向量的线性组合. 于是 $S \cup \{v\}$ 是真包含有 S 的线性无关集，但这是一个矛盾. 这就证明了向量空间基的存在性. ■

命题 A2.1.2 这样定义的维数是合理的.

证明 先来证明如下的结果:

若 V 是一向量空间，而向量 $v_1, \ldots, v_n$ 是线性无关的，向量 $s_1, \ldots, s_m$ 生成 V，则 $n \leqslant m$. 先列出这两个向量集

$$s_1, \ldots, s_m; \qquad v_1, \ldots, v_n.$$

将后一个的 v_n 移到前一个，成为

$$v_n, s_1, \ldots, s_m; \qquad v_1, \ldots, v_{n-1}.$$

由于 $s_1, \ldots, s_m$ 生成 V，故 v_n 可表为 $s_1, \ldots, s_m$ 的线性组合，故可以从众 s_i 中移走其中的一个，例如 s_j，这样使移走 s_j 后的前一个向量集仍能生成 V，得到新的两个向量集

$$v_n, s_1, \ldots, \hat{s}_j, \ldots, s_m; \qquad v_1, \ldots, v_{n-1}.$$

其中 $\hat{s}_j$ 表示 s_j 已被移走. 现在再将 v_{n-1} 从后一个集合移到前一个集合，得

$$v_{n-1}, v_n, s_1, \ldots, \hat{s}_j, \ldots, s_m; \qquad v_1, \ldots, v_{n-2}.$$

同样理由可从前一个集合移走某个 s_k，使得移走后的前一个集合仍可生成 V，得到

$$v_{n-1}, v_n, s_1, \ldots, \hat{s}_j, \ldots, \hat{s}_k, \ldots, s_m; \qquad v_1, \ldots, v_{n-2}.$$

① **Zorn 引理** 若 P 为一个偏序集合 (partially ordered set)，并且每个链都有上界，则 P 中有极大元. 详见 [1].

将以上这些步骤可以一直进行下去，直到所有的 v_i 或所有 s_ℓ 全部移完. 这一过程称为对向量集 $\{s_1,\ldots,s_m\}$ 进行 Steinitz 替换. 若所有的 s_ℓ 首先移完，即 $m<n$，则前一个集合只是后一个集合 $v_1,\ldots,v_n$ 的一个真子集，而这又生成 V，这与 $v_1,\ldots,v_n$ 是线性无关矛盾，故必须有 $m\leqslant n$.

由此结果立即得到：若 V 由有限个向量所生成，则 V 的任意两个基有相同的基数，即在此情形，维数的定义式合理的. 至于 V 由无限个向量所生成的情形，也可证明同样的结论，这里就从略了. ■

6. 若 S 是域 $\mathbb{F}$ 上的向量空间 V 的子空间，$u,v\in V$，若 $u-v\in S$，则称 u 与 v **同余模 S (congruent modulo S)**，记作

$$u\equiv v \pmod S.$$

将所有与 v 同余的元素的全体记为 $[v]$，即 $u\in[v]$ 当且仅当 $u\equiv v \pmod S$. 称 $[v]$ 为向量空间 V 的一个**陪集 (coset)**. 易知同余是一个等价关系，它将 V 进行划分，$[v]$ 是块. 若 V^* 是对同余关系而言的标准形式，则陪集全体可记作

$$V/S=\left\{v+S\,\middle|\,v\in V^*\right\}.$$

在 V/S 中定义加法为

$$(u+S)+(v+S)=(u+v)+S,$$

$\mathbb{F}$ 对 V/S 的数乘为

$$r(u+S)=ru+S,$$

则 V/S 成为一个向量空间，成为 V 模 S 的**商空间 (quotient space)**.

由以上这些定义，可以得到如下命题.

命题 A2.1.3 如果 S 是域 $\mathbb{F}$ 上 n 维向量空间 V 的集合，则以下叙述是等价的：

(1) S 是 V 的基；

(2) V 中的每一向量 v 可唯一地写成

$$v=r_1v_1+\cdots+r_nv_n,$$

这里 $v_i\in S$，$r_i\in\mathbb{F}$；

(3) S 是 V 中的极小生成元集； **(4)** S 是 V 中极大线性无关组.

命题 A2.1.4 若 S 与 T 为有限维线性空间 V 的两个子空间，则

$$\dim S+\dim T=\dim(S+T)+\dim(S\cap T). \tag{A2.1.1}$$

若 V 是域 $\mathbb{F}$ 上 n 维向量空间，$\mathcal{B}=\{b_1,\ldots,b_n\}$ 是 V 的一组基，则对每一个向量 $w\in V$，存在唯一的一组数 $(r_1,\ldots,r_n)$，使得 V 可以写成

$$w=r_1b_1+\cdots+r_nb_n=(b_1,\ldots,b_n)(r_1,\ldots,r_n)^{\mathrm{T}}.$$

故对基 $\mathcal{B}$ 来讲，w 可用列向量 $(r_1,\ldots,r_n)^{\mathrm{T}}$ 表示之，记作 $[w]_{\mathcal{B}}$，称为 w 在基 $\mathcal{B}$ 下的坐标. 如果 $\mathcal{C}=\{c_1,\ldots,c_n\}$ 也是 V 的一组基，则存在 n 阶可逆矩阵 $M_{\mathcal{B},\mathcal{C}}=(A_1,\ldots,$

$A_n)$,这里 A_i 表示一个 n 维列向量,使得

$$[w]_{\mathcal{C}} = M_{\mathcal{B},\mathcal{C}}[w]_{\mathcal{B}}.$$

取 $w = b_i$,则得到 $A_i = [b_i]_{\mathcal{C}}$,即 $M_{\mathcal{B},\mathcal{C}} = \big([b_1]_{\mathcal{C}}, \ldots, [b_n]_{\mathcal{C}}\big)$.

若 V 是域 $\mathbb{F}$ 上 n 维向量空间,$\mathcal{B}$ 是 V 的一组基,考虑映射

$$\varphi_{\mathcal{B}}: \begin{array}{ccc} V & \longrightarrow & \mathbb{F}^n, \\ v & \longmapsto & [v]_{\mathcal{B}}. \end{array}$$

易证: $\varphi_{\mathcal{B}}$ 是 V 到 $\mathbb{F}^n$ 的同构映射,即 $\varphi_{\mathcal{B}}$ 是双射且是一个线性变换. 因此 V 与 $\mathbb{F}^n$ 同构. 这就得到如下定理.

定理 A2.1.1 域 $\mathbb{F}$ 上 n 维向量空间 V 同构于 $\mathbb{F}^n$. 域 $\mathbb{F}$ 上两个向量空间同构当且仅当它们的维数相等.

这个定理告诉我们: 在同构意义下,n 维向量空间只有一个,即 $\mathbb{F}^n$.

A2.2 对偶空间

有了线性空间,即向量空间,首先要讨论的是其上最简单的一类函数——线性函数.

定义 A2.2.1 若 V 是域 $\mathbb{F}$ 上的向量空间,函数 $f: V \to V$ 称为 V 上的**线性函数 (linear function)** 或**线性泛函 (linear functional)**,如果对任意 $r, s \in \mathbb{F}$ 和 $u, v \in V$ 有

$$f(ru + sv) = rf(u) + sf(v).$$

V 上所有线性泛函的全体记为 V^*. 若 $f, g \in V^*$,定义加法为: 对任意 $v \in V$,

$$(f + g)(v) = f(v) + g(v);$$

$\mathbb{F}$ 对 V^* 的数乘为: 对任意 $r \in \mathbb{F}, v \in V$,

$$(rf)(v) = rf(v).$$

易见这样定义了加法与数乘之后,V^* 也是一个向量空间,这时称其为 V 的**对偶空间 (dual space)**.

设 V 是一个 n 维向量空间,$\mathcal{B} = \{v_1, \ldots, v_n\}$ 是 V 的一组基. 对每个 v_i,可以定义一个线性泛函 $v_i^* \in V^*$,使得

$$v_i^*(v_j) = \delta_{ij}, \tag{A2.2.1}$$

这里 δ_{ij} 是 Kronecker 函数. 易证 $\mathcal{B}^* = \{v_1^*, \ldots, v_n^*\}$ 是 V^* 的一组基,这时称 $\mathcal{B}^*$ 为 $\mathcal{B}$ 的**对偶基 (dual basis)**. 由此立即可以得到

$$\dim V = \dim V^*.$$

由于 V^* 也是向量空间,故 V^* 有对偶空间 $V^{**} = (V^*)^*$. 若 V 是有限维向量空

间，则

$$\dim V^{**} = \dim V^* = \dim V.$$

因此由定理 A2.1.1 可知：对有限维向量空间 V，有 $V^{**} \approx V$. 考虑映射

$$\tau_1\colon \begin{array}{ccc} V & \longrightarrow & V^*, \\ v = \sum\limits_{i=1}^n x_i v_i & \longmapsto & v^* = \sum\limits_{i=1}^n x_i v_i^*. \end{array}$$

易知 τ_1 是一个同构映射. 对任意 $u = \sum\limits_{i=1}^n y_j v_j \in V$，由 (A2.2.1)，

$$v^*(u) = \sum_{i=1}^n x_i v_i^*(u) = \sum_{i=1}^n x_i v_i^*\Big(\sum_{i=1}^n y_j v_j\Big) = \sum_{i=1}^n x_i y_i.$$

同样对每个 v_i^*，可以定义一个线性泛函 $v_i^{**} \in V^{**}$，使得

$$v_i^{**}(v_j^*) = \delta_{ij},$$

这里 δ_{ij} 也是 Kronecker 函数. 易证 $\mathcal{B}^{**} = \{v_1^{**}, \ldots, v_n^{**}\}$ 是 V^{**} 的一组基，它为 $\mathcal{B}^*$ 的对偶基，$V^* \approx V^{**}$. 考虑映射

$$\tau_2\colon \begin{array}{ccc} V^* & \longrightarrow & V^{**}, \\ v^* = \sum\limits_{i=1}^n x_i v_i^* & \longmapsto & v^{**} = \sum\limits_{i=1}^n x_i v_i^{**}. \end{array}$$

易知 τ_2 是一个同构映射. 对任意 $w = \sum\limits_{i=1}^n z_i v_i^* \in V^*$，由 (A2.2.1)，

$$w(v_j) = \sum_{i=1}^n z_i v_i^*(v_j) = z_j.$$

故 $w = \sum\limits_{i=1}^n w(v_i) v_i^*$. 于是

$$\begin{aligned} v^{**}(w) &= \sum_{i=1}^n x_i v_i^{**}(w) = \sum_{i=1}^n x_i v_i^{**}\Big(\sum_{i=1}^n w(v_i) v_i^*\Big) \\ &= \sum_{i=1}^n x_i w(v_i) = w\Big(\sum_{i=1}^n x_i v_i\Big) = w(v). \end{aligned}$$

现在令 $\tau = \tau_2\tau_1$，则易知 $\tau\colon V \to V^{**}$ 是一个同构映射，且

$$\tau(v) = \tau_2\big(\tau_1(v)\big) = \tau_2(v^*) = v^{**}$$

对每个 $v \in V$ 都成立. 已证 $v^{**}(w) = w(v)$ 对每个 $w \in V^*$ 都成立. 由此可见，v 在 V 到 V^{**} 的同构映射 τ 下的像不依赖于 V 中基的选取. 称这样的同构映射为**自然同构映射**. 在这样的自然同构映射下，可以把 v 与 $\tau(v) = v^{**}$ 等同，从而把 V 与 V^{**} 等同起来. 也就是可以把 V 看成 V^* 的对偶空间，这样 V 与 V^* 互为对偶空间. 这就是把 V^* 称为 V 的对偶空间的原因.

一个十分重要的线性泛函是零化子.

定义 A2.2.2 若 M 是向量空间 V 的非空集合，V^* 中的集合

$$M^\circ = \{f \in V^* \mid f(M) = 0\}$$

称为 M 的**零化子 (annihilator)**,这里 $f(M)=\{f(v)\mid v\in M\}$.

关于零化子有如下一些结论.

命题 A2.2.1 M° 是 V^* 的子空间,即使 M 不是 V 的子空间.

命题 A2.2.2 当 M 是 n 维向量空间的子空间时,有

$$\dim M+\dim M^\circ=n.$$

证明 若 $\mathcal{U}=\{u_1,\ldots,u_k\}$ 是 M 的一组基,将 $\mathcal{U}$ 扩充为

$$\mathcal{B}=\{u_1,\ldots,u_k,v_1,\ldots,v_{n-k}\},$$

使 $\mathcal{B}$ 成为 V 的一组基,则

$$\mathcal{B}^*=\{u_1^*,\ldots,u_k^*,v_1^*,\ldots,v_{n-k}^*\}$$

是 $\mathcal{B}$ 的对偶基. 往证 $\{v_1^*,\ldots,v_{n-k}^*\}$ 是 M° 的一组基.

显然它们是线性无关的,只需证它们张成 M°.

若 $f\in M^\circ$,则 $f\in V^*$,故 f 可写成

$$f=r_1u_1^*+\cdots+r_ku_k^*+s_1v_1^*+\cdots+s_{n-k}v_{n-k}^*,$$

这里 $r_i\in\mathbb{F},s_j\in\mathbb{F}$. 由于 $f\in M^\circ$,故 $f(u_i)=0$,但 $f(u_i)=r_i$,故 $r_i=0$. 因此

$$f=s_1v_1^*+\cdots+s_{n-k}v_{n-k}^*.$$

于是 $\{v_1^*,\ldots,v_{n-k}^*\}$ 张成 M°. ■

命题 A2.2.3 若 M,N 是向量空间 V 的子集,且 $M\subseteq N$,则

$$N^\circ\subseteq M^\circ.$$

命题 A2.2.4 若 V 是有限维向量空间,如视 V^{**} 与 V 等同,则对 V 的任一子集 M,都有

$$M^{\circ\circ}=\operatorname{span}M;$$

若 S 为 V 的子空间,则

$$S^{\circ\circ}=S.$$

命题 A2.2.5 若 S,T 是有限维向量空间的子空间,则

$$(S\cap T)^\circ=S^\circ+T^\circ,\qquad (S+T)^\circ=S^\circ\cap T^\circ.$$

命题 A2.2.6 若向量空间 V 是它的两个子空间 S 与 T 的直和,则

(1) $S^*\approx T^\circ$ 及 $T^*\approx S^\circ$; **(2)** $(S\oplus T)^*=S^\circ\oplus T^\circ$.

证明 **(1)** 若 $f\in T^\circ\subseteq V^*$,则 $f(T)=0$,定义映射

$$\tau: f\to f\big|_S,$$

即将 $f\in T^\circ$ 映为 f 在 S 上的限制,显然,$f\big|_S\in S^*$,故这确实是 T° 到 S^* 的映射,易知这是线性的.

若 $f\big|_S=0$,则 $f(S)=0$,而已知 $f(T)=0$,这导出 $f=0$,故映射 τ 是单射.

若 $g \in S^*$,定义 f 为

$$f(s+t) = g(s),$$

这里 $s \in S$,$t \in T$,显然 $f \in V^*$. 由于 $f(0+t) = g(0) = 0$ 对所有 $t \in T$ 都成立,故 $f \in T^\circ$. 而 $f\big|_S = g$,故对任意 $g \in S^*$,都有 $f \in T^\circ \subseteq V^*$,使得 $f\big|_S = g$,故 τ 是满射.

综上所述,$T^\circ \approx S^*$. 同理可证 $T^* \approx S^\circ$.

(2) 若 $f \in S^\circ \cap T^\circ$,则 $f(S) = 0$ 及 $f(T) = 0$,故 $f = 0$,即 $S^\circ \cap T^\circ = \{0\}$. 而 S°, T° 是 V^* 的子空间,故 $(S \oplus T)^* \supseteq S^\circ \oplus T^\circ$.

若 $f \in (S \oplus T)^*$,定义

$$g(s+t) = f(t), \qquad h(s+t) = f(s),$$

这里 $s \in S$,$t \in T$. 显然,$g, h \in (S \oplus T)^*$. 由于 $g(S) = 0$ 及 $h(T) = 0$,故 $g \in S^\circ$,$h \in T^\circ$,而

$$f(s+t) = f(s) + f(t) = g(s+t) + h(s+t) = (g+h)(s+t).$$

因此,$f = g + h \in S^\circ \oplus T^\circ$,于是 $(S \oplus T)^* \subseteq S^\circ \oplus T^\circ$. ■

A2.3 双线性形式

在上一节中,讨论了向量空间上最简单的一类函数——线性函数,即线性泛函,对有限维向量空间证明了它的对偶空间的对偶空间同构于它自己. 还定义与讨论了对偶空间中一类重要的子空间——零化子空间,这在今后十分有用.

1. 讨论了线性函数,顺理成章的是讨论向量空间上的双线性形式及二次型. 在这一节中,讨论的向量空间全是有限维的.

定义 A2.3.1 若 V 是域 $\mathbb{F}$ 上的向量空间,映射 $\langle\,,\rangle: V \to V$ 称为**双线性形式 (bilinear form)**,若其对每个坐标而言都是线性函数,即对任意 $\alpha, \beta \in \mathbb{F}$,$x, y, z \in V$ 有

$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$$

及

$$\langle z, \alpha x + \beta y\rangle = \alpha\langle z, x\rangle + \beta\langle z, y\rangle.$$

特别的,$\langle x, x\rangle$,$x \in V$ 称为 V 上的**二次型 (quadratic form)**.

如果对任意 $x, y \in V$,有

$$\langle x, y\rangle = \langle y, x\rangle,$$

则称 $\langle\,,\rangle$ 为**对称 (symmetric)** 的双线性形式.

如果对任意 $x, y \in V$,有

$$\langle x, y\rangle = -\langle y, x\rangle,$$

则称 $\langle\,,\rangle$ 为**斜对称 (skew-symmetric)** 的双线性形式.

命题 A2.3.1　设 $\mathbb{F}$ 的特征不为 2，$\langle\,,\rangle$ 是斜对称的当且仅当对任意的 $z\in V$，有 $\langle z,z\rangle=0$.

证明　若对任意的 $z\in V$ 有 $\langle z,z\rangle=0$，任取 $x,y\in V$，则

$$0=\langle x+y,x+y\rangle=\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle=\langle x,y\rangle+\langle y,x\rangle,$$

即 $\langle x,y\rangle=-\langle y,x\rangle$，故 $\langle\,,\rangle$ 斜对称. 这部分对任意特征都对.

若 $\langle\,,\rangle$ 斜对称，则对任意 $x\in V$，有 $\langle x,x\rangle=-\langle x,x\rangle$，即 $2\langle x,x\rangle=0$，由于 $\mathbb{F}$ 的特征不等于 2，从而 $\langle x,x\rangle=0$. ■

2. 在向量空间 V 上，如果定义了双线性形式 $\langle\,,\rangle$，则称 $(V,\langle\,,\rangle)$ 为**度量向量空间 (metric vector space)**，有时候也简写成 V. 而取定的双线性形式 $\langle\,,\rangle$ 称为度量空间 V 的度量. 一个度量向量空间称为**非奇异 (non-singular)**的，若对任意 $v\in V$，$\langle x,v\rangle=0$ 蕴涵 $x=0$. 若 $(V,\langle\,,\rangle)$ 是非奇异度量向量空间，且 $\langle\,,\rangle$ 是对称的，则称 $(V,\langle\,,\rangle)$ 为域 $\mathbb{F}$ 上的**对称度量向量空间**，这时也称 V 是 $\mathbb{F}$ 上的**正交几何 (orthogonal geometry)**. 若 $(V,\langle\,,\rangle)$ 是非奇异度量向量空间，且 $\langle\,,\rangle$ 是斜对称的，则称 $(V,\langle\,,\rangle)$ 为域 $\mathbb{F}$ 上的**斜对称向量空间**，这时也称 V 是域 $\mathbb{F}$ 上的**辛几何 (symplectic geometry)**.

我们只讨论正交几何与辛几何. 先来证明重要的秩与零度定理.

若 V,W 为两个向量空间，$\tau\in\mathcal{L}(V,W)$，则有 $\operatorname{Ker}\tau$ 与 $\operatorname{Im}\tau$. 称 $\dim\operatorname{Ker}\tau$ 为 τ 的**零度 (nullity)**，记作 $\operatorname{null}\tau$；称 $\dim\operatorname{Im}\tau$ 为 τ 的**秩 (rank)**，记作 $\operatorname{rank}\tau$.

定理 A2.3.1（秩与零度定理） 若 $\tau\in\mathcal{L}(V,W)$，则

$$\operatorname{rank}\tau+\operatorname{null}\tau=\dim V.$$

证明　由于 $\tau\in\mathcal{L}(V,W)$，故 $\operatorname{Ker}\tau$ 是 V 的一个子空间，于是有补空间 $(\operatorname{Ker}\tau)^c$，即

$$V=\operatorname{Ker}\tau\oplus(\operatorname{Ker}\tau)^c.$$

设 $\mathcal{K}$ 是 $\operatorname{Ker}\tau$ 的基，$\mathcal{C}$ 是 $(\operatorname{Ker}\tau)^c$ 的基. 由于 $\mathcal{K}\cap\mathcal{C}=\varnothing$ 及 $\mathcal{K}\cup\mathcal{C}$ 是 V 的基，故

$$\dim V=\dim\operatorname{Ker}\tau+\dim(\operatorname{Ker}\tau)^c.$$

将 τ 限制在 $(\operatorname{Ker}\tau)^c$ 上，记作 τ^c，则易证

$$\tau^c:(\operatorname{Ker}\tau)^c\to\operatorname{Im}\tau$$

是同构映射，因此定理成立. 事实上，若 $v\in(\operatorname{Ker}\tau)^c$，且 $\tau^c(v)=0$，由于 τ^c 是 τ 在 $(\operatorname{Ker}\tau)^c$ 上的限制，故 $\tau(v)=0$. 于是 $v\in\operatorname{Ker}\tau\cap(\operatorname{Ker}\tau)^c$，从而 $v=0$，这表明 τ^c 是单射. 若 $\tau(v)\in\operatorname{Im}\tau$，则 $v=u+w$，这里 $u\in\operatorname{Ker}\tau$，$w\in(\operatorname{Ker}\tau)^c$. 于是

$$\tau(v)=\tau(u)+\tau(w)=\tau(w)=\tau^c(w),$$

从而 $\tau(v)\in\operatorname{Im}\tau^c$，即 $\operatorname{Im}\tau\subseteq\operatorname{Im}\tau^c$，而 $\operatorname{Im}\tau^c\subseteq\operatorname{Im}\tau$ 是显然的，故 $\operatorname{Im}\tau=\operatorname{Im}\tau^c$. 因此 τ^c 是将 $(\operatorname{Ker}\tau)^c$ 映到 $\operatorname{Im}\tau$ 上的满射. 而 τ^c 显然是线性的，从而

$$(\operatorname{Ker}\tau)^c \approx \operatorname{Im}\tau. \qquad \blacksquare$$

有定理 A2.3.1 可以得出一些了重要推论.

推论 A2.3.1　若 $\tau \in \mathcal{L}(V, W)$,且 $\dim V = \dim W < \infty$,那么 τ 为单射当且仅当 τ 为满射.

推论 A2.3.2(第一同构定理) 若 $\tau \in \mathcal{L}(V, W)$, $V/\operatorname{Ker}\tau$ 是 V 模 $\operatorname{Ker}\tau$ 的商空间,则

$$V/\operatorname{Ker}\tau \approx \operatorname{Im}\tau.$$

证明　定义映射 $\tau': V/\operatorname{Ker}\tau \to W$ 为

$$\tau'(v + \operatorname{Ker}\tau) = \tau(v).$$

这样定义的 τ' 是有意义的. 事实上,设 $u, v \in V$,

$$\begin{aligned} & v + \operatorname{Ker}\tau = u + \operatorname{Ker}\tau \implies \tau'(v + \operatorname{Ker}\tau) = \tau'(u + \operatorname{Ker}\tau);\\ \iff & v + \operatorname{Ker}\tau = u + \operatorname{Ker}\tau \implies \tau(v) = \tau(u);\\ \iff & u - v \in \operatorname{Ker}\tau \implies \tau(u - v) = 0; \end{aligned}$$

因此上述定义的 τ' 是有意义的,且 τ' 是单射.

显然 τ' 是一个线性变换,由定理 A2.3.1 及 τ' 是单射可知,

$$\dim \operatorname{Im}\tau' = \dim(V/\operatorname{Ker}\tau),$$

但

$$\operatorname{Im}\tau' = \left\{\tau'(v + \operatorname{Ker}\tau) \,\middle|\, v + \operatorname{Ker}\tau \in V/\operatorname{Ker}\tau\right\} = \left\{\tau(v) \,\middle|\, v \in V\right\} = \operatorname{Im}\tau,$$

故 τ' 为 $V/\operatorname{Ker}\tau$ 到 $\operatorname{Im}\tau$ 的满射.

综上所述, $V/\operatorname{Ker}\tau \approx \operatorname{Im}\tau$. ■

推论 A2.3.3　设 S 是向量空间 V 的一个子空间, S^c 是 S 的补,则

$$V/S \approx S^c,$$

且

$$\dim S + \dim S^c = \dim V.$$

证明　V 中任一向量都可以唯一地写成 $v = s + s^c$,这里 $s \in S, s \in S^c$. 现在定义线性算子 $\rho: V \to V$ 为

$$\rho(s + s^c) = s^c,$$

这样定义的 ρ 是有意义的,显然

$$\operatorname{Im}\rho = S^c, \quad \operatorname{Ker}\rho = \left\{s + s^c \in V \,\middle|\, s^c = 0\right\} = S,$$

故由第一同构定理,得 $V/S \approx S^c$. 由定理 A2.3.1,得

$$\dim S + \dim S^c = \dim V. \qquad \blacksquare$$

推论 A2.3.4(第二同构定理) 若 V 是一个向量空间, S, T 为 V 的子空间,则

$$(S + T)/T \approx S/(S \cap T).$$

推论 A2.3.5（第三同构定理）若 V 是一个向量空间，$S\subseteq T\subseteq V$ 均为 V 的子空间，则

$$(V/S)\big/(T/S)\approx V/T.$$

推论 A2.3.4 与推论 A2.3.5 的证明从略.

