线性代数 B2

Fir1247

2022 年 09 月-2022 年 12 月

前言

本文档为 2022 年秋季学期陈发来老师的《线性代数 B2》课程笔记, 内容主要为上课板书、习题课、部分课后习题。

为避免歧义，本文档中的 $\subset$，相当于 $\subseteq, \subseteqq$，均表示“包含于”的意思，且只使用第一种形式；本文档中的 $\subsetneq$，相当于 $\subsetneqq$，均表示“真包含于”的意思，且只使用第一种形式。

Fir1247
2022 年 09 月

目录

第一章 多项式 **1**

第二章 行列式与矩阵 **2**
2.1 行列式的定义与性质 2
2.2 行列式的 laplace 展开 3
2.3 行列式的计算 4
2.4 矩阵的基本运算 4
2.5 矩阵的秩与相抵 4
2.6 矩阵的广义逆 6
2.7 多项式矩阵相抵 6
2.8 矩阵的相合与正定 9

第三章 线性空间 **16**
3.1 线性空间基本理论 16
3.2 线性空间同构与同态 18
3.3 子空间运算 19
3.4 商空间 20
3.5 对偶空间 21

第四章 线性变换 **23**
4.1 基础概念 23
4.2 线性映射的像与核 25
4.3 线性变换 27
4.4 最小多项式 29

第五章 Jordan 标准型 **33**
5.1 Jordan 标准型的计算 33
5.2 矩阵相似与多项式矩阵相抵 35
5.3 根子空间分解 36
5.4 循环子空间分解 38
5.5 实方阵的实相似标准型 42
5.6 一些例子 43

第六章 内积空间 **47**
6.1 欧氏空间 47
6.2 正交变换与对称变换 48
6.3 规范变换 49

6.4 酉空间 51
6.5 酉空间上的线性变换 51

第七章 多重线性函数 **53**
7.1 二重线性函数 53
7.2 多重线性函数 55

第一章　多项式

定理 1.0.1 如果多项式 $f(x), g(x)$ 互素，则存在多项式 $u(x), v(x)$ 使得 $u(x)f(x)+v(x)g(x) = 1$. 推广形式：如果多项式 $f_1(x), \cdots, f_n(x)$ 互素，则存在 $u_1(x), \cdots, u_n(x)$ 使得

$$\sum_{i=1}^{n} u_i(x)f_i(x) = 1$$

这是一个后面被反复使用的定理，本章其他内容不是特别重要，所以这里就省略了。

第二章　行列式与矩阵

2.1　行列式的定义与性质

定义 2.1.1 行列式的递归展开式定义

$$A=(a_{ij})\in F_{n\times n},n=1,det(A)=a_{11};n\geq 2,det(A)=\sum_{j=1}^{n}a_{ij}A_{ij}$$

A_{ij} 是 a_{ij} 的代数余子式。

$$A=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix},\alpha_i=(a_{i1},\cdots,a_{in})\in F^n$$

记 $det(A)=det(\alpha_1,\cdots,\alpha_n)$.

定理 2.1.1 上述定义下的行列式的性质

1. $det(\cdots,\alpha_i,\cdots,\alpha_j,\cdots)=-det(\cdots,\alpha_j,\cdots,\alpha_i,\cdots)$
2. $det(\cdots,\lambda\alpha_i,\cdots)=\lambda det(\cdots,\alpha_i,\cdots)$
3. $det(\cdots,\alpha+\beta,\cdots)=det(\cdots,\alpha,\cdots)+det(\cdots,\beta,\cdots)$
4. $det(A^T)=det(A)$

定义 2.1.2 行列式的公理化定义

$$A=(a_{ij})\in F_{n\times n},A=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix},det(A)=det(\alpha_1,\cdots,\alpha_n)$$

是满足下列性质的函数：

1. 反对称性：$det(\cdots,\alpha_i,\cdots,\alpha_j,\cdots)=-det(\cdots,\alpha_j,\cdots,\alpha_i,\cdots)$
2. 线性性：$det(\cdots,\lambda\alpha_i,\cdots)=\lambda det(\cdots,\alpha_i,\cdots),det(\cdots,\alpha+\beta,\cdots)=det(\cdots,\alpha,\cdots)+det(\cdots,\beta,\cdots)$
3. 规范性：$det(e_1,\cdots,e_n)=1$

定理 2.1.2 满足上述定义的函数存在且唯一。

证明 行列式的递归展开式显然满足此定义，所以存在性得证；
设 f,g 是满足上述性质的两个函数，要证 $f=g$，令 $h=f-g$

$$\begin{aligned}
h(e_1,\cdots,e_n) &= 1-1=0 \\
\alpha_i &= \sum_{j_i=1}^{n} a_{ij_i}e_{j_i} \\
h(\alpha_1,\cdots,\alpha_n) &= h(\sum_{j_1=1}^{n} a_{1j_1}e_{j_1},\cdots,\sum_{j_n=1}^{n} a_{nj_n}e_{j_n}) \\
&= \sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n} a_{1j_1}\cdots a_{nj_n}h(e_{j_1},\cdots,e_{j_n})
\end{aligned}$$

如果 $j_1,\cdots,j_n$ 中有两个以上相同，则 $f(e_{j_1},\cdots,e_{j_n})=g(e_{j_1},\cdots,e_{j_n})=h(e_{j_1},\cdots,e_{j_n})=0$
如果 $j_1,\cdots,j_n$ 两两不同，即为 $1,\cdots,n$ 的一个排列，
则 $h(e_{j_1},\cdots,e_{j_n})=(-1)^{\tau(e_{j_1},\cdots,e_{j_n})}h(e_1,\cdots,e_n)=0$。 □

由公理化定义自然导出：

定义 2.1.3 行列式的完全展开式
$A=(a_{ij})\in F_{n\times n}$,

$$det(A)=\sum_{\begin{pmatrix}1 & 2 & \cdots & n\\ j_1 & j_2 & \cdots & j_n\end{pmatrix}}(-1)^{\tau(e_{j_1},\cdots,e_{j_n})}a_{1j_1}\cdots a_{nj_n}$$

定义 2.1.4 定义行列式为方阵特征值的乘积。(特征值可以绕开行列式从映射的角度来定义)

几何含义：

1. 平行多面体的体积；
2. 映射的区域体积之比。

2.2 行列式的 laplace 展开

定理 2.2.1 $A=(a_{ij})\in F_{n\times n}$, 设 $1\leqslant i_1\leqslant\cdots\leqslant i_r\leqslant n$ 则

$$det(A)=\sum_{1\leqslant j_1\leqslant\cdots\leqslant j_r\leqslant n}A\begin{pmatrix}i_1\cdots i_r\\ j_1\cdots j_r\end{pmatrix}(-1)^{i_1+\cdots+i_r+j_1+\cdots+j_r}A\begin{pmatrix}i_{r+1}\cdots i_n\\ j_{r+1}\cdots j_n\end{pmatrix}$$

这里 $i_1\cdots i_n, j_1\cdots j_n$ 都是 $1,\cdots,n$ 的一个全排列。

证明 详见 *9.26* 课程前 *45* 分钟。 □

例 2.2.1 $A=(a_{ij})\in F_{n\times n}, rank(A)=r<n$, 则 $det(A)=0$

证明 A 的所有 $r+1$ 阶子式为 0，所以

$$det(A)=\sum_{1\leqslant j_1\leqslant\cdots\leqslant j_{r+1}\leqslant n} A\begin{pmatrix} i_1\cdots i_{r+1}\\ j_1\cdots j_{r+1}\end{pmatrix}(-1)^{i_1+\cdots+i_{r+1}+j_1+\cdots+j_{r+1}}A\begin{pmatrix} i_{r+2}\cdots i_n\\ j_{r+2}\cdots j_n\end{pmatrix}$$

显然就等于 0。 □

特别地，当 r 取 1 时，laplace 展开就是行列式的递推展开式

$$det(A)=\sum_{j=1}^{n}a_{ij}A_{ij}$$

$$\sum_{j=1}^{n}a_{kj}A_{ij}(k\neq i)=0$$

$$\Leftrightarrow AA^*=A^*A=det(A)I$$

定理 2.2.2 *Binet-Cauthy* 公式

详见 9.26 课程后 45 分钟.

2.3 行列式的计算

方法：化为三角阵，写出递推公式，加边，把行列式看作某些变量的多项式，拆和，拆积以及其他方法。
本节详见 9.28 课程和 9.29 课程前 75 分钟。

2.4 矩阵的基本运算

加法、数乘、乘积、逆矩阵、初等变换。

2.5 矩阵的秩与相抵

定义 2.5.1 $A\in F^{m\times n}$，A 的非零子式的最高阶数称为 A 的秩。

定理 2.5.1 关于秩的一些结论

1. $r(AB)\leqslant min r(A), r(B)$;
2. P，Q 可逆，$r(PAQ)=r(PA)=r(AQ)=r(A)$;
3. 初等变换不改变 A 的秩。
4. $r(A)=r,\exists$可逆$P,Q\ st.\ A=P\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}Q$

定义 2.5.2 $A,B\in F^{m\times n}$，如果 A 经过有限次初等变换化为 B，称 A 与 B 相抵。

注 相抵是一种等价关系。不难得出相抵的矩阵秩也相等，$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ 被称为相抵标准型。

例 2.5.1（满秩分解）
$A \in F^{m\times n}, r(A) = r, \exists B \in F^{m\times r}, C \in F^{r\times n}\ st.\ A = BC$ 且 $r(B) = r(C) = r.$

证明

$$\begin{aligned} A &= P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q \\ &= P \begin{pmatrix} I_r \\ 0 \end{pmatrix} \cdot \begin{pmatrix} I_r & 0 \end{pmatrix} Q \\ &= B \cdot C \end{aligned}$$

□

例 2.5.2 $A, B \in F^{m\times n}, r(A+B) \leqslant r(A) + r(B)$

证明

$$r(A) + r(B) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & A+B \\ 0 & B \end{pmatrix} \geqslant r(A+B)$$

□

例 2.5.3（*Frobenius* 秩不等式）$r(ABC) \geqslant r(AB) + r(BC) - r(B)$
证明略。

例 2.5.4 $A \in F^{n\times n}$ 且 $r(A^k) = r(A^{k+1})$，则 $r(A^k) = r(A^{k+1}) = r(A^{k+2}) = \cdots$

证明 $r(A^{k+2}) = r(AA^kA) \geq r(AA^k) + r(A^k - A) - r(A^k) = r(A^{k+1})$ 且
$r(A^{k+2}) \leq r(A^{k+1})$，所以 $r(A^{k+2}) = r(A^{k+1})$，归纳可证。 □

例 2.5.5 $A \in F^{n\times n}, A^2 = A, \exists P\ st.\ P^{-1}AP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, r = r(A)$

证明

$$
\begin{aligned}
A &= P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q \\
A^2 &= P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} QP \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q \\
\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} QP \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \\
\text{设} QP &= \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \\
\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R_{11} & 0 \\ 0 & 0 \end{pmatrix} \\
R_{11} = I_r, Q &= \begin{pmatrix} I_r & R_{12} \\ R_{21} & R_{22} \end{pmatrix} P^{-1} \\
A &= P\begin{pmatrix} I_r & R_{12} \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} I_r & -R_{12} \\ 0 & I_{n-r} \end{pmatrix} \begin{pmatrix} I_r & R_{12} \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} I_r & -R_{12} \\ 0 & I_{n-r} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_r & R_{12} \\ 0 & I_{n-r} \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} I_r & -R_{12} \\ 0 & I_{n-r} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_r & -R_{12} \\ 0 & I_{n-r} \end{pmatrix}^{-1} P^{-1} \\
P_1 = P\begin{pmatrix} I_r & -R_{12} \\ 0 & I_{n-r} \end{pmatrix}&, \text{则} A = P_1 \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P_1^{-1}
\end{aligned}
$$

□

例 2.5.6 S 为 n 阶实对称方阵的秩为 r，则 S 至少有一个 r 阶主子式不为 *0*.

例 2.5.7 Ax=b 有解的充要条件为 $r(A, b) = r(A)$.

2.6 矩阵的广义逆

详见 10.10 课程后 15 分钟、10.12 课程前 50 分钟。

2.7 多项式矩阵相抵

定义 2.7.1 F 是数域，λ 为未定元，称 $A(\lambda) = (a_{ij}(\lambda)) \in F[\lambda]^{n\times n}$ 为多项式矩阵，$a_{ij}(\lambda) \in F[\lambda]$ 为多项式，简称为 λ 矩阵。$rank(A(\lambda))$ 同样定义为非零子式的最高阶数。

注 子式为 λ 的多项式，不等于 *0* 时即为非零子式。

定义 2.7.2 $A(\lambda)\in F[\lambda]^{n\times n},\exists B(\lambda)\in F[\lambda]^{n\times n}\ st.\ A(\lambda)B(\lambda)=I_n$，称 $A(\lambda)$ 可逆，$A(\lambda)^{-1}=B(\lambda)$

定理 2.7.1 $A(\lambda)\in F[\lambda]^{m\times n}$ 则 $A(\lambda)$ 经过有限次初等变换化为

$$B(\lambda)=\begin{pmatrix}D(\lambda)&0\\0&0\end{pmatrix},D(\lambda)=diag(d_1(\lambda),\cdots,d_r(\lambda)),d_i(\lambda)|d_{i+1}(\lambda),d_i(\lambda)\text{为首一多项式},$$

$r=rank(A(\lambda))$.

证明 不妨设 $A(\lambda)\neq 0$，首先证 $A(\lambda)$ 可通过初等变换化为 $\tilde{A}=(\tilde{a_{ij}}(\lambda))$，满足 $\tilde{a_{11}}(\lambda)|\tilde{a_{ij}}(\lambda)$.
不妨设 $a_{11}\neq 0$，设 $k=deg(a_{11}(\lambda))$.
$k=0$ 时，命题显然成立；
假设结论对 $<k$ 的情况都成立，下面证明对 k 成立：
当某个 $a_{ij}(\lambda)$ 不被 $a_{11}(\lambda)$ 所整除，

(1) $j=1,a_{i1}=q_ia_{11}+r_i,deg(r_i)<k,r_i\neq 0$. 把第一行乘上 $-q_i$ 加到第 i 行，再把第 $1,i$ 行交换得到 $A_1(\lambda)$，$A_1(\lambda)$ 第一行第一列元素 r_i 满足次数小于 k，所以 $A_1(\lambda)$ 可进一步初等变换化为 $\tilde{A}$；

$$P_{1i}T_{i1}(-q_i(\lambda))A(\lambda)=A_1(\lambda)=\begin{pmatrix}r_i(\lambda)&\cdots\\\vdots&\end{pmatrix}$$

(2) $i=1$, 同理有

$$A(\lambda)T_{j1}(-q_j(\lambda))P_{1j}=A_2(\lambda)=\begin{pmatrix}r_j(\lambda)&\cdots\\\vdots&\end{pmatrix}$$

(3) $i\neq 1,j\neq 1$, 先通过行列变换将第一行第一列的所有元素（除了 a_{11}）化为 *0*，再将第 i 行加到第一行，最后分别重复一次（*1*）、（*2*）步骤即可。

□

推论 2.7.1 $A(\lambda)\in F[\lambda]^{m\times n},\exists$一系列初等方阵$P_1(\lambda),\cdots,P_i(\lambda),Q_1(\lambda),\cdots,Q_i(\lambda)$,
$st.\ P_1(\lambda)\cdots P_i(\lambda)A(\lambda)Q_1(\lambda)\cdots Q_i(\lambda)=diag(D(\lambda),0)$.

推论 2.7.2 $A(\lambda)$ 可逆 $\Leftrightarrow A(\lambda)$ 是一系列初等方阵的乘积。

推论 2.7.3 $A(\lambda)\in F[\lambda]^{m\times n},\exists$可逆方阵$P(\lambda),Q(\lambda),\ st.\ P(\lambda)A(\lambda)Q(\lambda)=diag(D(\lambda),0)$.

定理 2.7.2 $D(\lambda)$ 由 $A(\lambda)$ 唯一确定。

定义 2.7.3 $A(\lambda)\in F[\lambda]^{m\times n}$ 的所有 k 阶子式最大公因子称为 $A(\lambda)$ 的 k 阶行列式因子，记为 $D_k(\lambda)$. 规定 $D_0(\lambda)=1$.

定理 2.7.3 $A(\lambda)\in F[\lambda]^{m\times n}$ 的秩为 r，$d_1(\lambda)\cdots d_k(\lambda)=D_k(\lambda),d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)},k=1,\cdots,r$.

证明 只需证 $D_k(\lambda)|d_1(\lambda)\cdots d_k(\lambda)$，
记 $B(\lambda)=\begin{pmatrix}D(\lambda)&0\\0&0\end{pmatrix}=P(\lambda)A(\lambda)Q(\lambda)$

$$\begin{aligned}B(\lambda)\begin{pmatrix}1\cdots k\\1\cdots k\end{pmatrix}&=\sum_{1\leqslant p_1\leqslant\cdots\leqslant p_k\leqslant n}\sum_{1\leqslant q_1\leqslant\cdots\leqslant q_k\leqslant n}P(\lambda)\begin{pmatrix}1\cdots k\\p_1\cdots p_k\end{pmatrix}A(\lambda)\begin{pmatrix}p_1\cdots p_k\\q_1\cdots q_k\end{pmatrix}Q(\lambda)\begin{pmatrix}q_1\cdots q_k\\1\cdots k\end{pmatrix}\\&=d_1(\lambda)\cdots d_k(\lambda)\end{aligned}$$

注意到式中 $A(\lambda)\begin{pmatrix}p_1\cdots p_k\\q_1\cdots q_k\end{pmatrix}$ 是 A 的一个 k 阶子式，因此可以被 $D_k(\lambda)$ 所整除，因此 $D_k(\lambda)|d_1(\lambda)\cdots d_k(\lambda)$。 □

定义 2.7.4 称 $d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)}$ 为 $A(\lambda)$ 的不变因子，$k=1,\cdots,r.B(\lambda)=\begin{pmatrix}D(\lambda)&0\\0&0\end{pmatrix}$ 称为相抵标准型，也叫 *Smith* 标准型。

定理 2.7.2 自然得证。

定理 2.7.4 $A(\lambda),B(\lambda)\in F[\lambda]^{m\times n},A(\lambda),B(\lambda)$ 相抵的充要条件为行列式因子都相等。

证明 充分性显然，只说明必要性：$A(\lambda)$ $B(\lambda),r(A(\lambda))=r(B(\lambda))=r.$ 分别用 $D_k(\lambda),\tilde{D}_k(\lambda)$ 来表示 $A(\lambda),B(\lambda)$ 的行列式因子。当 $k>r$ 时 $D_k(\lambda)=\tilde{D}_k(\lambda)=0$，所以下面只需考虑 $k\leqslant r$

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$$
$$B(\lambda)\begin{pmatrix}i_1\cdots i_k\\j_1\cdots j_k\end{pmatrix}=\sum\sum\cdots A(\lambda)\begin{pmatrix}p_1\cdots p_k\\q_1\cdots q_k\end{pmatrix}\cdots$$

注意到式中 $A(\lambda)\begin{pmatrix}p_1\cdots p_k\\q_1\cdots q_k\end{pmatrix}$ 是 A 的一个 k 阶子式，因此可以被 $D_k(\lambda)$ 所整除，因此 $D_k(\lambda)|B(\lambda)\begin{pmatrix}i_1\cdots i_k\\j_1\cdots j_k\end{pmatrix}$。所以 $D_k(\lambda)|\tilde{D}_k(\lambda)$，反过来也成立，所以 $D_k(\lambda)=\tilde{D}_k(\lambda)$。 □

定义 2.7.5 把矩阵 A 的每个次数大于零的不变因子分解成互不相同的首项为 *1* 的一次因式方幂的乘积，所有这些一次因子方幂（相同的必须按出现的次数计算）称为矩阵 A 的初等因子。例如矩阵 $A(\lambda)$ 的不变因子 $d_i(\lambda)=\prod_j(\lambda-\lambda_{ij})^{r_{ij}}$，初等因子组由所有的 $(\lambda-\lambda_{ij})^{r_{ij}}(r_{ij}>0)$ 组成，重复出现的元素不合并。

推论 2.7.4 $A(\lambda),B(\lambda)\in F[\lambda]^{m\times n},A(\lambda),B(\lambda)$ 相抵 $\Leftrightarrow$ 行列式因子都相等 $\Leftrightarrow$*Smith* 标准型相等 $\Leftrightarrow$ 不变因子都相等 $\Leftrightarrow$ 初等因子组相等。

推论 2.7.5 $J=diag(J_1,\cdots,J_s),J_i$ 是 m_i 阶若尔当方阵 $\begin{pmatrix}\lambda_i&1&&\\&\ddots&\ddots&\\&&\ddots&1\\&&&\lambda\end{pmatrix}_{m_i\times m_i}$，则 $\lambda I-J$ 的初等因子组为 $\{(\lambda-\lambda_i)^{m_i}\}$.