在非奇异的度量空间上，上一节讨论的线性泛函，都可以用双线性形式来表示之.

定理 A2.3.2（Riesz 表示定理）若 $(V,\langle\,,\rangle)$ 是有限维非奇异的度量向量空间，任取 $f\in V^*$，则一定存在唯一的向量 $x\in V$ 使得

$$f(v)=\langle v,x\rangle$$

对所有的 $v\in V$ 都成立.

证明　设 $x\in V$，定义映射 $\varphi_x\colon V\to\mathbb{F}$ 为

$$\varphi_x(v)=\langle v,x\rangle.$$

易证 $\varphi_x\in V^*$，故可定义函数 $\tau\colon V\to V^*$ 为

$$\tau(x)=\varphi_x.$$

显然，这是线性的. 由于 V 是非奇异的，故其核

$$\left\{x\in V\,\middle|\,\varphi_x=0\right\}=\left\{x\in V\,\middle|\,\text{对所有的 } v\in V\text{，有 }\langle v,x\rangle=0\right\}$$

是 V 的只含有零向量的子集，故 τ 是单射.

τ 可以在整个 V 上定义，且为单射，而易知 $\dim V=\dim V^*$，故由推论 A2.3.1，τ 是 V 上的满射. 因此，τ 是一个同构映射，将 V 映到 V^*，即 V 的任一线性泛函都是形如 φ_x，这里 $x\in V$. ■

Riesz 表示定理告诉我们，**在有限维非奇异的度量向量空间，其上的线性泛函只有一类，那就是定义度量向量空间的双线性形式.**

3. 若 $(V,\langle\,,\rangle)$ 是 n 维度量向量空间，$\mathcal{B}=\{b_1,\ldots,b_n\}$ 是 V 的一组基，于是 $\langle\,,\rangle$ 完全可以由 $n\times n$ 矩阵

$$M_{\mathcal{B}}=(a_{ij})=(\langle b_i,b_j\rangle)$$

来决定，$M_{\mathcal{B}}$ 称为双线性形式 $\langle\,,\rangle$ 在基 $\mathcal{B}$ 下的矩阵.

若 $x,y\in V$，且

$$x=\sum_{i=1}^n x_ib_i,\quad y=\sum_{i=1}^n y_ib_i,$$

则

$$\langle x,y\rangle=\sum_{i=1}^n\sum_{j=1}^n x_iy_j\langle b_i,b_j\rangle=[x]_{\mathcal{B}}^{\mathrm{T}}M_{\mathcal{B}}[y]_{\mathcal{B}},$$

这里 $[x]_{\mathcal{B}},[y]_{\mathcal{B}}$ 表示 x,y 在基 $\mathcal{B}$ 下的坐标，即

$$[x]_{\mathcal{B}}=(x_1,\ldots,x_n)^{\mathrm{T}},\quad [y]_{\mathcal{B}}=(y_1,\ldots,y_n)^{\mathrm{T}}.$$

$\langle\,,\rangle$ 是对称的当且仅当 $M_{\mathcal{B}}$ 是对称矩阵，即 $a_{ij}=a_{ji}$；$\langle\,,\rangle$ 是斜对称的当且仅当

$M_{\mathcal{B}}$ 是斜对称矩阵,即 $a_{ij} = -a_{ji}$.

若 $\mathcal{C} = (c_1, \ldots, c_n)$ 是 V 的另一组基,则由 A2.1 的最后可知,对任意 $v \in V$,

$$[v]_{\mathcal{C}} = M_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} \quad \text{及} \quad [v]_{\mathcal{B}} = M_{\mathcal{C},\mathcal{B}}[v]_{\mathcal{C}}.$$

于是

$$\langle x, y\rangle = [x]_{\mathcal{B}}^{\mathrm{T}} M_{\mathcal{B}} [y]_{\mathcal{B}} = [x]_{\mathcal{C}}^{\mathrm{T}} M_{\mathcal{C},\mathcal{B}}^{\mathrm{T}} M_{\mathcal{B}} M_{\mathcal{C},\mathcal{B}} [y]_{\mathcal{C}} = [x]_{\mathcal{C}}^{\mathrm{T}} M_{\mathcal{C}} [y]_{\mathcal{C}}.$$

这就得到

$$M_{\mathcal{C}} = M_{\mathcal{C},\mathcal{B}}^{\mathrm{T}} M_{\mathcal{B}} M_{\mathcal{C},\mathcal{B}},$$

也就是说,矩阵 $M_{\mathcal{C}}$ 与 $M_{\mathcal{B}}$ 是相合的.

4. 要弄清楚对称的、斜对称的双线性形式一共有多少,也就是在相合的意义下双线性形式的矩阵有多少标准形式,这是线性代数最基本问题之一. 为此要引入正交的概念.

向量 x 与向量 y 称为**正交的 (orthogonal)**,记作 $x \perp y$,若 $\langle x, y\rangle = 0$. 对于对称双线性形式及斜对称双线性形式,显然有 $x \perp y$ 当且仅当 $y \perp x$. 称 S, T 是度量向量空间 $(V, \langle\,,\rangle)$ 的两个子空间,称它们是正交的,记作 $S \perp T$,若对所有 $s \in S, t \in T$,都有 $\langle s, t\rangle = 0$.

集合 $\{v \in V \mid v \perp s\}$ 称为子空间 S 的**正交补 (orthogonal complement)**,记作 $S^{\perp}$. 若 $(V, \langle\,,\rangle)$ 是度量向量空间,S, T 是它的子空间,并且 $V = S \oplus T$ 及 $S \perp T$,则称 V 是 S 与 T 的**正交直和 (orthogonal direct sum)**,记作 $V = S \operatorname{\textcircled{\scriptsize$\perp$}} T$.

定理 A2.3.3　若 $(V, \langle\,,\rangle)$ 是非奇异的度量向量空间,S 是 V 的子空间,则 $V = S \operatorname{\textcircled{\scriptsize$\perp$}} S^{\perp}$ 当且仅当 S 非奇异.

现在来证明定理 A2.3.3,为此先证明如下引理.

引理 A2.3.1　若 S 是非奇异的度量空间 V 的一个子空间,则

$$\dim S + \dim S^{\perp} = \dim V. \tag{A2.3.1}$$

证明　对每个 $v \in V$,在 S 上定义线性泛函 $\varphi_v: S \to \mathbb{F}$ 为

$$\varphi_v(u) = \langle u, v\rangle,$$

这里 $u \in V$. 显然 $\varphi_v \in S^*$. 定义映射 $\tau: V \to S^*$ 为

$$\tau(v) = \varphi_v,$$

显然这是一个线性映射,且

$$\operatorname{Ker} \tau = \{v \in V \mid \varphi_v = 0\} = \{v \in V \mid \text{对所有的 } u \in S, \langle u, v\rangle = 0\} = S^{\perp}. \tag{A2.3.2}$$

此外,由定理 A2.3.2,S^* 中任一线性泛函均可用 S 上的双线性函数表示之,故 $\tau|_S: S \to S^*$ 是满射,从而 $\operatorname{Im} \tau = S^*$. 由定理 A2.3.1 可知

$$\dim \operatorname{Ker} \tau + \dim \operatorname{Im} \tau = \dim V.$$

而 $\dim \operatorname{Im} \tau = \dim S^* = \dim S$,由式 (A2.3.2),$\operatorname{Ker} \tau = S^{\perp}$,这就完成了证明. ∎

定理 A2.3.3 的证明 由 (A2.1.1) 及 (A2.3.1) 可知,

$$\dim(S+S^{\perp})=\dim S+\dim S^{\perp}-\dim(S\cap S^{\perp})=\dim V-\dim(S\cap S^{\perp}).$$

若 S 是非奇异的,则 $S\cap S^{\perp}=0$,因此,$V=S\oplus S^{\perp}$,这就证明了 $V=S\perp S^{\perp}$.

反之,易证若 S 不是非奇异的,则 $V=S\perp S^{\perp}$ 不成立, ■

5. 有了这些准备,就可以讨论正交几何与辛几何的正交分解,也就是要确定正交几何与辛几何的标准形式.

先来讨论辛几何. 若 $(V,\langle\,,\rangle)$ 为辛几何,由于 $\langle\,,\rangle$ 是斜对称的,故对每个 $x\in V$ 都有 $\langle x,x\rangle=0$,取定一个 $u(\neq 0)\in V$,由于 V 是非奇异的,故一定存在一个 v 使得 $\langle u,v\rangle\neq 0$. 考虑以 (u,v) 为一组基的二维子空间 H,则

$$\langle u,u\rangle=\langle v,v\rangle=0.$$

而 $\langle u,v\rangle=a\neq 0$,以 $a^{-1}v$ 来代替 v,就有

$$\langle u,v\rangle=1,\quad \langle v,u\rangle=-1.$$

于是在 H 的基 (u,v) 下 $\langle u,v\rangle$ 的矩阵为

$$M=\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

由于 V 非奇异,故由定理 A2.3.3,可将 V 进行正交分解: $V=H\perp H^{\perp}$,而 $H^{\perp}$ 仍是非奇异的斜对称度量向量空间,仍可将 $H^{\perp}$ 进行这样的正交分解. 重复这样的步骤,由于 V 是有限维的,故 V 最终可正交分解为

$$V=H_1\perp H_2\perp\cdots\perp H_k.$$

归纳起来为如下结论.

定理 A2.3.4 若 $(V,\langle\,,\rangle)$ 为非奇异斜对称度量向量空间,则 V 可正交分解为

$$V=H_1\perp H_2\perp\cdots\perp H_k,$$

这里 H_i 都为二维斜对称度量子空间,其 $\langle\,,\rangle$ 对应的矩阵为

$$\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

也就是说,在 V 中取到一组基,使得 $\langle\,,\rangle$ 的矩阵为

$$M=\operatorname{diag}\left(\begin{pmatrix}0&1\\-1&0\end{pmatrix},\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right).$$

因此,非奇异的斜对称度量向量空间都是偶数维的.

用矩阵的语言表示为:

若 P 是一个 n 阶非奇异的斜对称矩阵,则 P 相合于 M,即存在 n 阶非奇异矩阵 Q,使得

$$P = Q^{\mathrm{T}}\operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)Q,$$

非奇异斜对称矩阵一定是偶数阶,即 n 是偶数.

6. 再来对正交几何进行正交分解.

若 $(V, \langle\,,\rangle)$ 是一个非奇异的对称度量向量空间,则存在 $u(\neq 0) \in V$ 使得 $\langle u, u\rangle \neq 0$. 这样的 u 一定存在,否则 $\langle\,,\rangle$ 是斜对称的. 由 u 生成的子空间 $S = \operatorname{span}\{u\}$ 是非奇异的. 由于 V 是非奇异的,由定理 A2.3.3,有正交分解 $V = S \odot S^{\perp}$,而 $S^{\perp}$ 仍为非奇异的对称度量向量空间,仍可以对 $S^{\perp}$ 进行这样的正交分解

$$V = S \odot T \odot T^{\perp},$$

这里 S, T 均为一维子空间,重复这样的步骤,可得

$$V = S_1 \odot S_2 \odot \cdots \odot S_n,$$

这里 S_i 是由向量 u_i 生成,且 $\langle u_i, u_i\rangle \neq 0$,故 $(u_1, \ldots, u_n)$ 是 V 的一组正交基(即基中向量相互正交). 若 $\langle u_i, u_i\rangle = a_i$,则有如下结论.

若 $(V, \langle\,,\rangle)$ 是 n 维非奇异的对称度量向量空间,则 V 有一组正交基 $\mathcal{B} = (u_1, \ldots, u_n)$,使得在基 $\mathcal{B}$ 下,$\langle\,,\rangle$ 所对应的矩阵为

$$M_{\mathcal{B}} = \operatorname{diag}(a_1, \ldots, a_n).$$

若取 $r_i(\neq 0) \in \mathbb{F}$,则

$$\mathcal{C} = (r_1 u_1, \ldots, r_n u_n)$$

也是 V 上的一组正交基,$\langle\,,\rangle$ 在这组基下的矩阵为

$$M_{\mathcal{C}} = \operatorname{diag}(r_1^2 a_1, \ldots, r_n^2 a_n).$$

若 $\mathbb{F}$ 为**代数封闭域 (algebraically closed field)**,即 $\mathbb{F}[x]$ 中任一多项式均可在 $\mathbb{F}$ 上分裂为一次因子的乘积,这时,可取 $r_i = 1/\sqrt{a_i}$,这里 $\sqrt{a_i}$ 为方程 $x^2 - a_i = 0$ 的根,这样

$$M_{\mathcal{C}} = \operatorname{diag}(1, \ldots, 1) = I_n,$$

这里 I_n 为 n 阶单位矩阵.

以上所述归纳起来就有如下定理.

定理 A2.3.5　若 $(V, \langle\,,\rangle)$ 为域 $\mathbb{F}$ 上的 n 维非奇异对称度量向量空间,则 V 有一组正交基 $\mathcal{U} = (u_1, \ldots, u_n)$,即 V 可正交分解为

$$V = S_1 \odot S_2 \odot \cdots \odot S_n,$$

这里 S_i 是由 u_i 生成; 若 $\langle u_i, u_i\rangle = a_i$,则 $\langle\,,\rangle$ 相对于基 $\mathcal{U}$ 有矩阵

$$M_{\mathcal{U}} = \operatorname{diag}(a_1, \ldots, a_n).$$

若 $\mathbb{F}$ 为代数封闭域,则 V 有一组**正规正交基 (orthonormal basis)** (即若基为 $\mathcal{C} = (c_1, \ldots, c_n)$,则 $\langle c_i, c_j\rangle = \delta_{ij}$),使得 $\langle\,,\rangle$ 相对于基 $\mathcal{C}$ 有矩阵

$$M_{\mathcal{C}} = \operatorname{diag}\big(1,\ldots,1\big) = I_n.$$

用矩阵的语言表述为:

若 P 是 n 阶非奇异的对称矩阵,则 P 相合于对角矩阵 $\operatorname{diag}\big(a_1,\ldots,a_n\big)$,这里 $a_i \neq 0$. 即存在 n 阶非奇异矩阵 Q,使得

$$P = Q^{\mathrm{T}} \operatorname{diag}\big(a_1,\ldots,a_n\big)Q.$$

若 $\mathbb{F}$ 为代数封闭域,则 P 相合于 I_n,即 P 可以写成

$$P = Q^{\mathrm{T}}Q,$$

这里 Q 为 n 阶非奇异矩阵.

若 $\mathbb{F}$ 为实数域 $\mathbb{R}$,$\mathbb{R}$ 不是代数封闭域,但可取 $r_i = 1/\sqrt{|a_i|}$,于是 $M_{\mathcal{C}}$ 成为

$$M_{\mathcal{C}} = \operatorname{diag}\big(\underbrace{1,\ldots,1}_{k},\underbrace{-1,\ldots,-1}_{n-k}\big),$$

即在主对角线上的元素,一部分为 1,另一部分为 -1,也就是 V 有一组正交基 $(u_1,\ldots,u_k,v_1,\ldots,v_{n-k})$,而 $\langle u_i,u_i\rangle = 1$,$\langle v_j,v_j\rangle = -1$.

要证明 k 由 $\langle\,,\rangle$ 唯一确定,而与 V 的基的选取无关. 事实上,记

$$\mathcal{P} = \operatorname{span}\big\{u_1,\ldots,u_k\big\},\quad \mathcal{V} = \operatorname{span}\big\{v_1,\ldots,v_{n-k}\big\}.$$

若 $v = \sum\limits_{i=1}^{k} r_i u_i \in \mathcal{P}$,则

$$\begin{aligned}\langle v,v\rangle &= \Big\langle \sum_{i=1}^{k} r_i u_i, \sum_{j=1}^{k} r_j u_j\Big\rangle = \sum_{i=1}^{k}\sum_{j=1}^{k} r_i r_j \langle u_i,u_j\rangle \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} r_i r_j \delta_{ij} = \sum_{i=1}^{k} r_i^2 \geqslant 0.\end{aligned}$$

同样可证:若 $v \in \mathcal{V}$,则 $\langle v,v\rangle \leqslant 0$.

如果 V 有另一组正交基 $\mathcal{C} = (\overline{u}_1,\ldots,\overline{u}_\ell,\overline{v}_1,\ldots,v_{n-\ell})$,这里 $\langle\overline{u}_i,\overline{u}_i\rangle = 1$,$\langle\overline{v}_j,\overline{v}_j\rangle = -1$,记

$$\overline{\mathcal{P}} = \operatorname{span}\big\{\overline{u}_1,\ldots,\overline{u}_\ell\big\},\quad \overline{\mathcal{V}} = \operatorname{span}\big\{\overline{v}_1,\ldots,\overline{v}_{n-\ell}\big\},$$

则 $\mathcal{P}\cap\overline{\mathcal{V}} = \{0\}$. 这是因为,若 $v \in \mathcal{P}\cap\overline{\mathcal{V}}$,则由于 $v \in \mathcal{P}$,则 $\langle v,v\rangle \geqslant 0$; 由于 $v \in \overline{\mathcal{V}}$,则 $\langle v,v\rangle \leqslant 0$,因此,$\langle v,v\rangle = 0$,故 $v = 0$. 由于 $\mathcal{P}$,$\overline{\mathcal{V}}$ 均为 V 的子空间,且交为 $\{0\}$,故由 (A2.1.1),

$$\dim\mathcal{P} + \dim\overline{\mathcal{V}} \leqslant \dim V,$$

即 $k + (n-\ell) \leqslant n$,即 $k \leqslant \ell$. 同样可证 $\ell \leqslant k$,故 $k = \ell$.

以上所述归纳起来有如下定理.

定理 A2.3.6(Sylvester 惯性定理) 若 $(V,\langle\,,\rangle)$ 为 n 维实数域 $\mathbb{R}$ 上的非奇异对称度量向量空间,则 V 有正交基 $\mathcal{B} = (u_1,\ldots,u_k,v_1,\ldots,v_{n-k})$,使得 $\langle u_i,u_i\rangle = 1$,$\langle v_j,v_j\rangle = -1$,在基 $\mathcal{B}$ 下,$\langle\,,\rangle$ 的矩阵为

$$M_{\mathcal{B}} = \operatorname{diag}\big(I_k,-I_{n-k}\big),$$

这里 k 是由 $\langle\,,\rangle$ 唯一确定的,而与 V 的基的选取无关.

用矩阵的语言表述为:

若 P 为实数域上 n 阶非奇异对称矩阵,则 P 相合于 $\operatorname{diag}\big(I_k,-I_{n-k}\big)$,这里 k 是由 P 唯一确定的,即存在非奇异的 n 阶矩阵 Q,使得

$$P=Q^{\mathrm{T}}\operatorname{diag}\big(I_k,-I_{n-k}\big)Q.$$

如果用双线性形式的语言来说,第 5、6 条可以总结为如下的结论.

设 $(V,\langle\,,\rangle)$ 为域 $\mathbb{F}$ 上的 n 维非奇异度量空间.

(1) 若 $\langle\,,\rangle$ 为斜对称,则存在 V 的一组基 $\mathcal{B}$,使得对任意 $x,y\in V$,有

$$\langle x,y\rangle=\xi_1\eta_2-\xi_2\eta_1+\cdots+\xi_{n-1}\eta_n+\xi_n\eta_{n-1},$$

这里 $(\xi_1,\ldots,\xi_n)^{\mathrm{T}}$ 与 $(\eta_1,\ldots,\eta_n)^{\mathrm{T}}$ 分别为 x,y 在基 $\mathcal{B}$ 下的坐标.

(2) 若 $\langle\,,\rangle$ 为对称,则存在 V 的一组基 $\mathcal{B}=(v_1,\ldots,v_n)$,使得对任意 $x,y\in V$,有

$$\langle x,y\rangle=a_1\xi_1\eta_1+\cdots+a_n\xi_n\eta_n,$$

这里 $(\xi_1,\ldots,\xi_n)^{\mathrm{T}}$ 与 $(\eta_1,\ldots,\eta_n)^{\mathrm{T}}$ 分别为 x,y 在基 $\mathcal{B}$ 下的坐标,而 $a_i=\langle v_i,v_i\rangle$.

(3) 若域 $\mathbb{F}$ 为代数封闭域,则存在 V 的一组正规正交基 $\mathcal{B}=(v_1,\ldots,v_n)$,使得对任意 $x,y\in V$,有

$$\langle x,y\rangle=\xi_1\eta_1+\cdots+\xi_n\eta_n,$$

这里 $(\xi_1,\ldots,\xi_n)^{\mathrm{T}}$ 与 $(\eta_1,\ldots,\eta_n)^{\mathrm{T}}$ 分别为 x,y 在基 $\mathcal{B}$ 下的坐标.

(4) 若域 $\mathbb{F}$ 为实数域,则存在 V 的一组正交基 $\mathcal{B}=(u_1,\ldots,u_k,v_1,\ldots,v_{n-k})$,使得对任意 $x,y\in V$,有

$$\langle x,y\rangle=\xi_1\eta_1+\cdots+\xi_k\eta_k-\xi_{k+1}\eta_{k+1}-\cdots-\xi_n\eta_n,$$

这里 $(\xi_1,\ldots,\xi_n)^{\mathrm{T}}$ 与 $(\eta_1,\ldots,\eta_n)^{\mathrm{T}}$ 分别为 x,y 在基 $\mathcal{B}$ 下的坐标,而 k只与 $\langle\,,\rangle$ 有关,而与基的选取无关.

特别对二次型 $\langle x,x\rangle$ 可表为

$$\langle x,x\rangle=a_1\xi_1^2+\cdots+a_n\xi_n^2;$$

当 $\mathbb{F}$ 为代数封闭域时,

$$\langle x,x\rangle=\xi_1^2+\cdots+\xi_n^2;$$

当 $\mathbb{F}$ 为实数域 $\mathbb{R}$ 时,

$$\langle x,x\rangle=\xi_2+\cdots+\xi_k^2-\xi_{k+1}^2-\cdots-\xi_n^2.$$

A2.4 内积空间

当域 $\mathbb{F}$ 为实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$ 时,还有重要的内积空间.

定义 A2.4.1 若 V 是域 $\mathbb{F}$ 上的向量空间,这里 $\mathbb{F}$ 是实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$,若存在映射 $\langle\,,\rangle: V\times V\to\mathbb{F}$ 满足

(1) 正定性 (positive definiteness) 对所有 $v\in V$ 有

$$\langle v,v\rangle\geqslant 0.$$

而 $\langle v,v\rangle=0$,当且仅当 $v=0$;

(2) 共轭对称(Hermite 对称)性或对称性 当 $\mathbb{F}=\mathbb{C}$ 时,有

$$\langle v,v\rangle=\overline{\langle v,v\rangle}.$$

当 $\mathbb{F}=\mathbb{R}$ 时,有

$$\langle v,v\rangle=\langle v,v\rangle;$$

(3) 第一坐标是线性 对所有的 $u,v,w\in V$ 及 $r,s\in\mathbb{F}$,有

$$\langle ru+sv,w\rangle=r\langle u,w\rangle+s\langle v,w\rangle,$$

则称 $\langle\,,\rangle$ 为 V 上的**内积 (inner product)**,有内积的向量空间称为**内积空间 (inner product space)**.

当 $\mathbb{F}$ 为 $\mathbb{R}$ 时,称内积空间为**实欧几里得空间 (Euclidean space)**,显然这是一个正定的对称非奇异的度量空间.

当 $\mathbb{F}$ 为 $\mathbb{C}$ 时,称内积空间为**复欧几里得空间**,也称**酉空间 (unitary space)**. 此时,由定义 A2.4.1 之条件 (2) 和 (3) 可得: 对所有 $u,v,w\in V$ 及 $r,s\in\mathbb{F}$,有

$$\langle w,ru+sv\rangle=\bar r\langle w,u\rangle+\bar s\langle w,v\rangle,$$

称为**共轭线性 (conjugate linearity)**. 因此,此时 $\langle\,,\rangle$ 不是双线性形式,故 $(V,\langle\,,\rangle)$ 不是度量向量空间.

若 $v\in V$,称

$$\|v\|=\sqrt{\langle v,v\rangle}$$

为向量 v 的**长度 (length)** 或**范数 (norm)**.

若 $u,v\in V$,称

$$\|u-v\|$$

为向量 u,v 之间的**距离 (distance)**,记作 $d(u,v)$. 有了距离的概念,就可以在 V 上定义向量序列的收敛,集合的并、开、闭包、邻域、完备性以及连续等概念. 还可以有:

(1) Cauchy 不等式 对所有 $u,v\in V$,有

$$|\langle u,v\rangle|\leqslant\|u\|\|v\|;$$

(2) 三角不等式 对所有 $u, v \in V$,有

$$\|u+v\| \leqslant \|u\| + \|v\|;$$

(3) 平行四边形法则 对所有 $u, v \in V$,有

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2;$$

(4) 距离的三角不等式 对所有 $u, v, w \in V$,有

$$d(u,v) \leqslant d(u,w) + d(w,v),$$

等等. 还可以由范数直接定义范数线性空间.

若 V 是一个向量空间,且在 V 上有函数 $\|\cdot\|: V \to \mathbb{R}$ 满足:

(1) 若 $v \in V$,则 $\|v\| \geqslant 0$,且 $\|v\| = 0$ 当且仅当 $v = 0$;

(2) 对所有 $r \in \mathbb{F}, v \in \mathbb{F}$,有 $\|rv\| = |r|\|v\|$;

(3) 对所有 $u, v \in V$,有

$$\|u+v\| \leqslant \|u\| + \|v\|,$$

则称 $\|\cdot\|$ 为 V 上的一个范数,这时 $(V, \|\cdot\|)$ 称为**范数线性空间 (normed linear space)**. 这是内积空间的一个推广.

对内积空间,也可以像上一节中那样来定义正交的概念,只是用内积来替代双线性形式,于是可以有正交补、正交基及 Riesz 表示定理等. 这里只叙述 Riesz 表示定理:

若 $(V, \langle\,,\rangle)$ 是一个有限维内积空间,$f \in V^*$,则存在唯一的向量 $x \in V$,使得对任意的 $v \in V$,有 $f(v) = \langle v, x\rangle$.

对于内积空间将在 A3.3 和 A5.5 中进一步讨论之.

第三讲

线性变换

如前所述,线性代数是研究线性空间(向量空间)、模和其上线性变换以及与之相关的问题(如线性、双线性、二次函数等)的数学学科. 在上一讲中讨论了向量空间以及其上的线性泛函及对偶空间; 其上的双线性形式、二次型及度量向量空间,正交几何与辛几何的分类; 还有大家十分熟悉的内积空间. 在这一讲中将讨论向量空间上的线性变换以及与之相关的共轭算子及伴随算子.

A3.1 线性变换的矩阵表示

设 V, W 分别是域 $\mathbb{F}$ 上的 n 维与 m 维向量空间. 在这一节中要证明每个 $\tau \in \mathcal{L}(V, W)$ 都与 $\mathcal{M}_{m,n}(\mathbb{F})$ 中的一个矩阵相对应,这就是 τ 在 $\mathcal{M}_{m,n}(\mathbb{F})$ 中的矩阵表示. 不但如此,还要证明: $\mathcal{L}(V, W)$ 与 $\mathcal{M}_{m,n}(\mathbb{F})$ 是同构的. 故对 $\mathcal{L}(V, W)$ 的讨论就是对 $\mathcal{M}_{m,n}(V, W)$ 的讨论.

设 $\tau \in \mathcal{L}(V, W)$, $\mathcal{B} = (b_1, \ldots, b_n)$ 与 $\mathcal{C} = (c_1, \ldots, c_m)$ 分别是 V 与 W 的基,则向量 $v \in V$ 在基 $\mathcal{B}$ 下的坐标为 $[v]_{\mathcal{B}} = (r_1, \ldots, r_n)^{\mathrm{T}}$, $\tau(v)$ 在基 $\mathcal{C}$ 下的坐标为 $[\tau(v)]_{\mathcal{C}} = (s_1, \ldots, s_m)^{\mathrm{T}}$. 于是对应于 τ, 在 $\mathbb{F}^n$ 与 $\mathbb{F}^m$ 之间有一线性变换 $\tau_A\colon [v]_{\mathcal{B}} \to [\tau(v)]_{\mathcal{C}}$, $\tau_A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$,即对应于 τ,有一 $m \times n$ 矩阵 A,使得

$$[\tau(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}}.$$

现在来确定 A,记 $A = (A_1, \ldots, A_n)$,这里 A_i 为列向量. 取 $v = b_i$,则立得 $A_i = [\tau(b_i)]_{\mathcal{C}}$,故

$$A = \big([\tau(b_1)]_{\mathcal{C}}, \ldots, [\tau(b_n)]_{\mathcal{C}}\big).$$

记 $A = [\tau]_{\mathcal{B},\mathcal{C}}$,于是

$$[\tau(v)]_{\mathcal{C}} = [\tau]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}.$$

$[\tau]_{\mathcal{B},\mathcal{C}}$ 即为当 V 的基为 $\mathcal{B}$, W 的基为 $\mathcal{C}$ 时, τ 的矩阵表示. 于是可以定义映射

$$\varphi\colon \begin{array}{ccc} \mathcal{L}(V, W) & \longrightarrow & \mathcal{M}_{m,n}(\mathbb{F}), \\ \tau & \longmapsto & [\tau]_{\mathcal{B},\mathcal{C}}. \end{array}$$

现在来证明这是一个同构映射.

先证 φ 为线性映射. 事实上若 $s, t \in \mathbb{F}$, $\sigma, \tau \in \mathcal{L}(V, W)$,则

$$[(s\sigma + t\tau)(b_i)]_{\mathcal{C}} = [s\sigma(b_i) + t\tau(b_i)]_{\mathcal{C}} = s[\sigma(b_i)]_{\mathcal{C}} + t[\tau(b_i)]_{\mathcal{C}}.$$

故

$$\varphi(s\sigma + t\tau) = [s\sigma + t\tau]_{\mathcal{B},\mathcal{C}} = s[\sigma]_{\mathcal{B},\mathcal{C}} + t[\tau]_{\mathcal{B},\mathcal{C}} = s\varphi(\sigma) + t\varphi(\tau),$$

即 φ 为线性映射.

再证 φ 为满射. 事实上若 A 为一个 $m\times n$ 矩阵,且可写成 $(A_1,\dots,A_n)$,这里 A_i 为列向量,定义 $\tau\colon V\to W$,使得 $[\tau(b_i)]_{\mathcal{C}}=A_i$,这是可以做到的,故 φ 为满射.

再证 φ 为单射. 由于 $[\tau]_{\mathcal{B},\mathcal{C}}=0$ 蕴涵 $[\tau(b_i)]_{\mathcal{C}}=0$,又蕴涵 $\tau(b_i)=0$,故 $\tau=0$.

综上所述 φ 为同构映射,于是有如下定理.

定理 A3.1.1 $\qquad \mathcal{L}(V,W)\approx\mathcal{M}_{m,n}(\mathbb{F}).$

由 τ 的矩阵表示还可以导出: 若 $\sigma\colon U\to V$, $\tau\colon V\to W$, 且 $\mathcal{B},\mathcal{C}$ 与 $\mathcal{D}$ 分别为 U,V,W 的基,则

$$[\tau\sigma]_{\mathcal{B},\mathcal{D}}=[\tau]_{\mathcal{C},\mathcal{D}}[\sigma]_{\mathcal{B},\mathcal{C}}.$$

因此,$\tau\sigma$ 的矩阵表示为 τ 与 σ 的矩阵表示的乘积.

验证如下. 由于当 $v\in U$,$w\in V$ 时,有

$$[\sigma(v)]_{\mathcal{C}}=[\sigma]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} \quad 及 \quad [\tau(w)]_{\mathcal{D}}=[\tau]_{\mathcal{C},\mathcal{D}}[w]_{\mathcal{C}},$$

故

$$[\tau]_{\mathcal{C},\mathcal{D}}[\sigma]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}=[\tau]_{\mathcal{C},\mathcal{D}}[\sigma(v)]_{\mathcal{C}}=[\tau(\sigma(v))]_{\mathcal{D}}=[\tau\sigma]_{\mathcal{B},\mathcal{D}}[v]_{\mathcal{B}}.$$

若 $\tau\in\mathcal{L}(V,W)$,现在来讨论当 V 与 W 的基变换时,τ 的相应矩阵之间的关系.

若 $\mathcal{B},\mathcal{C}$ 分别是 V,W 的基,$\mathcal{B}',\mathcal{C}'$ 也分别是 V,W 的基,τ 对基 $\mathcal{B},\mathcal{C}$ 及 $\mathcal{B}',\mathcal{C}'$ 分别有矩阵表示 $[\tau]_{\mathcal{B},\mathcal{C}}$ 及 $[\tau]_{\mathcal{B}',\mathcal{C}'}$,于是又: 对于任意 $v\in V$,有

$$[\tau(v)]_{\mathcal{C}}=[\tau]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}},\quad [\tau(v)]_{\mathcal{C}'}=[\tau]_{\mathcal{B}',\mathcal{C}'}[v]_{\mathcal{B}'}.$$

在第二讲中已经谈到,对于任意 $v\in V$,有

$$[v]_{\mathcal{B}'}=M_{\mathcal{B},\mathcal{B}'}[v]_{\mathcal{B}} \quad 及 \quad [\tau(v)]_{\mathcal{C}'}=M_{\mathcal{C},\mathcal{C}'}[\tau(v)]_{\mathcal{C}}.$$

将这些结果带入上式,得到

$$[\tau(v)]_{\mathcal{C}'}=M_{\mathcal{C},\mathcal{C}'}[\tau(v)]_{\mathcal{C}}=M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}},$$

而

$$[\tau(v)]_{\mathcal{C}'}=[\tau]_{\mathcal{B}',\mathcal{C}'}[v]_{\mathcal{B}'}=[\tau]_{\mathcal{B}',\mathcal{C}'}M_{\mathcal{B},\mathcal{B}'}[v]_{\mathcal{B}}.$$

由于 v 可取 V 中任意向量,故

$$[\tau]_{\mathcal{B}',\mathcal{C}'}M_{\mathcal{B},\mathcal{B}'}=M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}},$$

即为

$$[\tau]_{\mathcal{B}',\mathcal{C}'}=M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M^{-1}_{\mathcal{B},\mathcal{B}'},$$

即 $[\tau]_{\mathcal{B}',\mathcal{C}'}$ 与 $[\tau]_{\mathcal{B},\mathcal{C}}$ 是等价的.

特别当 $W=V$ 及 $\tau\in\mathcal{L}(V)$,且 $\mathcal{B}=\mathcal{C}$,$\mathcal{B}'=\mathcal{C}'$,$[\tau]_{\mathcal{B},\mathcal{B}}=[\tau]_{\mathcal{B}}$,$[\tau]_{\mathcal{B}',\mathcal{B}'}=[\tau]_{\mathcal{B}'}$,于是有

$$[\tau]_{\mathcal{B}'}=M_{\mathcal{B},\mathcal{B}'}[\tau]_{\mathcal{B}}M^{-1}_{\mathcal{B},\mathcal{B}'},$$

即 $[\tau]_{\mathcal{B}'}$ 与 $[\tau]_{\mathcal{B}}$ 是相似的.

在第五讲中将讨论 τ 在相似意义下的分类.

A3.2　伴随算子

由线性变换 $\tau \in \mathcal{L}(V, W)$，可以导出各种与之相关的线性变换来. 在这一节中先来定义与讨论在一般向量空间上的线性变换的伴随算子.

若 $\tau \in \mathcal{L}(V, W)$，可定义 W 的对偶空间 W^* 到 V 的对偶空间 V^* 映射

$$\tau^{\times}: \begin{array}{ccc} W^* & \longrightarrow & V^*, \\ f & \longmapsto & f \circ \tau = f\tau. \end{array}$$

这是有意义的，因为映射 $f\tau: V \to \mathbb{F}$ 确实是属于 V^*，即对任意 $v \in V$，有

$$\tau^{\times}(f)(v) = f(\tau(v)).$$

$\tau^{\times}$ 称为 τ 的**伴随算子 (adjoint operator)**.

易证如下命题.

命题 A3.2.1　**(1)** 对任意 $\tau, \sigma \in \mathcal{L}(V, W)$，有 $(\tau + \sigma)^{\times} = \tau^{\times} + \sigma^{\times}$；

(2) 对任意 $r \in \mathbb{F}, \sigma \in \mathcal{L}(V, W)$，有 $(r\tau)^{\times} = r\tau^{\times}$；

(3) 对 $\tau \in \mathcal{L}(V, W), \sigma \in \mathcal{L}(W, U)$，有 $(\sigma\tau)^{\times} = \tau^{\times}\sigma^{\times}$；

(4) 对任意可逆 $\tau \in \mathcal{L}(V)$，有 $(\tau^{-1})^{\times} = (\tau^{\times})^{-1}$.

证明　(1)、(2) 是显然那成立的. 对于 $f \in U^*$，

$$(\sigma\tau)^{\times}(f) = f\sigma\tau = \tau^{\times}(f\sigma) = \tau^{\times}\big(\sigma^{\times}(f)\big) = (\tau^{\times}\sigma^{\times})(f),$$

故得 (3). 由 (3)，

$$(\tau^{\times})(\tau^{-1})^{\times} = (\tau^{-1}\tau)^{\times} = \varepsilon^{\times} = \varepsilon,$$

这里 ε 为恒等映射. 同样 $(\tau^{-1})^{\times}\tau^{\times} = \varepsilon$，故得 (4). ■

命题 A3.2.2　若 $\dim V < \infty$，$\tau \in \mathcal{L}(V, W)$，且等同 V^{**} 为 V，W^{**} 为 W，则

$$\tau^{\times\times} = \tau.$$

证明　由定义，映射 $\tau^{\times\times}: V^{**} \to W^{**}$. 对任意 $f \in W^*$，有

$$\tau^{\times\times}(v^{**})(f) = v^{**}\tau^{\times}(f) = v^{**}(f\tau) = f\tau(v) = \big(\tau(v)\big)^{**}(f),$$

这里 v^{**} 是由 A2.2 中定义：v^{**} 由 v 而来，$v^{**} \in V^{**}$，定义为 $v^{**}(g) = g(v)$，这里 $g \in V^*$. 由上式即得

$$\tau^{\times\times}(v^{**}) = \big(\tau(v)\big)^{**}.$$

如果 V^{**} 与 V 等同，W^{**} 与 W 等同，则上式即为

$$\tau^{\times\times}(v) = \tau(v),$$

这对所有 $v \in V$ 都成立，故 $\tau^{\times\times} = \tau$. ■

此外，伴随算子与 A2.2 中定义的零化子还有以下一些结果.

命题 A3.2.3　若 $\tau \in \mathcal{L}(V, W)$，则

(1) $\operatorname{Ker}\tau^{\times}=(\operatorname{Im}\tau)^{\circ}$; **(2)** $(\operatorname{Im}\tau^{\times})^{\circ}=\operatorname{Ker}\tau$;

(3) 当 $\dim V,\dim W<\infty$ 时，$\operatorname{Im}\tau^{\times}=(\operatorname{Ker}\tau)^{\circ}$.

证明 (1) 由定义，映射 $\tau: V\to W$，$\tau^{\times}: W^*\to V^*$，故对所有 $v\in V$，

$$\begin{aligned} f\in\operatorname{Ker}\tau^{\times} &\iff \tau^{\times}(f)=0 \iff f\tau=0 \iff f(\tau(v))=0\\ &\iff f(\operatorname{Im}\tau)=0 \iff f\in(\operatorname{Im}\tau)^{\circ}. \end{aligned}$$

(2) 由于对所有 $f\in W^*$，

$$\begin{aligned} v\in\operatorname{Ker}\tau &\iff \tau(v)=0 \iff f(\tau(v))=0 \iff \tau^{\times}(f)(v)=0\\ &\iff v^{**}(\tau^{\times}(f))=0 \iff v^{**}\in(\operatorname{Im}\tau^{\times})^{\circ}. \end{aligned}$$

若 V^{**} 与 V 等同，即得 (2).

(3) 对所有 $v\in\operatorname{Ker}\tau$，$f\in W^*$，有

$$\tau^{\times}(f)(v)=f(\tau(v))=0,$$

故 $\tau^{\times}(f)(\operatorname{Ker}\tau)=0$，即 $\tau^{\times}(f)\in(\operatorname{Ker}\tau)^{\circ}$，这对所有的 $f\in W^*$ 都成立，故

$$\operatorname{Im}\tau^{\times}\subseteq(\operatorname{Ker}\tau)^{\circ}.$$

若空间是有限维的，则由命题 A2.2.4 以及上述 (2)，

$$\operatorname{Im}\tau^{\times}\approx(\operatorname{Im}\tau^{\times})^{\circ\circ}\approx(\operatorname{Ker}\tau)^{\circ},$$

故 $\operatorname{Im}\tau^{\times}=(\operatorname{Ker}\tau)^{\circ}$. ■

由此还可得到如下命题.

命题 A3.2.4 设 $\dim V,\dim W<\infty$，$\tau\in\mathcal{L}(V,W)$，则 $\operatorname{rank}\tau=\operatorname{rank}\tau^{\times}$.

证明 由命题 A3.2.3 之 (3)，知 $\operatorname{Im}\tau^{\times}=(\operatorname{Ker}\tau)^{\circ}$. 由命题 A2.2.6 之 (1)，知

$$(\operatorname{Ker}\tau)^{\circ}=\big((\operatorname{Ker}\tau)^{c}\big)^{*},$$

这里 $(\operatorname{Ker}\tau)^{c}$ 为 $\operatorname{Ker}\tau$ 在 V 中的补，故

$$\dim(\operatorname{Ker}\tau)^{\circ}=\dim\big((\operatorname{Ker}\tau)^{c}\big)^{*}=\dim(\operatorname{Ker}\tau)^{c}=\dim\operatorname{Im}\tau,$$

这是因为 $(\operatorname{Ker}\tau)^{c}\approx\operatorname{Im}\tau$. 于是 $\operatorname{rank}\tau^{\times}=\operatorname{rank}\tau$. ■

若 V,W 均为有限维的向量空间，$\tau\in\mathcal{L}(V,W)$，$\tau^{\times}\in\mathcal{L}(W^*,V^*)$，$\mathcal{B}=(b_1,\dots,b_n)$ 与 $\mathcal{C}=(c_1,\dots,c_n)$ 分别为 V 和 W 的一组基，而 $\mathcal{B}^*=(b_1^*,\dots,b_n^*)$ 与 $\mathcal{C}^*=(c_1^*,\dots,c_n^*)$ 分别为其对偶基. 于是 τ 有矩阵表示 $[\tau]_{\mathcal{B},\mathcal{C}}$，$\tau^{\times}$ 有矩阵表示 $[\tau^{\times}]_{\mathcal{C}^*,\mathcal{B}^*}$. 这两个矩阵之间关系如何?

已知

$$[\tau]_{\mathcal{B},\mathcal{C}}=\big([\tau(b_1)]_{\mathcal{C}},\dots,[\tau(b_n)]_{\mathcal{C}}\big),\quad [\tau^{\times}]_{\mathcal{C}^*,\mathcal{B}^*}=\big([\tau^{\times}(c_1^*)]_{\mathcal{B}^*},\dots,[\tau^{\times}(c_n^*)]_{\mathcal{B}^*}\big),$$

由于 $\tau(b_i)\in W$，故在基 $\mathcal{C}$ 下，其可表示为 $\tau(b_i)=\beta_1^{(i)}c_1+\cdots+\beta_n^{(i)}c_n$，即

$$[\tau(b_i)]_{\mathcal{C}}=\big(\beta_1^{(i)},\dots,\beta_n^{(i)}\big)^{\mathrm{T}}.$$

于是

$$[\tau]_{\mathcal{B},\mathcal{C}}=\begin{pmatrix}\beta_1^{(1)} & \cdots & \beta_1^{(n)}\\ \vdots & \ddots & \vdots\\ \beta_n^{(1)} & \cdots & \beta_n^{(n)}\end{pmatrix}.$$

由于 $\tau^{\times}(c_i^*)\in V^*$,故在基 $\mathcal{B}^*$ 下,其可表示为 $\tau^{\times}(c_i^*)=\alpha_1^{(i)}b_1^*+\cdots+\alpha_n^{(i)}b_n^*$,即

$$[\tau^{\times}(c_i^*)]_{\mathcal{B}^*}=\left(\alpha_1^{(i)},\ldots,\alpha_n^{(i)}\right)^{\mathrm{T}}.$$

于是

$$[\tau^{\times}]_{\mathcal{C}^*,\mathcal{B}^*}=\begin{pmatrix}\alpha_1^{(1)} & \cdots & \alpha_1^{(n)}\\ \vdots & \ddots & \vdots\\ \alpha_n^{(1)} & \cdots & \alpha_n^{(n)}\end{pmatrix}.$$

由 $\tau^{\times}$ 的定义可知 $\tau^{\times}(c_j^*)(b_i)=c_j^*(\tau(b_i))$,而这就是 $\alpha_i^{(j)}=\beta_j^{(i)}$. 于是得到如下定理.

定理 A3.2.1 τ 的伴随算子 $\tau^{\times}$ 所对应的矩阵是 τ 所对应的矩阵的转置,即

$$[\tau^{\times}]_{\mathcal{C}^*,\mathcal{B}^*}=\left([\tau]_{\mathcal{B},\mathcal{C}}\right)^{\mathrm{T}}.$$

A3.3 共轭算子

在上一节中对于一般的向量空间上的线性变换,定义并讨论了其伴随算子. 当向量空间是内积空间,其上的线性变换,则可定义并讨论其共轭算子,这是内积空间中十分重要的算子,由此可以导出一系列的结果,这是本节的内容.

由内积空间的 Riesz 表示定理(详见 A2.4),若 V 是有限维内积空间,$f\in V^*$,则存在唯一的 $x\in V$,使得

$$f(v)=\langle v,x\rangle$$

对所有 $v\in V$ 都成立. 由此可以定义映射 $\varphi\colon V^*\to V$ 为 $\varphi(f)=x$,即 $\varphi(f)$ 定义为

$$f(v)=\langle v,\varphi(f)\rangle.$$

由于对 $v\in V$,$f,g\in V^*$,$r,s\in\mathbb{F}$,

$$\begin{aligned}\langle v,\varphi(rf+sg)\rangle&=(rf+sg)(v)=rf(v)+sg(v)\\&=\langle v,\overline{r}\varphi(f)\rangle+\langle v,\overline{s}\varphi(g)\rangle=\langle v,\overline{r}\varphi(f)+\overline{s}\varphi(g)\rangle,\end{aligned}$$

即

$$\varphi(rf+sg)=\overline{r}\varphi(f)+\overline{s}\varphi(g),$$

于是 φ 为共轭线性. φ 显然是满射. 由于 $\varphi(f)=0$ 蕴涵 $f=0$,故 φ 也是单射,从而 $\varphi\colon V^*\to V$ 为**“共轭同构”(conjugate isomorphism)**.

若 V,W 为域 $\mathbb{F}$ 上的有限维内积空间,$\tau\in\mathcal{L}(V,W)$,对于固定的 $w\in W$,考虑函数 $\theta_w\colon V\to\mathbb{F}$ 定义为

$$\theta_w(v)=\langle\tau(v),w\rangle.$$

易证,θ_w 是 V 上的线性泛函. 由 Riesz 表示定理,存在唯一的 $x\in V$,使得

$$\theta_w(v) = \langle \tau(v), w \rangle = \langle v, x \rangle$$

对所有的 $v \in V$ 都成立,令 $\tau^*(w) = x$,则

$$\langle \tau(v), w \rangle = \langle v, \tau^*(w) \rangle$$

对所有的 $v \in V$ 都成立,故存在唯一的 τ^* 使上式成立. 可证 τ^* 是线性的. 由于对任意的 $v \in V, w, w' \in W$,

$$\begin{aligned}\langle v, \tau^*(rw + sw') \rangle &= \langle \tau(v), rw + sw' \rangle = \overline{r}\langle \tau(v), w \rangle + \overline{s}\langle \tau(v), w' \rangle \\ &= \overline{r}\langle v, \tau^*(w) \rangle + \overline{s}\langle v, \tau^*(w') \rangle = \langle v, r\tau^*(w) \rangle + \langle v, s\tau^*(w') \rangle \\ &= \langle v, r\tau^*(w) + s\tau^*(w') \rangle,\end{aligned}$$

故

$$\tau^*(rw + sw') = r\tau^*(w) + s\tau^*(w').$$

因此,$\tau^* \in \mathcal{L}(W, V)$,这时称 τ^* 为 τ 的**共轭 (adjoint)**.

现在来讨论 τ^* 与 $\tau^\times$ 之间的关系. 若 V, W 是域 $\mathbb{F}$ 上的两个向量空间,$\tau \in \mathcal{L}(V, W)$,则存在映射 $\tau^\times: W^* \to V^*$ 及 $\tau^*: W \to V$. 再由前面定义的映射 $\varphi_1: V^* \to V$ 及 $\varphi_2: W^* \to W$,可得

$$\begin{tikzcd} V^* \arrow[d, "\varphi_1"] & W^* \arrow[l, "\tau^\times"'] \arrow[d, "\varphi_2"] \\ V & W \arrow[l, "\tau^*"'] \end{tikzcd}$$

定义映射 $\sigma: W^* \to V^*$ 为

$$\sigma = \varphi_1^{-1}\tau^*\varphi_2,$$

显然这是一个线性映射. 若 $x \in V$, $\varphi_1^{-1} = g$,则 $\varphi_1(g) = x$,

$$\varphi_1^{-1}(x)(v) = g(v) = \langle v, \varphi_1(g) \rangle = \langle v, x \rangle.$$

因此对任意 $f \in W^*, x \in V$,有

$$\begin{aligned}\big(\sigma(f)\big)(v) &= \big(\varphi_1^{-1}\tau^*\varphi_2(f)\big)(v) = \varphi_1^{-1}\big(\tau^*\varphi_2(f)\big)(v) = \langle v, \tau^*\varphi_2(f) \rangle \\ &= \langle \tau(v), \varphi_2(f) \rangle = f(\tau(v)) = \tau^\times(f)(v),\end{aligned}$$

故 $\sigma = \tau^\times$,即

$$\tau^\times = \varphi_1^{-1}\tau^*\varphi_2,$$

上图为交换图.