2.8 矩阵的相合与正定

定义 2.8.1 $A, B \in F^{n\times n}$, 如果存在可逆方阵 P 使得 $A = P^TBP$，则称矩阵 A 与 B 相合。

注 讨论矩阵相合时，我们一般只关注对称方阵。

定理 2.8.1 对称方阵一定相合与对角阵。

推论 2.8.1 对称方阵 $A \in C^{n\times n}$, 则 A 相合与 $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, r = rank(A).$

推论 2.8.2 对称方阵 $A \in R^{n\times n}$, 则 A 相合与 $\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}, p + q = rank(A), p, q$ 分别被称为正负惯性指数，A 的惯性指数记作 (p, q)。

定义 2.8.2 如果 n 阶对称方阵 A 的惯性指数为 $(n, 0)$，称 A 正定；惯性指数为 $(p, 0)$，称 A 半正定；类似可定义负定、半负定。

定义 2.8.3 A 是 n 阶实对称方阵，对于任意 n 维列向量 X 有 $X^TAX > 0$，则称 A 正定；$X^TAX \geqslant 0$，则称 A 半正定；类似可定义负定、半负定。

注 A 正定、半正定常常记作 $A > 0, A \geqslant 0.$

定理 2.8.2 A 是 n 阶实对称方阵，则下列命题等价：

(1) A 正定；

(2) A 的特征值全为正；

(3) 存在正定方阵 B，$A = B^2$；

(4) 存在可逆方阵 P，$A = P^TP$；

(5) A 的顺序主子式全为正.

证明 只说明一个最不显然的 $(5) \Rightarrow (1)$:
归纳法，假设结论对 $n-1$ 成立，即 $D_1, \cdots, D_{n-1} > 0$，那么

$$A = \begin{pmatrix} A_1 & C \\ C^T & a_{nn} \end{pmatrix}$$

中的 $A_1 > 0$，如果 $D_n > 0, det(A) > 0$，令 $P = \begin{pmatrix} I_{n-1} & -A_1^{-1}C \\ 0 & 1 \end{pmatrix}, P^TAP = \begin{pmatrix} A_1 & 0 \\ 0 & a \end{pmatrix}$，其中 $a = a_{nn} - C^TA_1^{-1}C$, 显然 $a \cdot det(A_1) = det(P^TAP) = det(P)^2det(A) > 0, a > 0$, 所以 A 正定，命题得证。 □

定理 2.8.3 A 是 n 阶实对称方阵，存在唯一正定方阵 B，$A = B^2$，记作 $B = \sqrt{A}$.

证明 存在性易证。下面证明唯一性：
假设 A 的特征值为 $\lambda_1 \geqslant \cdots \geqslant \lambda_n > 0$, 设 $B_1 > 0, B_2 > 0, B_1^2 = B_2^2 = A$, 要证 $B_1 = B_2$, 首先可知 B_1, B_2 的特征值均为 $\sqrt{\lambda_1} \geqslant \cdots \geqslant \sqrt{\lambda_n} > 0$, 存在正交阵 O_1, O_2 $st.$

$$B_1 = O_1 \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} O_1^T, B_2 = O_2 \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} O_2^T$$

$$O_1 \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} O_1^T = O_2 \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} O_2^T$$

$$O_2^T O_1 \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} O_2^T O_1$$

令 $P = O_2^T O_1$，显然 P 是一个对角阵，因此 P 满足

$$P \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} P$$

进而得到 $B_1 = B_2$. □

定理 2.8.4 A 是 n 阶实对称方阵，则下列命题等价：

(1) A 半正定；

(2) A 的特征值非负；

(3) 存在方阵 P，$A = P^T P$；

(4) 存在半正定方阵 B，$A = B^2$；

(5) A 的 k 阶主子式之和非负，$k = 1, \cdots, n$.

定义 2.8.4 $A, B \in F^{m\times n}$，如果存在正交方阵 U, V $st.$ $UAV = B$，称 A 与 B 正交相抵。

定义 2.8.5 $A \in R^{m\times n}, AA^T$ 的非零特征值 $\lambda_1, \cdots, \lambda_r$ 都为正，令 $\sigma_i = \sqrt{\lambda_i}, i = 1, \cdots, r$，称为矩阵 A 的奇异值。

推论 2.8.3 $A, B \in F^{m\times n}$，如果 A 与 B 正交相抵，则 AA^T, BB^T 正交相似。

定理 2.8.5 $A \in R^{m\times n}$, 存在正交方阵 U, V 使得

$$A = U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V, D = diag(\sigma_1, \cdots, \sigma_r), \sigma_1 \geqslant \cdots \geqslant \sigma_r > 0$$

.

证明 AA^T 是实对称阵，其正交相似于对角阵，

$$AA^T = U\begin{pmatrix}\sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots\end{pmatrix}U^T$$

$$\begin{aligned}(U^TA)(U^TA)^T &= \begin{pmatrix}\sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots\end{pmatrix}\\ &= BB^T\\ \text{令}B &= \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix}, BB^T = (\beta_i\beta_j^T)_{n\times n}\end{aligned}$$

$\beta_i\beta_i^T = \sigma_i^2(i \leqslant r), \beta_i\beta_i^T = 0(r < i \leqslant n), \beta_i\beta_j^T = 0(i \neq j)$，所以 $\beta_1, \cdots, \beta_r$ 相互正交，$\beta_i = 0(r < i \leqslant n)$，令 $\gamma_i = \frac{\beta_i}{\sigma_i}(i \leqslant r)$，

$$\begin{aligned}B = \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix} &= \begin{pmatrix}\sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots\end{pmatrix}\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}\\ &= \begin{pmatrix}D & 0\\ 0 & 0\end{pmatrix}V, \text{且 } V \text{ 是正交阵}\\ A = UB &= U\begin{pmatrix}D & 0\\ 0 & 0\end{pmatrix}V\end{aligned}$$

□

定理 2.8.6 $A \in R^{n\times n}$，则存在半正定方阵 S 及正交阵 O 使得 $A = SO$，且 S 由 A 唯一确定。

证明 $A = U\begin{pmatrix}D & 0\\ 0 & 0\end{pmatrix}V = U\begin{pmatrix}D & 0\\ 0 & 0\end{pmatrix}U^T \cdot UV, S = U\begin{pmatrix}D & 0\\ 0 & 0\end{pmatrix}U^T$是半正定方阵，
$O = UV$是正交阵，$AA^T = SOO^TS^T = SS^T = S^2$，根据定理 *2.8.3*，$S$ 由 A 唯一确定。 □

例 2.8.1 $A > 0$，存在可逆上三角阵 P 使得 $A = P^TP$.

证明 方法一：由于 P^{-1} 也是上三角矩阵，所以原命题等价于存在可逆上三角阵 P 使得 $P^TAP=I$. 归纳：$n=1$ 时命题成立，假设命题对 $n-1$ 成立，

$$A=\begin{pmatrix}A_1 & C\\ C^T & a_{nn}\end{pmatrix}>0$$
$$A_1>0,\text{存在上三角阵}P_1, P_1^TA_1P_1=I$$
$$\tilde{A}=\begin{pmatrix}P_1 & 0\\ 0 & 1\end{pmatrix}^TA\begin{pmatrix}P_1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}I_{n-1} & P_1^TC\\ C^TP_1 & a_{nn}\end{pmatrix}$$
$$\begin{pmatrix}I_{n-1} & -P_1^TC\\ 0 & 1\end{pmatrix}^T\tilde{A}\begin{pmatrix}I_{n-1} & -P_1^TC\\ 0 & 1\end{pmatrix}=\begin{pmatrix}I_{n-1} & 0\\ 0 & a\end{pmatrix}, a>0$$
$$\text{令}P=\begin{pmatrix}P_1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}I_{n-1} & -P_1^TC\\ 0 & 1\end{pmatrix}\begin{pmatrix}I_{n-1} & 0\\ 0 & a^{-\frac{1}{2}}\end{pmatrix}$$

则有 $P^TAP=I$，得证。

方法二：对任意方阵 P 的列向量作施密特正交化，得到 $P=QR$，Q 是正交阵，R 是对应的上三角系数阵。A 正定，所以存在可逆 P，使得 $A=P^TP=R^TR$，结论得证。 □

注 线性方程组 $Ax=b\Leftrightarrow P^TPx=b\Leftrightarrow P^Ty=b, Px=y$，系数矩阵为上三角阵，非常好解。

例 2.8.2 $A>0, K$ 是反对称阵，则 $det(A+K)>0.$

证明 首先有 $det(A+K)\neq 0$，否则存在 $x\neq 0,(A+K)x=0, x^T(A+K)x=0,$ 但 $x^TAx>0, x^TKx=0,$ 矛盾。
令 $f(\lambda)=det(A+\lambda K)\neq 0, f(0)>0$ 所以 $f(\lambda)>0, f(1)=det(A+K)>0$ 得证。 □

例 2.8.3 $A=(a_{ij})\in R^{n\times n}, A>0,$ 则 $det(A)\leqslant a_{11}a_{22}\cdots a_{nn}.$

证明 归纳，$n=1$ 成立，假设对 $n-1$ 成立，

$$A=\begin{pmatrix}A_1 & C\\ C^T & a_{nn}\end{pmatrix}, P_1=\begin{pmatrix}I_{n-1} & -A^TC\\ 0 & 1\end{pmatrix}$$

则 $P_1^TAP_1=diag(A_1, a_{nn}-C^TA_1^{-1}C), det(A)=det(P_1^TAP_1)=det(A_1)(a_{nn}-C^TA_1^{-1}C)\leqslant det(A_1)a_{nn}\leqslant a_{11}a_{22}\cdots a_{nn}.$ □

例 2.8.4（*Hadamade* 不等式）对 n 阶方阵 $A=\{a_{ij}\}_{n\times n}\in M_n(R)$，有

$$|A|^2\leqslant\prod_{i=1}^n(\sum_{j=1}^n a_{ij}^2)$$

证明

$$B=AA^T, |A|^2=|B|\leqslant b_{11}\cdots b_{nn}=\prod_{i=1}^n(\sum_{j=1}^n a_{ij}^2)$$

□

例 2.8.5 $A \geqslant B > 0$，则 $det(A) \geqslant det(B) > 0$

证明 存在可逆方阵 P 使得 $P^TAP = I, B_1 = P^TBP$ 仍正定且 $I - B_1 = P^T(A-B)P \geqslant 0$，要证 $det(A) \geqslant det(B)$，只需证 $1 \geqslant det(B_1).B_1$ 可正交相似于对角阵，

$$B_1 = U^TDU, I - B_1 = U^T(D-I)U \geqslant 0$$

故 B_1 的特征值 $\mu_i \leqslant 1, det(B_1) \leqslant 1$ 得证。 □

注 题目条件改为 $A \geqslant B \geqslant 0$，结论还成立吗？提示：令 $A_\lambda = A + \lambda I > 0$。

例 2.8.6 $A \in R^{m\times n}, \sigma_1, \cdots, \sigma_r$ 是 A 的奇异值，则 $\sum_{i,j} a_{ij}^2 = \sum_i \sigma_i^2$.

证明

$$A = U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}V, AA^T = U\begin{pmatrix} D^2 & 0 \\ 0 & 0 \end{pmatrix}U^T$$

$$\sum_{i,j} a_{ij}^2 = tr(AA^T) = tr(D^2) = \sum_i \sigma_i^2$$

□

例 2.8.7 $A \in R^{n\times n}$，则 $tr(A) \leqslant tr(\sqrt{AA^T})$，且等号成立当且仅当 A 半正定。

证明

$$A = U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}V, \sqrt{AA^T} = U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}U^T$$

$$\begin{aligned} tr(A) &= tr(\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}VU) \\ O &= VU = (p_{ij})_{n\times n} \\ tr(A) &= tr(\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}O) \\ &= \sum_{i=1}^r \sigma_i p_{ii} \leqslant \sum_{i=1}^r \sigma_i = tr(\sqrt{AA^T}) \end{aligned}$$

等号成立当且仅当

$$\begin{aligned} p_{ii} &= 1, r = 1, \cdots, r \\ O &= \begin{pmatrix} I_r & 0 \\ 0 & R \end{pmatrix} \\ A &= U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}OU^T \\ &= U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & R \end{pmatrix}U^T \\ &= U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}U^T \end{aligned}$$

即 A 是半正定的。 □

定义 2.8.6 $H \in C^{n\times n}, H^*$ 是 H 的共轭转置，若 $H^* = H$，则称 H 为 *Hermite* 方阵。
$A, B \in C^{n\times n}$, 若存在复方阵 P 使得 $P^*AP = B$，则称 A，B 共轭相合。
酉空间：

$$\alpha = \sum_{i=1}^{n} x_i\alpha_i, \beta = \sum_{i=1}^{n} y_i\beta_i, (\alpha, \beta) = x^*Hy$$

$H = (\alpha_i, \alpha_j)_{n\times n}$ 称为度量矩阵。

定理 2.8.7 H 是 *Hermite* 方阵，则 H 共轭相合于对角阵 $\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}$

. 此标准型唯一，$p+q = rank(H), (p,q)$ 被称为 H 的惯性指数。

定义 2.8.7 H 为 *Hermite* 方阵，类似实对称方阵，可定义 H 正定、半正定等。

定理 2.8.8 H 是 *Hermite* 方阵，则下列命题等价：

(1) H 正定；

(2) H 的特征值全为正；

(3) 存在唯一 *Hermite* 方阵 H_1，$H = H_1^2$；

(4) 存在可逆方阵 P，$H = P^TP$；

(5) H 的顺序主子式全为正.

定理 2.8.9 $A \in C^{m\times n}$, 存在酉矩阵 U, V 使得

$$A = U\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}V, D = diag(\sigma_1, \cdots, \sigma_r), \sigma_1 \geqslant \cdots \geqslant \sigma_r > 0$$

.σ_i 同样被称为 A 的奇异值，是 A^*A 的非零特征值的平方根。

注（*M-P* 广义逆）

$$A^{-1} = V^*\begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}U^*$$

定理 2.8.10 $A \in C^{n\times n}$，则存在半正定方阵 S 及酉矩阵 O 使得 $A = SO$，且 S 由 A 唯一确定。

例 2.8.8 $H = A + Bi$ 是正定 *Hermite* 方阵，A，B 为实方阵，则 $A > 0$.

证明 $H^* = A^T - B^Ti = A + Bi, A = A^T, B = -B^T$, 所以 $xHx^T = xAx^T > 0, A > 0.$ □

例 2.8.9 A，B 都是 *Hermite* 方阵，$A > 0$，则 AB 的特征值都是实数。

证明 $A = P^*P, P$ 可逆，$AB = P^*PB = P^*PBP^*(P^*)^{-1}$，所以 AB 相似于 *Hermite* 方阵 PBP^*，所以特征值都是实数。 □

例 2.8.10 H_1, H_2 都是 *Hermite* 方阵，$H_1 > 0$，则 $H_1 + H_2 > 0 \Leftrightarrow H_1^{-1}H_2$ 的特征值均大于 -1.

证明 $H_1^{-1} > 0, H_1^{-1}H_2$ 的特征值均是实数。$H_1 = P^*P$,

$$
\begin{aligned}
H_1 + H_2 > 0 &\Leftrightarrow P^*P + H_2 > 0 \\
&\Leftrightarrow I + (P^{-1})^*H_2P^{-1} > 0 \\
&\Leftrightarrow (P^{-1})^*H_2P^{-1}\text{特征值均大于 } -1 \\
&\Leftrightarrow H_1^{-1}H_2 = P^{-1} \cdot (P^{-1})^*H_2P^{-1} \cdot P\text{特征值均大于 } -1
\end{aligned}
$$

□

第三章　线性空间

3.1　线性空间基本理论

主要是复习内容，只挑比较重要的概念记了。

定义 3.1.1 V 是非空集合，F 是数域，对 V 中的元素定义运算：

1. 加法：任意 $\alpha, \beta \in V$ 构成的有序对 (α, β) 存在唯一 $\gamma \in V$ 与之对应，简记为 $\alpha + \beta = \gamma$.
2. 数乘：任意 $\alpha \in V, \lambda \in F$ 存在唯一 $\gamma \in V$ 与之对应，简记为 $\lambda\alpha = \gamma$.

且满足如下规律：

(A1) $\alpha + \beta = \beta + \alpha$.

(A2) $(\alpha + \beta) + \gamma = \beta + (\alpha + \gamma)$.

(A3) $\exists \theta \in V, \alpha + \theta = \theta + \alpha = \alpha$，称 θ 为零元素，记作 *0*.

(A4) $\exists \alpha' \in V, \alpha + \alpha' = \alpha' + \alpha = 0$，称为 α 的负元素，记作 $-\alpha$，并定义 $\beta - \alpha = \beta + (-\alpha)$.

(D1) $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$.

(D2) $(\lambda + \mu)\alpha = \lambda\alpha + \mu\alpha$.

(M1) $\lambda(\mu\alpha) = (\lambda\mu)\alpha$.

(M2) $1\alpha = \alpha$.

则称 V 是 F 上的线性空间，简记为 $V(F)$ 或 V. 线性空间 V 中的元素称为向量。

定义 3.1.2 V 是 F 上的线性空间，W 是 V 的非空子集，且 W 关于线性空间 V 的加法与数乘运算保持封闭，则称 W 是 V 的一个子空间。

定义 3.1.3 V 是 F 上的线性空间，S 是 V 的非空子集，则

$$\langle S \rangle = \{\lambda_1\alpha_1 + \cdots + \lambda_m\alpha_m | \lambda_i \in F, \alpha_i \in V\}$$

是 V 的一个子空间，称为 V 的生成子空间，S 称为生成子空间的生成元。

定义 3.1.4 V 是 F 上的线性空间，如果向量组 $S, T \subset V$，T 中的每一个向量都可以由向量组 S 线性表示，则称 T 可以由 S 线性表示。如果 T 和 S 可以相互线性表示，称 S 与 T 等价。

定义 3.1.5 V 是 F 上的线性空间，S 是 V 中一组向量。如果 S 中的某个向量能用 S 中其他向量线性表示，则称 S 线性相关，否则称线性无关。

注 一个向量组成的向量组线性相关当且仅当该向量为零向量。

定义 3.1.6 S 是线性空间 V 中的向量组，若 S 的子集 S_1 线性无关，且任加 S 中的一个其他向量 α 后 $S_1 \cup \{\alpha\}$ 线性相关，则 S_1 称为向量组 S 的极大无关组。

定理 3.1.1 两个等价的向量组 $\{\alpha_1, \cdots, \alpha_r\}, \{\beta_1, \cdots, \beta_s\}$ 分别向量无关，则 $r = s$.

证明

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix} = A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_s \end{pmatrix}, A \in F^{r \times s}$$

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_s \end{pmatrix} = B \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}, B \in F^{s \times r}$$

则 $AB = I_r, BA = I_s.$，因此 A 和 B 都只能是方阵且互为逆矩阵，所以 $r = s$. □

定义 3.1.7 向量组 S 所含极大无关组的向量的个数称为向量组的秩，记为 $rank(S)$ 或 $r(S)$.

定义 3.1.8 设 n 维线性空间 V 在两组基 $\{\alpha_1, \cdots, \alpha_n\}, \{\beta_1, \cdots, \beta_n\}$ 下的关系式为

$$(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)T,$$

矩阵 T 称为从基 $\{\alpha_1, \cdots, \alpha_n\}$ 到基 $\{\beta_1, \cdots, \beta_n\}$ 的过渡矩阵。

$v \in V$ 在两组基下的坐标分别为 x, y，则 $x = Ty$.

定义 3.1.9

$$A_{m \times n} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \beta_1, \cdots, \beta_n \end{pmatrix}$$

A 的行空间 $R(A) = \langle \alpha_1, \cdots, \alpha_m \rangle$ 和列空间 $\langle \beta_1, \cdots, \beta_n \rangle$ 的秩相同，均等于 A 矩阵的秩。
A 的零空间 $N(A) = \{x \in F^n | Ax = 0\}$，秩等于 $n - r(A)$.

例 3.1.1 求 *Fibonacci* 数列通项公式：$F_1 = 1, F_2 = 1, F_{n+2} = F_{n+1} + F_n$.
设 $V = \{(a_1, \cdots, a_n) | a_1, a_2 \in R, a_{i+2} = a_{i+1} + a_i\}$，则 V 是 R 上的线性空间，且 V 中的每个向量 $(a_1, \cdots, a_n)$ 由 a_1, a_2 唯一确定，所以 V 是二维线性空间。
显然，$u = (1, 0, \cdots), v = (0, 1, \cdots) \in V$ 是 V 的一组平凡基，设 $(1, q, \cdots, q^{n-1}) \in V, q^2 = q + 1 \Rightarrow q_{1,2} = \frac{1 \pm \sqrt{5}}{2}$，不难验证 $\alpha = (1, q_1, \cdots, q_1^{n-1}), \beta = (1, q_2, \cdots, q_2^{n-1})$ 是 V 的一组基，且

$$(u, v) = (\frac{\alpha - \beta}{q_1 - q_2}, \frac{q_1\beta - q_2\alpha}{q_1 - q_2})$$

所以

$$u + v = (1, 1, F_3, \cdots, F_n) = \frac{\alpha - \beta + q_1\beta - q_2\alpha}{q_1 - q_2}$$

$$F_n = \frac{q_1^{n-1} - q_2^{n-1} + q_1 q_2^{n-1} - q_2 q_1^{n-1}}{q_1 - q_2} = \frac{(\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n}{\sqrt{5}}$$

例 3.1.2 $A \in F^{m\times n}, B \in F^{n\times p}$, 求证 $r(A)+r(B)-n \leqslant r(AB)$.

证明 设

$$
\begin{aligned}
V_A &= \{x \in F^n | Ax = 0\} \\
V_B &= \{y \in F^p | By = 0\} \\
V_{AB} &= \{z \in F^p | ABz = 0\}
\end{aligned}
$$

注意到 $\forall z \in V_{AB}, Bz \in V_A$, 所以设 $V_B' = \{y = Bz | z \in V_{AB}\} \subset V_A$，现求 V_B' 的维数：设 V_B 的一组基为 $\{\alpha_1, \cdots, \alpha_r\}$，由于 $V_B \subset V_{AB}$，所以把基 $\{\alpha_1, \cdots, \alpha_r\}$ 扩充为 V_{AB} 的一组基 $\{\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s\}$，

$$
\begin{aligned}
\forall z \in V_{AB}, z &= \lambda_1\alpha_1 + \cdots + \lambda_r\alpha_r + \mu_1\beta_1 + \cdots + \mu_s\beta_s \\
Bz &= \lambda_1 B\alpha_1 + \cdots + \lambda_r B\alpha_r + \mu_1 B\beta_1 + \cdots + \mu_s B\beta_s \\
&= \mu_1 B\beta_1 + \cdots + \mu_s B\beta_s
\end{aligned}
$$

所以

$$V_B' = \langle B\beta_1, \cdots, B\beta_s \rangle$$

，下证 $B\beta_1, \cdots, B\beta_s$ 线性无关：
否则，$\exists \mu_1, \cdots, \mu_s$

$$
\begin{aligned}
&\mu_1 B\beta_1 + \cdots + \mu_s B\beta_s = B(\mu_1\beta_1 + \cdots + \mu_s\beta_s) = 0 \\
&\mu_1\beta_1 + \cdots + \mu_s\beta_s \in V_B \\
&\exists \lambda_1\alpha_1 + \cdots + \lambda_r\alpha_r = \mu_1\beta_1 + \cdots + \mu_s\beta_s
\end{aligned}
$$

这与 $\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s$ 线性无关矛盾。
因此 $B\beta_1, \cdots, B\beta_s$ 是 V_B' 的一组基，所以 $dim(V_B') = s = dim(V_{AB}) - dim(V_B) \leqslant dim(V_A)$
即 $p - r(AB) - p + r(B) \leqslant n - r(A)$，命题得证。 □

例 3.1.3 无限域 F 上的线性空间 V 不能被它的有限个真子空间覆盖。即设 $V_1, \cdots, V_k$ 是 V 的真子空间，则存在向量 $\alpha \notin \bigcup_{i=1}^k V_i$.

3.2 线性空间同构与同态

定义 3.2.1 设 V_1, V_2 都是 F 上的线性空间，如果存在一一映射 $\sigma : V_1 \to V_2$ 满足：

1. $\sigma(x+y) = \sigma(x) + \sigma(y)$
2. $\sigma(\lambda x) = \lambda\sigma(x)$

称 V_1, V_2 同构，σ 称为同构映射。显然，同构是一种等价关系。同构映射的逆、复合均是同构映射。

定理 3.2.1 V 是 F 上的线性空间，$\alpha_1, \cdots, \alpha_n$ 是 V 上一组线性无关的向量，令

$$\beta_i = \sum_{j=1}^n a_{ij}\alpha_j, i = 1, \cdots, m, A = (a_{ij}) \in F^{m\times n}$$

则 $dim\langle \beta_1, \cdots, \beta_m \rangle = rank(A)$.

证明 由题意知

$$\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix}=A\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}$$

先设 $V=F^n$，于是 $\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}$ 为满秩方阵，自然 $rank\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix}=dim\langle\beta_1,\cdots,\beta_m\rangle=rank(A)$.
对于一般的线性空间，构造同构映射

$$\sigma:\langle\alpha_1,\cdots,\alpha_n\rangle\to F^n,\sigma(\alpha_i)=e_i$$

设 A 的行向量为 A_i，则有 $\sigma(\beta_i)=A_i$，于是 $dim\langle\beta_1,\cdots,\beta_m\rangle=dim\langle\sigma(\beta_1),\cdots,\sigma(\beta_m)\rangle=dim\langle A_1,\cdots,A_m\rangle=rank(A)$. □

定义 3.2.2 设 V_1,V_2 都是 F 上的线性空间，如果存在映射 $\phi:V_1\to V_2$ 满足：

1. $\phi(x+y)=\phi(x)+\phi(y)$
2. $\phi(\lambda x)=\lambda\phi(x)$

ϕ 称为同态映射。

例 3.2.1 V 是 F 上的线性空间，$\alpha_1,\cdots,\alpha_n$ 是 V 上一组线性无关的向量，设 V 中的一组向量组 $\beta_1,\cdots,\beta_m$ 可以由 $\alpha_1,\cdots,\alpha_n$ 线性表示，且 $m<n$，求证 $\beta_1,\cdots,\beta_m$ 线性相关。

证明 由题意知

$$\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix}=A\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}$$

构造同态映射：

$$\phi:F^n\to V,\phi(x_1,\cdots,x_n)=x_1\alpha_1+\cdots+x_n\alpha_n$$

设 A 的行向量为 A_i，则 $\phi(A_i)=\beta_i$. 由于 $rank(A)\leqslant m<n$，所以 $A_1,\cdots,A_m$ 线性相关，从而 $\beta_1,\cdots,\beta_m$ 也线性相关。 □

3.3 子空间运算

定义 3.3.1 V_1,V_2 是线性空间 V 的子空间，则

$$\{\alpha+\beta|\alpha\in V_1,\beta\in V_2\}$$

称为子空间 V_1,V_2 的和，记作 V_1+V_2.

定理 3.3.1

$$dim(V_1+V_2)=dim(V_1)+dim(V_2)-dim(V_1\cap V_2)$$

证明 设 $dim(V_1 \cap V_2) = r, dim(V_1) = r+s, dim(V_2) = r+t$, 如果 $\{\alpha_1, \cdots, \alpha_r\}$ 是 $V_1 \cap V_2$ 的一组基，则可分别扩充为 $\{\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s\}$ 和 $\{\alpha_1, \cdots, \alpha_r, \gamma_1, \cdots, \gamma_t\}$，使其分别为 V_1, V_2 的一组基。于是

$$V_1 + V_2 = \langle \alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s, \gamma_1, \cdots, \gamma_t \rangle$$

$\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s, \gamma_1, \cdots, \gamma_t$ 线性无关，否则：
存在不全为 *0* 的一系列常数 λ_i 使得

$$\begin{aligned}
&\lambda_1\alpha_1 + \cdots + \lambda_{r+s+t}\gamma_s = 0 \\
&\lambda_1\alpha_1 + \cdots + \lambda_{r+s}\beta_s = -(\lambda_{r+t+1}\gamma_1 + \cdots + \lambda_{r+s+t}\gamma_s) \in V_2, \text{从而} \in V_1 \cap V_2 \\
&\beta_1, \cdots, \beta_s \text{线性无关，则} \lambda_{r+1}, \cdots, \lambda_{r+s} = 0 \\
&\lambda_1\alpha_1 + \cdots + \lambda_r\alpha_r + \lambda_{r+s+1}\gamma_1 + \cdots + \lambda_{r+s+t}\gamma_t = 0
\end{aligned}$$

与 $\alpha_1, \cdots, \alpha_r, \gamma_1, \cdots, \gamma_t$ 线性无关矛盾。
所以 $dim(V_1 + V_2) = r + s + t$，命题得证。 □

定义 3.3.2 V_1, V_2 是线性空间 V 的子空间，如果任意 $\alpha \in V_1 + V_2$ 可以唯一地写成

$$\alpha = \alpha_1 + \alpha_2, \alpha_1 \in V_1, \alpha_2 \in V_2,$$

称 $V_1 + V_2$ 为直和，记作 $V_1 \oplus V_2$. 如果 $V = V_1 \oplus V_2$，则称 V_1, V_2 互为补空间。

例 3.3.1 设 V 是 F 上的 n 维线性空间，$S_1, \cdots, S_k$ 是 V 的 m 维子空间，$m < n$. 求证：存在 V 的子空间 T，使得 $\forall i, V = S_i \oplus T$.

证明 根据例 *3.1.3*，存在 $\alpha_1 \notin \bigcup_{i=1}^k S_i$. 如果 $n = m+1$，则取 $T_1 = \langle \alpha_1 \rangle$，则 $V = S_i \oplus T_1$；否则，取 $\alpha_2 \notin T_1 \cup \bigcup_{i=1}^k S_i, T_2 = \langle \alpha_1, \alpha_2 \rangle$，重复此步骤直至取 $T_{n-m} = \langle \alpha_1, \cdots, \alpha_{n-m} \rangle$ ，此时一定有 $V = S_i \oplus T_{n-m}$. □

3.4　商空间

定义 3.4.1 设 V 是线性空间，W 是 V 的子空间，如果 $\alpha, \beta \in V, \alpha - \beta \in W$，称 α, β 模 W 同余。同余是一种等价关系。按照同余关系，可将 V 中的向量划分为同余类，α 的同余类可以表示为 $\tilde{\alpha} = \{\alpha + W\} = \{\alpha + \gamma | \gamma \in W\}$. 用 V/W 表示所有模 W 同余类的集合，定义线性运算 $\tilde{\alpha} + \tilde{\beta} = \widetilde{\alpha + \beta}, \lambda\tilde{\alpha} = \widetilde{\lambda\alpha}$，易验证 V/W 在上述运算下构成线性空间，称为商空间，W 是此空间的零元素。

定理 3.4.1 商空间 V/W 与 W 的补空间 W' 同构。

证明 考虑同构映射 $\sigma : V/W \to W', \sigma(\alpha) = \alpha_2$，其中 $\alpha = \alpha_1 + \alpha_2, \alpha_1 \in W, \alpha_2 \in W'$.

$$\begin{aligned}
&\sigma(\tilde{\alpha}) = \alpha_2, \sigma\text{是满射；} \\
&\sigma(\widetilde{\alpha_1}) = \sigma(\widetilde{\alpha_2}) \Rightarrow \widetilde{\alpha_1} = \widetilde{\alpha_2}, \sigma\text{是单射；} \\
&\sigma(\widetilde{\alpha_1} + \widetilde{\alpha_2}) = \sigma(\widetilde{\alpha_1 + \alpha_2}), \sigma(\lambda\widetilde{\alpha}) = \lambda\sigma(\widetilde{\alpha}), \sigma\text{保持线性性。}
\end{aligned}$$

□

推论 3.4.1 $dim\ V/W = dim\ V - dim\ W$，记 $dim\ V/W = codim\ W$，称为 W 的余维数。

例 3.4.1 $V = F[t]$，W 是所有能被 t^n 整除的全体多项式构成的子空间，求 $codim\ W$.

例 3.4.2 $codim\ V_1 + V_2 + codim\ V_1 \cup V_2 = codim\ V_1 + codim\ V_2$.

证明 见本章补充内容。 □

3.5 对偶空间

定义 3.5.1 设 V 是 F 上的线性空间，如果 $f:V\to F$ 满足

$$f(\lambda x+\mu y)=\lambda f(x)+\mu f(y),\lambda,\mu\in F,x,y\in V,$$

则称 f 是 V 上的线性函数。V 中的任意两个线性函数 f,g 可定义加法和数乘运算：

$$(f+g)(x)=f(x)+g(x),(\lambda f)(x)=\lambda f(x),$$

V 上所有的线性函数在上述运算下构成一个线性空间，称为 V 的对偶空间，记作 V^*.

定理 3.5.1 设 V 是 F 上的 n 维线性空间，则 V^* 也是 n 维线性空间。设 $e_1,\cdots,e_n$ 是 V 的一组基，$e^1,\cdots,e^n$ 是 V^* 中满足 $e^i(e_j)=\delta_{ij}$ 的线性函数，则 $e^1,\cdots,e^n$ 是 V^* 的一组基，称为 $e_1,\cdots,e_n$ 的对偶基。

证明 首先证 $e^1,\cdots,e^n$ 线性无关，否则存在一系列常数 λ_i 使得

$$\lambda_1e^1(x)+\cdots+\lambda_ne^n(x)=0$$

令 $x=e_1$ 得到 $1=0$，矛盾。接下来只需证明，对于任意 $f\in V^*,f=f(e_1)e^1+\cdots+f(e_n)e^n$. 设

$$f(e_1)e^1+\cdots+f(e_n)e^n=\tilde{f}$$

于是 $\tilde{f}(e_i)=f(e_i),i=1,\cdots,n$, 从而任意 $x=\lambda_1e_1+\cdots+\lambda_ne_n,\tilde{f}(x)=\lambda_1f(e_1)+\cdots+\lambda_nf(e_n)=f(x)$. □

注 实际上，由于 $dim\,V=dim\,V^*,\sigma:e_i\to e^i$ 确定了 V 到 V^* 的一个同构映射，$(V^*)^*=V^{**}$ 也与 V 同构。实际上，定义 V^* 上的线性函数 f_x,

$$f_x(\alpha)=\alpha(x),\forall\alpha\in V^*$$

容易验证 $f_x\in V^{**}$.

定理 3.5.2 映射 $f:x\to f_x$ 是 V 到 V^{**} 的同构映射。

证明 首先验证 f 是线性的，实际上：

$$x+y\to f_{x+y}=\alpha(x+y)=\alpha(x)+\alpha(y)=f_x+f_y,$$
$$\lambda x\to f_{\lambda x}=\alpha(\lambda x)=\lambda\alpha(x)=\lambda f_x.$$

下面说明 f 是双射，取 V 的一组基 $e_1,\cdots,e_n$, 在 V^* 的对应的对偶基是 $e^1,\cdots,e^n$, 则

$$f_{e_i}(e^j)=e^j(e_i)=\delta_{ij},$$

于是 $f_{e_1},\cdots,f_{e_n}$ 就成为 V^{**} 的一组基，并且是 $e^1,\cdots,e^n$ 的一组对偶基，所以 f 把 V 的一组基映到 V^{**} 的一组基，从而是双射。 □

注 V^{**} 和 V 自然同构，不依赖于选取的基，所以可以把 V 和 V^{**} 等同起来，即 V 中的元素是 V^* 上的线性函数。

定义 3.5.2 W 是 V 的子空间，定义 W 的零化子：

$$W^0 = \{\alpha \in V^* | \alpha(v) = 0, \forall v \in W\}$$

W^0 是 V^* 的子空间。

定理 3.5.3 $dim\ W^0 = dim\ V - dim\ W.$

证明 取 W 的一组基 $e_1, \cdots, e_k$，将其扩充为 V 的一组基 $e_1, \cdots, e_n$，其对应的对偶基为 $e^1, \cdots, e^n$，下面证明 $W^0 = \langle e^{k+1}, \cdots, e^n \rangle$：
实际上，对于任意 $\alpha \in W^0, \alpha = \lambda_1 e^1 + \cdots + \lambda_n e^n$，由 $\alpha(v) = 0, \forall v \in W$ 可知，$\alpha(e_i) = 0, i = 1, \cdots, k$，从而 $\lambda_i = 0, i = 1, \cdots, k, \alpha = \lambda_{k+1} e^{k+1} + \cdots + \lambda_n e^n$，又因为 $e^j(e_i) = 0, 1 \leqslant i \leqslant k, k+1 \leqslant j \leqslant n$，所以 $e^j \in W^0, k+1 \leqslant j \leqslant n$，至此 $W^0 = \langle e^{k+1}, \cdots, e^n \rangle$ 得证，原命题得证。 □

对于 V^* 的子空间 U, U 的零化子就是 V^{**} 的子空间，把 V^{**} 和 V 等同，则

$$U^0 = \{v \in V | \alpha(v) = 0, \forall \alpha \in U\}$$

推论 3.5.1 对 V 的任意子空间 W，有 $(W^0)^0 = W.$

推论 3.5.2 V 的任意子空间都是 V^* 某个子空间的零化子，反过来，V^* 的任意子空间都是 V 的某个子空间的零化子。

例 3.5.1 设 $V = F^{n \times n}$ 是 F 上全体 n 阶方阵构成的线性空间，f 是 V 上的线性函数，证明存在唯一矩阵 $A \in V$ 使得 $f(X) = tr(AX), X \in V.$

证明 用 E_{ij} 表示第 i 行第 j 列元素为 *1*，其余为 *0* 的 n 阶矩阵，则 V 的一组平凡基为 $E_{11}, \cdots, E_{1n}, \cdots, E_{nn}$. 设其对应的对偶基为 $E^{11}, \cdots, E^{nn}$, 则 f 可以写为

$$f = \lambda_{11} E^{11} + \cdots + \lambda_{nn} E^{nn} = \sum_{i=1}^{n} \sum_{j=i}^{n} \lambda_{ij} E^{ij}$$

设 $A = (a_{ij})_{n \times n}, X = (x_{ij})_{n \times n}$，则有

$$f(X) = \sum_{i=1}^{n} \sum_{j=i}^{n} \lambda_{ij} x_{ij}$$
$$tr(AX) = \sum_{i=1}^{n} \sum_{j=i}^{n} a_{ji} x_{ij}$$

对比两式可知，$f(X) = tr(AX)$ 当且仅当 $a_{ij} = \lambda_{ji}, \forall i, j.$ 由于与 f 对应的一组 λ_{ij} 唯一存在，所以对应的矩阵 A 唯一存在。 □

第四章　线性变换

4.1　基础概念

定义 4.1.1 设 U，V 是 F 上的线性空间，若映射 $\mathcal{A}: U \to V$ 满足

$$\mathcal{A}(\alpha+\beta) = \mathcal{A}(\alpha) + \mathcal{A}(\beta), \mathcal{A}(\lambda\alpha) = \lambda\mathcal{A}(\alpha)$$

则称 $\mathcal{A}$ 是从 U 到 V 的一个线性映射或线性算子。U 到 V 的全体线性映射记作 $L(U,V)$. 特别地，当 $U = V$ 时，V 到 V 的线性映射也叫 V 上的线性变换，V 上的全体线性变换记作 $L(V)$.

一类重要的映射就是矩阵 $A \in F^{m\times n}$ 确定的从 F^n 到 F^m 的线性映射：

$$\mathcal{A}: F^n \to F^m, x \mapsto Ax, x \in F^n$$

定理 4.1.1 $\mathcal{A}$ 是单射当且仅当对于任意 $\alpha \neq \theta_1, \mathcal{A}\alpha \neq \theta_2$；$\mathcal{A}$ 是满射当且仅当存在 $S \subset U$，使得 $\mathcal{A}(S)$ 是 V 的基.

证明 假设 $\mathcal{A}$ 不是单射，则存在 $\alpha \neq \beta$ 使得 $\mathcal{A}\alpha = \mathcal{A}\beta, \mathcal{A}\gamma = \theta_2, \gamma = \alpha - \beta \neq \theta_1$. 反之，设 $\mathcal{A}\gamma = \theta_2$, 但 $\gamma \neq \theta_1$, 由于 $\mathcal{A}\theta_1 = \theta_2$，因此 $\mathcal{A}$ 不是单射；
设 T 是 U 的一组基，则 $\mathcal{A}(U) = \langle\mathcal{A}(T)\rangle$. 因此，$\mathcal{A}$ 是满射当且仅当 $V = \langle\mathcal{A}(T)\rangle$. 设 $\mathcal{A}(S)$ 是 $\mathcal{A}(T)$ 的极大无关组，$S \subset T$. 则 $\mathcal{A}$ 是满射当且仅当 $\mathcal{A}(S)$ 是 V 的一组基。 □

定理 4.1.2 设 U，V 是 F 上的线性空间，$\{\alpha_i | i \in I\}$ 是 U 的基，I 是指标集合，$\{\beta_i | i \in I\}$ 是 V 中任意一组向量，则存在唯一地线性映射 $\mathcal{A} \in L(U,V)$，使得 $\mathcal{A}\alpha_i = \beta_i.\forall i \in I$.