设 $\mathcal{B} = (b_1, \ldots, b_n)$ 与 $\mathcal{C} = (c_1, \ldots, c_m)$ 分别是 V 与 W 的一组正规正交基. 已知

$$[\tau]_{\mathcal{B},\mathcal{C}} = \big([\tau(b_1)]_{\mathcal{C}}, \ldots, [\tau(b_n)]_{\mathcal{C}}\big) = \begin{pmatrix} \beta_1^{(1)} & \cdots & \beta_1^{(n)} \\ \vdots & & \vdots \\ \beta_m^{(1)} & \cdots & \beta_m^{(n)} \end{pmatrix},$$

于是 $\beta_j^{(i)} = \langle \tau(b_i), c_j \rangle$. 而

$$[\tau^*]_{\mathcal{C},\mathcal{B}} = \big([\tau^*(c_1)]_{\mathcal{B}},\ldots,[\tau^*(c_m)]_{\mathcal{B}}\big) = \begin{pmatrix} \gamma_1^{(1)} & \cdots & \gamma_1^{(m)} \\ \vdots & & \vdots \\ \gamma_n^{(1)} & \cdots & \gamma_n^{(m)} \end{pmatrix},$$

于是 $\gamma_j^{(i)} = \langle \tau^*(c_i), b_j\rangle$. 但是

$$\langle \tau^*(c_i), b_j\rangle = \overline{\langle b_j, \tau^*(c_i)\rangle} = \overline{\langle \tau(b_j), c_i\rangle} = \overline{\beta}_i^{(j)},$$

因此得到如下定理.

定理 A3.3.1 τ 的共轭所对应矩阵为 τ 所对应矩阵的共轭转置,也用 $\square^*$ 来记之,即

$$[\tau^*]_{\mathcal{C},\mathcal{B}} = \overline{[\tau]}_{\mathcal{B},\mathcal{C}}^{\mathrm{T}} = [\tau]_{\mathcal{B},\mathcal{C}}^*.$$

对于 τ 的共轭,易证有以下的这些性质.

命题 A3.3.1 若 V, W 为有限维内积空间,$\sigma,\tau\in\mathcal{L}(V,W)$,则

(1) $\langle \tau^*(w), v\rangle = \langle w, \tau(v)\rangle$; **(2)** $(\sigma+\tau)^* = \sigma^*+\tau^*$;

(3) $(r\tau)^* = \overline{r}\tau^*$, $r\in\mathbb{F}$; **(4)** $\tau^{**} = \tau$;

(5) 若 $V = W$,则 $(\sigma\tau)^* = \tau^*\sigma^*$; **(6)** 若 τ 可逆,则 $(\tau^{-1})^* = (\tau^*)^{-1}$.

进一步,可给出如下的一些定义.

定义 A3.3.1 设 V 是内积空间,$\tau\in\mathcal{L}(V)$,则

(1) 称 τ 是**自共轭 (self-adjoint)** 或**埃尔米特 (Hermitian)**,若 $\tau^* = \tau$;

(2) 称 τ 是**酉 (unitary)**,若 τ 是双射且 $\tau^* = \tau^{-1}$;

(3) 称 τ 是**正规 (normal)**,若 $\tau\tau^* = \tau^*\tau$.

相应的,当 A 是复矩阵时,可定义:

(1) A 是**埃尔米特**,若 $A^* = A$; **(2)** A 是**斜埃尔米特**,若 $A^* = -A$;

(3) A 是**酉**,若 A 可逆且 $A^* = A^{-1}$; **(4)** A 是**正规**,若 A 可逆且 $AA^* = A^*A$;

当 A 是实矩阵时,可定义:

(1) A 是**对称**,若 $A^{\mathrm{T}} = A$; **(2)** A 是**斜对称**,若 $A^{\mathrm{T}} = -A$;

(3) A 是**正交**,若 A 可逆且 $A^* = A^{-1}$.

易证 τ 是正规、自共轭、酉算子当且仅当 τ 在正规正交基下对应的矩阵是正规、埃尔米特、酉阵. 下面来讨论着这三类算子.

1. 自共轭算子 由定义 $\tau\in\mathcal{L}(V)$ 是自共轭,所对所有 $v, w\in V$ 有

$$\langle \tau(v), w\rangle = \langle b, \tau(w)\rangle,$$

即 $\tau^* = \tau$. 易证有如下的性质.

命题 A3.3.2 若 V 是内积空间,且 $\sigma,\tau\in\mathcal{L}(V)$,

(1) 若 σ,τ 自共轭,则 $\sigma+\tau$ 也是自共轭;

(2) 若 τ 自共轭,r 为实数,则 $r\tau$ 是自共轭;

(3) 若 σ,τ 为自共轭,则 $\sigma\tau$ 为自共轭当且仅当 $\sigma\tau = \tau\sigma$;

(4) 若 τ 为自共轭且可逆，则 τ^{-1} 也是自共轭；

(5) 若 τ 自共轭，则 $\langle\tau(v),v\rangle$ 是实的，对所有的 $v\in V$ 成立；

(6) 若 V 是酉空间，对所有的 $v\in V$，$\langle\tau(v),v\rangle$ 是实的，则 τ 是自共轭；

(7) 若 τ 是自共轭，且 $\langle\tau(v),v\rangle=0$ 对所有的 v 都成立，则 $\tau=0$；

(8) 若 τ 是自共轭，则对任意给定的 $k>0$，$\tau^k=0$ 蕴涵 $\tau=0$.

证明　由于 τ 自共轭，所对应的矩阵为埃尔米特阵，即 $A=A^*$，故立即可得 (1)—(4).

由于 τ 自共轭，所以

$$\langle\tau(v),v\rangle=\langle v,\tau(v)\rangle=\overline{\langle\tau(v),v\rangle},$$

故 $\langle\tau(v),v\rangle$ 是实的，这就证明了 (5).

为了证明 (6)、(7)，先证明：若 V 是酉空间，且对所有的 $v\in V$ 有 $\langle\tau(v),v\rangle=0$，则 $\tau=0$.

令 $v=rx+y$，这里 $x,y\in V$，$r\in\mathbb{C}$，则

$$\begin{aligned}0&=\langle\tau(rx+y),rx+y\rangle=|r|^2\langle\tau(x),x\rangle+\langle\tau(y),y\rangle+r\langle\tau(x),y\rangle+\bar{r}\langle\tau(y),x\rangle\\&=r\langle\tau(x),y\rangle+\bar{r}\langle\tau(y),x\rangle.\end{aligned}$$

取 $r=1$，则

$$\langle\tau(x),y\rangle+\langle\tau(y),x\rangle=0;$$

取 $r=\mathrm{i}$，则

$$\langle\tau(x),y\rangle-\langle\tau(y),x\rangle=0.$$

故 $\langle\tau(x),y\rangle=0$ 对所有的 $x,y\in V$ 都成立，因此 $\tau=0$.

用这个结果来证明 (6)，

$$\begin{aligned}\langle(\tau-\tau^*)(v),v\rangle&=\langle\tau(v),v\rangle-\langle\tau^*(v),v\rangle=\langle\tau(v),v\rangle-\langle v,\tau(v)\rangle\\&=\langle\tau(v),v\rangle-\overline{\langle\tau(v),v\rangle}=0.\end{aligned}$$

由上述结果可知 $\tau-\tau^*=0$，即 τ 是自共轭算子.

再来证明 (7). 当 $\mathbb{F}=\mathbb{C}$ 时，由上述结果立得 (7)，故只要证明 $\mathbb{F}=\mathbb{R}$ 的情形. 此时，

$$\begin{aligned}0&=\langle\tau(x+y),x+y\rangle=\langle\tau(x),x\rangle+\langle\tau(y),y\rangle+\langle\tau(x),y\rangle+\langle\tau(y),x\rangle\\&=\langle\tau(x),y\rangle+\langle\tau(y),x\rangle=\langle\tau(x),y\rangle+\langle x,\tau(y)\rangle=2\langle\tau(x),y\rangle.\end{aligned}$$

由上述结果可知 $\tau=0$.

最后证明 (8). 若 $\tau^k(v)=0$ 对所有 $v\in V$ 成立，则存在 m 使得 $2^m\geqslant k$. 于是 $\tau^{2^m}(v)=0$，因此

$$0=\langle\tau^{2^m}(v),v\rangle=\langle\tau^{2^{m-1}}\tau^{2^{m-1}}(v),v\rangle=\langle\tau^{2^{m-1}}(v),\tau^{2^{m-1}}(v)\rangle.$$

因此 $\tau^{2^{m-1}}(v)=0$，重复这样的步骤，最后得 $\tau=0$. ■

对于内积空间上的任意一个线性变换 $\tau\in\mathcal{L}(V,W)$ 均可写成

$$\tau = \tau_1 + \mathrm{i}\tau_2,$$

这里 τ_1, τ_2 为自共轭变换. 事实上,令

$$\tau_1 = \frac{\tau + \tau^*}{2}, \quad \tau_2 = \frac{\tau - \tau^*}{2\mathrm{i}},$$

即得上式. 这与"任意一个复数 z 均可写成 $z = x + \mathrm{i}y$,这里 x, y 均为实数"相类似. 所以,对于内积空间,自共轭变换是十分基本的.

2. 酉算子 由定义,$\tau \in \mathcal{L}(V)$ 是酉算子,若对所有的 $v, w \in V$,有

$$\langle \tau(v), w \rangle = \langle v, \tau^{-1}(w) \rangle.$$

易证有如下的性质.

命题 A3.3.3 若 V 是内积空间,$\sigma, \tau \in \mathcal{L}(V)$,

(1) 若 τ 是酉变换,则 τ^{-1} 也是酉变换;

(2) 若 σ, τ 是酉变换,则 $\sigma\tau$ 也是酉变换;

(3) τ 是酉变换当且仅当 τ 是满射且保距;($\tau \in \mathcal{L}(V, W)$ 称为**保距 (isometry)**,若对 $u, v \in V$,有 $\langle \tau(u), \tau(v) \rangle = \langle u, v \rangle$.)

(4) 若 V 为有限维,则 τ 是酉变换当且仅当 τ 将正交基映为正交基.

证明 (1)、(2) 是显然的,来证 (3). 若 τ 是满射,则对所有 $v, w \in V$,

$$\begin{aligned} \tau \text{ 是保距} &\iff \langle \tau(v), \tau(w) \rangle = \langle v, w \rangle \iff \tau^*\tau(w) = w \iff \tau^*\tau = \varepsilon \\ &\iff \tau^* = \tau^{-1} \iff \tau \text{ 是酉变换}. \end{aligned}$$

若 τ 是酉变换,$\mathcal{O} = (u_1, \ldots, u_n)$ 是 V 的正规正交基,于是

$$\langle \tau(u_i), \tau(u_j) \rangle = \langle u_i, u_j \rangle = \delta_{ij}.$$

故 $\tau(\mathcal{O})$ 是 V 的正规正交基. 反之,若 $\mathcal{O}$ 及 $\tau(\mathcal{O})$ 是 V 的正规正交基,则

$$\langle \tau(u_i), \tau(u_j) \rangle = \delta_{ij} = \langle u_i, u_j \rangle.$$

若 $v = \sum_{i=1}^n r_i u_i$, $w = \sum_{j=1}^n s_j u_j$,则

$$\begin{aligned} \langle \tau(v), \tau(w) \rangle &= \left\langle \sum_{i=1}^n r_i \tau(u_i), \sum_{j=1}^n r_j \tau(u_j) \right\rangle = \sum_{i=1}^n \sum_{j=1}^n r_i \bar{s}_j \langle \tau(u_i), \tau(u_j) \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n r_i \bar{s}_j \langle u_i, u_j \rangle = \left\langle \sum_{i=1}^n r_i u_i, \sum_{j=1}^n r_j u_j \right\rangle = \langle u, w \rangle, \end{aligned}$$

故 τ 为酉变换. ■

与酉算子对应的是酉矩阵,对酉矩阵显然有以下性质.

命题 A3.3.4 若 A 为矩阵,

(1) $n \times n$ 矩阵 A 为酉的当且仅当 A 的所有的列组成 $\mathbb{C}^n$ 的一组正规正交基;

(2) $n \times n$ 矩阵 A 为酉的当且仅当 A 的所有的行组成 $\mathbb{C}^n$ 的一组正规正交基;

(3) 若 A 为酉,则 $|\det A| = 1$; 若 A 为正交,则 $\det A = \pm 1$.

3. 正规算子 由定义 $\tau \in \mathcal{L}(V)$ 是正规算子,若 $\tau\tau^* = \tau^*\tau$. 易证有如下的性质.

命题 A3.3.5 若 V 是内积空间,τ 是 V 上的正规算子,则

(1) 对任意多项式 $p(x) \in \mathbb{F}[x]$,$p(\tau)$ 也是正规算子;

(2) $\tau(v) = 0$ 蕴涵 $\tau^*(v) = 0$;

(3) 对任意给定的 $k > 0$,$\tau^k = 0$ 蕴涵 $\tau = 0$;

(4) 对任意 $\lambda \in \mathbb{F}$ 和 $k > 0$,$(\tau - \lambda\varepsilon)^k = 0$ 蕴涵 $\tau - \lambda\varepsilon = 0$;

(5) 若 $\tau(v) = \lambda v$,则 $\tau^*(v) = \overline{\lambda} v$.

证明 (1)、(2) 是显然的,来证 (3). $\sigma = \tau\tau^*$ 显然是自共轭的,由于 τ 为正规的,故

$$\sigma^k(v) = (\tau^*)^k \tau^k(v) = 0.$$

由命题 A3.3.2 之 (8),得 $\sigma(v) = 0$,即 $\tau\tau^*(v) = 0$,但

$$0 = \langle \tau^*\tau(v), v\rangle = \langle \tau(v), \tau(v)\rangle,$$

故 $\tau(v) = 0$.

由 (1) 及 (3) 得 (4),来证 (5). 若 $\tau(v) = \lambda v$,这里 $v \neq 0$,则 $(\tau - \lambda\varepsilon)(v) = 0$. 由 (2),$(\tau - \lambda\varepsilon)^*(v) = 0$,但 $(\tau - \lambda\varepsilon)^* = \tau^* - \overline{\lambda}\varepsilon$,故得 (5). ■

在最后一讲中,要对着三个算子作进一步深入的讨论,尤其是分解定理.

第四讲

主理想整环上的模及其分解

在第二讲及第三讲中讨论了向量空间及其上的线性变换,在这一讲及下一讲中将从模的观点来重新认识之,这是这本小书的主要部分. 在这一讲中,将介绍模的定义与基本性质,尤其是在主理想整环上的模及其分解.

A4.1 环上的模的基本概念

1. 若 V 是域 $\mathbb{F}$ 上的一个向量空间,$\tau\in\mathcal{L}(V)$. 对 $\mathbb{F}[x]$ 中任一多项式 $p(x)$,对任意 $v\in V$,可定义

$$p(x)v=p(\tau)(v).$$

这就是我们要讨论线性算子作用在 V 上. 显然,对任意 $r(x),s(x)\in\mathbb{F}[x]$,$u,v\in V$,有

$$\begin{aligned}
r(x)(u+v)&=r(x)u+r(x)v, & \big(r(x)+s(x)\big)u&=r(x)u+s(x)u,\\
\big(r(x)s(x)\big)u&=r(x)\big(s(x)u\big) & 1u&=u,
\end{aligned}$$

等等. 但是 $\mathbb{F}[x]$ 不是域而是环,所以将 $\mathbb{F}[x]$ 中元素对 V 作数乘,V 并不能称为一个向量空间. 于是引入了比向量空间更一般的概念——模.

定义 A4.1.1 若 R 是有单位元的交换环,其元素称为纯量. 一个 $\boldsymbol{R}$**–模 (**$\boldsymbol{R}$**–module)**或 $\boldsymbol{R}$ **上的一个模**是一个非空集合 M,有运算加法,记作 $+$,对 $(u,v)\in M\times M$,有 $u+v\in M$; 另一个是 R 与 M 的运算是数乘,用毗连来表示,对 $(r,u)\in\mathbb{F}\times M$,有 $ru\in M$,而且有

(1) M 对 $+$ 作成 Abel 群;

(2) R 对 M 的数乘满足: 对所有的 $r,s\in R$,$u,v\in M$ 有

分配律

$$r(u+v)=ru+rv,\qquad (r+s)u=ru+su;$$

结合律

$$(rs)u=r(su);$$

及

$$1u=u.$$

显然,当 R 是域时,则模为向量空间,即**域上的模就是向量空间**. 当 $R=\mathbb{Z}$(整数环)时,则 **$\mathbb{Z}$– 模就是 Abel 群**. 故模也是 Abel 群的概念的扩充.

特别重要的是在第一讲一开始就说到的 $R=\mathbb{F}[x]$,若 $\mathbb{F}$ 是域,则由定理 A1.2.1,

$\mathbb{F}[x]$ 是主理想整环,于是可以定义 $\mathbb{F}[\boldsymbol{x}]$- 模,这是我们今后主要讨论的对象.

若 R 是环,则所有 $m\times n$ 矩阵的集合 $\mathcal{M}_{m,n}(R)$ 是一个 R- 模,其加法与数乘就是矩阵的加法与数乘. 当 $R=\mathbb{F}[x]$ 时,$\mathcal{M}_{m,n}(\mathbb{F}[x])$ 是矩阵中的元素全为多项式的矩阵的全体,它成为一个 $\mathbb{F}[x]$- 模.

另一个重要的模是环 R 自己可以成为 R- 模. 若 R 是有单位元的交换环,定义数乘为环乘法,这就是一个 R- 模. 今后会用到这个模.

在今后讨论线性算子 $\tau\in\mathcal{L}(V)$ 作用在向量空间 V 上时,**$\boldsymbol{V}$ 可以看成 $\mathbb{F}$ 上的一个向量空间,也可以看成在 $\mathbb{F}[\boldsymbol{x}]$ 上的模.**

2. 一些向量空间中的概念可以推广到模上.

先定义 R- 模 M 的子模如下. 若 S 为 R- 模 M 的一个子集,本身是一个 R- 模,称 S 为 M 的**子模**. S 上的运算就是 M 的运算在 S 上的限制. 显然有如下结论成立.

(1) R- 模 M 的一个非空子集 S 成为子模当且仅当对任意 $r,s\in R,u,v\in S$ 有

$$ru+sv\in S;$$

(2) 若 S,T 是 M 的子模,则 $S\cap T$ 及

$$S+T=\left\{u+v \mid u\in S,v\in T\right\}$$

也是 M 的子模;

(3) 有单位元的交换环 R 就是 R 上的模,则 **$\boldsymbol{R}$- 模 $\boldsymbol{R}$ 的子模就是环 $\boldsymbol{R}$ 的理想**.

由子模的概念,可以定义模的直和. 若 M 是 R- 模,称 M 的子模 $S_1,\ldots,S_n$ 的直和,若每个 $v\in M$ 都可以唯一地(不及前后次序)写成子模 S_i 中元素之和. 即对任意 $v\in M$,有 $u_i\in S_i$,使得

$$v=u_1+\cdots+u_n,$$

并且若还有 $w_i\in S_i$,使得

$$v=w_1+\cdots+w_n,$$

则经过适当排列后,有 $u_i=w_i$, $i=1,\ldots,n$.

当 M 是 $S_1,\ldots,S_n$ 的直和,写成

$$M=S_1\oplus\cdots\oplus S_n.$$

若 $M=S\oplus S^c$,则 S^c 称为 S 在 M 中的补.

显然,M 是子模 $S_1,\ldots,S_n$ 的直和当且仅当 $M=S_1+\cdots+S_n$ 及对每个 i,有

$$S_i\cap\Big(\sum_{j\neq i}S_j\Big)=\{0\}.$$

可以定义**生成集 (spanning set)** 如下. 若 M 是 R- 模,S 是一个子集,由 S 生成的子模为 S 中元素的所有 R- 线性组合的集合,即

$$\langle S\rangle=\operatorname{span}S=\left\{r_1v_1+\cdots+r_nv_n \mid r_i\in R,v_i\in S\right\}.$$

M 中的一个子集 S 称为生成 M,若 $M=\operatorname{span}S$,即每个 $v\in M$,可写成

$$v = r_1v_1 + \cdots + r_nv_n,$$

这里 $r_1, \cdots, r_n \in R, v_1, \cdots, v_n \in S$.

特别由一个元素生成的子模,即

$$\langle v \rangle = Rv = \left\{ rv \mid r \in R \right\}, \quad v \in M$$

称为由 v 生成的**循环子模 (cyclic submodule)**. 需要注意的是,这是一种十分重要的子模,今后要不断出现.

如果 R– 模 M 可以由有限集生成,则称 M 是**有限生成的 (finitely generated)**.

3. 同样可以在模中定义子集的 R– 线性无关性及 R– 基.

称 M 的子集 S 为 ***R*– 线性无关的 (linearly independent)**,若对任意 $v_1, \ldots, v_n \in S$ 及 $r_1, \ldots, r_n \in R$,

$$r_1v_1 + \cdots + r_nv_n = 0$$

蕴涵 $r_1 = \cdots = r_n = 0$. 若集 S 不是线性无关的,则称为线性相关.

若 M 是 R– 模,M 的子集 $\mathcal{B}$ 称为 M 的基,若 $\mathcal{B}$ 线性无关且生成 M. 由此易见,

(1) 模 M 的子集 $\mathcal{B}$ 是一组基当且仅当对每个 $v(\neq 0) \in M$, $\mathcal{B}$ 有唯一的子集 $\left\{v_1, \ldots, v_n\right\}$ 及非零纯量 $r_1, \ldots, r_n \in R$ 使得

$$v = r_1v_1 + \ldots + r_nv_n;$$

(2) 若 $\mathcal{B}$ 是 R– 模 M 的基,则 $\mathcal{B}$ 是 M 的极小生成元集,是极大线性无关集.

对于向量空间,有线性变换,对于模有模同态.

定义 A4.1.2 若 M, N 为 R– 模,映射 $\tau: M \to N$ 称为 ***R*– 同态 (homomorphism)**,若对所有 $r, s \in R, u, v \in M$,有

$$\tau(ru + sv) = r\tau(u) + s\tau(v).$$

所有从 M 到 N 的 R– 同态记作 $\mathrm{Hom}_R(M, N)$.

显然 R– 同态是线性变换的推广. 称 M 到 M 的 R– 同态为**自同态 (endomorphism)**; 称单射的同态为**单同态 (monomorphism)**; 称满射的同态为**满同态 (epimorphism)**; 称双射的同态为**同构 (isomorphism)**.

若 $\tau \in \mathrm{Hom}_R(M, N)$,定义 τ 的核与像为

$$\mathrm{Ker}\,\tau = \left\{ v \in M \mid \tau(v) = 0 \right\} \quad 及 \quad \mathrm{Im}\,\tau = \left\{ \tau(v) \mid v \in M \right\},$$

易知它们分别是 M 与 N 的子模.

由于不是所有的模都有 R– 基,故有以下的定义.

定义 A4.1.3 R– 模 M 称为**自由的 (free)**,若 M 有 R– 基. 若 $\mathcal{B}$ 是 M 的基,称 M 在 $\mathcal{B}$ 上自由. M 的基的基数称为 M 的秩,记作 $\mathrm{rank}\, M$.

下面来证明这样的定义是有意义的.

4. 若 M 为 R– 模,S 是它的子模,称

$$v+S=\left\{v+s \mid s \in S\right\}, \quad v \in M$$

为 S 在 M 中的一个陪集. 所有 S 在 M 中的陪集作成的集合记作 M/B. 这是一个 $R-$ 模,其运算定义为

$$(u+S)+(v+S)=(u+v)+S \quad 及 \quad r(u+S)=ru+S,$$

而 M/S 的零元素为 $s+S=0+S=S, s \in S$. 这个 $R-$ 模称为 M 关于 S 的**商模 (quotient module)**.

现在来证明,若 M 是自由模,则 M 的任意两个基有相同的基数. 有了这个结果,自由模 M 的秩才有意义.

若 R 是有单位元的交换环,S 是 R 的理想,S 在 R 中的陪集的集合 R/S 作成一个环,称为 R 关于 S 的**商环 (quotient ring)**,其加法与乘法定义为,设 $a, b \in R$,

$$(a+S)+(b+S)=(a+b)+S \quad 及 \quad (a+S)(b+S)=ab+S.$$

要证明乘法有意义,就要证明

$$b+S=b'+S \Longrightarrow ab+S=ab'+S,$$

也就是

$$b-b' \in S \Longrightarrow a(b-b') \in S,$$

由于 S 是理想,故上式成立.

现在来证明如下的结果,以说明极大理想(详见 A1.2)的重要性.

引理 A4.1.1 若 R 是有单位元的交换环,商环 $R/\mathcal{Q}$ 是域当且仅当 $\mathcal{Q}$ 是极大理想(即不存在 R 的理想 $\mathcal{I}$,使得 $\mathcal{Q} \subsetneq \mathcal{I} \subsetneq R$).

证明 若 $R/\mathcal{Q}$ 是域,且 $\mathcal{Q}$ 不是极大,则存在理想 $\mathcal{I}$ 适合 $\mathcal{Q} \subsetneq \mathcal{I} \subsetneq R$. 设 $i \in \mathcal{I}-\mathcal{Q}$,考虑由 i 及 $\mathcal{Q}$ 生成的理想

$$\mathcal{K}=\langle i, \mathcal{Q}\rangle \subseteq \mathcal{I}.$$

由于 $i \notin \mathcal{Q}$,故 $i+\mathcal{Q} \neq 0$. 由于 $R/\mathcal{Q}$ 是域,故 $i+\mathcal{Q}$ 有逆,不妨设为 $i'+\mathcal{Q}$,则

$$(i+\mathcal{Q})(i'+\mathcal{Q})=ii'+\mathcal{Q}=1+\mathcal{Q}.$$

故 $1-ii' \in \mathcal{Q} \subseteq \mathcal{K}$. 由于 $ii' \in \mathcal{K}$,故 $1 \in \mathcal{K}$,这就导出 $\mathcal{K}=R$. 但 $\mathcal{K} \subseteq \mathcal{I}$,$\mathcal{I}$ 是 R 的真子集,这个矛盾导出 $\mathcal{Q}$ 是极大理想.