证明（存在性）构造映射 $\mathcal{A}: U \to V$ 如下：对于任意向量 $\alpha \in U$，设

$$\alpha = x_1\alpha_{i_1} + \cdots + x_k\alpha_{i_k}.$$

令 $\mathcal{A}(\alpha) = x_1\beta_{i_1} + \cdots + x_k\beta_{i_k}$. 显然，$\mathcal{A}\alpha_i = \beta_i, \forall i \in I$，且 $\mathcal{A}$ 是线性映射。
（唯一性）设线性映射 $\mathcal{B}$ 也满足 $\mathcal{B}\alpha_i = \beta_i, \forall i \in I$, 对于 $\forall \alpha \in V$,

$$\mathcal{B}\alpha = x_1\mathcal{B}\alpha_{i_1} + \cdots + x_k\mathcal{B}\alpha_{i_k} = \mathcal{A}\alpha.$$

因此 $\mathcal{B} = \mathcal{A}$. □

若线性空间 U,V 都是有限维的，则线性变换 $\mathcal{A}: U \to V$ 可以通过矩阵来表示。

定义 4.1.2 取定线性空间 U 的一组基 $S=\{\alpha_1,\cdots,\alpha_n\}$ 和 V 的一组基 $T=\{\beta_1,\cdots,\beta_m\}$，设

$$\mathcal{A}\alpha_j=\sum_{i=1}^{m}a_{ij}\beta_i, j=1,\cdots,n$$

或者形式上记作

$$(\mathcal{A}\alpha_1,\cdots,\mathcal{A}\alpha_n)=(\beta_1,\cdots,\beta_n)A$$

则矩阵 $A=(a_{ij})\in F^{m\times n}$ 被称为线性映射 A 在基 S 下的矩阵表示。

$\mathcal{A}$ 与 A 的关系如下：

$$\begin{array}{ccc} \alpha\in U & \xrightarrow{\mathcal{A}} & \mathcal{A}\alpha\in V \\ \downarrow_S & & \downarrow_T \\ x\in F^n & \xrightarrow{A} & Ax\in F^m \end{array}$$

实际上，对于任意 $\alpha\in U$，设 $\alpha=x_1\alpha_1+\cdots+x_n\alpha_n,\mathcal{A}\alpha=y_1\beta_1+\cdots+y_m\beta_m$. 由

$$\mathcal{A}\alpha=\sum_{j=1}^{n}x_j\mathcal{A}\alpha_j=\sum_{j=1}^{n}x_j\left(\sum_{i=1}^{m}a_{ij}\beta_i\right)=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}x_j\right)\beta_i$$

可得 $y_1=\sum_{j=1}^{n}a_{ij}x_j, i=1,\cdots,m$. 即 $y=Ax$.

定理 4.1.3 设 U，V 是 F 上的有限维线性空间，S_1,S_2 是 U 的两组基，P 是从 S_1 到 S_2 的过渡矩阵，T_1,T_2 是 V 的两组基，Q 是从 T_1 到 T_2 的过渡矩阵，线性映射 $\mathcal{A}:U\to V$ 在 $\{S_i,T_i\}$ 下的矩阵表示为 $A_i, i=1,2$. 则有 $A_2=Q^{-1}A_1P$. 即 A_1 与 A_2 相抵，反之亦然。特别地，$U=V,S_1=T_1,S_2=T_2$ 时，$A_2=P^{-1}A_1P$，即 A_1 与 A_2 相似。

定理 4.1.4 设 $\mathcal{A}$ 是 $U\to V$ 的线性变换，则存在 U 的一组基 $\alpha_1,\cdots,\alpha_n$ 及 V 的一组基 $\beta_1,\cdots,\beta_m$ 使得

$$\mathcal{A}(\alpha_1,\cdots,\alpha_n)=(\beta_1,\cdots,\beta_m)\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}$$

其中 r 由 A 唯一决定，称为 $\mathcal{A}$ 的秩，记为 $rank(\mathcal{A})$.

证明 设 $\mathcal{A}$ 在 U 的一组基 $\tilde{\alpha_1},\cdots,\tilde{\alpha_n}$ 及的一组基 $\tilde{\beta_1},\cdots,\tilde{\beta_m}$ 下的矩阵为 A，则存在可逆方阵 P，Q 使得 $PAQ=\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}, r=rank(A)$. 令

$$(\alpha_1,\cdots,\alpha_n)=(\tilde{\alpha_1},\cdots,\tilde{\alpha_n})Q,$$

$$(\beta_1,\cdots,\beta_m)=(\tilde{\beta_1},\cdots,\tilde{\beta_m})P^{-1}.$$

由定理 *4.1.3* 可知结论成立。实际上，由于线性变换 $\mathcal{A}$ 在不同基下的矩阵 A 相似，秩不变，所以 r 由 $\mathcal{A}$ 唯一确定。 □

定义 4.1.3 U，V 都是 F 上的线性空间，$\mathcal{A},\mathcal{B}\in L(U,V)$，定义线性映射的加法运算和数乘运算如下：

$$(\mathcal{A}+\mathcal{B})\alpha=\mathcal{A}\alpha+\mathcal{B}\alpha, (\lambda\mathcal{A})\alpha=\lambda(\mathcal{A}\alpha), \forall\alpha\in U.$$

易验证 $L(U,V)$ 或者 $L(V)$ 均构成 F 上的线性空间。

定理 4.1.5 设 U，V 分别是 F 上的 n，m 维线性空间，S，T 分别是 U，V 的一组基。对于任意线性映射 $\mathcal{A}\in L(U,V)$，设 A 是 $\mathcal{A}$ 在 S，T 下的矩阵表示，则映射：

$$\sigma : L(U,V) \to F^{m\times n}, \mathcal{A} \mapsto A$$

是同构映射，从而 $dim\ L(U,V) = mn.$

定义 4.1.4 设 U，V，W 是 F 上的线性空间，$\mathcal{A}\in L(U,V),\mathcal{B}\in L(V,W)$ 则复合映射

$$\mathcal{B}\circ\mathcal{A} : U \to W, \alpha \mapsto \mathcal{B}(\mathcal{A}(\alpha))$$

也是线性映射，称为 $\mathcal{B}$ 与 $\mathcal{A}$ 的乘积，在不引起歧义的情况下，常把 $\mathcal{B}\circ\mathcal{A}$ 记作 $\mathcal{B}\mathcal{A}$. 进一步可以定义线性映射 $\mathcal{A}$ 的乘方，规定 $\mathcal{A}^0=\mathcal{I}$，即恒等映射。

定理 4.1.6 设 S_U,S_V,S_W 分别是 F 上有限维线性空间 U,V,W 的基，A 是 $\mathcal{A}\in L(U,V)$ 在 S_U,S_V 下的矩阵表示，B 是 $\mathcal{B}\in L(V,W)$ 在 S_V,S_W 下的矩阵表示，则 BA 是 $\mathcal{B}\circ\mathcal{A}$ 在 S_U,S_W 下的矩阵表示。

定义 4.1.5 设 U，V 是 F 上的线性空间，$\mathcal{A}\in L(U,V),\mathcal{B}\in L(V,U)$，若 $\mathcal{A}\mathcal{B}=\mathcal{I}_V$，则称 $\mathcal{A}$ 是 $\mathcal{B}$ 的左逆，$\mathcal{B}$ 是 $\mathcal{A}$ 的右逆。若 $\mathcal{A}$ 既有左逆 $\mathcal{B}$ 也有右逆 $\mathcal{C}$，则 $\mathcal{B}=\mathcal{B}(\mathcal{A}\mathcal{C})=(\mathcal{B}\mathcal{A})\mathcal{C}=\mathcal{C}$，$\mathcal{B}$ 称为 $\mathcal{A}$ 的逆映射，记作 $\mathcal{A}^{-1}$，并称 $\mathcal{A}$ 是可逆的。

定理 4.1.7 $\mathcal{A}$ 有左逆/有右逆/可逆当且仅当 $\mathcal{A}$ 是单射/满射/一一映射。

定理 4.1.8 设 U，V 是 F 上的有限维夏宁空间，S，T 分别是 U，V 的基，$\mathcal{A}\in L(U,V),A$ 是 $\mathcal{A}$ 在 S，T 下的矩阵表示，则 $\mathcal{A}$ 有左逆/有右逆/可逆当且仅当 A 有左逆/有右逆/可逆。

4.2 线性映射的像与核

定义 4.2.1 $\mathcal{A}\in L(U,V)$，称 $Im\ \mathcal{A}=\mathcal{A}(U)=\{\mathcal{A}\alpha|\alpha\in U\}$ 称为 $\mathcal{A}$ 的像，称 $Ker\ \mathcal{A}=\{\alpha|\mathcal{A}\alpha=0,\alpha\in U\}$ 为 $\mathcal{A}$ 的核。

推论 4.2.1 $Im\ \mathcal{A}$ 是 U 的子空间，$Ker\ \mathcal{A}$ 是 V 的子空间。

定理 4.2.1 $dim(Im\ \mathcal{A})=rank(\mathcal{A})$

定理 4.2.2 $V/_{Ker\mathcal{A}}$ 与 $Im\ \mathcal{A}$ 同构。

定理 4.2.3 $\mathcal{A}\in L(U,V),\mathcal{A}$ 可逆当且仅当 $Im\ \mathcal{A}=V$ 且 $Ker\ \mathcal{A}=0.$

推论 4.2.2 $\mathcal{A}\in L(U,V),\mathcal{A}$ 可逆当且仅当下列条件中的任意两个成立：

(1). $dim\ U=dim\ V$;

(2). $Im\ \mathcal{A}=V$;

(3). $Ker\ \mathcal{A}=0.$

$\mathcal{A}\in L(V),\mathcal{A}$ 可逆当且仅当 $Im\ \mathcal{A}=V$ 或 $Ker\ \mathcal{A}=0.$

例 4.2.1 $A \in F^{m\times n}, B \in F^{n\times p}, rank(AB) \leqslant min\{rank(A), rank(B)\}$.

证明 设 $\mathcal{A}: x \to Ax, x \in F^n, \mathcal{B}: y \to By, y \in F^p$，则 $Im\ \mathcal{B} = \{By|y \in F^p\} \subset F^n$，所以

$$\mathcal{A}(Im\ \mathcal{B}) \subset \mathcal{A}(F^n)$$
$$Im\ \mathcal{AB} \subset Im\ \mathcal{A}$$

并且 $\forall x \in F^p, x \in Ker\ \mathcal{B}$ 有 $ABx = A0 = 0, x \in Ker\mathcal{AB}$，得到 $Ker\ \mathcal{B} \subset Ker\ \mathcal{AB}$. 所以 $dim\ Im\ \mathcal{AB} \leqslant dim\ Im\ \mathcal{A}, dim\ Ker\ \mathcal{B} \leqslant dim\ Ker\ \mathcal{AB}$. 命题得证。 □

例 4.2.2 A 是 n 阶方阵，且 $r(A^k) = r(A^{k+1})$，则 $r(A^k) = r(A^{k+1}) = r(A^{k+2}) = \cdots$.

证明 设 $\mathcal{A}: x \to Ax$.

$$\begin{aligned} r(A^k) = r(A^{k+1}) &\Rightarrow dim\ \mathcal{A}^k(V) = dim\ \mathcal{A}^{k+1}(V) \\ &\Rightarrow \mathcal{A}^k(V) = \mathcal{A}^{k+1}(V) \\ &\Rightarrow dim\ \mathcal{A}^{k+1}(V) = dim\ \mathcal{A}^{k+2}(V) \\ &\Rightarrow r(A^{k+1}) = r(A^{k+2}) \end{aligned}$$

递推可知原命题成立。 □

例 4.2.3 A 是 n 阶方阵，则 $r(A^k) - r(A^{k+1}) \geqslant r(A^{k+1}) - r(A^{k+2})$.

证明 设 $\mathcal{A}: x \to Ax, U_1 = \mathcal{A}^k(V), U_2 = \mathcal{A}^{k+1}(V), U_3 = \mathcal{A}^{k+2}(V)$,

$$dim\ U_1 - dim\ U_2 = dim\ U_1 - dim\ \mathcal{A}U_1 = dim(Ker\ \mathcal{A}(U_1))$$
$$dim\ U_2 - dim\ U_3 = dim\ U_2 - dim\ \mathcal{A}U_2 = dim(Ker\ \mathcal{A}(U_2)) = dim(Ker\ \mathcal{A}(\mathcal{A}U_1))$$

由于 $\mathcal{A}U_1 \subset U_1$，所以 $Ker\ \mathcal{A}(\mathcal{A}U_1) \subset Ker\ \mathcal{A}(U_1)$, 进而 $dim\ U_1 - dim\ U_2 \geqslant dim\ U_2 - dim\ U_3$. 结论得证。 □

例 4.2.4 设 $\mathcal{A}$ 是 V 上的线性变换，且 $\mathcal{A}^2 = \mathcal{I}$. 证明：$V = Im(\mathcal{A}+\mathcal{I}) \oplus Im(\mathcal{A}-\mathcal{I})$.

证明 设 $\mathcal{A}: x \to Ax$. 先证 $Im(\mathcal{A}+\mathcal{I}) \cap Im(\mathcal{A}-\mathcal{I}) = 0$，设 $\alpha \in Im(\mathcal{A}+\mathcal{I}) \cap Im(\mathcal{A}-\mathcal{I})$,

$$\begin{aligned} \alpha &= (\mathcal{A}+\mathcal{I})\beta_1 = (\mathcal{A}-\mathcal{I})\beta_2 \\ (\mathcal{A}+\mathcal{I})\alpha &= (\mathcal{A}^2-\mathcal{I})\beta_2 = 0 \\ (\mathcal{A}-\mathcal{I})\alpha &= (\mathcal{A}^2-\mathcal{I})\beta_1 = 0 \\ \alpha &= -\alpha, \alpha = 0 \end{aligned}$$

而 $\forall \alpha \in V, \alpha = (\mathcal{A}+\mathcal{I})\frac{\alpha}{2} + (\mathcal{A}-\mathcal{I})\frac{-\alpha}{2}$，所以结论得证。 □

4.3 线性变换

定义 4.3.1 W 是 V 的子空间，如果 V 上的线性变换 $\mathcal{A}$ 满足 $\mathcal{A}(W) \subset W$，则称 W 是 V 的属于 $\mathcal{A}$ 的不变子空间。可以定义 $\mathcal{A}|_W : W \to W$, 称为 $\mathcal{A}$ 在 W 上的限制。

例 4.3.1 V 上的线性变换 $\mathcal{A}, \mathcal{B}$ 满足 $\mathcal{AB} = \mathcal{BA}$, 则 $Ker\ \mathcal{B}$ 和 $Im\ \mathcal{B}$ 都是 V 的不变子空间。

定理 4.3.1 任意多个不变子空间的交，仍然是不变子空间；有限多个不变子空间的和，仍然是不变子空间。

定理 4.3.2 $\mathcal{A}$ 是 V 上的线性变换，W 是 V 的不变子空间，如果 $\{\alpha_1, \cdots, \alpha_r\}$ 是 W 的一组基，$\{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一组基，那么 $\mathcal{A}$ 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵为准上三角阵；反之，如果 $\mathcal{A}$ 在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵为上三角阵，则 $W = \langle \alpha_1, \cdots, \alpha_r \rangle$ 是 V 的不变子空间。

定理 4.3.3 $\mathcal{A}$ 是 V 上的线性变换，W_1, W_2 是 V 的不变子空间且 $V = W_1 \oplus W_2$，如果 $\{\alpha_1, \cdots, \alpha_r\}$ 是 W_1 的一组基，$\{\alpha_{r+1}, \cdots, \alpha_n\}$ 是 W_2 的一组基，则 $\mathcal{A}$ 在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵为准对角阵。

定义 4.3.2 $\mathcal{A}$ 是 V 上的线性变换，如果存在非零向量 $x \in V, \lambda \in F\ st.\ \mathcal{A}x = \lambda x$，则称 x 为 $\mathcal{A}$ 的特征向量，λ 为对应的特征值。称 $V_\lambda = \{x \in V | \mathcal{A}x = \lambda x\}$ 为 V 的特征子空间，$dim\ V_\lambda$ 被称为特征值 λ 的几何重数。特征子空间都是不变子空间。

定义 4.3.3 $\mathcal{A}$ 是 V 上的线性变换，A 是 $\mathcal{A}$ 在一组基下的矩阵，定义 $P_A(\lambda) = P_{\mathcal{A}}(\lambda) = det(\lambda I - A)$.

定理 4.3.4 $\mathcal{A}$ 是 V 上的线性变换，W 是 V 的不变子空间，则 $P_{\mathcal{A}|W}(\lambda) | P_{\mathcal{A}}(\lambda)$.

推论 4.3.1 任何一个特征值的几何重数不超过它的代数重数。

定理 4.3.5 $\mathcal{A}$ 是 V 上的线性变换, $W_i, i = 1, \cdots, r$ 是 V 的不变子空间且 $V = W_1 \oplus \cdots \oplus W_r$，则

$$P_{\mathcal{A}}(\lambda) = P_{\mathcal{A}|W_1}(\lambda) \cdots P_{\mathcal{A}|W_r}(\lambda)$$

引理 4.3.1 $V_{\lambda_1} + \cdots + V_{\lambda_s}$ 是直和。

证明 原命题 $\Leftrightarrow \mathcal{A}$ 不同特征值对应的特征向量线性无关，否则，

$$\begin{aligned}
0 &= \alpha_1 + \cdots + \alpha_s, \alpha_i \in V_{\lambda_i}, i = 1, \cdots, s \\
\mathcal{A}0 = 0 &= \lambda_1\alpha_1 + \cdots + \lambda_s\alpha_s \\
0 &= \lambda_1^2\alpha_1 + \cdots + \lambda_s^2\alpha_s \\
&\vdots \\
0 &= \lambda_1^{s-1}\alpha_1 + \cdots + \lambda_s^{s-1}\alpha_s
\end{aligned}$$

即

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_s \\ \vdots & \vdots & & \vdots \\ \lambda_1^{s-1} & \lambda_2^{s-1} & \cdots & \lambda_s^{s-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{pmatrix} = 0$$

左边这个方阵是范德蒙方阵，由于特征值互不相同，所以此方阵是可逆方阵，因此

$$0=\alpha_1+\cdots+\alpha_s \Leftrightarrow \alpha_i=0, i=1,\cdots,s$$

所以不同特征值对应的特征向量线性无关。 □

定理 4.3.6 $\mathcal{A}$ 是 V 上的线性变换，$P_{\mathcal{A}}(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$，下列命题等价：

1. $\mathcal{A}$ 可对角化；
2. $V=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}$；
3. $\mathcal{A}$ 的任意特征值的代数重数和其对应的几何重数相等，即 $dim\ V_{\lambda_i}=n_i, i=1,\cdots,s.$
4. $\mathcal{A}$ 有 n 个线性无关的特征向量；

证明 $1\Rightarrow 2:$ 设 $\mathcal{A}$ 在基 $S=\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵为

$$A=\begin{pmatrix}\lambda_1 I_{r_1} & & \\ & \ddots & \\ & & \lambda_s I_{r_s}\end{pmatrix}$$

设 $T_1=\{\alpha_1,\cdots,\alpha_{r_1}\}, T_2=\{\alpha_{r_1+1},\cdots,\alpha_{r_1+r_2}\},\cdots,T_s=\{\alpha_{n-r_s+1},\cdots,\alpha_n\}$，显然，$V_{\lambda_i}=\langle T_i\rangle, i=1,\cdots,s.$，于是 $V=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}.$
$2\Rightarrow 3:$

$$dim\ V=\sum_{i=1}^{s} dim\ V_{\lambda_i}=\sum_{i=1}^{s} n_i=n$$

根据推论 *4.3.1*，$dim\ V_{\lambda_i}\leqslant n_i$，求和后又相等，所以几何重数和对应代数重数相等。 □

例 4.3.2 $\mathcal{A}$ 是 V 上的线性变换，$\mathcal{A}^2=\mathcal{I}$，则 $\mathcal{A}$ 可对角化。

证明 设 λ 是 $\mathcal{A}$ 的特征值，α 是特征向量，则

$$\mathcal{A}^2\alpha=\lambda^2\alpha=\alpha,$$

得 $\lambda=\pm1$. 说明特征值要么为 *1* 要么为 -1. 如果全为 *1* 或者全为 -1, 则 $\mathcal{A}=\mathcal{I}$ 或 $\mathcal{A}=-\mathcal{I}$; 如果特征值既有 *1* 也有 -1，设其对应的特征子空间为 V_1 和 V_{-1}，$\forall\alpha\in V$,

$$\alpha=\frac{1}{2}(\alpha+\mathcal{A}\alpha)+\frac{1}{2}(\alpha-\mathcal{A}\alpha)$$

其中，

$$\mathcal{A}(\alpha+\mathcal{A}\alpha)=\mathcal{A}\alpha+\alpha\Rightarrow\alpha+\mathcal{A}\alpha\in V_1;$$
$$\mathcal{A}(\alpha-\mathcal{A}\alpha)=\mathcal{A}\alpha-\alpha\Rightarrow\alpha+\mathcal{A}\alpha\in V_{-1}.$$

所以 $V=V_1\oplus V_{-1}$，因此 $\mathcal{A}$ 可对角化。 □

4.4 最小多项式

定理 4.4.1 $A \in C^{n\times n}, A$ 相似于上三角阵。

证明 归纳：若命题对 $n-1$ 成立，则：
假设 A 的一个特征值为 λ_1，对应的特征向量为 x_1，$Ax_1 = \lambda_1 x_1$. 把 x 扩充成一组 $C^{n\times n}$ 的基 $\{x_1, \cdots, x_n\}, X = (x_1, \cdots, x_n)$，则

$$A(x_1, \cdots, x_n) = (x_1, \cdots, x_n) \begin{pmatrix} \lambda_1 & b_{1\times(n-1)} \\ 0_{(n-1)\times 1} & A_1 \end{pmatrix}$$

$$\begin{aligned} A &= X \begin{pmatrix} \lambda_1 & b \\ 0 & A_1 \end{pmatrix} X^{-1} \\ &= X \begin{pmatrix} \lambda_1 & & b & \\ & & & \\ 0 & P_1 & \begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix} & P_1^{-1} \\ & & & \end{pmatrix} X^{-1} \\ &= XP \begin{pmatrix} \lambda_1 & bP_1 \\ 0 & \begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \end{pmatrix} P^{-1} X^{-1} \end{aligned}$$

其中 $P = \begin{pmatrix} 1 & \\ & P_1 \end{pmatrix}$，所以命题对 n 也成立。 □

定理 4.4.2 $(Hamilton - Carley$定理$)$ $A \in F^{n\times n}, P_A(A) = 0.$

证明 只给出一个 $F = C$ 时的证明。根据定理 *4.4.1*，我们可以假定 A 是上三角阵，

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$P_A(\lambda) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

$$P_A(A) = (A - \lambda_1 I)\cdots(A - \lambda_n I)$$

只需证 $(A - \lambda_1 I)\cdots(A - \lambda_n I) = 0$，归纳：若命题对 $n-1$ 成立，设 $A = \begin{pmatrix} A_1 & * \\ 0 & \lambda_n \end{pmatrix}$，

$$\begin{aligned} P_A(A) &= \begin{pmatrix} A_1 - \lambda_1 I_{n-1} & * \\ 0 & \lambda_n - \lambda_1 \end{pmatrix} \cdots \begin{pmatrix} A_1 - \lambda_{n-1} I_{n-1} & * \\ 0 & \lambda_n - \lambda_{n-1} \end{pmatrix} \begin{pmatrix} A_1 - \lambda_n I_{n-1} & * \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} (A_1 - \lambda_1 I_{n-1})\cdots(A_1 - \lambda_{n-1} I_{n-1}) & * \\ 0 & (\lambda_n - \lambda_1)\cdots(\lambda_n - \lambda_{n-1}) \end{pmatrix} \begin{pmatrix} A_1 - \lambda_n I_{n-1} & * \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} P_{A_1}(A_1) & * \\ 0 & (\lambda_n - \lambda_1)\cdots(\lambda_n - \lambda_{n-1}) \end{pmatrix} \begin{pmatrix} A_1 - \lambda_n I_{n-1} & * \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix} \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} \\ &= 0 \end{aligned}$$

即命题对 n 也成立。 □

定义 4.4.1 $A \in F^{n\times n}, f(\lambda) \in F[\lambda]$，如果 $f(A)=0$，则 $f(\lambda)$ 称为 A 的化零多项式。次数最低的首一化零多项式称为 A 的最小多项式，记作 $d_A(\lambda)$。

定理 4.4.3 $d_A(\lambda)|P_A(\lambda)$，且 $d_A(\lambda)$ 存在唯一。

定理 4.4.4 相似矩阵的最小多项式相等。

定理 4.4.5 特征多项式的根也是最小多项式的根，但重数可能不同。

证明 设 λ 是 A 的特征值，$Ax=\lambda x$，则 $d_A(A)x = d_A(\lambda)x = 0 \Rightarrow d_A(\lambda)=0$，得证。 □

定理 4.4.6 $A \in C^{n\times n}, A$ 相似于对角阵 $\Leftrightarrow d_A(\lambda)$ 无重根。

证明 $(\Rightarrow)$：设 $A = P\begin{pmatrix}\lambda_1 I_{r_1} & & \\ & \ddots & \\ & & \lambda_s I_{r_s}\end{pmatrix}P^{-1}$，则

$$\begin{aligned}
f(A) = P &\begin{pmatrix}f(\lambda_1) I_{r_1} & & \\ & \ddots & \\ & & f(\lambda_s) I_{r_s}\end{pmatrix}P^{-1} = 0 \\
&\Leftrightarrow f(\lambda_i)=0, i=1,\cdots,s \\
&\Leftrightarrow f(\lambda) = (\lambda-\lambda_1)\cdots(\lambda-\lambda_s)h(\lambda) \\
&\Leftrightarrow d_A(\lambda) = (\lambda-\lambda_1)\cdots(\lambda-\lambda_s)
\end{aligned}$$

即 $d_A(\lambda)$ 无重根。
$(\Leftarrow)$：$d_A(\lambda) = (\lambda-\lambda_1)\cdots(\lambda-\lambda_s)$，设 $f_i = \frac{d_A(\lambda)}{\lambda-\lambda_i}, i=1,\cdots,s$，则 $gcd(f_1,\cdots,f_s)=1$，于是存在一系列多项式 μ_i 使得

$$\begin{aligned}
&\mu_1 f_1 + \cdots + \mu_s f_s = 1 \\
&\Rightarrow \mu_1(A)f_1(A) + \cdots + \mu_s(A)f_s(A) = I \\
&\Rightarrow \alpha = \mu_1(A)f_1(A)\alpha + \cdots + \mu_s(A)f_s(A)\alpha
\end{aligned}$$

下证 $\alpha_i = \mu_i(A)f_i(A)\alpha_i \in V_{\lambda_i}$.

$$(A-\lambda_i I)\alpha_i = (A-\lambda_i I)\mu_i(A)f_i(A)\alpha_i = \mu_i(A)d_A(A)\alpha_i = 0.$$

因此有 $V = V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}$，进而 A 可对角化。 □

例 4.4.1 $A = \begin{pmatrix}2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 2\end{pmatrix}$，求 A 的最小多项式和 A^{-1}.

例 4.4.2 $A \in C^{m\times n}, B \in C^{n\times m}$，则 AB 和 BA 有相同的非零特征值。

证明

$$det(I_m - AB) = det(I_n - BA)$$
$$\lambda^n det(\lambda I_m - AB) = \lambda^m det(\lambda I_n - BA)$$

λ 非零时，$P_{AB}(\lambda) = 0 \Leftrightarrow P_{BA}(\lambda) = 0.$ □

定理 4.4.7 $A, B \in C^{n\times n}, AB = BA$，则 A，B 可同时对角化。

证明 设 $A = P\Lambda P^{-1}, \Lambda = \begin{pmatrix} \lambda_1 I_{r_1} & & \\ & \ddots & \\ & & \lambda_s I_{r_s} \end{pmatrix}$，则

$$P\Lambda P^{-1}B = BP\Lambda P^{-1}$$
$$\Lambda P^{-1}BP = P^{-1}BP\Lambda$$

记 $\tilde{B} = P^{-1}BP$，设 $\tilde{B} = (B_{ij})_{s\times s}$，由 $A\tilde{B} = \tilde{B}A$ 可知 $B_{ij} = 0, B = \begin{pmatrix} B_{11} & & \\ & \ddots & \\ & & B_{ss} \end{pmatrix}, B_{ii} = P_i\Lambda_i P_i^{-1}, i = 1, \cdots, s.$ 那么

$$\begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_s \end{pmatrix} P^{-1}BP \begin{pmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_s^{-1} \end{pmatrix} = \begin{pmatrix} \Lambda_1 & & \\ & \ddots & \\ & & \Lambda_s \end{pmatrix}$$

$$\begin{aligned} \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_s \end{pmatrix} P^{-1}AP \begin{pmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_s^{-1} \end{pmatrix} &= \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_s \end{pmatrix} \Lambda \begin{pmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_s^{-1} \end{pmatrix} \\ &= \begin{pmatrix} P_1\lambda_1 I_{r_1} P_1^{-1} & & \\ & \ddots & \\ & & P_s\lambda_s I_{r_1} P_s^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 I_{r_1} & & \\ & \ddots & \\ & & \lambda_s I_{r_1} \end{pmatrix} = \Lambda \end{aligned}$$

所以 A 和 B 可同时对角化。 □

定理 4.4.8 $A \in C^{m\times m}, B \in C^{n\times n}, \mathcal{A} : C^{m\times n} \to C^{m\times n}, x \mapsto Ax - xB.A$ 与 B 无相同特征值，求证：$\mathcal{A}$ 可逆。

证明 只需证 $Ker\ \mathcal{A} = 0.$

$$\begin{aligned} & x \in Ker\ \mathcal{A} \\ & \Rightarrow Ax = xB \\ & \Rightarrow A^2x = AxB = xB^2 \\ & \Rightarrow A^kx = xB^k \\ & \Rightarrow \forall f(\lambda) \in C[\lambda], f(A)x = xf(B) \\ & \Rightarrow P_A(A)x = xP_A(B) = 0 \end{aligned}$$

A 与 B 无相同特征值，所以 $P_A(B)$ 可逆 $\Rightarrow x = 0.$ □

定理 4.4.9 $A\in C^{n\times n}, P_A(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$，则 A 相似于 $\begin{pmatrix}A_1 & & \\ & \ddots & \\ & & A_s\end{pmatrix}, A_i=\begin{pmatrix}\lambda_i & & * \\ & \ddots & \\ & & \lambda_i\end{pmatrix}_{n_i\times n_i}, i=1,\cdots,s.$

证明 归纳：若命题对 $s-1$ 成立，设 A 相似于

$$\tilde{A}=\begin{pmatrix}\lambda_1 & & & & & & * \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_s & & \\ & & & & & \ddots & \\ & & & & & & \lambda_s\end{pmatrix}=\begin{pmatrix}A_{11} & B \\ 0 & A_s\end{pmatrix}, A_s\in C^{n_s\times n_s}$$

其中 A_{11} 相似于上三角阵 $\tilde{A}_1=PA_{11}P^{-1}$. 则

$$\begin{pmatrix}P & \\ & I_s\end{pmatrix}\tilde{A}\begin{pmatrix}P^{-1} & \\ & A_s\end{pmatrix}=\begin{pmatrix}\tilde{A}_1 & PB \\ 0 & I_s\end{pmatrix}$$

记 $C=PB$，下证存在 $S\in C^{(n-n_s)\times n_s}$ 使得

$$\begin{aligned}&\begin{pmatrix}I & -S \\ & I\end{pmatrix}\begin{pmatrix}\tilde{A}_1 & C \\ 0 & A_s\end{pmatrix}\begin{pmatrix}I & S \\ & I\end{pmatrix}\\ &=\begin{pmatrix}\tilde{A}_1 & \tilde{A}_1S-SA_s+C \\ 0 & A_s\end{pmatrix}\\ &=\begin{pmatrix}\tilde{A}_1 & 0 \\ 0 & A_s\end{pmatrix}\end{aligned}$$

即 $\tilde{A}_1S-SA_s=-C$. 只需证线性变换

$$\mathcal{A}: C^{(n-n_s)\times n_s}\to C^{(n-n_s)\times n_s}, S\mapsto \tilde{A}_1S-SA_s$$

可逆即可，由定理 *4.4.8* 可知的确是可逆的，所以 S 的存在性得证。 □

第五章 Jordan 标准型

5.1 Jordan 标准型的计算

定理 5.1.1 $A \in C^{n\times n}, P_A(\lambda) = (\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}, d_A(\lambda) = (\lambda-\lambda_1)^{k_1}\cdots(\lambda_s)^{k_s}$. 则 A 相似于 *Jordan* 标准型

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}, J_i \in C^{n_i\times n_i},$$

$$J_i = \begin{pmatrix} J_{i1} & & & \\ & J_{i2} & & \\ & & \ddots & \\ & & & J_{im_i} \end{pmatrix}, m_i = dim\ V_{\lambda_i},$$

$$J_{ij} = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \in C^{l_{ij}\times l_{ij}}, \text{称为 } Jordan \text{ 块}.$$

并且在不考虑 *Jordan* 块顺序的情况下，*Jordan* 标准型唯一。
设 J 的化零多项式为 $f(\lambda), f(J) = 0 \Leftrightarrow \forall f(J_i) = 0 \Leftrightarrow \forall f(J_{ij}) = 0$，设 $P_{J_{ij}}(\lambda) = (\lambda-\lambda_i)^{l_{ij}}, d_{J_{ij}}(\lambda) = (\lambda-\lambda_i)^{d_{ij}}, d_{ij} \leqslant l_{ij}$.

$$\begin{aligned}(J_{ij}-\lambda_i I)^{d_{ij}} = 0 &\Leftrightarrow \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}^{d_{ij}} = 0 \\ &\Leftrightarrow d_{ij} \geqslant l_{ij} \\ &\Rightarrow d_{ij} = l_{ij}, d_{J_{ij}}(\lambda) = P_{J_{ij}}(\lambda) = (\lambda-\lambda_i)^{l_{ij}}.\end{aligned}$$

所以 $(\lambda-\lambda_i)^{l_{ij}}|f(\lambda) \Rightarrow (\lambda-\lambda_i)^{max_j\{l_{ij}\}}|f(\lambda) \Rightarrow$

$$d_J(\lambda) = (\lambda-\lambda_1)^{max_j\{l_{1j}\}}\cdots(\lambda-\lambda_s)^{max_j\{l_{sj}\}}$$

因此 $k_i = max_j\{l_{ij}\}, i = 1,\cdots,s$.

记 δ_p^i 为 J_i 中 p 阶 Jordan 块的个数，$r_k^i = rank(A-\lambda_i I)^k$.

$$\begin{aligned}
r_1^i &= rank(J-\lambda_i I) = \sum_{j=1}^{s} rank(J_j-\lambda_i I) = n - n_i + rank(J_i - \lambda_i I) \\
&= n - n_i + \sum_{p=2}^{m_i}(p-1)\delta_p^i \\
r_2^i &= n - n_i + \sum_{p=2}^{m_i}(p-2)\delta_p^i \\
&\vdots \\
r_k^i &= n - n_i + \sum_{p=k+1}^{m_i}(p-k)\delta_p^i
\end{aligned}$$

记 $d_k^i = r_{k-1}^i - r_k^i = \sum_{p=k}^{m_i}\delta_p^i$
，于是 $d_k^i - d_{k+1}^i = \delta_k^i$. 可求出所有的 δ_p^i.

$$\begin{array}{c|cccccc}
r^i & r_0^i & r_1^i & \cdots & r_k^i & \cdots & 1\text{ 阶差分} \\
d^i & & d_1^i & \cdots & d_k^i & \cdots & 2\text{ 阶差分} \\
\delta^i & & \delta_1^i & \cdots & \delta_k^i & \cdots & 3\text{ 阶差分}
\end{array}$$

求 Jordan 标准型的基本步骤：

1. $P_A(\lambda)=0$ 求出所有的特征值，每个特征值 λ_i 对应了一个 J_i；
2. 计算所有的 $r_k^i = rank(A-\lambda_i I)^k$；
3. 计算所有的 δ_k^i，写出 J_i；
4. 计算所有的 J_i，写出 Jordan 标准型。

注 一般把 Jordan 块

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in C^{n\times n}$$

记作 $J_n(\lambda)$.

例 5.1.1 $A = I + N^2, N = J_n(0)$，求 A 的 Jordan 标准型并求 P 使得 $P^{-1}AP = J$.

$$P_A(\lambda) = (\lambda-1)^n, r_k = rank(A-I)^k = rank\ N^{2k} = n-2k, k \leqslant \frac{n}{2}\text{或}r_k = 0\ k \geqslant \frac{n}{2}.$$

n 为偶数时，J 有两个 $\frac{n}{2}$ 阶的 Jordan 块，

$$J = \begin{pmatrix} J_{\frac{n}{2}}(1) & 0 \\ 0 & J_{\frac{n}{2}}(1) \end{pmatrix}$$

k	0	1		$\frac{n}{2}-1$	$\frac{n}{2}$	$\frac{n}{2}+1$	
r_k	n	$n-2$	$\cdots$	2	0	0	$\cdots$
d_k		2	$\cdots$	2	2	0	$\cdots$
δ_k		0	$\cdots$	0	2	0	$\cdots$

n 为奇数时，J 有一个 $\frac{n-1}{2}$ 阶和 $\frac{n+1}{2}$ 阶的 $Jordan$ 块，

$$J = \begin{pmatrix} J_{\frac{n-1}{2}}(1) & 0 \\ 0 & J_{\frac{n+2}{2}}(1) \end{pmatrix}$$

下面求 P，设 $P = (x_1, \cdots, x_n).n = 2m$：

k	0	1		$\frac{n-3}{2}$	$\frac{n-1}{2}$	$\frac{n+1}{2}$	
r_k	n	$n-2$	$\cdots$	3	1	0	$\cdots$
d_k		2	$\cdots$	2	2	1	$\cdots$
δ_k		0	$\cdots$	0	1	1	$\cdots$

表 5.1: 奇数

$$\begin{aligned} AP = PJ &\Rightarrow A(x_1, \cdots, x_n) = (x_1, \cdots, x_n)J \\ &\Rightarrow (A - I)(x_1, \cdots, x_n) = (x_1, \cdots, x_n) \begin{pmatrix} J_m(0) & 0 \\ 0 & J_m(0) \end{pmatrix} \\ &\Rightarrow N^2(x_1, \cdots, x_n) = x_1 + \cdots + x_{m-1} + x_{m+1} + \cdots + x_{2m-1} \\ &\Rightarrow \cdots \end{aligned}$$

P 的一个简单解为 $P = (e_1, e_3, \cdots, e_{2m-1}, e_2, \cdots, e_{2m})$.

5.2 矩阵相似与多项式矩阵相抵

定理 5.2.1 $A, B \in F^{n\times n}$，则 A 与 B 相似当且仅当 $\lambda I - A$ 与 $\lambda I - B$ 相抵。

证明 $(\Rightarrow)$：显然；
$(\Leftarrow)$：$\lambda I - B = P(\lambda)(\lambda I - A)Q(\lambda), P(\lambda), Q(\lambda)$ 可逆。于是

$$P(\lambda)^{-1}(\lambda I - B) = Q(\lambda)\lambda - AQ(\lambda)$$

令 $\lambda = B, 0 = Q(B)B - AQ(B) \Rightarrow Q(B)B^k = A^kQ(B)$，下证 $P = Q(B)$ 可逆。

$$Q(\lambda) = Q_k\lambda^k + \cdots + Q_1\lambda + Q_0, Q(\lambda)^{-1} = R_m\lambda^m + \cdots + R_0, Q_i, R_i \in F^{n\times n}$$

$Q(\lambda)^{-1}Q(\lambda) = I$，得到

$$\begin{aligned} R_mQ(\lambda)\lambda^m + \cdots + R_0Q(\lambda) &= I \\ R_mPB^m + \cdots + R_1PB + R_0P &= I \\ R_mA^mP + \cdots + R_1AP + R_0P &= I \\ (R_mA^m + \cdots + R_1A + R_0)P &= I \end{aligned}$$

即 P 可逆，于是 A 与 B 相似。 □

定理 5.2.2 $A\in C^{n\times n},\lambda I-A$ 的初等因子组 $\{(\lambda-\lambda_i)^{l_{ij}}|1\leqslant i\leqslant s,1\leqslant j\leqslant m_i\}$，则 A 相似于 $J=diag(J_1,\cdots,J_s),J_i=diag(J_{i1},\cdots,J_{im_i}),J_ij=J_{l_{ij}}(\lambda_i)$.

证明 证 $\lambda I-A$ 与 $\lambda I-J$ 相抵即可，即证 $r(\lambda I-A)=r(\lambda I-J)=n$ 且二者初等因子组相同。
$\lambda I-J$ 的初等因子组是所有块 $\lambda I-J_{ij}$ 的初等因子组的并，

$$\lambda I-J_{ij}=\begin{pmatrix}\lambda-\lambda_i & -1 & & \\ & \ddots & \ddots & \\ & & \lambda-\lambda_i & -1\\ & & & \lambda-\lambda_i\end{pmatrix}$$

的初等因子组为 $\{(\lambda-\lambda_i)^{l_{ij}}\}$，所以所有块的初等因子组的并就是 $\lambda I-A$ 的初等因子组。□

定理 5.2.3 $A,B\in F^{n\times n},F\subset K,$ 则 A 与 B 在 F 上相似当且仅当 A 与 B 在 K 上相似。

证明 $(\Rightarrow)$：显然；$(\Leftarrow)$：A 与 B 在 K 上相似 $\Leftrightarrow\lambda I-A$ 与 $\lambda I-B$ 在 $K[\lambda]$ 相抵

$$\Rightarrow\lambda I-A \text{ 与 } \lambda I-B \text{ 有相同的不变因子 } d_i(\lambda)$$

$d_i(\lambda)\in F[\lambda]\Rightarrow\lambda I-A,\lambda I-B$ 在 $F[\lambda]$ 相抵 $\Rightarrow A$ 与 B 在 F 上相似。 □

例 5.2.1 $A=\begin{pmatrix}0&0&1&&\\&0&0&1&\\&&0&0&1\\&&&0&0\\&&&&0\end{pmatrix}$，求 A 的 *Jordan* 标准型。

例 5.2.2 $A\in F^{n\times n}$，证明 A^T 与 A 相似。

例 5.2.3 $A\in F^{n\times n},\lambda I-A$ 的初等因子为 $d_1(\lambda),\cdots,d_n(\lambda)$，则

(1) $P_\lambda(\lambda)=d_1(\lambda)\cdots d_n(\lambda)$；

(2) $d_A(\lambda)=d_n(\lambda)$.