反之,若 $\mathcal{Q}$ 是极大理想,$0 \neq r+\mathcal{Q}$,则 $r \notin \mathcal{Q}$,故 $\mathcal{I}=\langle r, \mathcal{Q}\rangle$ 是严格地包含 $\mathcal{Q}$. 因为 $\mathcal{Q}$ 是极大,故 $\mathcal{I}=R$. 这导出 $1 \in \mathcal{I}$,故有 $s \in R$,使得 $1=sr+i$ 对某个 $i \in Q$ 成立. 故

$$(s+\mathcal{Q})(r+\mathcal{Q})=sr+\mathcal{Q}=(1-i)+\mathcal{Q}=1+\mathcal{Q},$$

即 $(r+\mathcal{Q})^{-1}=s+\mathcal{Q}$. 故 $R/\mathcal{Q}$ 是域. ■

引理 A4.1.2 任意有单位元的交换环 R 一定有极大理想.

证明 R 不是零环,则一定有真理想,例如 $\{0\}$. 若 $\mathcal{I}$ 为 R 的所有真理想的集合,则 $\mathcal{I}$ 非空. 若

$$\mathcal{Q}_1 \subseteq \mathcal{Q}_2 \subseteq \cdots$$

是 R 中真理想链,则 $I=\bigcup_j Q_j$ 也是一个理想. 因为 $1\notin I$,故 $I\in\mathcal{I}$. 因此 $\mathcal{I}$ 中任何链都有上界,由 Zorn 引理,$\mathcal{I}$ 有极大元(对照 A2.1 中向量空间基的存在性的证明). ■

定理 A4.1.1 若 M 是自由 R- 模,则 M 的任意两个基有相同的基数.

证明 由引理 A4.1.2,R 有极大理想 Q,再由引理 A4.1.1,R/Q 是域. 令

$$QM=\left\{a_1v_1+\cdots+a_nv_n \,\middle|\, a_i\in Q, v_i\in M\right\},$$

则 QM 是 M 的一个子模,作商模 M/QM. 现在来证明 M/QM 是域 R/Q 上的一个向量空间,为此定义其数乘为

$$(r+Q)(u+QM)=ru+QM,$$

这里 $r\in R, v\in M$. 当然要来验证这样定义的数乘是有意义的. 这就要证明:若

$$r+Q=r'+Q \quad 及 \quad u+QM=u'+QM,$$

这里 $r,r'\in R, u,u'\in M$,则

$$ru+QM=r'u'+QM.$$

也就是要证:若 $r-r'\in Q, u-u'\in QM$,则 $ru-r'u'\in QM$.

事实上,由于 $r-r'\in Q, u-u'\in QM$,则 $(r-r')u'\in QM$ 及 $r(u-u')\in QM$. 于是

$$(r-r')u'+r(u-u')=ru-r'u'\in QM.$$

因此,这样定义的数乘是有意义的. 可以直接验证,这样定义的数乘满足定义 A1.3.1 (向量空间的定义)中数乘必须满足的 4 个条件,而 M/QM 显然对加法成 Abel 群. 故 M/QM 确实是 R/Q 上的一个向量空间.

若 $\mathcal{B}$ 是自由 R- 模 M 上的一组基,且 $b_i,b_j\in\mathcal{B}\,(i\neq j)$,则 b_i+QM 与 b_j+QM 是不相同的. 事实上,若 $b_i+QM=b_j+QM$ 成立,则 $b_i-b_j\in QM$,故

$$b_i-b_j=a_1v_1+\cdots+a_nv_n,$$

这里 $a_k\in Q, v_k\in M$. 由于每个 v_k 都是 $\mathcal{B}$ 元素的线性组合,因此可设 v_k 中 b_i 的系数为 r_k,比较上式两边的 b_i 的系数,得到

$$1=a_1r_1+\cdots+a_nr_n.$$

而 $a_k\in Q$,故上式右边属于 Q,即 $1\in Q$,这与 Q 是极大理想相矛盾. 故 b_i+QM 与 b_j+QM 当 $i\neq j$ 时是不相同的. 因此,

$$\mathcal{B}'=\left\{b+QM \,\middle|\, b\in\mathcal{B}\right\}$$

与 M 的基 $\mathcal{B}$ 有相同的基数. 现在来证明,$\mathcal{B}'$ 是 R/Q 上的向量空间 M/QM 的一组基.

事实上,由于 $\mathcal{B}$ 生成 M,故 $\mathcal{B}'$ 生成 M/QM,要证 $\mathcal{B}'$ 线性无关. 若

$$\sum_j (r_j+Q)(b_j+QM)=0,$$

则 $\sum_j(r_jb_j+QM)=0$,也就是 $\sum_j r_jb_j\in QM$,于是 $\sum_j r_jb_j=\sum_i a_ib_i$,这里 $a_i\in Q$. 两

边相等蕴涵 $r_j \in \mathcal{Q}$. 于是 $r_j + \mathcal{Q} = 0$, 即 $r_j + \mathcal{Q}$ 是 $R/\mathcal{Q}$ 中的零元素. 故 $\mathcal{B}'$ 线性无关, $\mathcal{B}'$ 的确是 $M/\mathcal{Q}M$ 的基.

终上所述, $|\mathcal{B}| = \dim(M/\mathcal{Q}M)$, 不依赖于 $\mathcal{B}$ 的选取. ■

定理 A2.1.1 告诉我们, 域 $\mathbb{F}$ 上两个向量空间同构当且仅当它们的维数相同. 在模的情形, 有如下定理.

定理 A4.1.2　两个自由 R– 模同构当且仅当它们有相同的秩.

证明　设 M, N 为两个自由 R– 模. 若 $M \approx N$, 则从 M 到 N 的任意同构映射 τ 将 M 的基映为 N 的基. 由于 τ 是双射, 故 $\operatorname{rank} M = \operatorname{rank} N$.

反之, 若 $\operatorname{rank} M = \operatorname{rank} N$, $\mathcal{B}$ 是 M 的基, $\mathcal{C}$ 是 N 的基, 由于 $\mathcal{B}, \mathcal{C}$ 的基数相同, 故有双射 $\tau: \mathcal{B} \to \mathcal{C}$, 这个映射可线性扩充到整个 M 到整个 N 之上的同构, 故 $M \approx N$. ■

当 R– 模的秩数 n 为有限时, 易见 $M \approx R^n$.

5. 由于有限生成 R– 模 M 的子模未必是有限生成的, 因此, 要讨论在什么情况下有限生成 R– 模 M 的子模也是有限生成的. 这个条件是升链条件.

定义 A4.1.4　一个 R– 模 M 称为满足**子模的升链条件 (ascending chain condition)**, 如果对 M 的任意上升子模序列

$$S_1 \subseteq S_2 \subseteq S_3 \subseteq \cdots$$

存在指标 k, 使得 $S_k = S_{k+1} = S_{k+2} = \cdots$. 子模升链条件简记为 $a.c.c.$

定理 A4.1.3　R– 模 M 的每个子模是有限生成的当且仅当 M 满足子模的升链条件.

定理中的模称为 **Noether 模 (Noetherian module)**.

证明　设 M 的所有子模都是有限生成的, 而 M 有无穷上升子模序列

$$S_1 \subseteq S_2 \subseteq S_3 \subseteq \cdots$$

易见 $S = \bigcup_j S_j$ 也是 M 的子模, 故 S 也是有限生成的. 设 $S = \langle u_1, \cdots, u_n \rangle$, $u_i \in M$. 由于 $u_i \in S$, 故有指标 k_i, 使得 $u_i \in S_{k_i}$. 令 $k = \max\{k_1, \cdots, k_n\}$, 则 $u_i \in S_k$. 因此

$$S = \langle u_1, \cdots, u_n \rangle \subseteq S_k \subseteq S_{k+1} \subseteq \cdots \subseteq S.$$

这表明上升子模序列 $S_1 \subseteq S_2 \subseteq S_3 \subseteq \cdots$ 从 S_k 起全是相同的.

反之, 若 M 满足子模的升链条件, 且 S 是 M 的子模. 取 $u_1 \in S$, 考虑子模 $S_1 = \langle u_1 \rangle \subseteq S$. 若 $S_1 = S$, 则 S 就是有限生成的. 若 $S_1 \neq S$, 于是有 $u_2 \in S - S_1$. 令 $S_2 = \langle u_1, u_2 \rangle$. 若 $S_2 = S$, 则 S 是有限生成的. 若 $S_2 \neq S$, 则有 $u_3 \in S - S_2$. 考虑子模 $S_3 = \langle u_1, u_2, u_3 \rangle$. 一直这样进行下去, 就得到一个子模的上升链

$$\langle u_1 \rangle \subseteq \langle u_1, u_2 \rangle \subseteq \langle u_1, u_2, u_3 \rangle \subseteq \cdots \subseteq S.$$

如果这样的子模没有一个等于 S, 就得到一个子模的无穷上升序列, 前一个为后一个所真包含, 这与 M 满足子模的升链条件相矛盾. 故有某个 n, 使得

$S=\langle u_1,\cdots,u_n\rangle$，也就是 S 是有限生成的的. ■

由于环 R 是自己上的 $R-$ 模，且模 R 的子模就是环 R 的理想，故定义 A4.1.4 与定理 A4.1.3 成为如下定义.

定义 A4.1.5　环 R 称为满足**理想的升链条件 (ascending chain condition)**，如果对 R 的任意上升理想序列

$$Q_1\subseteq Q_2\subseteq Q_3\subseteq\cdots$$

存在指标 k，使得 $Q_k=Q_{k+1}=Q_{k+2}=\cdots$.

在 A1.2 中证明了主理想整环一定满足升链条件.

定理 A4.1.4　环 R 的每个理想（作为 $R-$ 模）是有限生成的当且仅当 R 满足理想的升链条件.

定理中的环称为 **Noether 环 (Noetherian ring)**. 下面要证明一条重要定理.

定理 A4.1.5　若 R 是 Noether 环，则任意有限生成的 $R-$ 模是 Noether 模.

这条定理说，若 R 是 Noether 环，即每个理想是有限生成的，则有限生成的 $R-$ 模的每个子模也是有限生成的. 这就给出了条件使有限生成的模的子模依然是有限生成的.

证明　若 M 是有限生成 $R-$ 模，$M=\langle u_1,\ldots,u_n\rangle$. 考虑满同态 $\tau\colon R^n\to M$ 定义为，设 $r_i\in R$，

$$\tau(r_1,\ldots,r_n)=r_1u_1+\cdots+r_nu_n.$$

若 S 是 M 的一个子模，则

$$\tau^{-1}(S)=\left\{r=(r_1,\ldots,r_n)\in R^n\ \middle|\ \tau(r)\in S\right\}$$

是 R^n 的一个子模，且 $\tau\left(\tau^{-1}(S)\right)=S$. 假设 R^n 只有有限生成子模，则 $\tau^{-1}(S)$ 是有限生成的，$\tau^{-1}(S)=\langle t_1,\ldots,t_k\rangle$. 于是若 $w\in S$，则有 $t\in\tau^{-1}(S)$，使得 $w=\tau(t)$. 由于

$$t=a_1t_1+\cdots+a_kt_k,$$

这里 $a_i\in R$，故

$$w=\tau(t)=a_1\tau(t_1)+\cdots+a_k\tau(t_k).$$

于是 S 由 $\left\{\tau(t_1),\ldots,\tau(t_k)\right\}$ 所有限生成. 所以余下要证得只是 R^n 的每个子模是有限生成的. 采用数学归纳法.

当 $n=1$ 时，这是对的，因为 R 是 Noether 环. 假设对所有的 $1\leqslant k<n$，R^k 只有有限生成子模，若 S 是 R^n 的子模，令

$$S_1=\left\{\alpha\in S\ \middle|\ \alpha=(s_1,\ldots,s_{n-1},0),\ 而\ s_1,\ldots,s_{n-1}\in R\right\}$$

及

$$S_2=\left\{(0,\ldots,0,s_n)\ \middle|\ (s_1,\ldots,s_{n-1},s_n)\in S,\ 而\ s_1,\ldots,s_{n-1}\in R\right\}.$$

于是 S_1 同构于 S^{n-1} 的一个子模（只要将 S_1 的每个元素的最后一个坐标去掉）. 由归纳假设，S_1 是有限生成的，令 $S_1=\langle\mathcal{B}\rangle$，而 $\mathcal{B}=\left\{\alpha_1,\ldots,\alpha_k\right\}$，$0\leqslant k\leqslant n-1$（若 S_1 是空

集，则 $\mathcal{B}$ 为空集而 $k=0$）.

同样 S_2 同构于 R 的一个子模（即理想），因此是有限生成的．若 S_2 是平凡的，则 S 的每个元素的最后一个坐标为 0，故 $S=S_1$ 是有限生成的．若 S_2 非平凡，而由 $(0,\ldots,0,b_n)$，$b_n\neq 0$ 所生成．在 S 中有 $\beta=(b_1,\ldots,b_{n-1},b_n)\in S$，则 $\mathcal{B}'=\{\beta\}\cup\mathcal{B}$ 生成 S。这是因为，若 $\alpha=(s_1,\ldots,s_n)\in S$，则 $(0,\ldots,0,s_n)\in S_2$，故有 $r\in R$，使得 $(0,\ldots,0,s_n)=r(0,\ldots,0,b_n)$，即 $s_n=rb_n$，因此 $\alpha-rb\in S_1$，于是 $\alpha\in rb+S_1$，这就是 $\mathcal{B}'$ 生成 S. ■

定理 A4.1.5 导致去研究哪些环是 Noether 环．可以证明，有单位元的交换环 R 是域当且仅当它只有理想 $\{0\}$ 及 R 而无其它理想．因此，由定理 A4.1.4，**域当然是 Noether 的**．当 R 是主理想整环时，由于所有的理想都是主理想，因此，**主理想整环也是 Noether 的**．下面给出十分重要的 Hilbert 基定理，这是一条十分有用的基本定理.

定理 A4.1.6（Hilbert 基定理） 若环 R 是一个 Noether 环，则 R 上的多项式环 $R[x]$ 也是一个 Noether 环.

证明 要证的是 $R[x]$ 的任意一个理想 Q 是有限生成的．令 L_j 为 Q 中所有 j 次多项式的最高项的系数及 R 中的零元素所作成的集合，$j=0,1,\ldots$．则容易看出，L_j 都是 R 的理想．事实上，对 $a,b\in L_j$，$a\neq 0$，$b\neq 0$，存在 j 次多项式 $f(x)=ax^j+\cdots$ 和 $g(x)=bx^j+\cdots$ 都属于 Q．于是 $f(x)-g(x)=(a-b)x^j+\cdots$ 必属于 Q．若 $a-b\neq 0$，则 $a-b\in L_j$；若 $a-b=0$，由于 $0\in L_j$，故必有 $a-b\in L_j$．对于任意 $c\in R$，$cf(x)=cax^j+\cdots\in Q$．若 $ca\neq 0$，则 $ca\in L_j$；若 $ca=0$，由于 $0\in L_j$，故必有 $ca\in L_j$，$a\in L_j$ 是显然的．于是 L_j 是 R 的一个理想.

若 $f(x)=ax^j+\cdots\in Q$，由于 Q 是 $R[x]$ 的一个理想，$xf(x)=ax^{j+1}+\cdots\in Q$．因此，若 $a\in L_j$，则 $a\in L_{j+1}$．于是得到一个理想升链 $L_0\subseteq L_1\subseteq L_2\subseteq\cdots$．因为 R 是 Noether 环，故存在 d 使得

$$L_d=L_{d+1}=\cdots=L:=\bigcup_{j\geqslant 0}L_j.$$

因为 R 是 Noether 环，故每个 L_j，$j=0,1,\ldots,d$ 均是有限生成的．设

$$L_j=\left\langle a_j^{(1)},\cdots,a_j^{(k_j)}\right\rangle,$$

从而由 L_j 的定义可知，有 Q 中的 j 次多项式 $f_j^{(1)},\ldots,f_j^{(k_j)}$ 使得 $f_j^{(i)}$ 的首项系数为 $a_j^{(i)}$，这里 $i=1,\ldots,k_j$.

可以证明，

$$S=\left\{f_0^{(1)},\ldots,f_0^{(k_0)},f_1^{(1)},\ldots,f_1^{(k_1)},\ldots\ldots,f_d^{(1)},\ldots,f_d^{(k_d)}\right\}$$

是理想 Q 的生成集．为此，设 $f\in Q$，$\deg f=n$．对 n 用数学归纳法．当 $n=0$ 时，显然 S 中的元生成 f．设

$$f=a_nx^n+\cdots+a_1x+a_0,\quad a_n\neq 0,$$

则 $a_n \in L_n \subseteq L = L_d$.

若 $n < d$,则

$$a_n = \sum_{1\leqslant i\leqslant k_n} r_i a_n^{(i)}, \quad r_i \in R.$$

于是多项式

$$h = f - \sum_{1\leqslant i\leqslant k_n} r_i f_n^{(i)}$$

的次数小于 n,从而由归纳假设 h 可由 S 中的元生成,进而 f 可由 S 中的元生成.

若 $n \geqslant d$,则 $a_n \in L_n = L_d$,故

$$a_n = \sum_{1\leqslant i\leqslant k_d} r_i a_d^{(i)}.$$

于是多项式

$$h = f - \sum_{1\leqslant i\leqslant k_d} r_i x^{n-d} f_d^{(i)}$$

的次数小于 n,同样由归纳假设 h 可由 S 中的元生成,从而 f 可由 S 中的元生成. ■

A4.2 主理想整环上的模

1. 在上一节中给出了有单位元的交换环 R 上的模的定义以及它的一些基本性质. 当环 R 是域时,模就是向量空间,至于一些向量空间中的一些基本概念与定理,有些可以移植到模上来. 如子空间、商空间、直和、生成集、线性无关、基、维数、线性变换、核与像等等,到了模理论中有相应的名词与定义. 例如,与子空间、商空间、维数与线性变换相对应的为子模、商模、秩与模同态等. 但是摸与向量空间,从表面上看,只是建立在环与域上的差别,但相互之间有着本质的差异. 这里列举一些事实,而不加以证明.

(1) 存在这样的模,它无线性无关的元素,当然也就没有基. 这就是引入自由模的原因;

(2) 存在这样的模,它的子模无补集;

(3) 存在这样的有限生成模,其子模不是有限生成的. 这就是引入 Noether 模的原因;

(4) 存在这样的模,它的线性相关集 S 中的元素不能用 S 中其它元素的线性组合来表示;

(5) 存在这样的模,其极小生成元集不是模的基,其极大线性无关集不是模的基;

(6) 存在这样的自由模,其子模不自由,其商模不自由;

(7) 若 V 是域 $\mathbb{F}$ 上的向量空间,$r \in \mathbb{F}, v \in V, r, v \neq 0$,则 $rv \neq 0$. 但在模的情形,这不再永远成立;

(8) 存在这样的自由模,有线性无关集不包含在基中,有生成集不包含有基.

由于在一般的有单位元的交换环 R 上的模有种种不很理想的性质,这导致我们专门讨论主理想整环上的模的理论. 这是因为主理想整环有很多很好的性质,这就保证了在其上的模也有很多很好的性质,情况大为改观. 例如上述 (6) 中,在主理想整环上自由模的子模也是自由的,等等. 在这一讲中对主理想整环上的模进行研究之后,在下一讲中以此来考虑向量空间中的一些问题,可以得到一系列重要的结果.

2. 来证明如下定理.

定理 A4.2.1 主理想整环 R 上的自由模 M 的任意子模 S 也是自由的,且

$$\operatorname{rank} S \leqslant \operatorname{rank} M.$$

证明 只证模的秩是有限的情形,尽管秩为无限时这也成立.

由定理 A4.1.2 可知,若 $\operatorname{rank} M = n$,则 $M \approx R^n$. 因此,直接来讨论 R^n 即可.

对 n 用归纳法. 当 $n=1$ 时,$M=R$,R 的任意子模 S 就是 R 的理想. 由于 R 是主理想整环,故 S 是主理想,即 $S=\langle a\rangle$. 设 $S\neq\{0\}$. 因 R 是整环,故对所有 $r\neq 0$ 有 $ra\neq 0$. 因此,映射

$$\sigma:\begin{array}{ccc} R & \longrightarrow & S; \\ r & \longmapsto & ra, \end{array}$$

是 R 到 S 的同构,故 S 是自由的.

假设当 $k<n$ 时 R^k 的子模是自由的,S 是 R^n 的子模. 考虑

$$S_1=\left\{\alpha\in S \,\middle|\, 有\ s_1,\ldots,s_{n-1}\in R,\ \alpha=(s_1,\ldots,s_{n-1},0)\right\}$$

及

$$S_2=\left\{(0,\ldots,0,s_n) \,\middle|\, 有\ s_1,\ldots,s_{n-1}\in R,\ (s_1,\ldots,s_{n-1},s_n)\in S\right\}.$$

因为 S_1 同构于 R^{n-1} 的一个子模,由归纳假设,这是自由的. 设 $\mathcal{B}=\left\{\alpha_1,\ldots,\alpha_k\right\}$ 是 S_1 的基,而 $k\leqslant n-1$(若 S_1 是平凡的,则 $\mathcal{B}$ 是空集).

同样,S_2 同构于 R 的一个子模(理想). 若 S_2 是平凡的,则 S 中每个元素的最后坐标为 0,故 $S=S_1$ 是自由的. 若 S_2 非平凡,则有秩 1,基为 $\left\{(0,\ldots,0,t_n)\right\}$,其中 $t_n\neq 0$ 且 $\tau=(t_1,\ldots,t_n)\in S$.

来证 $\mathcal{B}'=\{\tau\}\cup\mathcal{B}$ 是 S 的基. 先证 $\mathcal{B}'$ 是线性无关的. 若有 $r,r_1,\ldots,r_k\in R$,使

$$r\tau+r_1\alpha_1+\cdots+r_k\alpha_k=0,$$

则 $r\tau=-(r_1\alpha_1+\cdots+r_k\alpha_k)$. 比较两边的最后坐标,有 $rt_n=0$,故 $r=0$. 而 $\alpha_1,\ldots,\alpha_k$ 的线性无关性导出 $r_i=0$,故 $\mathcal{B}'$ 是线性无关的. 其次,若 $\alpha=(s_1,\ldots,s_n)\in S$,则 $(0,\ldots,0,s_n)\in S_2$,故

$$(0,\ldots,0,s_n)=r(0,\ldots,0,t_n),$$

这里 $r\in R$,故 $s_n=rt_n$. 于是 $\alpha-r\tau\in S_1$,即 $\alpha=r\tau+S_1$,因此 $\mathcal{B}'$ 生成 S. ■

在第 1 条之 (7) 中已经说到,域 $\mathbb{F}$ 上的向量空间 V,若 $r(\neq 0)\in\mathbb{F}$,$v(\neq 0)\in V$,则

$rv \neq 0$. 但在模的情形,这不一定成立. 于是有如下的定义.

定义 A4.2.1 设 R 是整环, M 是 R– 模. 对 $v \in M$, 如果有非零的 $r \in R$ 使得 $rv = 0$, 则称 v 是 M 的一个**挠元 (torsion element)**. 一个模如果无挠元则称为**无挠的 (torsion free)**. 如果模的所有元素都是挠元, 则称 M 是**挠模 (torsion module)**.

若 M 为模, 其所有的挠元组成的集合记作 $\operatorname{Tor} M$, 易见这是 M 的一个子模, 并且 $M/\operatorname{Tor} M$ 是无挠模. 事实上, 若 $a, b \in \operatorname{Tor} M$, 则有 $\alpha, \beta \in R$, $\alpha \neq 0$, $\beta \neq 0$ 使得 $\alpha a = 0$ 及 $\beta b = 0$. 于是对任意的 $r, s \in R$, 有 $\alpha\beta(ra + sb) = 0$, 即 $ra + sb \in \operatorname{Tor} M$. 由 A4.1 第 2 条之 (1) 可知, $\operatorname{Tor} M$ 是一个子模. 再证 $M/\operatorname{Tor} M$ 是无挠模. 事实上, 若 $a + \operatorname{Tor} M$ 是 $M/\operatorname{Tor} M$ 中的一个挠元, 则有 $r(\neq 0) \in R$, 使得 $r(a + \operatorname{Tor} M) = 0$, 即 $ra \in \operatorname{Tor} M$. 于是有 $s(\neq 0) \in R$, 使得 $s(ra) = (sr)a = 0$, 且 $sr \neq 0$, 故 $a \in \operatorname{Tor} M$, 于是 $M/\operatorname{Tor} M$ 中的挠元是 0, 即 $M/\operatorname{Tor} M$ 是无挠模.

其次, 也易见, 主理想整环上任意自由模是无挠的. 反之不真, 但是有如下的定理.

定理 A4.2.2 主理想整环上的模 M 如果是无挠的, 且是有限生成的, 则模是自由的.

证明 由于 M 是有限生成的, 故有 $0 \neq v_i \in M$, 使得 $M = \langle v_1, \ldots, v_n \rangle$. 在这些生成元中取极大线性无关子集 $S = \{u_1, \ldots, u_k\}$, 将 M 的生成元重新写成

$$M = \langle u_1, \ldots, u_k, v_1, \ldots, v_{n-k} \rangle.$$

于是对每个 v_i, 集合 $\{u_1, \ldots, u_k, v_i\}$ 是线性相关的, $i = 1, \ldots, n-k$. 故对每个 v_i, 有 $a_i \neq 0$ 及 $r_1^{(i)}, \ldots, r_k^{(i)}$ 使得

$$a_i v_i + r_1^{(i)} u_1 + \cdots + r_k^{(i)} u_k = 0.$$

令 $a = a_1 \cdots a_{n-k}$, 则 $av_i = \operatorname{span} S$, $i = 1, \ldots, n-k$. 于是 $aM = \{av \mid v \in M\}$ 是 $\operatorname{span} S$ 的一个子模. 但 $\operatorname{span} S$ 是自由模, 基为 S, 故由定理 A4.2.1, aM 也是自由的. 由于 $M \approx aM$, 这是因为

$$\tau(v) = av$$

是满射同态, 由于 M 是无挠模, 故 τ 也是单射. 故 $M \approx aM$, M 是自由的. ■

A4.3　主理想整环上的有限生成模的分解定理

有了以前这些准备之后，要进入本讲的主题，给出主理想整环上有限生成模的分解定理. 这需要三个步骤来建立起这些重要的定理.

第一步，将主理想整环上有限生成模分解为挠模与自由模之直和.

定理 A4.3.1　若 M 是主理想整环 R 上的有限生成模，则

$$M = \operatorname{Tor} M \oplus \operatorname{Free} M,$$

这里 $\operatorname{Free} M$ 是一个自由 R– 模. 并且这种分解是唯一的，即若还有分解 $M = T \oplus N$，这里 T 是 M 的挠子模，而 N 是 M 的自由子模，则 $T = \operatorname{Tor} M$，$N \cong \operatorname{Free} M$.

证明　由于商模 $M/\operatorname{Tor} M$ 是无挠的以及当 M 是有限生成时，$M/\operatorname{Tor} M$ 也是有限生成的，故由定理 A4.2.2 可知 $M/\operatorname{Tor} M$ 是自由模.

考虑将 M 映到自由模 $M/\operatorname{Tor} M$ 的满同态 $\pi: M \to M/\operatorname{Tor} M$. 若 $\mathcal{B}$ 是 $M/\operatorname{Tor} M$ 的基，对每个 $b \in \mathcal{B}$，取定一个 $b' \in M$ 使得 $\pi(b') = b$. 令 $\mathcal{B}'$ 为 M 中所有这样的元素的集合. 显然 $\mathcal{B}'$ 是线性无关的，于是 $S = \operatorname{span} \mathcal{B}'$ 是 M 的子模，且同构于 $M/\operatorname{Tor} M$. 显然 $\operatorname{Ker} \pi = \operatorname{Tor} M$，$\operatorname{Im} \pi = M/\operatorname{Tor} M$. 而

$$\begin{aligned} v \in \operatorname{Tor} M \cap S = \operatorname{Ker} \pi \cap S &\Longrightarrow \pi(v) = 0,\ \ v = \sum r_i b_i',\ \ r_i \in R \\ &\Longrightarrow 0 = \sum r_i \pi(b_i') = \sum r_i b_i \\ &\Longrightarrow \text{对所有 } i,\ \ r_i = 0 \Longrightarrow v = 0, \end{aligned}$$

即 $\operatorname{Tor} M \cap S = \{0\}$. 若 $v \in M$，则 $\pi(v) = \sum s_i b_i$，而 $s_i \in R$. 令 $u = \sum s_j b_j' \in S$，则

$$\pi(v - u) = \pi(v) - \pi\Big(\sum s_j b_j'\Big) = \sum s_j b_j - \sum s_j b_j = 0.$$

于是 $x = v - u \in \operatorname{Ker} \pi = \operatorname{Tor} M$. 故 $v = x + u \in \operatorname{Tor} M + S$. 即得 $M = \operatorname{Tor} M \oplus S$. 由于 $S \approx M/\operatorname{Tor} M$，因此，$M = \operatorname{Tor} M \oplus \operatorname{Free} M$，这里 $\operatorname{Free} M$ 为 S.

若 $M = T \oplus N$，这里 T 为 M 的挠子模，N 是 M 的自由子模，则由定义 $T \subseteq \operatorname{Tor} M$. 设 $v \in \operatorname{Tor} M$，则 $v = t + n$，其中 $t \in T$，$n \in N$. 于是 $n = v - t \in \operatorname{Tor} M$，但 n 属于自由模 N，故 $n = 0$，即 $v \in T$. 于是 $T = \operatorname{Tor} M$. 从而

$$N \approx M/\operatorname{Tor} M \approx \operatorname{Free} M. \qquad \blacksquare$$

第二步，将主理想整环上有限生成模分解为准素子模的直和.

在 A2.2 中，曾引入了向量空间的零化子的概念. 现将此概念扩充到模上.

定义 A4.3.1　若 M 是一个 R– 模，$v \in M$ 的**零化子**为

$$\operatorname{ann}(v) = \{r \in R \mid rv = 0\},$$

M 的**零化子**为

$$\operatorname{ann} M = \left\{ r \in R \,\middle|\, rM = \{0\} \right\},$$

这里 $rM = \left\{ rv \,\middle|\, v \in M \right\}$.

显然 $\operatorname{ann}(v)$ 及 $\operatorname{ann} M$ 是 R 中的理想.

若 M 是主理想整环上的有限生成模, $\operatorname{ann}(v)$ 的生成元称为 v 的阶 **(order)**, $\operatorname{ann} M$ 的生成元称为 M 的阶.

显然, 若 μ, v 是 M (或 $v \in M$) 的两个阶, 则它们是**相伴的 (associate)**, 即

$$\operatorname{ann} M = \langle \mu \rangle = \langle v \rangle \implies \mu = uv, \text{ 存在某个可逆元 } u \in R.$$

故 M 的阶, 除去可逆元外, 是唯一确定的. μ, v 除去可逆元外, 由定理 A1.2.3 有相同的素元乘积分解.

定义 A4.3.2　模 M 称为**准素模 (primary module)**, 若其零化子 $\operatorname{ann} M = \langle p^e \rangle$, 这里 p 为素元, 而 e 是正整数. 即 M 是准素的, 若其阶是一个素元的正次方.

显然, 主理想整环上有限生成模 M 是准素的当且仅当 M 中每个元素的阶是一个固定的素元的幂.

分解定理的第二步是将挠模 M 分解为准素子模的直和.

定理 A4.3.2 (准素唯一分解定理) 若 M 是一个主理想整环上非零有限生成挠模, 阶为

$$\mu = p_1^{e_1} \cdots p_n^{e_n},$$

这里 p_i 为互不相伴的素元, 则 M 可分解为直和

$$M = M_{p_1} \oplus \cdots \oplus M_{p_n},$$

这里

$$M_{p_i} = \left\{ v \in M \,\middle|\, p_i^{e_i} v = 0 \right\}$$

为准素子模, 阶为 $p_i^{e_i}$. 进一步, 这样的分解是唯一的, 即, 若还有分解

$$M = N_1 \oplus \cdots \oplus N_m,$$

其中 N_i 是阶为 $q_i^{f_i}$ 的准素模, 则 $m = n$, 并且可适当安排下标 i 使得 $N_i = M_{p_i}$, q_i 与 p_i 相伴, 而 $e_i = f_i$.

证明　设 $\mu = pq$, 且 p, q 的最大公因子 $\gcd(p, q) = 1$. 考虑集合

$$M_p = \{ v \in M \mid pv = 0 \} \quad \text{及} \quad M_q = \{ v \in M \mid qv = 0 \}.$$

来证 $M = M_p \oplus M_q$, 及 M_p 与 M_q 分别有零化子 $\langle p \rangle$ 与 $\langle q \rangle$.

由于 p 与 q 互素, 故理想 $\langle p, q \rangle$ 由 $\gcd(p, q) = 1$ 生成 (证明如同命题 A1.2.1 的证明), 故 $1 \in \langle p, q \rangle$. 因此, 存在 $a, b \in R$, 使得

$$ap + bq = 1.$$

若 $v \in M_p \cap M_q$, 则 $pv = qv = 0$, 故

$$v = 1v = (ap + bq)v = 0.$$

因此，$M_p \cap M_q = \{0\}$.

对任意 $v \in M$，也有

$$v = 1v = apv + bqv,$$

而 $q(apv) = a(pq)v = a\mu v = 0$，故 $apv \in M_q$，同样 $bqv \in M_p$. 因此 $M = M_p \oplus M_q$.

若 $rM_p = 0$，则对任意 $v = v_1 + v_2 \in M_p \oplus M_q = M$，有

$$rqv = rq(v_1 + v_2) = qrv_1 + rqv_2 = 0,$$

因此，$rq \in \operatorname{ann} M$，这蕴涵 $\mu = pq \mid rq$，即 $p \mid r$，这说明 $\operatorname{ann} M_p = \langle p \rangle$. 同样可证 $\operatorname{ann} M_q = \langle q \rangle$.

若 μ 是素元的乘积

$$\mu = p_1^{e_1} \cdots p_n^{e_n},$$

由刚才的证明可知

$$M = M_{p_1^{e_1}} \oplus N,$$

这里 N 为具有零化子 $\langle \mu / p_1^{e_1} \rangle$ 的子模. 重复这个步骤，记 $M_{p_i^{e_i}}$ 为 M_{p_i}，就得到准素分解.

至于分解的唯一性，注意到由 $M = N_1 \oplus \cdots \oplus N_m$ 可知 $\operatorname{ann} M = \langle q_1^{f_1}, \ldots, q_n^{f_n} \rangle$. 因此 $q_1^{f_1} \cdots q_m^{f_m}$ 与 $p_1^{e_1} \cdots p_1^{e_n}$ 相伴. 由定理 A1.2.3，R 是唯一因子分解环，故 $n = m$，且可适当安排下标 i 使得 q_i 与 p_i 相伴. 从而

$$N_i = \left\{ v \in M \mid q_i^{f_i} v = 0 \right\} = \left\{ v \in M \mid p_i^{e_i} v = 0 \right\} = M_{p_i}. \qquad \blacksquare$$

第三步，由定理 A4.3.2 可知，下一步就应该对定理 A4.3.2 中的那些准素子模 M_{p_i} 进行分解，$i = 1, \ldots, n$.

定理 A4.3.3（循环分解定理） 若 M 是主理想整环 R 上非零准素有限生成模，其阶为 p^e，则 M 可分解为循环子模的直和

$$M = C_1 \oplus \cdots \oplus C_n. \tag{A4.3.1}$$

其中诸 C_i 为有阶 p^{e_i} 的循环子模，且满足

$$e = e_1 \geqslant e_2 \geqslant \cdots \geqslant e_n \geqslant 1,$$

或等价地

$$p \mid p^{e_n}, p^{e_n} \mid p^{e_{n-1}}, \ldots, p^{e_2} \mid p^{e_1}. \tag{A4.3.2}$$

证明 先来证明在 M 中一定存在一个元素 v_1，使得

$$\operatorname{ann}(v_1) = \operatorname{ann} M = \langle p^e \rangle.$$

事实上，如果这样的 v_1 不存在，那么对所有的 $v \in M$，都有 $\operatorname{ann}(v) = \langle p^k \rangle$，而 $k < e$，故 $p^{e-1} \in \operatorname{ann} M$. 这将导致 $p^e \mid p^{e-1}$，矛盾.

如果能证明循环子模 $\langle v_1 \rangle$ 是 M 分解中的一个被加项，即

$$M = \langle v_1 \rangle \oplus S_1, \tag{A4.3.3}$$

这里 S_1 是 M 中的某个子模，于是 S_1 也是一个在 R 上的有限生成准素挠模，以致可

以重复这个步骤,得到

$$M = \langle v_1 \rangle \oplus \langle v_2 \rangle \oplus S_2.$$

这里 $\mathrm{ann}(v_2) = \langle p^{e_2} \rangle$,而 $e_2 \leqslant e_1$. 这样一直进行下去,得到一个上升子模序列

$$\langle v_1 \rangle \subseteq \langle v_2 \rangle \subseteq \langle v_3 \rangle \subseteq \cdots.$$

由于 R 是主理想整环,故 R 是 Noether 环. 由于 M 是有限生成的,根据定理 A4.1.3 可知 M 是 Noether 模. 由 定理 A4.1.4,M 满足升链条件,于是上述子模链到有限步停止. 这就证明了 M 可以分解为循环子模 $\langle v_i \rangle$ 的直和,其相应的阶为 p^{e_i},且 $e = e_1 \geqslant e_2 \geqslant \cdots \geqslant 1$.

现在就着手证明 M 确实可以分解为 $M = \langle v_1 \rangle \oplus S_1$. 事实上,由于 M 是有限生成的,故有 $M = \langle v_1, v_2, \ldots, v_k \rangle$. 对 k 进行归纳法. 若 $k = 0$,则只要令 $S_1 = \{0\}$ 即可. 假设结论对 k 成立,设

$$M = \langle v_1, u_1, \ldots, u_k, u \rangle.$$

由归纳假设

$$\langle v_1, u_1, \ldots, u_k \rangle = \langle v_1 \rangle \oplus S_0,$$

而 S_0 是一个子模. 将 $u - \alpha v_1, \alpha \in R$,替代 u,不会影响生成模 M,即

$$\langle v_1, u_1, \ldots, u_k, u - \alpha v_1 \rangle = \langle v_1, u_1, \ldots, u_k, u \rangle = M.$$

于是可以寻求 $\alpha \in R$,使得

$$\langle v_1 \rangle \cap \langle u - \alpha v_1, S_0 \rangle = \{0\},$$

这样就得到

$$M = \langle v_1 \rangle \oplus \langle u - \alpha v_1, S_0 \rangle =: \langle v_1 \rangle \oplus S_1.$$

也就是令 $S_1 = \langle u - \alpha v_1, S_0 \rangle$ 就可以了. $\langle u - \alpha v_1, S_0 \rangle$ 中的元素形为 $r(u - \alpha v_1) + s_0$,于是 $\langle v_1 \rangle \cap \langle u - \alpha v_1, S_0 \rangle = \{0\}$ 等价于对于 $r \in R, s_0 \in S_0$,有

$$r(u - \alpha v_1) + s_0 \in \langle v_1 \rangle \implies r(u - \alpha v_1) + s_0 = 0,$$

这也等价于

$$r(u - \alpha v_1) \in \langle v_1 \rangle \oplus S_0 \implies r(u - \alpha v_1) \in S_0.$$

即

$$ru \in \langle v_1 \rangle \oplus S_0 \implies r(u - \alpha v_1) \in S_0, \tag{A4.3.4}$$

易证

$$\mathcal{I} = \left\{ r \in R \,\middle|\, ru \in \langle v_1 \rangle \oplus S_0 \right\}$$

是 R 的一个理想,故为主理想,因此 $\mathcal{I} = \langle a \rangle$. 但是

$$p^e u = 0 \in \langle v_1 \rangle \oplus S_0,$$

故 $p^e \in \langle a \rangle$,这蕴涵 $a \mid p^e$,故有 $f \leqslant e$,使得 $a = p^f$. 于是有

$$\begin{aligned} ru \in \langle v_1 \rangle \oplus S_0 &\implies r \in \mathcal{I} \implies r = qp^f, \quad \text{其中 } q \in R, \\ &\implies r(u - \alpha v_1) = qp^f(u - \alpha v_1). \end{aligned}$$

因此,如能找到 $\alpha \in R$ 使得

$$p^f(u-\alpha v_1) \in S_0, \tag{A4.3.5}$$

则 (A4.3.4) 得证.

由于 $p^f \in \mathcal{I}$,故 $p^f u \in \langle v_1 \rangle \oplus S_0$ 可写为

$$p^f u = tv_1 + s_0, \tag{A4.3.6}$$

这里 $t \in R$ 而 $s_0 \in S_0$. 于是 (A4.3.5) 成为

$$tv_1 + s_0 - \alpha p^f v_1 \in S_0,$$

即

$$(t - \alpha p^f)v_1 \in S_0.$$

上式成立当且仅当 $t - \alpha p^f = 0$,即当且仅当 $p^f \mid t$.

由 (A4.3.6),有

$$0 = p^{e-f} p^f u = p^{e-f} t v_1 + p^{e-f} s_0.$$

由于 $\langle v_1 \rangle \cap S_0 = \{0\}$,故 $p^{e-f} t v_1 = 0$. 由于 v_1 的阶为 p^e,故 $p^e \mid p^{e-f} t$,即 $p^f \mid t$,这正是我们所需要的. 于是 (A4.3.3) 得证,从而 (A4.3.1) 得证.

(A4.3.1) 蕴涵 (A4.3.2). 事实上,由于

$$p^e \in \operatorname{ann} M \subseteq \operatorname{ann} C_i,$$

故若 $\operatorname{ann} C_i = \langle \alpha_i \rangle$,则 $\alpha_i \mid p^e$. 于是 $\alpha_i = p^{e_i}$,而 $e_i \leqslant e$. 对 $e_i, i = 1, \ldots, n$,重新排列,就可以得到 (A4.3.2). ■

从定理 A4.3.3 证明的过程中可以看出这样的分解是不唯一的,虽然如此,除去乘上可逆元,阶 p^{e_i} 是唯一确定的,素元 p 也是唯一确定的,因为它要除尽 M 的阶 p^e. 于是有如下的唯一性定理.

定理 A4.3.4(循环分解唯一性定理) 若 M 是主理想整环 R 上一个非零的有限生成挠模,其阶为 p^e. 若 M 可分解为

$$M = C_1 \oplus \cdots \oplus C_n,$$

这里 C_i 是阶为 p^{e_i} 的非零循环子模,且 $e_1 \geqslant \cdots \geqslant e_n \geqslant 1$.

若 M 还可以分解为

$$M = D_1 \oplus \cdots \oplus D_m,$$

这里 D_i 是阶为 p^{f_i} 的非零循环子模,且 $f_1 \geqslant \cdots \geqslant f_m \geqslant 1$,则有 $n = m$ 及

$$e_1 = f_1, \ldots, e_n = f_n.$$

为了证明这条唯一性定理,需要用到以下这些容易证明的结果. 设 R 是主理想整环.

(1) 在 A2.3 中第 3 条关于向量空间的同构定理都可以推广到主理想整环 R 上的 R– 模上来. 例如 R– 模上的第一同构定理为:若 M, N 为主理想整环 R 上的两个 R– 模,映射 $\tau \in \operatorname{Hom}_R(M, N)$,则 $M/\operatorname{Ker}\tau \approx \operatorname{Im}\tau$;

(2) 若 $\langle v\rangle$ 是一个循环 R- 模, $\mathrm{ann}(v)=\langle a\rangle$, 则映射

$$\tau:\begin{array}{ccc} R & \longrightarrow & \langle v\rangle, \\ r & \longmapsto & rv, \end{array}$$

是满射同态, 其核为 $\langle a\rangle$, 故由 (1) 中的第一同构定理, 有

$$\langle v\rangle\approx R/\langle a\rangle.$$

若 a 是素元, 则 $\langle a\rangle$ 是 R 中的极大理想, 故由引理 A4.1.1, $R/\langle a\rangle$ 是域;

(3) 若 $p\in R$ 是素元, M 是这样的一个 R- 模, 使得 $pM=\{0\}$, 则 M 是 $R/\langle p\rangle$ 上的一个向量空间, 其数乘定义为, 对所有的 $v\in M$,

$$(r+\langle p\rangle)v=rv;$$

(4) 若 $p\in R$ 是素元, 对于 R- 模 M 的任意子模 S, 集合

$$S^{(p)}=\left\{v\in S\,\middle|\,pv=0\right\}$$

是 M 的一个子模. 若 $M=S\oplus T$, 则 $M^{(p)}=S^{(p)}\oplus T^{(p)}$.

定理 A4.3.4 的证明 先证 $n=m$. 由 (4),

$$M^{(p)}=C_1^{(p)}\oplus\cdots\oplus C_n^{(p)} \quad 及 \quad M^{(p)}=D_1^{(p)}\oplus\cdots\oplus D_m^{(p)}.$$

由于 $pM^{(p)}=\{0\}$, 故由 (3), $M^{(p)}$ 是 $R/\langle p\rangle$ 上的一个向量空间. 由于 C_i, D_j 均为循环子模, 故 $C_i^{(p)}, D_j^{(p)}$ 均为 $M^{(p)}$ 这个向量空间的一维向量子空间, 故 $n=m$.

再证 $e_i=f_i, i=1,\ldots,n$. 对 e_1 进行数学归纳法.

设 $e_1=1$, 则所有的 $e_i=1$, 故 $pM=\{0\}$. 这样所有的 $f_i=1$, 因为若 $f_1>1$, 而 $D_1=\langle w\rangle$, 则 $pw\neq 0$, 矛盾.

若结论对 $e_1\leqslant k-1$ 都成立, 来证明当 $e_1=k$ 时结论也成立. 设

$$(e_1,\ldots,e_n)=(e_1,\ldots,e_s,1,\ldots,1), \quad e_s>1$$

及

$$(f_1,\ldots,f_n)=(f_1,\ldots,f_t,1,\ldots,1), \quad f_t>1.$$

则

$$pM=pC_1\oplus\cdots\oplus pC_s \quad 及 \quad pM=pD_1\oplus\cdots\oplus pD_t.$$

易见 pC_i 是 M 的循环子模及 $\mathrm{ann}(pC_i)=\langle p^{e_i-1}\rangle$. 这是因为, 若 $C_i=\langle v_i\rangle$, 则

$$pC_i=\left\{pc\,\middle|\,c\in C_i\right\}=\left\{prv_i\,\middle|\,r\in R\right\}=\left\{r(pv_i)\,\middle|\,r\in R\right\}=\langle pv_i\rangle,$$

而 pv_i 的阶为 p^{e_i-1}. 同样 pD_i 是 M 的循环子模及 $\mathrm{ann}(pD_i)=\langle p^{f_i-1}\rangle$. 特别 $\mathrm{ann}(pC_1)=\langle p^{e_1-1}\rangle$, 由归纳假设, 有

$$s=t \quad 及 \quad e_1=f_1,\ldots,e_s=f_s. \qquad ■$$

总结以上的三步, 得到

(1) 先将主理想整环 R 上的有限生成模分解为挠模与自由模之直和(定理 A4.3.1), 即

$$M = \operatorname{Tor} M \oplus \operatorname{Free} M,$$

这里 Free M 为 M 的一个自由子模,而 Tor M 为 M 中所有挠元组成的挠模;

(2) 若 Tor M 的阶为

$$\mu = p_1^{e_1} \cdots p_n^{e_n},$$

这里 p_i 为互不相伴的素元,则有准素分解(定理 A4.3.2):

$$\operatorname{Tor} M = M_{p_1} \oplus \cdots \oplus M_{p_n},$$

这里 M_{p_i} 为准素模,其阶为 $p_i^{e_i}$. 于是 M 有分解

$$M = M_{p_1} \oplus \cdots \oplus M_{p_n} \oplus \operatorname{Free} M.$$

(3) 由定理 A4.3.3 再将准素模 M_{p_i} 分解为循环子模的直和.

归纳起来有这样重要的两个不同形式的定理.

定理 A4.3.5(主理想整环上有限生成模的循环分解定理——初等因子形式) 若 M 是主理想整环 R 上的一个非零有限生成模,则

$$M = \operatorname{Tor} M \oplus \operatorname{Free} M,$$

这里 Tor M 是 M 中所有挠元的集合,Free M 是一个自由模,其秩由模 M 所唯一确定. 若 Tor M 有阶

$$\mu = p_1^{e_1} \cdots p_n^{e_n},$$

这里 p_i 为互不相伴的素元,则

$$\operatorname{Tor} M = M_{p_1} \oplus \cdots \oplus M_{p_n},$$

这里

$$M_{p_i} = \left\{ v \in M \,\middle|\, p_i^{e_i} v = 0 \right\}$$

为准素模,其阶为 $p_i^{e_i}$.

每个 M_{p_i} 可以分解为循环子模的直和

$$M_{p_i} = C_{i,1} \oplus \cdots \oplus C_{i,k_i},$$

这里 $C_{i,j}$ 的阶为 $p_i^{e_{i,j}}$,且

$$e_i = e_{i,1} \geqslant e_{i,2} \geqslant \cdots \geqslant e_{i,k_i} \geqslant 1,$$

将 M 的循环子模直和项 $C_{i,j}$ 的阶 $p_i^{e_{i,j}}$ 称为 M 的**初等因子 (elementary divisors)**. 除了乘以可逆元外,M 的初等因子由模 M 所唯一确定.

最终得到 M 可以分解为循环子模及一个自由模的直和

$$M = \left(C_{1,1} \oplus \cdots \oplus C_{1,k_1}\right) \oplus \cdots \oplus \left(C_{n,1} \oplus \cdots \oplus C_{n,k_n}\right) \oplus \operatorname{Free} M. \tag{A4.3.7}$$

定理 A4.3.5 的分解部分前面已经证完. 下面说明一下初等因子的唯一性. 根据定理 A4.3.1 中的唯一性部分,不妨设 Free $M = \{0\}$. 令

$$D_i = D_{i,1} \oplus \cdots \oplus D_{i,\ell_i},$$

则 D_i 的阶为 $q_i^{f_{i,1}}$ 的准素模. 于是 M 有如下两种准素分解

$$M = D_1 \oplus \cdots \oplus D_m = M_{p_1} \oplus \cdots \oplus M_{p_n}.$$

故由定理 A4.3.2 的唯一性部分可知，$n = m$，且不妨设 $D_i = M_{p_i}$，从而 p_i 与 q_i 相伴，而 $e_{i,1} = f_{i,1}$. 再由定理 A4.3.4 可知，$\ell_i = k_i$，$f_{i,j} = e_{i,j}$. 这就证明了分解的唯一性，即 M 的初等因子是由 M 唯一确定的.

这种分解还可以写成另一种形式. 设 S, T 是 M 的循环子模. 若 $\operatorname{ann} S = \langle a \rangle$ 及 $\operatorname{ann} T = \langle b \rangle$，且 $\gcd(a, b) = 1$，则 $S \cap T = \{0\}$，于是 $S \oplus T$ 也是一个循环子模，且 $\operatorname{ann}(S \oplus T) = \langle ab \rangle$.