5.3 根子空间分解

定义 5.3.1 $\mathcal{A}$ 是 V 上的线性变换，λ 是 $\mathcal{A}$ 的特征值，称

$$W_{\lambda_i}=\{\alpha\in V|\exists k,(\mathcal{A}-\lambda\mathcal{I})^k\alpha=0\}$$

为 λ 的根子空间，$\alpha\in W_\lambda$ 称为根向量。

推论 5.3.1 根据例 1.1.2，如果 k 就是使 $Ker(\mathcal{A}-\lambda_i\mathcal{I})^k=Ker(\mathcal{A}-\lambda_i\mathcal{I})^{k+1}$ 成立的最小的 k，那么就有

$$Ker(\mathcal{A}-\lambda_i\mathcal{I})^k=Ker(\mathcal{A}-\lambda_i\mathcal{I})^{k+1}=\cdots=Ker(\mathcal{A}-\lambda_i\mathcal{I})^{n_i}=C^{n_i}$$

当然，k 不小于最小的 k 时，上式也是成立的。$W_{\lambda_i}=Ker(\mathcal{A}-\lambda_i\mathcal{I})^k$，且 W_{λ_i} 是 $\mathcal{A}$ 的不变子空间，这是因为 $\mathcal{A}$ 和 $(\mathcal{A}-\lambda_i\mathcal{I})$ 可交换。

定理 5.3.1 $\mathcal{A}$ 是 V 上的线性变换，$P_{\mathcal{A}}(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s},\sum_{i=1}^{s}n_i=n$，则

$$W_{\lambda_i}=(\mathcal{A}-\lambda_i\mathcal{I})^{n_i}.dim\ W_{\lambda_i}=n_i,\mathcal{A}_i=\mathcal{A}|_{W_{\lambda_i}},P_{\mathcal{A}_i}(\lambda)=(\lambda-\lambda_i)^{n_i}$$

证明 由定理 *4.4.1*，存在一组基 $\alpha_1,\cdots,\alpha_n$ 使得 $\mathcal{A}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)A$，其中矩阵 A 有这样的形式：

$$A=\begin{pmatrix}\lambda_i & & * & & & \\ & \ddots & & * & & \\ & & \lambda_i & & & \\ & & & \lambda_j & & * \\ & & & & \ddots & \\ 0 & & & & & \lambda_s\end{pmatrix}=\begin{pmatrix}A_{11} & * \\ 0 & A_{22}\end{pmatrix}$$

$$A_{11}=\begin{pmatrix}\lambda_i & & * \\ & \ddots & \\ & & \lambda_i\end{pmatrix}\in C^{n_i\times n_i}\text{，特征值均为}\lambda_i$$

$$A_{22}=\begin{pmatrix}\lambda_j & & * \\ & \ddots & \\ & & \lambda_s\end{pmatrix}\in C^{(n-n_i)\times(n-n_i)}\text{，特征值均不是}\lambda_i$$

设 $\mathcal{A}_i=\mathcal{A}|_{W_{\lambda_i}}$，则 $\mathcal{A}_i$ 在基 $\alpha_1,\cdots,\alpha_{n_i}$ 下的矩阵为 $A_{11}-\lambda_iI_{n_i}$，则

$$(A-\lambda_iI_n)^k=\begin{pmatrix}(A_{11}-\lambda_iI_{n_i})^k & * \\ 0 & (A_{22}-\lambda_iI_{(n-n_i)})^k\end{pmatrix}$$

$k\geqslant n_i$ 时，$(A_{11}-\lambda_iI_{n_i})^k=0,r(A-\lambda_iI_n)^k=n-n_i\Rightarrow dim\ Ker(A-\lambda_iI_n)^k=n_i,k\geqslant n_i$，也就是 $dim\ W_{\lambda_i}=dim\ Ker(A-\lambda_iI_n)^{n_i},P_{\mathcal{A}_i}(\lambda)=|\lambda I_{n_i}-A_{11}|=(\lambda-\lambda_i)^{n_i}$. □

定理 5.3.2 $\mathcal{A}$ 是 V 上的线性变换，$P_{\mathcal{A}}(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s},\sum_{i=1}^{s}n_i=n$，则

$$V=W_{\lambda_1}\oplus\cdots\oplus W_{\lambda_s}$$

证明 先证直和：设 $0=\alpha_1+\cdots+\alpha_s,\alpha_i\in W_{\lambda_i},i=1,\cdots,s$. 设多项式 $f_i(\lambda)=\frac{P_{\mathcal{A}}(\lambda)}{(\lambda-\lambda_i)^{n_i}}$，于是 $f_i(\lambda)$ 与 $(\lambda-\lambda_i)^{n_i}$ 互素，存在多项式 $u(\lambda),v(\lambda)$ 使得

$$\begin{aligned}&u(\lambda)f_i(\lambda)+v(\lambda)(\lambda-\lambda_i)^{n_i}=1\\&\Rightarrow u(\mathcal{A})f_i(\mathcal{A})+v(\mathcal{A})(\mathcal{A}-\lambda_i\mathcal{I})^{n_i}=\mathcal{I}\\&\Rightarrow u(\mathcal{A})f_i(\mathcal{A})\alpha_i+v(\mathcal{A})(\mathcal{A}-\lambda_i\mathcal{I})^{n_i}\alpha_i=\alpha_i\\\alpha_i\in W_{\lambda_i}&\Rightarrow(\mathcal{A}-\lambda_i\mathcal{I})^{n_i}\alpha_i=0\Rightarrow f_i(\mathcal{A})\alpha_j=0,i\neq j\\f_i(\mathcal{A})(\alpha_1+\cdots+\alpha_s)=0&\Rightarrow f_i(\mathcal{A})\alpha_i=0\\&\Rightarrow\alpha_i=u(\mathcal{A})f_i(\mathcal{A})\alpha_i+v(\mathcal{A})(\mathcal{A}-\lambda_i\mathcal{I})^{n_i}\alpha_i=0\end{aligned}$$

所以直和得证，又因为 $\sum dim\ W_{\lambda_i}=n$，所以 $V=W_{\lambda_1}\oplus\cdots\oplus W_{\lambda_s}$. □

定理 5.3.3 $A \in C^{n\times n}, P_{\mathcal{A}}(\lambda) = (\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$，则 A 相似于 $diag(A_{11},\cdots,A_{ss})$，其中

$$A_{ii} = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ & & \lambda_i \end{pmatrix} \in C^{n_i\times n_i}$$

证明 设线性变换为 $\mathcal{A}: x \mapsto Ax$，取 W_{λ_i} 的基为 T_i，则这一系列 T_i 的并构成 V 的基 T，$\mathcal{A}|_{W_{\lambda_i}}$ 在 T_i 下的矩阵为上三角阵 A_{ii} ，所以 $\mathcal{A}$ 在基 T 下的矩阵就是 $diag(A_{11},\cdots,A_{ss})$. □

5.4 循环子空间分解

定义 5.4.1 $B \in C^{n\times n},\exists m, B^m = 0$，称 B 为幂零矩阵，满足 $B^m = 0$ 的最小的 m 称为 B 的幂次。相应地，如果 $\mathcal{B}$ 是 V 上的线性变换，$\exists m, \mathcal{B}^m = 0$，称 $\mathcal{B}$ 为幂零变换。

推论 5.4.1 B 是幂零矩阵当且仅当 B 的特征值都为 *0.*

定义 5.4.2 $\mathcal{A}$ 是 V 上的线性变换，$\alpha \in V$，称 $C = \langle \alpha, \mathcal{A}\alpha, \cdots\rangle$ 为循环子空间，显然 C 是关于 $\mathcal{A}$ 的不变子空间。

定义 5.4.3 $\mathcal{A}$ 是 V 上的线性变换，$\alpha \in V$，设非零多项式 $f(\lambda)$ 满足 $f(\mathcal{A})\alpha = 0$，称 f 为 α 的关于 $\mathcal{A}$ 的化零多项式。最小次数的首一化零多项式称为最小多项式，记作 $d_{\mathcal{A},\alpha}(\lambda)$，或者 $d_\alpha(\lambda)$.

定理 5.4.1 α 的关于 $\mathcal{A}$ 的最小多项式存在唯一，并且 $d_{\mathcal{A},\alpha}(\lambda)$ 能够整除任意化零多项式。

定理 5.4.2 $\mathcal{A}$ 是 V 上的线性变换，C 是由 α 生成的循环子空间，$d_\alpha(\lambda) = a_0 + a_1\lambda + \cdots + a_k\lambda^k (a_k = 1)$. 则

(1) $\alpha, \mathcal{A}\alpha, \cdots, \mathcal{A}^{k-1}\alpha$ 是 C 的一组基，$dim\ C = k, \mathcal{A}|_C$ 在基 $\{\mathcal{A}^{k-1}\alpha, \cdots, \alpha\}$ 下的矩阵为

$$A = \begin{pmatrix} -a_{k-1} & 1 & & \\ \vdots & & \ddots & \\ \vdots & & & 1 \\ -a_0 & 0 & \cdots & 0 \end{pmatrix}$$

(2) A 的特征多项式 $P_A(\lambda)$ 和最小多项式 $d_A(\lambda)$ 都是 $d_\alpha(\lambda)$.

证明 *(1)* $\alpha, \mathcal{A}\alpha, \cdots, \mathcal{A}^{k-1}\alpha$ 线性无关，否则 $\exists \mu_0\alpha + \cdots + \mu_{k-1}\mathcal{A}^{k-1}\alpha = 0$，设 $g(\lambda) = \mu_0 + \cdots + \mu_{k-1}\lambda^{k-1}$，则 $g(\mathcal{A})\alpha = 0 \Rightarrow g(\lambda)$ 是化零多项式，但是它的次数小于最小多项式，矛盾。而

$$\mathcal{A}^k\alpha = -a_0\alpha - \cdots - a_{k-1}\mathcal{A}^{k-1}\alpha$$

即 $\mathcal{A}^k\alpha$ 可以由其他 k 个向量线性表示，进而 $\mathcal{A}^{k+i}\alpha, i>0$ 都可以被线性表示，因此 $C=\langle\alpha,\mathcal{A}\alpha,\cdots,\mathcal{A}^{k-1}\alpha\rangle$.

$$\begin{aligned}\mathcal{A}(\mathcal{A}^{k-1}\alpha,\cdots,\alpha)&=(\mathcal{A}^{k-1}\alpha,\cdots,\alpha)A\\&=(\mathcal{A}^{k}\alpha,\cdots,\mathcal{A}\alpha)\\&=(-a_0\alpha-\cdots-a_{k-1}\mathcal{A}^{k-1}\alpha,\cdots,\mathcal{A}\alpha)\\&=(\mathcal{A}^{k-1}\alpha,\cdots,\alpha)\begin{pmatrix}-a_{k-1}&1&&\\ \vdots&&\ddots&\\ \vdots&&&1\\ -a_0&0&\cdots&0\end{pmatrix}\end{aligned}$$

所以结论得证。

(2) $P_A(\lambda)=|\lambda I-A|=d_\alpha(\lambda)$, 所以 $d_A(\lambda)|P_A(\lambda)=d_\alpha(\lambda)$. 而 $d_\alpha(\lambda)|d_A(\lambda)$，所以二者相等，结论得证。

□

推论 5.4.2 $\mathcal{B}$ 是 V 上的幂零变换, 幂次是 m, $\alpha\in V$, 则 $d_{\mathcal{B},\alpha}(\lambda)=\lambda^m$. 由 α 生成的循环子空间 C 的一组基为 $\mathcal{B}^{m-1}\alpha,\cdots,\alpha$, 且 $\mathcal{B}$ 在这租基下的矩阵为 $B=J_m(0), P_B(\lambda)=d_B(\lambda)=\lambda^m$.

推论 5.4.3 $\mathcal{A}$ 是 V 上的线性变换，α 是 $\mathcal{A}$ 的特征值 λ 的 m 次根向量，即 $(\mathcal{A}-\lambda\mathcal{I})^m\alpha=0, (\mathcal{A}-\lambda\mathcal{I})^{m-1}\alpha\neq 0$. 记 $\alpha_i=(\mathcal{A}-\lambda\mathcal{I})^{m-i}\alpha, i=1,\cdots,m.C=\langle\alpha_1,\cdots,\alpha_m\rangle$ 是 α 的关于 $\mathcal{A}-\lambda\mathcal{I}$ 的循环子空间，则 C 在基 $\alpha_1,\cdots,\alpha_m$ 下的矩阵为 $J_m(\lambda)$.

定理 5.4.3 $\mathcal{B}$ 是 n 维线性空间 W 上的幂零变换, 幂次为 m, 存在 W 上的一组基 $\alpha_1,\cdots,\alpha_k$ 使得

$$W=C_{\alpha_1}\oplus\cdots\oplus C_{\alpha_k}$$

其中 C_{α_i} 是 α_i 生成的循环子空间。

证明 首先有 $Ker\ \mathcal{B}\subset Ker\ \mathcal{B}^2\subset\cdots\subset Ker\ \mathcal{B}^m=W.$，现给出如下构造：

$$\begin{aligned}Ker\ \mathcal{B}^m&=Ker\ \mathcal{B}^{m-1}\oplus U^m,\\ Ker\ \mathcal{B}^{m-1}&=Ker\ \mathcal{B}^{m-2}\oplus\mathcal{B}U^m\oplus U^{m-1},\\ &\vdots\\ Ker\ \mathcal{B}^{2}&=Ker\ \mathcal{B}\oplus\mathcal{B}^{m-2}U^m\oplus\cdots\oplus U^2,\\ Ker\ \mathcal{B}&=\mathcal{B}^{m-1}U^m\oplus\cdots\oplus\mathcal{B}U^2\oplus U^1.\end{aligned}$$

其中 U_m 代表包含于 $Ker\ \mathcal{B}^{m-1}$ 但包含于 $Ker\ \mathcal{B}^m$ 的部分; U_{m-1} 是包含于 $Ker\ \mathcal{B}^{m-2}+\mathcal{B}U^m$ 但不包含于 $Ker\ \mathcal{B}^{m-1}$ 的部分，以此类推，所以每行中的最后一个和为直和，其他的和是直和仍未证明。此外，还需证明 $\mathcal{B}U^m\subset Ker\ B^{m-1}$。所以关于这个构造，我们补充以下几个细节的证明：

1° $\mathcal{B}^jU^i\subset Ker\mathcal{B}^{i-j}, j<i$.

这是因为 $\forall x\in\mathcal{B}^jU^i, x=\mathcal{B}^jy, y\in U^i\subset Ker\mathcal{B}^i$, 则 $\mathcal{B}^{i-j}x=\mathcal{B}^iy=0$.

2° $dim\ \mathcal{B}^j U^i = dim\ U^i, j < i.$

设 U^i 的一组基 $T_i = \{\alpha_1, \cdots, \alpha_s\}$，只需证明 $\mathcal{B}^j T_i = \{\mathcal{B}^j\alpha_1, \cdots, \mathcal{B}^j\alpha_s\}$ 是线性无关的即可。否则，

$$\begin{aligned}
&\exists \mu_1\mathcal{B}^j\alpha_1 + \cdots + \mu_s\mathcal{B}^j\alpha_s = 0 \\
&\Rightarrow \mathcal{B}^j(\mu_1\alpha_1 + \cdots + \mu_s\alpha_s) = 0 \\
&\Rightarrow \mu_1\alpha_1 + \cdots + \mu_s\alpha_s \in Ker\ \mathcal{B}^j \\
&\Rightarrow \mu_1\alpha_1 + \cdots + \mu_s\alpha_s \in Ker\ \mathcal{B}^j \cap U^i
\end{aligned}$$

$Ker\ \mathcal{B}^j \cap U^i = \emptyset$，这是因为 $Ker\ \mathcal{B}^j \subset Ker\ \mathcal{B}^{i-1} \cap U^i = \emptyset$，于是 $\mu_1\alpha_1 + \cdots + \mu_s\alpha_s = 0$，矛盾。

3° $Ker\ \mathcal{B}^{m-i-1}U^m + \cdots + \mathcal{B}U^{i+2}$ 是直和。

我们只证明 $Ker\ B^{m-2} + BU^m$ 是直和，其余证明过程同理。设 $\beta_1, \cdots, \beta_s$ 是 $Ker\ B^{m-2}$ 的一组基，$\gamma_1, \cdots, \gamma_t$ 是 BU^m 的一组基，存在一系列 $\gamma_i = \mathcal{B}\alpha_i$，并且 $\alpha_1, \cdots, \alpha_t$ 是 U^m 的一组基，只需证 $\beta_1, \cdots, \beta_s, \gamma_1, \cdots, \gamma_t$ 是线性无关的，否则

$$\begin{aligned}
&\exists \mu_1\beta_1 + \cdots + \mu_s\beta_s + c_1\gamma_1 + \cdots + c_t\gamma_t = 0 \\
&\Rightarrow \mathcal{B}^{m-2}(\mu_1\beta_1 + \cdots + \mu_s\beta_s + c_1\gamma_1 + \cdots + c_t\gamma_t) = 0 \\
&\Rightarrow \mathcal{B}^{m-2}(c_1\gamma_1 + \cdots + c_t\gamma_t) = 0 \\
&\Rightarrow \mathcal{B}^{m-1}(c_1\alpha_1 + \cdots + c_t\alpha_t) = 0 \\
&\Rightarrow c_1\alpha_1 + \cdots + c_t\alpha_t \in Ker\ \mathcal{B}^{m-1} \\
&\Rightarrow c_1\alpha_1 + \cdots + c_t\alpha_t \in Ker\ \mathcal{B}^{m-1} \cap U^m = \emptyset \\
&\Rightarrow c_1\alpha_1 + \cdots + c_t\alpha_t = 0
\end{aligned}$$

所以矛盾。

根据这一构造，我们得到 W 的一个分解：

$$\begin{aligned}
W = Ker\ \mathcal{B}^m &= Ker\ \mathcal{B}^{m-1} \oplus U^m \\
&= Ker\ \mathcal{B}^{m-2} \oplus (\mathcal{B}U^m \oplus U^{m-1}) \oplus U^m \\
&= Ker\ \mathcal{B}^{m-3} \oplus (\mathcal{B}^2U^m \oplus \mathcal{B}U^{m-1} \oplus U^{m-2}) \oplus (\mathcal{B}U^m \oplus U^{m-1}) \oplus U^m \\
&\ \ \vdots \\
&= (\mathcal{B}^{m-1}U^m \oplus \cdots \oplus \mathcal{B}U^2 \oplus U^1) \oplus (\mathcal{B}^{m-2}U^m \oplus \cdots \oplus U^2) \oplus \\
&\cdots \oplus (\mathcal{B}U^m \oplus U^{m-1}) \oplus U^m
\end{aligned}$$

更换一下顺序，得到

$$\begin{aligned}
W = &(U^m \oplus \mathcal{B}U^m \oplus \cdots \oplus \mathcal{B}^{m-1}U^m) \oplus (U^{m-1} \oplus \mathcal{B}U^{m-1} \oplus \cdots \oplus \mathcal{B}^{m-2}U^{m-1}) \oplus \\
&\cdots \oplus (U^2 \oplus \mathcal{B}U^2) \oplus U^1
\end{aligned}$$

设 U^i 的基为 T_i，$\mathcal{B}^jU^i$ 的基为 $\mathcal{B}^jT_i$. 因为是直和，于是这些基

$$T_m, \mathcal{B}T_m, \cdots, \mathcal{B}^{m-1}T_m, T_{m-1}, \mathcal{B}T_{m-1}, \cdots, \mathcal{B}^{m-2}T_{m-1}, \cdots, T_2, \mathcal{B}T_2, T_1$$

的并线性无关，而且是 $Ker\ \mathcal{B}^m$ 即 W 的一组基。记作 M。

注 其实老师在证明 M 是 W 的一组基的大致思路为：先证这些基的并线性无关（老师一句显然带过了），然后证明这些基的向量个数求和就是线性空间 W 的维数。这个思路涉及 $Ker\ \mathcal{B}^i$ 的维度等复杂的符号表示，成功让笔者没有理解为什么一系列求和最终变成了维数 n，而笔者整理思路的过程中意外发现了可以根据递推地把 W 一步步“化开”，写成一系列空间的直和后能够非常直白的看出来 M 就是 W 的一组基，而且这些基线性无关是非常显然的（因为是直和嘛）。也就是说上面这个思路并不是老师本来的思路，而是笔者的思路，并且这个思路似乎是没用到引理 *2* 的（老师的思路是用到了的），因此可能不对，仅供参考。

这个证明还没结束，后续还需要说明 $\alpha_1, \cdots, \alpha_k$ 的构造方法，老师确实是没讲明白，笔者说也说不清楚，还是看例子来理解吧。

□

例 5.4.1

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}，求 A 的 Jordan 标准型 J，和可逆矩阵 P 使得 AP = PJ.$$

解：$P_A(\lambda) = (\lambda - 1)(\lambda - 2)^5$，设 $\mathcal{A} : x \mapsto Ax$，

$$W_1 = Ker(A - I) = \langle x_1 \rangle, x_1 = (1, 0, 0, 0, 0, 0)$$

$r(A-2I) = 4, r(A-2I)^2 = 2, r(A-2I)^3 = 1$，所以 $dim\ Ker(A-2I) = 2, dim\ Ker(A-2I)^2 = 4, dim\ Ker(A-2I)^3 = 5, W_2 = Ker(A-2I)^3$. 设 $\mathcal{B} = \mathcal{A} - 2\mathcal{I}$，于是

$$Ker\ \mathcal{B} \subset Ker\ \mathcal{B}^2 \subset Ker\ \mathcal{B}^3$$

先找 $Ker\ \mathcal{B}$ 的一组基 $S_1 = \{y_1, y_2\}, Ker\ \mathcal{B}^2$ 的一组基为 $S_1 \cup S_2, S_2 = \{y_3, y_4\}, Ker\ \mathcal{B}^3$ 的一组基为 $S_1 \cup S_2 \cup S_3, S_3 = \{y_5\}$

固定 $\mathcal{B}^2 y_5$，在 $\mathcal{B}y_3, \mathcal{B}y_4$ 中尽可能多地保留一部分（两个，一个，或者全部舍弃），使此向量组线性无关，在本题中应当保留 $\mathcal{B}y_4$;

固定 $\mathcal{B}^2 y_5, \mathcal{B}y_4$, 在 y_1, y_2 中尽可能多地保留一部分，使此向量组线性无关，本题中两个都不保留。

最终我们留下了 $\mathcal{B}^2 y_5, \mathcal{B}y_4$，其中 y_5 是 *3* 次根向量，对应了一个 *3* 阶 *Jordan* 块，y_4 是 *2* 次根向量，对应了一个 *2* 阶 *Jordan* 块，令 $x_2 = y_5, x_3 = y_4$, 于是分别生成的循环子空间为

$$C_{x_2} = \left\langle x_2, \mathcal{B}x_2, \mathcal{B}^2 x_2 \right\rangle, C_{x_3} = \langle x_3, \mathcal{B}x_3 \rangle$$

那么就有

$$W_2 = C_{x_2} \oplus C_{x_3}$$

于是 $P = (x_1, x_2, \mathcal{B}x_2, \mathcal{B}^2 x_2, x_3, \mathcal{B}x_3), J = diag(1, J_3(2), J_2(2))$.

定理 5.4.4 $\mathcal{A}$ 是 W 上的线性变换，其特征值全为 a，则 $W = C_1 \oplus \cdots \oplus C_m, C_i$ 是关于 $\mathcal{A}$ 的 l_i 维循环子空间，并且存在 W 的一组基 M，$\mathcal{A}$ 在基 M 下的矩阵为 $diag(J_{l_1}(a), \cdots, J_{l_m}(a))$.

证明 令 $\mathcal{B} = \mathcal{A} - a\mathcal{I}$ 即可。 □

定理 5.4.5 $\mathcal{A}$ 是 V 上的线性变换，$P_{\mathcal{A}}(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}, \sum_{i=1}^{s} n_i = n$，则特征子空间可以分解为循环子空间的直和：

$$W_{\lambda_i} = C_{i1} \oplus \cdots \oplus C_{im_i}$$

5.5 实方阵的实相似标准型

引理 5.5.1 $J_m(\tau) \sim \tau J_m(1), \tau \neq 0.$

证明 实际上可证一个更强的结论：

$$J_m(\tau) \sim J = \begin{pmatrix} \tau & \epsilon & & \\ & \tau & \ddots & \\ & & \ddots & \epsilon \\ & & & \tau \end{pmatrix}$$

令 $P = diag(1, \epsilon, \cdots, \epsilon^{m-1})$，则 $P^{-1} J_m \tau P = J.$ □

引理 5.5.2 $\tau = a + bi \neq 0, a, b \in R, A = \begin{pmatrix} J_m(\tau) & \\ & J_m(\overline{\tau}) \end{pmatrix} \sim L = \begin{pmatrix} aJ_m(1) & bJ_m(1) \\ -bJ_m(1) & aJ_m(1) \end{pmatrix}$

证明 根据引理 *5.5.1*,

$$A \sim A_1 = \begin{pmatrix} \tau J_m(1) & \\ & \overline{\tau} J_m(1) \end{pmatrix}$$

设 $P = \begin{pmatrix} I & I \\ iI & -iI \end{pmatrix}$，则 $PA_1P^{-1} = L.$ □

定理 5.5.1 $A \in R^{n \times n}$，初等因子组为 $\{(\lambda - \lambda_j)^{m_j} | 1 \leqslant j \leqslant s, \lambda_j \in R\}$ 和 $\{(\lambda - a_j - ib_j)^{k_j}, (\lambda - a_j + ib_j)^{k_j}, a, b \neq 0 \in R, 1 \leqslant j \leqslant t\}$. 则 A 相似于实标准型

$$diag(J_{m_1}(\lambda_1), \cdots, J_{m_s}(\lambda_s), L_{k_1}(a_1 \pm ib_i), \cdots, L_{k_t}(a_t \pm ib_t))$$

其中，$L_{k_j}(a_j \pm ib_j) = \begin{pmatrix} a_j J_{k_j}(1) & b_j J_{k_j}(1) \\ -b_j J_{k_j}(1) & a_j J_{k_j}(1) \end{pmatrix}$

例 5.5.1 $\tau = a + bi, a, b \in R, A = \begin{pmatrix} J_m(\tau) & \\ & J_m(\overline{\tau}) \end{pmatrix}$，则

$$A \sim \begin{pmatrix} K & I_2 & & \\ & K & \ddots & \\ & & \ddots & I_2 \\ & & & K \end{pmatrix}$$

其中 $K = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

证明 设标准向量 $e_i \in R^n$，则有

$$\begin{aligned} Ae_1 &= \tau e_1 \\ Ae_j &= \tau e_j + e_{j-1}, 2 \leqslant j \leqslant n \\ Ae_{m+1} &= \overline{\tau} e_{m+1} \\ Ae_{m+j} &= \overline{\tau} e_{m+j} + e_{m+j-1} \end{aligned}$$

设 $P=(e_1,e_{m+1},\cdots,e_m,e_{2m})$，则

$$P^{-1}AP=\begin{pmatrix} L & I_2 & & \\ & L & \ddots & \\ & & \ddots & I_2 \\ & & & L \end{pmatrix}$$

其中 $L=diag(\tau,\overline{\tau}),L\sim K.$ □

例 5.5.2 $A\in R^{2n\times 2n},A^2+I_{2n}=0$，则 A 相似于 $\begin{pmatrix} & I_n \\ -I_n & \end{pmatrix}$

证明 $f(\lambda)=\lambda^2+1$ 是 A 的化零式，A 的最小多项式是实多项式，所以只能是 $d_A(\lambda)=\lambda^2+1,P_A(\lambda)=(\lambda^2+1)^n.$

$$A\sim\begin{pmatrix} iI_n & \\ & -iI_n \end{pmatrix}\sim\begin{pmatrix} & I_n \\ -I_n & \end{pmatrix}$$

□

5.6 一些例子

例 5.6.1 $A^2=I$，则 $A\sim\begin{pmatrix} I_r & \\ & I_{n-r} \end{pmatrix}.$

证明 $A=PJP^{-1},A^2=I\Rightarrow J^2=I$，设 J_i 是 J 的 *Jordan* 块，则 $J_i^2=I,J_i$ 只能是一阶矩阵 (λ). 于是 $\lambda^2=1,\lambda=\pm 1.$ □

例 5.6.2 $A\in C^{n\times n},\forall\epsilon>0$，存在上三角阵 $B=(b_{ij})$ 满足 $|b_{ij}|<\epsilon,i\neq j.$

证明 $A=PJP^{-1}$，设 J_i 是 J 的 *Jordan* 块，由引理 *5.5.1* 的强版本可知

$$J_i=\begin{pmatrix} \lambda & 1 & & \\ lambda & \ddots & & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}\sim\begin{pmatrix} \lambda & \epsilon & & \\ & \lambda & \ddots & \\ & & \ddots & \epsilon \\ & & & \lambda \end{pmatrix}$$

□

例 5.6.3 $A,B\in C^{n\times n}$，且无公共特征值，则矩阵方程 $AX=XB$ 只有零解，$X\in C^{n\times n}.$

证明 先证明 A 和 B 都是 *Jordan* 块的形式时结论是成立的，即

$$A=\begin{pmatrix} a & 1 & & \\ & a & \ddots & \\ & & \ddots & 1 \\ & & & a \end{pmatrix},B=\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b \end{pmatrix},a\neq b$$

设 X 由若干行向量 x_i 组成，$X=(x_{ij})$

$$\begin{aligned}AX=XB&\Rightarrow AX-aIX=XB-aIX\\&\Rightarrow (A-aI)X=(B-aI)X\\&\Rightarrow\begin{pmatrix}x_2\\ \vdots\\ x_n\\ 0\end{pmatrix}=\begin{pmatrix}x_1\\ \vdots\\ x_{n-1}\\ x_n\end{pmatrix}\begin{pmatrix}b-a&1&&\\&b-a&\ddots&\\&&\ddots&1\\&&&b-a\end{pmatrix}\end{aligned}$$

关注最后一行元素，有

$$\begin{aligned}&0=x_{n_1}(b-a)\Rightarrow x_{n_1}=0\\&0=x_{n_1}+x_{n_2}(b-a)\Rightarrow x_{n_2}=0\\&\ \vdots\\&0=x_{n,n-1}+x_{nn}(b-a)\Rightarrow x_{nn}=0\end{aligned}$$

于是得到 $x_n=0$，再关注倒数第二行元素可以得到 $x_{n-1}=0$，以此类推最终有 $X=0$.

然后我们考虑一般情况，假设 A，B 的 *Jordan* 标准型分别是 $J,\tilde{J}$.

$$\begin{aligned}&A=PJP^{-1},B=\tilde{P}\tilde{J}\tilde{P}^{-1},AX=BX\Rightarrow PJP^{-1}X=X\tilde{P}\tilde{J}\tilde{P}^{-1}\\&\text{令}Y=P^{-1}X\tilde{P},JY=Y\tilde{J}\end{aligned}$$

设 $J=diag(J_1,\cdots,J_s),J_i$ 是单独的 *Jordan* 块，把 Y 按照 J 的 *Jordan* 块模式分块，$Y=(Y_{ij}),\tilde{J}$ 同理，于是

$$\begin{pmatrix}J_1Y_{11}&J_1Y_{12}&\cdots&J_1Y_{1s}\\ \vdots&\vdots&&\vdots\\ J_sY_{s1}&J_sY_{s2}&\cdots&J_sY_{ss}\end{pmatrix}=\begin{pmatrix}Y_{11}\tilde{J}_1&Y_{12}\tilde{J}_2&\cdots&Y_{1s}\tilde{J}_s\\ \vdots&\vdots&&\vdots\\ Y_{s1}\tilde{J}_1&Y_{s2}\tilde{J}_2&\cdots&Y_{ss}\tilde{J}_s\end{pmatrix}$$

对应起来，得到 $J_iY_{jk}=Y_{jk}\tilde{J}_k$. 由于 $J_i,\tilde{J}_k$ 分别来自 A，B，所以也没有公共特征值。因此根据上面证明的特殊情况可知 $Y_{jk}=0$，进而 $Y=0,X=0$.

注 这里有一个不严谨的地方，就是 $\tilde{J}$ 的分块模式和 J 一定一样吗？暂时没有想到可以规避这个瑕疵的办法。

$\square$

例 5.6.4 $A\in R^{n\times n}$，若 A 相似于 $diag(\begin{pmatrix}0&1\\1&0\end{pmatrix},1,\cdots,1)$，则称 A 是反射方阵，证明：若 $A^2=I$，则 A 是有限个反射方阵之积。

证明 $A\sim A_1=\begin{pmatrix}-I_r&\\&I_{n-r}\end{pmatrix},A_1=D_1D_2\cdots D_r$，其中 D_r 为 I 的对角线上第 r 个元素替换为 -1 得到的方阵，显然 D_i 之间相似，而

$$D_1=diag(\begin{pmatrix}-1&0\\0&1\end{pmatrix},1,\cdots,1)\sim diag(\begin{pmatrix}0&1\\1&0\end{pmatrix},1,\cdots,1)$$

所以结论得证。 $\square$

例 5.6.5 $A \in C^{n\times n}, A$ 可逆，则存在 $B \in C^{n\times n}$ 使得 $B^2 = A$.

证明 设 A 的 *Jordan* 标准型为 J，$A = PJP^{-1}, J = diag(J_1, \cdots, J_s), J_i$ 是单独的 *Jordan* 块. 再设 $\tilde{J} = diag(\tilde{J}_1, \cdots, \tilde{J}_s), \tilde{J}_i$ 是单独的 *Jordan* 块，如果 ${\tilde{J}_i}^2 = J_i$，取 $B = P\tilde{J}P^{-1}$ 即可。下面证明对于每个 J_i 都存在对应的 $\tilde{J}_i$.

设 $J_i = J_m(a), a \neq 0$. 则 $J_i = aI + N = a(I + \frac{N}{a})$. 其中 $N = J_m(0)$, 是 m 次幂零阵。我们取 $\sqrt{1+x}$ 的泰勒展开：

$$\sqrt{1+x} = 1 + \frac{x}{2} + \cdots + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-m+2\right)}{(m-1)!}x^{m-1} + \cdots$$

把其中的 $1+x$ 替换为 $I + \frac{N}{a}$, 可知 m 次及以上的项都为零，得到

$$\sqrt{I + \frac{N}{a}} = I + \frac{\frac{N}{a}}{2} + \cdots + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-m+2\right)}{(m-1)!}\left(\frac{N}{a}\right)^{m-1}$$

这个和式有限，因此 $\sqrt{I + \frac{N}{a}}$ 一定存在，取 $\tilde{J}_i = \sqrt{a}\sqrt{I + \frac{N}{a}}$ 即可。 □

例 5.6.6 $A \in C^{n\times n}$，如果 A 的最小多项式和特征多项式相同，称 A 是“单纯方阵”。证明：A 是单纯方阵，如果 $AB = BA$，则存在多项式 $f(\lambda)$ 使得 $B = f(A)$，并求出这个多项式。

证明 最小多项式和特征多项式相同，说明除了最后一个不变因子 $d_n(\lambda) = (\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$，其他均为 *1*，所以 A 的不变因子组就是 $\{(\lambda-\lambda_1)^{n_1}, \cdots, (\lambda-\lambda_s)^{n_s}\}$，则 A 的 *Jordan* 标准型为 $J = diag(J_{n_1}(\lambda_1), \cdots, J_{n_s}(\lambda_s)) = (J_1, \cdots, J_s)$，设 $A = PJP^{-1}$，

$$\begin{aligned}&AB = BA \Rightarrow PJP^{-1}B = BPJP^{-1}\\ &\text{令}C = P^{-1}BP, JC = CJ\end{aligned}$$

把 C 按照 J 的 *Jordan* 块模式分块，$C = (C_{ij})$，于是

$$\begin{pmatrix} J_1C_{11} & J_1C_{12} & \cdots & J_1C_{1s} \\ \vdots & \vdots & & \vdots \\ J_sC_{s1} & J_sC_{s2} & \cdots & C_sY_{ss} \end{pmatrix} = \begin{pmatrix} C_{11}J_1 & C_{12}J_2 & \cdots & C_{1s}J_s \\ \vdots & \vdots & & \vdots \\ C_{s1}J_1 & C_{s2}J_2 & \cdots & C_{ss}J_s \end{pmatrix}$$

根据例 *5.6.3*，$C_{ij} = 0, i \neq j$. 现确定 C_{ii},

$$C_{ii}J_i = J_iC_{ii} \Rightarrow C_{ii}(J_i - \lambda_iI) = (J_i - \lambda_iI)C_{ii}$$

可以得出 C_{ii} 有如下形式：

$$C_{ii} = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ & c_0 & c_1 & \ddots & \vdots \\ & & c_0 & \ddots & c_2 \\ & & & \ddots & c_1 \\ & & & & c_0 \end{pmatrix}$$

设 $N = J_{n_i}(0)$, 于是 $C_{ii} = c_0I + c_1N + \cdots + c_{n-1}N^{n-1} = g_i(N)$. 于是

$$B = P\begin{pmatrix} g_1(N) & & \\ & \ddots & \\ & & g_s(N) \end{pmatrix}P^{-1}$$

如果存在多项式 $f(\lambda)$ 使得 $B=f(A)$，则 $f(A)=Pf(J)P^{-1},f(\lambda)$ 需要满足 $f(J)=\begin{pmatrix} g_1(N) & & \\ & \ddots & \\ & & g_s(N) \end{pmatrix}$，即 $f(J_i)=g_i(N)$. 利用拉格朗日插值法可以构造出这样的 $f(\lambda)$: 设 $f_i(\lambda)=\frac{P_A(\lambda)}{(\lambda-\lambda_i)^{n_i}},i=1,\cdots,s$. 由于 $(J_i-\lambda_i I)^{n_i}=N^{n_i}=0$，所以 $f_i(J_j)=0,i\neq j$. 然后，根据 $f_1(\lambda),\cdots,f_s(\lambda)$ 互素，可知存在多项式 $u_1(\lambda),\cdot,u_s(\lambda)$ 使得

$$\sum_{k=1}^{s} u_k(\lambda)f_k(\lambda)=1$$

将 J_i 代入，得到

$$\sum_{k=1}^{s} u_k(J_i)f_k(J_i)=u_i(J_i)f_i(J_i)=I$$

于是令 $f(\lambda)=\sum_{k=1}^{s}u_k(\lambda)f_k(\lambda)g_k(\lambda-\lambda_i)$，就有

$$\begin{aligned} f(J_i) &= \sum_{k=1}^{s} u_k(J_i)f_k(J_i)g_k(J_i-\lambda_i I) \\ &= \sum_{k=1}^{s} u_k(J_i)f_k(J_i)g_k(N) \\ &= u_i(J_i)f_i(J_i)g_i(N) \\ &= g_i(N) \end{aligned}$$

□

第六章　内积空间

6.1　欧氏空间

定义 6.1.1 V 是 R 上的线性空间，定义内积 $F: V\times V\to R$ 满足：

(1) 对称性：$(\alpha,\beta)=(\beta,\alpha)$；

(2) 双线性性：$(\lambda\alpha_1+\mu\alpha_2,\beta)=\lambda(\alpha_1,\beta)+\mu(\alpha_2,\beta)$；

(3) 正定性：$(\alpha,\alpha)\geqslant 0$，等号成立当且仅当 $\alpha=0$.

定义了内积的线性空间 V 称为欧氏空间，进一步可定义欧氏空间上的度量、距离和向量之间的夹角：

$$|\alpha|=\sqrt{(\alpha,\alpha)},\rho(\alpha,\beta)=|\alpha-\beta|,\theta=\arccos\frac{(\alpha,\beta)}{|\alpha||\beta|}$$

定义 6.1.2 V 是欧氏空间，W 是 V 的子空间，称

$$W^{\perp}=\{\alpha\in V|\alpha\perp W\text{，即 }\alpha\text{ 与 }W\text{ 中的每个向量都垂直}\}$$

为 W 的正交补空间。显然有

$$V=W\oplus W^{\perp}$$

定义 6.1.3 V_1,V_2 是都是空间，若映射 $\sigma: V_1\to V_2$ 满足：

(1) σ 是一一映射；

(2) σ 保线性运算，即 $\sigma(\lambda\alpha+\mu\beta)=\lambda\sigma(\alpha)+\mu\sigma(\beta)$；

(3) σ 保内积，即 $(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta)$.

则称 σ 是 V_1 到 V_2 的同构映射，并称 V_1,V_2 同构。

定理 6.1.1 欧氏空间 V_1,V_2 同构当且仅当 V_1,V_2 维数相等。

定理 6.1.2 V 是欧氏空间，V^* 是 V 的对偶空间，则映射 $\sigma:\beta\in V\mapsto f_\beta\in V^*$ 是 V 到 V^* 的同构映射，其中

$$f_\beta: V\to R, f_\beta(\alpha)=(\alpha,\beta)$$

证明 1° σ 是单射：$f_{\beta_1}=f_{\beta_2}\Rightarrow f_{\beta_1}(\alpha)=f_{\beta_1}(\alpha)\Rightarrow(\beta_1,\alpha)=(\beta_2,\alpha)\Rightarrow(\beta_1-\beta_2,\alpha)\Rightarrow \beta_1-\beta_2\perp V\Rightarrow\beta_1-\beta_2=0$.

2° σ 是满射：设 $\alpha_1, \cdots, \alpha_n$ 是 V 的一组标准正交基，$\forall f \in V^*$，取

$$\beta = f(\alpha_1)\alpha_1 + \cdots + f(\alpha_n)\alpha_n,$$

则 $f = f_\beta$.