在 (A4.3.7) 中，记

$$D_1 = C_{1,1} \oplus \cdots \oplus C_{n,1},$$

则 D_1 是一个循环子模，其阶为

$$q_1 = \prod_{i=1}^{n} p_i^{e_{i,1}}.$$

类似可以定义 $D_2, \ldots, D_m$，这里 $m = \max\limits_i(k_i)$.

于是有另一种形式的分解定理.

定理 A4.3.6（**主理想整环上有限生成模的循环分解定理——不变因子形式**）若 M 是主理想整环 R 上一个有限生成模，则

$$M = D_1 \oplus \cdots \oplus D_m \oplus \operatorname{Free} M,$$

这里 $\operatorname{Free} M$ 是 M 的一个自由模，而 D_i 是 M 的循环子模，其阶为 q_i，而且

$$q_m \mid q_{m-1}, q_{m-1} \mid q_{m-2}, \ldots, q_2 \mid q_1.$$

纯量 q_i 称为 M 的**不变因子 (invariant factor)**.

由定理 A4.3.5 的初等因子的唯一性部分容易看出除去可逆元，这些不变因子由 M 所唯一确定，并且 $\operatorname{Free} M$ 的秩由 M 所唯一确定.

向量空间在线性算子下的分解

上一讲研究模理论的目的是为了站在更高的层次上来认识线性代数,在这一讲中回到向量空间及线性变换,应用上一讲的有力的模理论来认识它们. 可以说在这一讲中只是将上一讲中的结果翻译成向量空间的语言. 这时 V 不仅仅是域 $\mathbb{F}$ 上的向量空间,更是多项式环 $\mathbb{F}[x]$ 上的模,其数乘定义为

$$p(x)v = p(\tau)(v),$$

这里 $p(x) \in \mathbb{F}[x], v \in V, \tau \in \mathcal{L}(V)$.

本讲中向量空间均是有限维的,以下不再每次注明.

A5.1 向量空间是主理想整环上的有限生成模

1. 若 V 是域 $\mathbb{F}$ 上的向量空间,线性算子 $\tau \in \mathcal{L}(V)$. 对于 V 中一个基,τ 对应于 $\mathbb{F}$ 上的一个矩阵. 对于 V 中另一个基,τ 对应于另一个矩阵,在 A3.1 中已经知道,这两个矩阵是相似的. 问题是: 对于一个固定的 $\tau \in \mathcal{L}(V)$,如何选取 V 的基,使得对应于 τ 的矩阵尽可能的简单. 当然最简单的矩阵是对角矩阵,但不是所有的 $\tau \in \mathcal{L}(V)$ 都能做到这点. 为此,只能退而求其次,找到另一种简单的矩阵.

上述的问题也可叙述为: 若 V 是域 $\mathbb{F}$ 上的向量空间,要找出所有与 $\mathcal{L}(V)$ 中给定的线性算子相对应的矩阵在相似意义下的标准形式.

这是线性代数中讨论的最基本问题之一.

首先,若 V 是域 $\mathbb{F}$ 上的 n 维向量空间,则 **V 作为 $\mathbb{F}[x]$– 模是挠模.**

显然 $\mathcal{L}(V)$ 同构于由所有 $n \times n$ 矩阵组成的向量空间 $\mathcal{M}_n(\mathbb{F})$. $\mathcal{M}_n(\mathbb{F})$ 的维数为 n^2,故对于 $\mathcal{L}(V)$ 中任一固定的 τ,n^2+1 个向量

$$\varepsilon, \tau, \tau^2, \ldots, \tau^{n^2}$$

是线性相关的,故在 $\mathbb{F}[x]$ 中存在多项式 $p(x)$,使得 $p(\tau) = 0$. 故 $p(x)v = \{0\}$. 因此,V 中所有元素是挠元.

其次,**V 作为 $\mathbb{F}[x]$– 模是有限生成模.**

若 $\mathcal{B} = (v_1, \ldots, v_n)$ 是向量空间 V 的一组基,则每个向量 $v \in V$ 有线性组合

$$v = r_1 v_1 + \cdots + r_n v_n,$$

这里 $r_i \in \mathbb{F} \subseteq \mathbb{F}[x]$,故 $\mathcal{B}$ 生成模 V.

在 A1.2 中已知 $\mathbb{F}[x]$ 是主理想整环,其中 $\mathbb{F}$ 是一个域. 因此,**向量空间 V 也是主理想整环 $\mathbb{F}[x]$ 上有限生成的挠模.** 所以上一讲中的分解定理能够应用.

2. 前面已经定义过,向量空间 V 的一个子空间 S,对一个固定的 $\tau \in \mathcal{L}(V)$ 来讲是不变的,如果 $\tau(S) \subseteq S$,S 称为关于 $\boldsymbol{\tau}$ **的不变子空间**.

易见,V 作为 $\mathbb{F}[x]$ 上的模,其子集 S 是子模当且仅当 S 是向量空间 V 关于 τ 的不变子空间.

固定 $\tau \in \mathcal{L}(V)$,模 V 的零化子为

$$\operatorname{ann} V = \{p(x) \in \mathbb{F}[x] \mid p(x)V = \{0\}\},$$

这是 $\mathbb{F}[x]$ 上的一个非零主理想(因为 $\mathbb{F}[x]$ 是主理想整环). 由于 V 的阶,即 $\operatorname{ann} V$ 的生成元,是相伴的,而 $\mathbb{F}[x]$ 的可逆元就是 $\mathbb{F}$ 中的非零元,故 V 有唯一的**首一阶 (monic order)**. 称这个唯一的首一阶,及生成 $\operatorname{ann} V$ 的唯一首一多项式,为 $\boldsymbol{\tau}$ **的极小多项式 (minimal polynomial)**,记作 $m_\tau(x)$,或 $\min \tau$. 于是

$$\langle \operatorname{ann} V \rangle = \langle m_\tau(x) \rangle$$

及

$$p(x)V = 0 \iff m_\tau(x) \mid p(x),$$

或

$$p(\tau) = 0 \iff m_\tau(x) \mid p(x),$$

在传统的线性代数教材中,并未引入模的概念,对线性算子 τ 的极小多项式定义为使得 $m_\tau(\tau) = 0$ 的最小次的唯一的首一多项式. 对矩阵也可定义极小多项式. 若 A 是域 $\mathbb{F}$ 上的一个方阵,A 的极小多项式 $m_A(x)$ 是使得 $p(A) = 0$ 的最小次的唯一的首一多项式 $p(x) \in \mathbb{F}[x]$.

显然如下的结论是成立的.

(1) 若 A 与 B 是相似矩阵,则 $m_A(x) = m_B(x)$,即极小多项式在相似下不变;

(2) $\tau \in \mathcal{L}(V)$ 的极小多项式和与 τ 相应的矩阵的极小多项式是相同的;

(3) 若 S 是模 V 的子模,则 S 的首一阶是限制 $\tau|_S$ 的极小多项式.

3. 固定 $\tau \in \mathcal{L}(V)$,考虑循环子模

$$\langle v \rangle = \{p(x)v \mid p(x) \in \mathbb{F}[x]\}.$$

设其首一阶为 $m(x)$. 于是 $m(x)$ 是限制 $\sigma = \tau|_{\langle v \rangle}$ 的极小多项式(见第 2 条之 (3)). 若

$$m(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x_n,$$

则可证

$$\mathcal{B} = \{v, xv, \ldots, x^{n-1}v\} = \{v, \sigma(v), \ldots, \sigma^{n-1}(v)\}$$

是向量空间 $\langle v \rangle$ 的一组基.

先来证明 $\mathcal{B}$ 是线性无关的. 事实上,若存在非零纯量 r_i,使得

$$r_0 v + r_1 xv + \cdots + r_{n-1}x^{n-1}v = 0,$$

即

$$(r_0 + r_1 x + \cdots + r_{n-1}x^{n-1})v = 0,$$

则

$$(r_0 + r_1 x + \cdots + r_{n-1} x^{n-1})\langle v\rangle = \{0\}.$$

故

$$m(x) \,\big|\, (r_0 + r_1 x + \cdots + r_{n-1} x^{n-1}),$$

这将导致 $r_i = 0$.

再证 $\mathcal{B}$ 生成 $\langle v\rangle$. $\langle v\rangle$ 中每个元素都是 $p(x)v, p(x) \in \mathbb{F}[x]$ 的形式,将 $p(x)$ 除以极小多项式 $m(x)$,得

$$p(x) = q(x)m(x) + r(x),$$

这里 $\deg r(x) < \deg m(x) = n$. 由于 $m(x)v = 0$,故 $p(x)v = r(x)v$,这表明 $\langle v\rangle$ 中所有元素都有 $r(x)v$ 的形式. 也就是

$$\langle v\rangle = \left\{ r(x)v \,\big|\, \deg r(x) < \deg m(x) \right\} = \operatorname{span}\mathcal{B}.$$

因此,$\mathcal{B}$ 确实是 $\langle v\rangle$ 的一组基.

现在来计算 σ 在基 $\mathcal{B}$ 之下的矩阵 $[\sigma]_{\mathcal{B}}$. 设 $i = 0, 1, \ldots, n-2$,则

$$\sigma\big(\sigma^i(v)\big) = \sigma^{i+1}(v).$$

而由于 $m(x)$ 是 σ 的极小多项式,故

$$\begin{aligned}\sigma\big(\sigma^{n-1}(v)\big) = \sigma^n(v) &= -\big(a_0 + a_1\sigma + \cdots + a_{n-1}\sigma^{n-1}\big)(v)\\ &= -a_0 v - a_1\sigma(v) - \cdots - a_{n-1}\sigma^{n-1}(v).\end{aligned}$$

于是 σ 在 $\mathcal{B}$ 下的矩阵为 $C[m(x)]$,

$$\sigma(\mathcal{B}) = \mathcal{B}C[m(x)],$$

而

$$C[m(x)] = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

$C[m(x)]$ 称为多项式

$$m(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x_n$$

的**友矩阵 (companian matrix)**. 这只对首一多项式定义.

若 $\tau \in \mathcal{L}(V)$,V 的一个子空间 S 称为 **τ– 循环 (τ–cyclic)**,若存在 $v \in S$ 使得

$$\left\{ v, \tau(v), \ldots, \tau^{m-1}(v) \right\}$$

是 S 的一组基,这里 $m = \dim S$.

于是有

(1) S 是 V 的子集,S 是 V 的循环子模当且仅当它是 V 的 τ 循环子空间;

(2) 若 $\langle v\rangle$ 是 V 中的一个循环子模,$\langle v\rangle$ 的首一阶(即 $\sigma = \tau\big|_{\langle v\rangle}$ 的极小多项式)是

$$m_\sigma(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x^n,$$

则

$$\mathcal{B} = \left(v, xv, \ldots, x^{n-1}v\right) = \left(v, \sigma(v), \ldots, \sigma^{n-1}(v)\right)$$

是 $\langle v\rangle$ 的一组基,σ 对 $\mathcal{B}$ 而言的矩阵 $[\sigma]_{\mathcal{B}}$ 是 $m_\sigma(x)$ 的友矩阵 $C([m_\sigma(x)])$.

A5.2 向量空间的分解

有了上一节作准备,就可以将上一讲中的分解定理翻译成向量空间中的结果. 定理 A4.3.5 成为如下定理.

定理 A5.2.1(向量空间关于线性变换的循环分解定理) 设 V 为有限维向量空间,并且 $\tau \in \mathcal{L}(V)$. 若 τ 的极小多项式为

$$m_\tau(x) = p_1^{e_1}(x)\cdots p_n^{e_n}(x),$$

这里 $p_i(x)$ 是相互不同的,不可约的首一多项式,则 V 可分解为直和

$$V = V_{p_1} \oplus \cdots \oplus V_{p_n},$$

这里

$$V_{p_i} = \left\{v \in V \;\middle|\; p_i^{e_i}(\tau)(v) = 0\right\}$$

是 V 的不变子空间(子模),而 $\tau|_{V_{p_i}}$ 的极小多项式为

$$\min \tau|_{V_{p_i}} = p_i^{e_i}(x).$$

进一步,V_{p_i} 可以再分解为 τ– 循环子空间(循环子模)的直和

$$V_{p_i} = \langle v_{i,1}\rangle \oplus \cdots \oplus \langle v_{i,k_i}\rangle,$$

这里 $\tau|_{\langle v_{i,j}\rangle}$ 的极小多项式为

$$\tau|_{\langle v_{i,j}\rangle} = p_i^{e_{i,j}}(x),$$

而

$$e_i = e_{i,1} \geqslant e_{i,2} \geqslant \cdots \geqslant e_{i,k_i} \geqslant 1,$$

V 的初等因子 $p_i^{e_{i,j}}(x)$,也就是 τ 的初等因子,由算子 τ 唯一确定.

归纳起来,V 可分解为 τ– 循环子空间的直和

$$V = \left(\langle v_{1,1}\rangle \oplus \cdots \oplus \langle v_{1,k_1}\rangle\right) \oplus \cdots \oplus \left(\langle v_{n,1}\rangle \oplus \cdots \oplus \langle v_{n,k_n}\rangle\right). \tag{A5.2.1}$$

在定理 A4.3.5 中,要求阶 $\mu = p_1^{e_1}\cdots p_n^{e_n}$,这里 p_i 为互不相伴的素元. 在定理 A5.2.1 中,说 $p_i(x)$ 为互不相同的不可约多项式. 由于 $\mathbb{F}[x]$ 是主理想整环,故素元与不可约元是一致的.

用循环分解定理可以确定相似意义下的标准形式.

若 $V = S \oplus T$,S, T 都是 $\tau \in \mathcal{L}(V)$ 之下的不变子空间,则称 $\boldsymbol{(S,T)}$ **约化** $\boldsymbol{\tau}$. 若 $V = S \oplus T$,且

$$\tau\big|_S: S \to S \quad 及 \quad \tau\big|_T: T \to T$$

分别是 S, T 上的线性算子.

记 $\tau = \rho \oplus \sigma$,若存在 V 的子空间 S 与 T,使得 (S, T) 约化 τ,及

$$\rho = \tau\big|_S, \quad \sigma = \tau\big|_T.$$

若 $\mathcal{C} = (c_1, \ldots, c_s)$ 是 S 的一组基,$\mathcal{D} = (d_1, \ldots, d_t)$ 是 T 的一组基,$s + t = n$,则

$$\mathcal{B} = (c_1, \ldots, c_s, d_1, \ldots, d_t),$$

是 V 的一组基. 于是矩阵 $[\tau]_{\mathcal{B}}$ 可以分解成分块对角矩阵 (block diagonal matrix)

$$[\tau]_{\mathcal{B}} = \begin{pmatrix} [\rho]_{\mathcal{C}} & 0 \\ 0 & [\sigma]_{\mathcal{D}} \end{pmatrix}.$$

这可推广到 τ 可分解为多个线性算子的的直和的情形.

回到 (A5.2.1). 若 $\mathcal{B}_{i,j}$ 是循环子模 $\langle v_{i,j}\rangle$ 的基,而

$$\mathcal{B} = \left(\mathcal{B}_{1,1}, \ldots, \mathcal{B}_{n,k_n}\right)$$

为 V 的基,则由定理 A5.2.1,

$$[\tau]_{\mathcal{B}} = \operatorname{diag}\left([\tau_{1,1}]_{\mathcal{B}_{1,1}}, \ldots, [\tau_{n,k_n}]_{\mathcal{B}_{n,k_n}}\right),$$

这里 $\tau_{i,j} = \tau\big|_{\langle v_{i,j}\rangle}$.

循环子模 $\langle v_{i,j}\rangle$ 有首一阶 $p_i^{e_{i,j}}(x)$,即限制 $\tau_{i,j}$ 有极小多项式 $p_i^{e_{i,j}}(x)$. 于是若

$$\deg p_i^{e_{i,j}}(x) = d_{i,j},$$

则

$$\mathcal{B}_{i,j} = \left(v_{i,j}, \tau_{i,j}(v_{i,j}), \ldots, \tau_{i,j}^{d_{i,j}-1}(v_{i,j})\right)$$

是 $\langle v_{i,j}\rangle$ 的一组基. 于是有如下定理.

定理 A5.2.2 假定 V 是有限维向量空间,$\tau \in \mathcal{L}(V)$ 的极小多项式为

$$m_\tau(x) = p_1^{e_1}(x) \cdots p_n^{e_n}(x),$$

这里首一多项式 $p_i(x)$ 互不相同且不可约,则 V 有分解

$$V = \left(\langle v_{1,1}\rangle \oplus \cdots \oplus \langle v_{1,k_1}\rangle\right) \oplus \cdots \oplus \left(\langle v_{n,1}\rangle \oplus \cdots \oplus \langle v_{n,k_n}\rangle\right),$$

这里 $\langle v_{i,j}\rangle$ 是 V 的 $\tau_{i,j}$- 循环子空间; $\tau_{i,j} = \tau\big|_{\langle v_{i,j}\rangle}$ 的极小多项式是 V 的初等因子

$$\min \tau_{i,j} = p_i^{e_{i,j}}(x),$$

这里

$$e_i = e_{i,1} \geqslant e_{i,2} \geqslant \cdots \geqslant e_{i,k_i} \geqslant 1.$$

初等因子由 τ 唯一确定. 若 $\deg p_i^{e_{i,j}}(x) = d_{i,j}$,则

$$\mathcal{B}_{i,j} = \left(v_{i,j}, \tau_{i,j}(v_{i,j}), \ldots, \tau_{i,j}^{d_{i,j}-1}(v_{i,j})\right)$$

是 $\langle v_{i,j}\rangle$ 的一组基,τ 相对于基

$$\mathcal{B} = \left(\mathcal{B}_{1,1}, \ldots, \mathcal{B}_{n,k_n}\right)$$

的矩阵是分块对角矩阵

$$[\tau]_{\mathcal{B}} = \mathrm{diag}\Big(C[p_1^{e_{1,1}}(x)],\ldots,C[p_1^{e_{1,k_1}}(x)],\ldots\ldots,C[p_n^{e_{n,1}}(x)],\ldots,C[p_n^{e_{n,k_n}}(x)]\Big).$$

上式右边的矩阵称为 τ 的**有理标准形式 (rational canonical form)**.

由向量空间的循环分解的唯一性定理,这样的有理标准形式是唯一的.

定理 A5.2.2 用矩阵的语言为:任意矩阵 A 唯一地(除去对角线上分块的次序)相似于一个有理标准形式的矩阵.

由此还可以有:域 $\mathbb{F}$ 上两个矩阵是相似的当且仅当它们有相同的初等因子.

定理 A5.2.1 及定理 A5.2.2 是线性代数理论的巅峰之一. 从几何上讲,它们彻底解决了一个域上的向量空间,在一个线性变换下的分解. 从代数上讲它们彻底地解决了在一个域上的矩阵在相似变换下的分类.

A5.3 特征多项式、特征值与特征向量

1. 假设 $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n$,而 $C[p(x)]$ 是其友矩阵,令

$$A = xI - C[p(x)] = \begin{pmatrix} x & 0 & \cdots & 0 & a_0 \\ -1 & x & \ddots & \vdots & a_1 \\ 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & x & a_{n-2} \\ 0 & \cdots & 0 & -1 & x + a_{n-1} \end{pmatrix}.$$

显然 A 是 $x, a_0, \ldots, a_{n-1}$ 的函数,记作 $A = A(x; a_0, \ldots, a_{n-1})$.

命题 A5.3.1 $\qquad \det\big(xI - C[p(x)]\big) = p(x).$

证明 当 $n = 2$ 时,

$$\det\big(A(x; a_0, a_1)\big) = \begin{vmatrix} x & a_0 \\ -1 & x + a_1 \end{vmatrix} = x(x + a_1) + a_0 = a_0 + a_1x + x^2 = p(x).$$

当 $n = 3$ 时,

$$\det\big(A(x; a_0, a_1, a_2)\big) = \begin{vmatrix} x & 0 & a_0 \\ -1 & x & a_1 \\ 0 & -1 & x + a_2 \end{vmatrix} = x\begin{vmatrix} x & a_1 \\ -1 & x + a_2 \end{vmatrix} + a_0\begin{vmatrix} -1 & x \\ 0 & -1 \end{vmatrix}$$
$$= a_0 + a_1x + a_2x^2 + x^3 = p(x).$$

对一般的 n,对行列式沿第一行展开,得到

$$\det\big(A(x; a_0, \ldots, a_{n-1})\big) = x\det\big(A(x; a_1, \ldots, a_{n-1})\big) + (-1)^{n+1}(-1)^{n-1}a_0.$$

由归纳假设这等于

$$x(x^{n-1} + a_{n-1}x^{n-2} + \cdots + a_2x + a_1) + a_0 = p(x). \qquad \blacksquare$$

由命题 A5.3.1 可得如下结论.

命题 A5.3.2　若 $\tau \in \mathcal{L}(V)$,R 是 τ 的有理标准形式,则

$$C_\tau(x) = \det(xI - R) = \prod_{i,j} p_i^{e_{i,j}}(x),$$

这个行列式称为 τ 的**特征多项式 (eigenpolynomial)**.

在线性代数通常的教材中,往往先定义矩阵的特征多项式,然后再定义线性算子的特征多项式. 方阵 A 的特征多项式定义为 $C_A(x) = \det(xI - A)$.

由此可得下面的结论.

(1) 若 A 与 B 相似,则 $C_A(x) = C_B(x)$. 即特征多项式在相似下不变;

(2) 线性算子 τ 的特征多项式和与 τ 相对应的矩阵的特征多项式相等;

(3) 线性算子 τ 的特征多项式是 τ 的初等因子的乘积.

2. $\lambda \in \mathbb{F}$ 是线性算子 $\tau \in \mathcal{L}(V)$ 的特征多项式 $C_\tau(x)$ 的根,当且仅当

$$\det(\lambda I - R) = 0,$$

即矩阵 $\lambda I - R$ 是奇异的. 若 $\dim V = d$,则 τ 的有理标准形式 R 为 $d \times d$ 的矩阵. 故 $\det(\lambda I - R) = 0$ 当且仅当存在非零向量 $x \in \mathbb{F}^d$,使得

$$(\lambda I - R)x = 0, \quad \text{即} \quad Rx = \lambda x.$$

若 $v \in V$ 是非零向量使得 $[v]_{\mathcal{B}} = x$,这里 $\mathcal{B}$ 是 V 的基使得 τ 的矩阵为 R,则上式等价于

$$\tau(v) = \lambda v.$$

定义 A5.3.1　若 $\tau \in \mathcal{L}(V)$,纯量 $\lambda \in \mathbb{F}$ 是 τ 的一个**特征值 (eigenvalue)**,若存在非零向量 $v \in V$,使得

$$\tau(v) = \lambda v.$$

这时称 v 为 τ 的以 λ 为特征值的**特征向量 (eigenvector)**.

若 A 为 $\mathbb{F}$ 上的矩阵,$\lambda \in \mathbb{F}$ 是 A 的特征值若存在非零列向量 x,使得

$$Ax = \lambda x.$$

这时称 x 为矩阵 A 的以 λ 为特征值的特征向量.

对一个给定的特征值 λ,所有以 λ 为特征值的特征向量加上零向量,组成 V 的一个子空间,称为 λ 的**特征空间 (eigenspace)**,记作 $\mathcal{E}_\lambda$.

由此可得如下结论.

命题 A5.3.3　**(1)** $\lambda \in \mathbb{F}$ 是 $\tau \in \mathcal{L}(V)$ 的特征值当且仅当它是特征多项式 $C_\tau(x)$ 的根;

(2) $\lambda \in \mathbb{F}$ 是 $\tau \in \mathcal{L}(V)$ 的特征值当且仅当它是 τ 的极小多项式 $m_\tau(x)$ 的根;

(3) $\lambda \in \mathbb{F}$ 是 $\tau \in \mathcal{L}(V)$ 的特征值当且仅当它是与 τ 相应的任何矩阵的特征值;

(4) 矩阵的特征值在相似意义下不变;

(5) 若 λ 是矩阵的特征值,则特征空间 $\mathcal{E}_\lambda$ 是齐次方程组

$$(\lambda I - A)x = 0$$

的解组成的空间.

证明 这里只证明 (3). λ 是 τ 的特征值当且仅当有 $0 \neq v \in V$,使得

$$\tau(v) = \lambda v.$$

设 $\dim V = d$,$\mathcal{B}$ 是 V 的一组基,令 $\varphi_{\mathcal{B}}: V \to \mathbb{F}^d$ 为由 $\varphi_{\mathcal{B}}(u) = [u]_{\mathcal{B}}$ 定义的同构. 若 $A = [\tau]_{\mathcal{B}}$,则

$$\tau = (\varphi_{\mathcal{B}})^{-1} A \varphi_{\mathcal{B}}.$$

于是 $\tau(v) = \lambda v$ 就是

$$(\varphi_{\mathcal{B}})^{-1} A \varphi_{\mathcal{B}}(v) = \lambda v = (\varphi_{\mathcal{B}})^{-1} \lambda \varphi_{\mathcal{B}}(v),$$

即

$$A \varphi_{\mathcal{B}}(v) = \lambda \varphi_{\mathcal{B}}(v).$$

这表明 λ 是 A 的一个特征值. ■

命题 A5.3.4 与不同的特征值对应的特征向量是线性无关的,即,若 $v_i \in \mathcal{E}_{\lambda_i}$,则 $\{v_1, \ldots, v_k\}$ 线性无关. 特别地,若 $\lambda_1, \ldots, \lambda_k$ 是线性算子 $\tau \in \mathcal{L}(V)$ 的不同的特征值,则 $\mathcal{E}_i \cap \mathcal{E}_j = \{0\}$.