3° σ 保加法和数乘：这是显然的。

所以 σ 是 V 到 V^* 的同构映射。 □

推论 6.1.1 如果 $\beta_1, \cdots, \beta_n$ 是 V 的一组标准正交基，则 $f_{\beta_1}, \cdots, f_{\beta_n}$ 是其对偶基。

如果在 V^* 上定义内积：$(f_{\beta_1}, f_{\beta_2}) = (\beta_1, \beta_2)$，则 V^* 也成为欧氏空间。

6.2 正交变换与对称变换

定义 6.2.1 $\mathcal{A}$ 是 V 上的线性变换，若满足 $\forall \alpha, \beta \in V, (\mathcal{A}\alpha, \mathcal{A}\beta) = (\alpha, \beta)$，称 $\mathcal{A}$ 为正交变换。

定理 6.2.1 关于正交变换，下列说法等价：

(1) $\mathcal{A}$ 为正交变换；

(2) $\forall \alpha \in V, (\mathcal{A}\alpha, \mathcal{A}\alpha) = (\alpha, \alpha)$；

(3) $\mathcal{A}$ 把标准正交基变换为标准正交基。

正交变换的一些性质如下：

(1) 正交变换 $\mathcal{A}, \mathcal{B}$ 的复合也是正交变换；

(2) 正交变换的逆也是正交变换；

(3) 正交变换的特征值模长都为 *1.*

定义 6.2.2 $\mathcal{A}$ 是 V 上的线性变换，若满足 $\forall \alpha, \beta \in V, (\mathcal{A}\alpha, \beta) = (\alpha, \mathcal{A}\beta)$，称 $\mathcal{A}$ 为对称变换。

定理 6.2.2 关于对称变换，下列说法等价：

(1) $\mathcal{A}$ 为对称变换；

(2) $\mathcal{A}$ 在标准正交基下的矩阵为对称阵。

定理 6.2.3 $\mathcal{A}$ 是 V 上的对称变换，则存在 V 的一组标准正交基 M，$\mathcal{A}$ 在基 M 下的矩阵为为对角阵 $diag(\lambda_1, \cdots, \lambda_n)$.

6.3 规范变换

定义 6.3.1 $\mathcal{A}$ 是 V 上的线性变换，存在另一个变换 $\mathcal{A}^*$ 满足 $\forall\alpha,\beta\in V,(\mathcal{A}\alpha,\beta)=(\alpha,\mathcal{A}^*\beta)$，称 $\mathcal{A}^*$ 是 $\mathcal{A}$ 的伴随变换。

定理 6.3.1 欧氏空间上线性变换 $\mathcal{A}$ 的伴随变换存在且唯一。

证明 唯一性：设 $\mathcal{A}_1^*,\mathcal{A}_2^*$ 是 $\mathcal{A}$ 的伴随变换，则

$$\begin{aligned}&(\mathcal{A}\alpha,\beta)=(\alpha,\mathcal{A}_1^*\beta)=(\alpha,\mathcal{A}_2^*\beta)\\&\Rightarrow(\alpha,(\mathcal{A}_1^*-\mathcal{A}_2^*)\beta)=0\\&\Rightarrow\mathcal{A}_1^*-\mathcal{A}_2^*=0\end{aligned}$$

存在性：设 M 是 V 的一组标准正交基，$\mathcal{A}$ 在 M 下的矩阵为 A，$\mathcal{A}^*$ 在 M 下的矩阵为 A^*，向量 α,β 在 M 下的坐标分别为 x,y，于是

$$(\mathcal{A}\alpha,\beta)=(Ax)^Ty=x^TA^Ty\qquad\qquad(\alpha,\mathcal{A}^*\beta)=x^T\mathcal{A}^*y\Rightarrow A^*=A^T$$

所以矩阵 A^T 对应的线性变换 $A^*:x\mapsto A^Tx$ 就是 $\mathcal{A}$ 的伴随变换。 □

推论 6.3.1 $\mathcal{A}$ 和其伴随变换 $\mathcal{A}^*$ 在同一组标准正交基下的矩阵互为转置，所以也有

$$(\mathcal{A}^*)^*=\mathcal{A}$$

定义 6.3.2 $\mathcal{A}$ 是 V 上的线性变换，若 $\mathcal{A}\mathcal{A}^*=\mathcal{A}^*\mathcal{A}$，则称 $\mathcal{A}$ 为规范变换。

若矩阵 A 满足 $AA^T=A^TA$，称 A 为规范矩阵。

定理 6.3.2 $\mathcal{A}$ 为规范变换 $\Leftrightarrow\forall\alpha,|\mathcal{A}\alpha|=|\mathcal{A}^*\alpha|\Leftrightarrow\mathcal{A}$ 在一组标准正交基下的矩阵式规范矩阵。

证明 $1\Rightarrow2$：

$$\begin{aligned}|\mathcal{A}\alpha|^2&=(\mathcal{A}\alpha,\mathcal{A}\alpha)\\&=(\alpha,\mathcal{A}^*\mathcal{A}\alpha)\\&=(\alpha,\mathcal{A}\mathcal{A}^*\alpha)\\&=(\mathcal{A}^*\alpha,\mathcal{A}^*\alpha)\\&=|\mathcal{A}^*\alpha|^2\end{aligned}$$

$$\Rightarrow|\mathcal{A}\alpha|=|\mathcal{A}^*\alpha|$$

$2\Rightarrow1$：

$$\begin{aligned}|\mathcal{A}\alpha|=|\mathcal{A}^*\alpha|&\Rightarrow(\alpha,\mathcal{A}\mathcal{A}^*\alpha)=(\alpha,\mathcal{A}^*\mathcal{A}\alpha)\\&\Rightarrow(\alpha,(\mathcal{A}^*\mathcal{A}-\mathcal{A}\mathcal{A}^*)\alpha)=0\\&\Rightarrow\mathcal{A}^*\mathcal{A}=\mathcal{A}\mathcal{A}^*\end{aligned}$$

$1\Leftrightarrow3$：设 $\mathcal{A}$ 在标准正交基 $M=(\alpha_1,\cdots,\alpha_n)$ 下的矩阵为 A，A^* 在基 M 下的矩阵为 A^T，

$$\mathcal{A}^*\mathcal{A}(\alpha_1,\cdots,\alpha_n)=\mathcal{A}^*(\alpha_1,\cdots,\alpha_n)A=(\alpha_1,\cdots,\alpha_n)A^TA$$
$$\mathcal{A}\mathcal{A}^*(\alpha_1,\cdots,\alpha_n)=\mathcal{A}(\alpha_1,\cdots,\alpha_n)A^T=(\alpha_1,\cdots,\alpha_n)AA^T$$

所以

$$\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^* \Leftrightarrow A^TA = AA^T.$$

□

定理 6.3.3 若实规范方阵 A 的特征值为 $a_1 \pm ib_1, \cdots, a_s \pm ib_s, \lambda_{s+1}, \cdots, \lambda_n$,其中 $a_k, b_k, \lambda_j \in R$. 则 A 正交相似于标准型

$$D = diag(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{s+1}, \cdots, \lambda_n)$$

推论 6.3.2 正交方阵的正交相似标准型为

$$diag(\begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix}, \cdots, \begin{pmatrix} \cos\theta_s & \sin\theta_s \\ -\sin\theta_s & \cos\theta_s \end{pmatrix}, 1, \cdots, 1, -1, \cdots, -1)$$

实对称方阵只有实特征值，其正交相似标准型为

$$diag(\lambda_1, \cdots, \lambda_n)$$

实反对称方阵只有纯虚数特征值，其正交相似标准型为

$$diag(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, 0, \cdots, 0)$$

定理 6.3.4 $A \in R^{n\times n}$ 的特征值为 $a_1 \pm ib_1, \cdots, a_s \pm ib_s, \lambda_{s+1}, \cdots, \lambda_n$，其中 $a_k, b_k, \lambda_j \in R$. 则 A 正交相似于准上对角阵

$$D = \begin{pmatrix} A_1 & & & & & * \\ & \ddots & & & & \\ & & A_s & & & \\ & & & \lambda_{s+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_n \end{pmatrix}$$

其中 $A_k \in R^{2\times 2}$，且特征值为 $a_k \pm ib_k$.

定理 6.3.5 $A \in R^{n\times n}$ 的特征值为 $\lambda_1, \cdots, \lambda_n$，则 $Tr(AA^T) \geqslant \sum_{i=1}^n |\lambda_i|^2$，且等号成立当且仅当 A 是规范矩阵。

例 6.3.1 $\mathcal{A}$ 是规范变换，则 $\mathcal{A}$ 不同特征值对应的特征向量相互正交。

证明 设 $M = \{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一组标准正交基，$\mathcal{A}$ 在 M 下的矩阵为标准型

$$D = diag(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{s+1}, \cdots, \lambda_n)$$

于是 $\mathcal{A}$ 的特征值 $a_k \pm ib_k$ 对应的特征向量为 $\alpha_{2k-1} \pm i\alpha_{2k}$，特征值 λ_j 对应的特征向量为 α_j，这些特征向量相互正交，结论得证。 □

例 6.3.2 $\mathcal{A}$ 是 V 上的规范变换，则 $(Im\ \mathcal{A})^\perp = Ker\ \mathcal{A}$.

6.4 酉空间

定义 6.4.1 V 是 C 上的线性空间，定义内积 $F: V \times V \to C$ 满足

(1) 共轭性：$\overline{(\alpha,\beta)} = (\beta,\alpha)$；

(2) 共轭线性性：$(\lambda_1\alpha_1 + \lambda_2\alpha_2, \beta) = \overline{\lambda_1}(\alpha_1,\beta) + \overline{\lambda_2}(\alpha_2,\beta)$；

(3) 正定性：$(\alpha,\alpha) \geqslant 0$.

定义了内积的线性空间 V 称为酉空间。

注 共轭线性性只对其中一个变量成立，我们一般默认对第一个变量成立，则对于第二个变量而言成立线性性：$(\beta, \lambda_1\alpha_1 + \lambda_2\alpha_2) = \lambda_1(\beta,\alpha_1) + \lambda_2(\beta,\alpha_2)$.

定义 6.4.2 V 是 C 上的线性空间，设 $M = \{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一组基，α, β 在 M 下的坐标分别为 x, y，则

$$(\alpha,\beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{x_i} y_j (\alpha_i, \alpha_j) = x^* H y.$$

其中 $H = ((\alpha_i, \alpha_j))_{n\times n}$，称为度量阵。$H$ 还满足 $H^* = H$，所以 H 也是 *Hermite* 阵。不同基的度量矩阵之间共轭相合。

注 共轭转置、共轭相合等详见定义 2.8.6

定义 6.4.3 $P \in C^{n\times n}, P^*P = I$，则称 P 为酉矩阵。

酉矩阵的一些性质：

(1) 酉矩阵的积、逆都是酉矩阵；

(2) 酉矩阵的行列式的模长为 1；

(3) 酉矩阵的所有特征值模长都为 1.

6.5 酉空间上的线性变换

定理 6.5.1 酉空间上的线性变换 $\mathcal{A}$ 的伴随变换存在且唯一，并且二者在同一组标准正交基下的矩阵互为共轭转置。

定义 6.5.1 $\mathcal{A}$ 是酉空间 V 上的线性变换，如果 $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$，称 $\mathcal{A}$ 是规范变换；$A \in C^{n\times n}, AA^* = A^*A$，称 A 是规范矩阵。

如果 $\forall \alpha,\beta \in V, (\mathcal{A}\alpha, \mathcal{A}\beta) = (\alpha,\beta)$，称 $\mathcal{A}$ 为酉变换。

如果 $\forall \alpha,\beta \in V, (\mathcal{A}\alpha, \beta) = (\alpha,\mathcal{A}\beta)$，称 $\mathcal{A}$ 为对称变换。

推论 6.5.1 1. $\mathcal{A}$ 为规范变换 $\Leftrightarrow$ $\mathcal{A}$ 在标准正交基下的矩阵为规范矩阵；

2. $\mathcal{A}$ 为对称变换 $\Leftrightarrow (\mathcal{A}\alpha, \mathcal{A}\alpha) = (\alpha,\alpha) \Leftrightarrow$ 存在一组标准正交基 M，使得 $\mathcal{A}$ 在 M 下的矩阵为酉矩阵 $\Leftrightarrow \mathcal{A}\mathcal{A}^* = \mathcal{I}$

3. $\mathcal{A}$ 为对称变换 $\Leftrightarrow$ 存在一组标准正交基 M，使得 $\mathcal{A}$ 在 M 下的矩阵为 *Hermite* 矩阵 $\Leftrightarrow \mathcal{A} = \mathcal{A}^*$

定理 6.5.2 A 是酉方阵，则 A 是规范方阵当且仅当 A 酉相似于对角阵。

引理 6.5.1 A 是规范方阵，U 是酉方阵，则 $B=U^*AU$ 是规范方阵。

引理 6.5.2 A 是上三角阵，也是规范方阵，则 A 一定是对角阵。

定理 6.5.3 $A\in C^{n\times n}$，则 A 酉相似于上三角阵。

定理 6.5.4 $A\in C^{n\times n}$ 的特征值为 $\lambda_1,\cdots,\lambda_n$，则 $Tr(AA^*)\geqslant\sum_{i=1}^n|\lambda_i|^2$，且等号成立当且仅当 A 是规范矩阵。

例 6.5.1 设 $Hermite$ 矩阵 $H=A+iB,A,B\in R^{n\times n}$，则 $det(A)\geqslant det(H)$，等号成立当且仅当 $H=0$.

例 6.5.2 $A\in C^{n\times n}$，A 是规范方阵当且仅当存在 $f(\lambda)\in C[\lambda]$ 使得 $A^*=f(A)$.

第七章　多重线性函数

7.1　二重线性函数

定义 7.1.1 V 是 F 上的线性空间，称 $f: V\times V\to F$ 为二重线性函数，如果

$$f(\lambda_1\alpha_1+\lambda_2\alpha_2,\beta)=\lambda_1 f(\alpha_1,\beta)+\lambda_2 f(\alpha_2,\beta)$$

$$f(\alpha,\lambda_1\beta_1+\lambda_2\beta_2)=\lambda_1 f(\alpha,\beta_1)+\lambda_2 f(\alpha,\beta_2)$$

设 $\{\alpha_1,\cdots,\alpha_n\}$ 是 V 的一组基，内积在这组基下的度量矩阵为 $A=(a_ij)_{n\times n}, a_{ij}=f(a_i,a_j)$，其中 f 可以是任何一个二重线性函数。容易验证这是符合内积定义的。

双重线性函数的全体记作 $L(V,V)$，在函数的加法和数乘运算下构成线性空间。

定理 7.1.1 设 V 是 F 上的 n 维线性空间，则 $L(V,V)$ 与 $F^{n\times n}$ 同构。

证明 定义映射 $\sigma: L(V,V)\to F^{n\times n}, f\mapsto A=(f(\alpha_i,\alpha_j))_{n\times n}$，下证它是同构映射：

1° σ 是单射：设 $\sigma(f_1)=\sigma(f_2)\Rightarrow \forall i,j, f_1(\alpha_i,\alpha_j)=f_2(\alpha_i,\alpha_j)$，$\{\alpha_i\}$ 是 V 的一组基，所以 $\forall x,y\in V, x=(\alpha_1,\cdots,\alpha_n)(x_1,\cdots,x_n)^T, y=(a_1,\cdots,a_n)(y_1,\cdots,y_n)^T$，有

$$\begin{aligned}f_1(x,y)=f_1(\sum_{i=1}^n x_i\alpha_i,\sum_{i=1}^n y_i\alpha_i)&=\sum_{i=1}^n\sum_{j=1}^n x_iy_jf_1(\alpha_i,\alpha_j)\\&=\sum_{i=1}^n\sum_{j=1}^n x_iy_jf_2(\alpha_i,\alpha_j)=f_2(x,y)\end{aligned}$$

即 $f_1=f_2$.

2° σ 是满射：设 $A=(a_ij)_{n\times n}$，V 的一组基为 $\{\alpha_i\}$，且 α,β 在这组基下的坐标分别为 x,y. 令 $f(\alpha,\beta)=x^TAy$，于是 $\sigma(f)=A$.

3° σ 是线性的：这是显然的。

□

注 V 的一组基为 $M=\{\alpha_i\}$，且 α,β 在这组基下的坐标分别为 $x,y.$，如果 $f(\alpha,\beta)=x^TAy$，称 A 是 f 在基 M 下的矩阵。

若 f 在基 $(\alpha_1,\cdots,\alpha_n),(\beta_1,\cdots,\beta_n)$ 下的矩阵分别为 A,B, $(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)P$，则 $B=P^TAP$.

例 7.1.1 章空间：$V=F^n,(x,y)=x^TAy,A^T=-A$. 在这个空间里，所有向量的模长都是 0.

定义 7.1.2 A 是 f 在基 $M=\{\alpha_i\}$ 下的矩阵,定义 $rank(f)=rank(A)$,如果 $rank(f)=dim\,V$ 则称 f 非退化，否则称 f 退化。

定义 7.1.3 若 $f(\alpha,\beta)=0$，称 α 与 β 左正交，记为 $\alpha\perp_L\beta$，β 与 α 右正交，记为 $\beta\perp_R\alpha$.
W 是 V 的子空间，则

$$W^{\perp_L}=\{\alpha\in V|f(\alpha,\beta)=0,\forall\beta\in W\},W^{\perp_R}=\{\beta\in V|f(\alpha,\beta)=0,\forall\alpha\in W\}$$

称为 W 的左、右正交补空间。

$V^{\perp_L},V^{\perp_R}$ 分别称为 f 的左根基和右根基。

定理 7.1.2 W_1,W_2 是 V 的子空间且 $W_1\subset W_2$，则有

(1) $W_2^{\perp_L}\subset W_1^{\perp_L},W_2^{\perp_R}\subset W_1^{\perp_R}$；

(2) $W_1\subset(W_1^{\perp_R})^{\perp_L},W_1\subset(W_1^{\perp_L})^{\perp_R}$.

$V=F^n,W=\langle y_1,\cdots,y_r\rangle,B=(y_1,\cdots,y_r),f(\alpha,\beta)=x^TAy$, 于是

$$\begin{aligned}W^{\perp_L}&=\{\alpha\in V|f(\alpha,\beta)=0,\forall\beta\in W\}\\&=\{\alpha\in V|\alpha^TA\beta=0,\forall\beta\in W\}\\&=\{\alpha\in V|\alpha^TAB=0\}=Ker\ (AB)^T\end{aligned}$$

同理，$W^{\perp_R}=Ker\ B^TA$. 特殊地，$V^{\perp_L}=Ker\ A^T,V^{\perp_R}=Ker\ A$.

定理 7.1.3 $dim\ V^{\perp_L}=dim\ V^{\perp_R}=dim\ V-rank(f)$.

推论 7.1.1 f 非退化 $\Leftrightarrow V^{\perp_L}=\{0\}\Leftrightarrow V^{\perp_R}=\{0\}$.

推论 7.1.2 f 非退化则 $dim\,W^{\perp_L}=dim\,W^{\perp_R}=dim\,V-dim\,W$, 且 $(W^{\perp_L})^{\perp_R}=(W^{\perp_R})^{\perp_L}=W$.

定理 7.1.4 W 是 V 的子空间，$f|_W$ 非退化，则 $V=W\oplus W^{\perp_L}=W\oplus W^{\perp_R}$.

定义 7.1.4 f 是 V 上的双线性函数，如果 $f(\alpha,\beta)=f(\beta,\alpha)$，称 f 对称。类似地可以定义 f 反对称。

引理 7.1.1 $f\in L(V,V)$，f 在一组基下的矩阵为 A，则 f 对称当且仅当 A 是对称阵。

定理 7.1.5 $f\in L(V,V)$，且 f 对称，则 f 在某一组基下的矩阵为对角阵 D。
V 是 R 上的线性空间，$D=diag(I_r,-I_s,0)$；V 是 C 上的线性空间，$D=diag(I_r,0)$.

定理 7.1.6 $f\in L(V,V)$，且 f 反对称，则 f 在某一组基下的矩阵为

$$D=diag(\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0)$$

证明 $f\neq0,\exists\alpha,\beta\in V$ 使得 $f(\alpha,\beta)\neq0$，不妨设 $f(\alpha,\beta)=1,f(\beta,\alpha)=-1$，令 $W=\langle\alpha,\beta\rangle$，则 f 在基 α,β 下的矩阵为 $\begin{pmatrix}0&1\\-1&0\end{pmatrix}$. 归纳即可。 □

定义 7.1.5 V 是 C 上的线性空间，$f:V\times V\to C$ 满足

$$f(\lambda_1\alpha_1+\lambda_2\alpha_2,\beta)=\overline{\lambda_1}f(\alpha_1,\beta)+\overline{\lambda_2}f(\alpha_2,\beta)$$

称 f 为共轭双线性函数。

当然，同样只对其中一个变量成立，详见定义 7.4.1 下注。

7.2 多重线性函数

本节不考，以后有机会再补。