证明 若众 v_i 线性相关,则在所有非平凡的线性组合为零的式子中,有一个最短的式子,为

$$r_1 v_1 + \cdots + r_j v_j = 0,$$

将 τ 作用上式,得

$$r_1 \tau(v_1) + \cdots + r_j \tau(v_j) = 0,$$

即

$$r_1 \lambda_1 v_1 + \cdots + r_j \lambda_j v_j = 0,$$

在 $r_1 v_1 + \cdots + r_j v_j = 0$ 上乘以 λ_1,再与上式相减得

$$r_2(\lambda_2 - \lambda_1)v_2 + \cdots + r_j(\lambda_j - \lambda_1)v_j = 0.$$

但这是一个更短的线性组合为零的式子,由此得到矛盾. 故所有的 $r_i = 0$. ■

由于 τ 的特征多项式 $C_\tau(x)$ 是所有初等因子的乘积,而 τ 的极小多项式为

$$m_\tau(x) = p_1^{e_1}(x) \cdots p_n^{e_n}(x),$$

故 $m_\tau(x) \mid C_\tau(x)$. 由此得到重要的定理.

定理 A5.3.1 若 $\tau \in \mathcal{L}(V)$,则

(1) 极小多项式 $m_\tau(x)$ 与特征多项式 $C_\tau(x)$ 有相同的素因子;

(2)(**Caley-Hamilton 定理**)$m_\tau(x) \mid C_\tau(x)$,即 $C_\tau(\tau) = 0$.

A5.4 Jordan 标准形式

有限维向量空间的每个线性算子 τ 都有有理标准形式,即所有有理标准形式的矩阵的全体组成一个标准形式集合. 显然,有理标准形式还不是我们指望的那样具有简单的形式. 对一些重要的特殊的情形,我们可以得到更为简单的标准形式. 这种重要的特殊情形是: 若 $\tau \in \mathcal{L}(V)$,它的极小多项式可以分解为线性因子的乘积,即

$$m_\tau(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_n)^{e_n}. \tag{A5.4.1}$$

当一个多项式在域 $\mathbb{F}$ 上可以分解为线性因子的乘积时,称多项式可以在 $\mathbb{F}$ 上**分裂 (split)**.

若域 $\mathbb{F}$ 上任一非常数的多项式的根仍在 $\mathbb{F}$ 中,称 $\mathbb{F}$ 为**代数封闭的 (algebraic closed)**. 因此,在代数封闭域上不可约多项式只有线性多项式. 故任意非常数多项式在 $\mathbb{F}$ 上分裂. 代数封闭域简单的例子是复数域.

回顾有理标准形式. $\langle v_{i,j}\rangle$ 是循环子模,其首一阶为初等因子 $p_i^{e_{i,j}}(x)$,由于对 $p_i^{e_{i,j}}$ 了解甚少,以致作为 V 的 τ– 循环子空间,选取基为

$$\mathcal{B}_{i,j} = \left(v_{i,j}, \tau_{i,j}(v_{i,j}), \ldots, \tau_{i,j}^{d_{i,j}-1}(v_{i,j})\right).$$

但当 τ 的极小多项式是 (A5.4.1) 时,其初等因子为

$$p_i^{e_{i,j}}(x) = (x - \lambda_i)^{e_{i,j}}.$$

这时,我们可以更明智地选取基.

由于 $\dim\langle v_{i,j}\rangle = \deg p_i^{e_{i,j}}(x)$,易见

$$\mathcal{G}_{i,j} = \left(v_{i,j}, (\tau_{i,j} - \lambda_i\varepsilon)(v_{i,j}), \ldots, (\tau_{i,j} - \lambda_i\varepsilon)^{e_{i,j}-1}(v_{i,j})\right)$$

也是 $\langle v_{i,j}\rangle$ 的一组基. 记 $\mathcal{G}_{i,j}$ 中第 k 个基向量为 b_k,则当 $k = 0, 1, \ldots, e_{i,j} - 2$ 时,

$$\begin{aligned}\tau_{i,j}(b_k) &= \tau_{i,j}\left((\tau_{i,j} - \lambda_i\varepsilon)^k(v_{i,j})\right) = (\tau_{i,j} - \lambda_i\varepsilon + \lambda_i\varepsilon)\left((\tau_{i,j} - \lambda_i\varepsilon)^k(v_{i,j})\right)\\ &= (\tau_{i,j} - \lambda_i\varepsilon)^{k+1}(v_{i,j}) + \lambda_i(\tau_{i,j} - \lambda_i\varepsilon)^k(v_{i,j}) = b_{k+1} + \lambda_i b_k;\end{aligned}$$

当 $k = e_{i,j} - 1$ 时,应用

$$(\tau_{i,j} - \lambda_i)^{k+1}(v_{i,j}) = (\tau_{i,j} - \lambda_i)^{e_{i,j}}(v_{i,j}) = 0,$$

可得

$$\tau_{i,j}(b_{e_{i,j}-1}) = \lambda_i b_{(e_{i,j}-1)}.$$

因此,相对于基 $\mathcal{G}_{i,j}$, $\tau_{i,j} = \tau|_{\langle v_{i,j}\rangle}$ 所对应的矩阵为 $e_{i,j} \times e_{i,j}$ 方阵

$$g(\lambda_i, e_{i,j}) = \begin{pmatrix} \lambda_i & & & \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_i \end{pmatrix}.$$

这个矩阵称为属于特征值 λ_i 的 **Jordan 块 (Jordan block)**. 即 Jordan 块为主对角线上元素为 λ_i，在次对角线上元素为 1，其余元素全为零的 $e_{i,j}$ 阶方阵.

于是在选取新的基后，类似于定理 A5.2.2，有如下的定理.

定理 A5.4.1 若算子 $\tau \in \mathcal{L}(V)$ 的极小多项式在域 $\mathbb{F}$ 上分裂，即

$$m_\tau(x) = (x-\lambda_1)^{e_1}\cdots(d-\lambda_n)^{e_n},$$

则 V 可分解为

$$V = \left(\langle v_{1,1}\rangle \oplus \cdots \oplus \langle v_{1,k_1}\rangle\right) \oplus \cdots \oplus \left(\langle v_{n,1}\rangle \oplus \cdots \oplus \langle v_{n,k_n}\rangle\right),$$

这里 $\langle v_{i,j}\rangle$ 是 V 的 τ– 循环子空间模，$\tau_{i,j} = \tau\big|_{\langle v_{i,j}\rangle}$ 的极小多项式为 V 的初等因子

$$\min \tau_{i,j} = (x-\lambda_i)^{e_{i,j}},$$

这里

$$e_i = e_{i,1} \geqslant e_{i,2} \geqslant \cdots \geqslant e_{i,k_i} \geqslant 1,$$

这些初等因子由 τ 唯一确定. 令

$$\mathcal{G}_{i,j} = \left(v_{i,j}, (\tau_{i,j}-\lambda_i)(v_{i,j}), \ldots, (\tau_{i,j}-\lambda_i)^{e_{i,j}-1}(v_{i,j})\right)$$

是 $\langle v_{i,j}\rangle$ 的一组基，则与 τ 对应的矩阵在基

$$\mathcal{G} = \left(\mathcal{G}_{1,1}, \ldots, \mathcal{G}_{n,k_n}\right)$$

下是分块对角矩阵

$$[\tau]_{\mathcal{G}} = \operatorname{diag}\left(g(\lambda_1, e_{1,1}), \ldots, g(\lambda_1, e_{1,k_1}), \ldots\ldots, g(\lambda_n, e_{n,1}), \ldots, g(\lambda_n, e_{n,k_n})\right).$$

上面的矩阵称为 τ 的 **Jordan 标准形式 (Jordan canonical form)**.

换成矩阵的语言：在代数封闭域 $\mathbb{F}$ 上每个矩阵都相似于唯一的一个 Jordan 标准形式. 也就是，所有的 Jordan 标准形式确实组成了在相似意义下的标准形式集合.

若 τ 有 Jordan 标准形式 g，则 g 中主对角线上元素就是特征多项式 $C_\tau(x)$ 的根 (包括重数). 也就是，g 中主对角线元素 λ_i 出现的次数就是特征多项式的根 λ_i 的重数.

定理 A5.4.1 是线性代数理论的另一个巅峰. 从几何上讲，它彻底解决了在代数封闭域上的一个向量空间在一个线性变换下的分解；从代数上讲，它彻底解决了在代数封闭域上的矩阵在相似变换下的分类.

A5.5 内积空间上算子的标准形式

1. 在 A2.4 及 A3.3 中介绍了内积空间以及其上的三种重要算子：自共轭算子、酉算子以及正规算子，还讨论了它们的一些简单性质. 现在来看看这些算子的特征值及特征空间.

命题 A5.5.1 若 τ 是自共轭算子，则 τ 的特征多项式的根都是实数. 也就是说，τ 的特征值全是实的.

证明 先设 V 是复向量空间，λ 是 τ 的特征多项式 $C_\tau(x)$ 的根，则有 $v \neq 0$，使得 $\tau(v) = \lambda v$. 于是

$$\langle \tau(v), v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle.$$

由于 τ 是自共轭的，故

$$\langle \tau(v), v \rangle = \langle v, \tau(v) \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle,$$

故 $\lambda = \overline{\lambda}$，即 λ 是实数.

若 V 是实向量空间，则 τ 对 V 的某一组基其对应的矩阵是实对称阵 A. 于是 $C_\tau(x) = C_A(x)$. 因 A 是实对称阵，可看作复空间 $\mathbb{C}^n$ 的一个自共轭线性算子，如上面所证，其特征多项式的根是实的. 将 A 看作实的或复的矩阵，其特征多项式是一样的，故得证. ■

命题 A5.5.2 若 τ 是酉算子，则 τ 的特征值的绝对值为 1.

证明 由于 τ 为酉算子及 $\tau(v) = \lambda v$，则

$$\lambda \overline{\lambda} \langle v, v \rangle = \langle \lambda v, \lambda v \rangle = \langle \tau(v), \tau(v) \rangle = \langle v, v \rangle,$$

故 $|\lambda|^2 = 1$，即 $|\lambda| = 1$. ■

命题 A5.5.3 若 τ 为正规算子，λ, μ 为 τ 的不同的特征值，则对应的特征子空间互相正交. 特别地，自共轭算子和酉算子的不同特征值对应的特征子空间互相正交.

证明 由于 $\tau(v) = \lambda v$，$\tau(w) = \mu w$，这里 $v, w \neq 0$，则

$$\lambda \langle v, w \rangle = \langle \tau(v), w \rangle = \langle v, \tau^*(w) \rangle = \langle v, \overline{\mu} w \rangle = \mu \langle v, w \rangle,$$

由 $\mu \neq \lambda$ 可知 $\langle v, w \rangle = 0$.

这里用到 $\tau^*(w) = \overline{\mu} w$，参见 A3.3 中正规算子的性质 (5). ■

前面的有理标准形式及 Jordan 形式一般不是对角矩阵，什么情况下可化为对角矩阵？

定义 A5.5.1 若 V 是有限维内积空间，$\tau \in \mathcal{L}(V)$. 若有 V 的正规正交基 $\mathcal{O}$ 使得 $[\tau]_{\mathcal{O}}$ 是一个对角矩阵，则称 τ **可正交对角化 (orthogonal diagonalizable)**.

定理 A5.5.1　假设 V 是有限维复内积空间,则

(1) V 上一个线性算子 τ 可以正交对角化当且仅当它是正规的;

(2) τ 是 V 上的一个正规算子,它是自共轭当且仅当它的特征值均是实数;

(3) τ 是 V 上的一个正规算子,它是酉当且仅当它的特征值的绝对值均为 1.

证明　**(1)** 若 τ 是复内积空间的一个正规算子,且 τ 的极小多项式的素因子分解为

$$m_\tau(x) = (x-\lambda_1)^{e_1}\cdots(x-\lambda_k)^{e_k},$$

则由准素分解定理,V 可分解为

$$V = V_1 \oplus \cdots \oplus V_k.$$

由 A3.3 中关于正规算子的命题 A3.3.5 之 (4),有

$$V_i = \left\{v \in V \,\middle|\, (\tau-\lambda_i)^{e_i}(v) = 0\right\} = \left\{v \in V \,\middle|\, (\tau-\lambda_i)(v) = 0\right\} = \mathcal{E}_{\lambda_i}.$$

故 $\tau\big|_{V_i}$ 的极小多项式为 $x-\lambda_i$,故 $e_i = 1$. 因此

$$V = \mathcal{E}_{\lambda_1} \oplus \cdots \oplus \mathcal{E}_{\lambda_k}.$$

有命题 A5.5.3 可知,V 可分解为正交直和

$$V = \mathcal{E}_{\lambda_1} \odot \cdots \odot \mathcal{E}_{\lambda_k}.$$

故将每个特征空间的正规正交基组合起来构造出由 τ 的特征向量组成的 V 的一个正规正交基,即 τ 是可以正交对角化的.

反之,若 τ 可正交对角化,则 V 有一个正规正交基 $\mathcal{O} = \left\{v_1, \ldots, v_k\right\}$,其中 $\tau(u_i) = \lambda_i \mu_i$. 于是

$$\langle u_i, \tau^*(u_j)\rangle = \langle \tau(u_i), u_j\rangle = \lambda_i\langle u_i, u_j\rangle = \lambda_i\delta_{ij} = \lambda_j\delta_{ij} = \langle u_i, \overline{\lambda}_j u_j\rangle,$$

故 $\tau^*(u_j) = \overline{\lambda}_j u_j$. 因此,

$$\tau\tau^*(u_j) = \overline{\lambda}_j\tau(u_j) = \overline{\lambda}_j\lambda_j u_j = \lambda_j\overline{\lambda}_j u_j = \lambda_j\tau^*(u_j) = \tau^*\tau(u_j),$$

故 τ 是正规算子.

(2) 已知自共轭算子是正规的,则特征值均是实的. 反之,若 τ 是正规的,且特征值均是实的,则对应于 λ_j 的任何特征向量 u_j,有

$$\tau^*(u_j) = \overline{\lambda}_j u_j = \lambda_j u_j = \tau(u_j).$$

由于这些 u_j 是特征向量构成的基,故 τ 自共轭.

(3) 若 τ 为酉算子,则它的特征值的绝对值均为 1. 反之,若 τ 是复内积空间 V 上的一个正规算子,且其特征值之绝对值均为 1,则对应于 λ_j 的任何特征向量 u_j,有

$$\langle u_j, u_j\rangle = \lambda_j\overline{\lambda}_j\langle u_j, u_j\rangle = \langle \lambda_j u_j, \lambda_j u_j\rangle = \langle \tau(u_j), \tau(u_j)\rangle = \langle u_j, \tau^*\tau(u_j)\rangle.$$

由于这些 u_j 是特征向量构成的基,故 $\tau^* = \tau^{-1}$,因此 τ 为酉算子. ■

上面给出了复内积空间上线性算子 τ 可正交对角化的充分必要条件是 τ 是正规的,下面给出实内积空间上线性算子可正交对角化的充分必要条件.

定理 A5.5.2 有限维实内积空间 V 上的一个线性算子 τ 可正交对角化当且仅当 τ 是自共轭的.

证明 若 τ 是 V 上的自共轭算子,则由命题 A5.5.1,τ 的极小多项式可在 $\mathbb{R}$ 上分裂. 由 A3.3 中关于自共轭算子的命题 A3.3.2 之 (8) 及命题 A5.5.3 得到,V 有 τ 的特征向量组成的正规正交基; 其证明类似于定理 A5.5.1.

必要性可用矩阵方法来证明. 事实上,若 τ 可正交对角化,则 V 有正规正交基 $\mathcal{O}$,使得 $[\tau]_{\mathcal{O}}$ 对角化. 由于 $[\tau]_{\mathcal{O}}$ 是实对称的,故

$$[\tau^*]_{\mathcal{O}} = [\tau]_{\mathcal{O}}^* = [\tau]_{\mathcal{O}}^{\mathrm{T}} = [\tau]_{\mathcal{O}}.$$

故 $\tau^* = \tau$. ■

定理 A5.5.1 以及定理 A5.5.2 的矩阵形式如下.

定理 A5.5.3 **(1)** 设 A 是一个复方阵,则存在酉阵 U 使得 UAU^{-1} 是对角阵当且仅当 A 是正规的; 一个正规复方阵 A 是埃尔米特的当且仅当 A 的特征值均为实数; 一个正规复方阵 A 是酉阵当且仅当 A 的特征值的绝对值均为 1;

(2) 设 A 是一个实方阵,则存在正交阵 O 使得 OAO^{-1} 是对角阵当且仅当 A 是对称的(见定理 A2.3.5).

2. 上面给出了正规算子与自共轭算子分别在复数域 $\mathbb{C}$ 及实数域 $\mathbb{R}$ 上的标准形式是对角阵. 现在来给出实数域 $\mathbb{R}$ 上的酉算子的标准形式.

若 τ 是实酉算子,则 $\sigma = \tau + \tau^* = \tau + \tau^{-1}$ 是自共轭的,故有一实特征值的完备集,如在定理 A5.5.1 中那样,V 可分解为

$$V = \mathcal{E}_{\lambda_1} \odot \cdots \odot \mathcal{E}_{\lambda_k},$$

这里

$$\mathcal{E}_{\lambda_i} = \left\{ v \in V \,\middle|\, (\tau + \tau^{-1} - \lambda_i)(v) = 0 \right\},$$

或乘以 τ,

$$\mathcal{E}_{\lambda_i} = \left\{ v \in V \,\middle|\, (\tau^2 - \lambda_i \tau + 1)(v) = 0 \right\}.$$

若 $\lambda_i = 2$,则由于 τ 是正规的,有

$$\mathcal{E}_2 = \left\{ v \in V \,\middle|\, (\tau - 1)^2(v) = 0 \right\} = \left\{ v \in V \,\middle|\, (\tau - 1)(v) = 0 \right\}.$$

若 $\lambda_i = -2$,有

$$\mathcal{E}_{-2} = \left\{ v \in V \,\middle|\, (\tau + 1)^2(v) = 0 \right\} = \left\{ v \in V \,\middle|\, (\tau + 1)(v) = 0 \right\}.$$

故算子 τ 在特征空间 $\mathcal{E}_2$ 及 $\mathcal{E}_{-2}$(如果存在的话)上的限制分别就是乘以 1 或 -1.

当 $\lambda_i \neq 2$ 时,若 $v \in \mathcal{E}_{\lambda_i}$,考虑 $\mathrm{span}\{v, \tau(v)\}$. 这是 $\mathcal{E}_{\lambda_i}$ 中的一个不变子空间,因为

$$\tau(\tau(v)) = \tau^2(v) = \lambda_i \tau(v) - v.$$

于是,

$$\mathcal{E}_{\lambda_i} = \mathrm{span}\{v, \tau(v)\} \odot \mathrm{span}\{v, \tau(v)\}^{\perp}.$$

继续这个步骤,每个 $\mathcal{E}_{\lambda_i}$ 可分解为二维子空间的正交直和,τ 在每个子空间上是实

酉的,而

$$V=\mathcal{E}_2\odot\mathcal{E}_{-2}\odot\mathcal{D}_1\odot\cdots\odot\mathcal{D}_m,$$

这里 $\dim\mathcal{D}_i=2$,每一项在 τ 下不变.

于是,只要给出在二维空间 $\mathcal{D}$ 上的实酉算子 τ 的矩阵即可. 由于对 $\mathcal{D}$ 的任意正规正交基 τ 的矩阵是正交的,故若

$$[\tau]=\begin{pmatrix}a & b\\ c & d\end{pmatrix},$$

则

$$a^2+b^2=1,\quad c^2+d^2=1,\quad ac+bd=0.$$

由于 $\det\big([\tau]\big)=1$,即

$$ad-bc=1.$$

解这些方程,得 $d=a$,$c=-b$,于是

$$[\tau]=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}.$$

由于 (a,b) 是 $\mathbb{R}^2$ 中的单位向量,因此 $(a,b)=(\cos\theta,\sin\theta)$,这里 θ 为实数,故

$$[\tau]=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}.$$

归纳起来,有如下定理.

定理 A5.5.4 若 τ 是有限维实内积空间 V 上的一个酉算子,则 V 有一个正规正交基,使得 τ 的矩阵有分块形式

$$\operatorname{diag}\left(1,\ldots,1,-1,\ldots,-1,\begin{pmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{pmatrix},\ldots,\begin{pmatrix}\cos\theta_k & \sin\theta_k\\ -\sin\theta_k & \cos\theta_k\end{pmatrix}\right). \tag{A5.5.1}$$

定理 A5.5.4 的矩阵形式为如下定理.

定理 A5.5.5 若 A 是一个正交矩阵,则存在正交矩阵 O,使得 OAO^{-1} 是形如 (A5.5.1) 的矩阵.

A5.6 附 记

在 A5.1 中已经说到,本讲是将域 $\mathbb{F}$ 上的向量空间 V 看成 $\mathbb{F}$ 上的多项式环 $\mathbb{F}[x]$ 上的模,于是上一讲中的结果可以翻译成向量空间的语言,这就得到了一系列十分重要的分解定理.

在 A4.1 中,还说到环 R 上的一个 R– 模,当 $R=\mathbb{Z}$ 时,则 $\mathbb{Z}$– 模就是 Abel 群,也就是可以将 Abel 群看作 $\mathbb{Z}$– 模. 于是也可以将上一讲中的结果翻译成为 Abel 群的语言,也可以得到十分重要的分解定理.

如果 G 是一个有限生成的 Abel 群,则 G 可以视为一个有限生成的 $\mathbb{Z}$– 模. 由定理 A4.3.1,G 可以分解为一个挠模 $\operatorname{Tor} G$ 和一个自由 $\mathbb{Z}$– 模 $\operatorname{Free} G$ 的直和. 设 $\operatorname{Free} G$ 的秩为 r,则 $\operatorname{Free} G \approx \mathbb{Z}^r$.

$\operatorname{Tor} G$ 也是有限生成的,设 $x_1,\ldots,x_m$ 是它的一组生成元. 于是 $\operatorname{Tor} G$ 的每个元素都可以用这组生成元来表示. 由于这些生成元的阶都是有限的,所以 $\operatorname{Tor} G$ 是一个有限 Abel 群.

由定理 A4.3.2,$\operatorname{Tor} G$ 的阶为 $\mu=p_1^{e_1}\cdots p_n^{e_n}$,这里 p_i 为互不相同的素数,则 $\operatorname{Tor} G$ 可分解为

$$\operatorname{Tor} G = G_{p_1}\oplus\cdots\oplus G_{p_n},$$

这里 $G_{p_i}=\left\{v\in\operatorname{Tor} G \,\middle|\, p_i^{e_i}v=0\right\}$,即 G_{p_i} 为 $\operatorname{Tor} G$ 的阶为 $p_i^{e_i}$ 的子群,即 G_{p_i} 为 $\operatorname{Tor} G$ 的 Sylow p_i– 子群.

这里要注意的是,对于有限 Abel 群 G 有两个阶的概念,其一指 $|G|$,即 G 中元素的个数; 其二是指 G 作为 $\mathbb{Z}$– 模的零化子的生成元. 一般而言,这两个阶是不同的. 在上一段的讨论中所讲的阶是指后者. 回顾群 G 的 Sylow p– 子群是指 G 的阶为 p^m 的子群,其中 $p^m \mid |G|$,且 $p^{m+1}\nmid |G|$. Sylow p– 子群总是存在的,因此在上一段的讨论中,G_{p_i} 必是 $\operatorname{Tor} G$ 的 Sylow p_i– 子群.

由定理 A4.3.3,有限 Sylow p_i– 子群又可以分解成一些循环 p_i– 子群的直和,即

$$G_{p_i}=G_{i,1}\oplus\cdots\oplus G_{i,k_i},$$

这里 $G_{i,j}$ 是阶为 $p_i^{e_{i,j}}$ 的循环 p_i– 子群,且满足

$$e_i=e_{i,1}\geqslant e_{i,2}\geqslant\cdots\geqslant e_{i,k_i},$$

或等价地为

$$p_i^{e_{i,k_i}}\mid p_i^{e_{i,k_i-1}},\ p_i^{e_{i,k_i-1}}\mid p_i^{e_{i,k_i-2}},\ldots,p_i^{e_{i,2}}\mid p_i^{e_{i,1}}.$$

归纳起来,有以下的定理.

定理 A5.6.1(有限生成的 Abel 群的分解定理) 若 G 是一个有限生成的 Abel 群,则 G 可以分解成 r 个无限循环子群及一些有限循环 p_i– 子群的直和.

r 和有限循环 p_i 子群的阶 $p_i^{e_{i,j}}$ 是 G 上一组完全不变量,即两个有限生成 Abel

群同构当且仅当它们的不变量完全相同.

这就是第四讲中主理想整环上有限生成模的分解定理译成有限生成 Abel 群时的语言. 由此可以看出,这完全解决了有限生成 Abel 群的分类问题. 当然这些内容不属于线性代数的范围,所以作为附记,用以显示模理论的有力作用.

参考文献

[1] 聂灵沼,丁石孙. 代数学引论(第二版). 北京: 高等教育出版社,2000

[2] 刘绍学. 近世代数基础. 北京: 高等教育出版社,1999

[3] 莫宗坚,蓝以中,赵春来. 代数学. 北京: 北京大学出版社,1986

[4] 许以超. 代数学引论. 上海: 上海科学技术出版社,1965

[5] 北京大学数学系几何与代数教研室前代数小组. 高等代数(第三版). 北京: 高等教育出版社,2003

[6] 李炯生,查建国. 线性代数. 合肥: 中国科学技术大学出版社,1989

[7] Jacobson N. *Basic Algebra*, Vol I. San Francisco: W. H. Freeman and Company, 1985

[8] Hugerford T. W. *Algebra*, GTM 73. New York: Springer-Verlag, 1974

[9] Roman S. *Advanced Linear Algebra*, GTM 135. New York: Springer-Verlag, 1992

[10] Blyth T. S. *Module Theory: An Approch to Linear Algebra*. London: Oxford University Press, 1990