线性代数

孙天阳

目录

Chapter 1

线性空间

1 线性函数与对偶空间

有限维线性空间 V 到其对偶空间的同构.

给定 V 的一组基 $\{v_1, \cdots, v_n\}$ 后，可定义 V^* 中的一组对偶基 $f_i(v_j) = \delta_{ij}$.

需注意的是，对偶基的定义是依赖于基的选取的，没有给出完整的一组基的时候是无法说单个 v_i 的对偶基是谁的.

当 V 和 V^* 都各自有了一组基之后，他们便分别同构于数组空间，从而彼此也是同构的，显式地写出同构映射就是

$$\varphi : V \longrightarrow V^*, \quad v = a_1 v_1 + \cdots + a_n v_n \longmapsto a_1 f_1 + \cdots + a_n f_n.$$

因此 $\varphi(v)(v) = a_1^2 + \cdots + a_n^2$.

2 自由线性空间

设 X 是任意非空集合，考虑映射 $f: X \to F$ 满足只有有限多个 $x \in X$ 使得 $f(x) \neq 0$，将这样的映射的全体记作 $C(X)$. 那么，如果 $f \in C(X), g \in C(X)$，$\lambda, \mu \in F$，则 $\lambda f + \mu g$ 也是 $C(X)$ 中元素. 所以 $C(X)$ 是一个线性空间.

对于任意 $a \in X$，记

$$f_a(x) = \begin{cases} 1 & x = a \\ 0 & x \neq a. \end{cases}$$

那么可验证 $\{f_a\}$ 构成 $C(X)$ 的一组基.

现在考虑嵌入映射

$$\iota_X : X \to C(X), \quad a \mapsto f_a.$$

该映射显然是 X 与基 $\{f_a\}$ 之间的双射，因此我们可将 X 视作 $C(X)$ 的基. 称 $C(X)$ 为 X 上的自由线性空间或由 X 生成的线性空间.

Chapter 2

Jordan 标准形

1 第 5 讲

定义 1.1. $a \in \mathbb{C}$,$m \in N_+$, 称矩阵

$$J_m(a) := \begin{pmatrix} a & 1 & & \\ & a & \ddots & \\ & & \ddots & 1 \\ & & & a \end{pmatrix}_{m\times m}$$

为 *Jordan* 块；称以 *Jordan* 块为对角块的准对角阵为 *Jordan* 形矩阵.

注记. 对角阵是特殊的 *Jordan* 形矩阵，它的每个 *Jordan* 块的阶数都为 *1*.

定理 1.1 (Jordan 标准形定理)**.** $A \in \mathbb{C}^{n\times n}$, $P_A(\lambda) = (\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$, 则 A 相似于 *Jordan* 标准形 $J = diag(J_1, J_2, \cdots, J_s)$, $J_i \in \mathbb{C}^{n_i\times n_i}$,$n_i$ 是 λ_i 的代数重数.

其中 $J_i = diag(J_{i1}, J_{i2}, \cdots, J_{im_i})$,$m_i$ 是 λ_i 的几何重数.

其中 $J_{ij} = J_{l_{ij}}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{l_{ij}\times l_{ij}}$.

并且 *Jordan* 标准形在不记 *Jordan* 块的排列次序下是唯一的.

注记. 注意到 J_i 是以代数重数 n_i 为阶数、被分为几何重数 m_i 块的 *Jordan* 形矩阵；它的 *Jordan* 块们的对角元都是 λ_i，但是阶数不同.

一般来说 J_i 与一整块以 n_i 为阶数的 *Jordan* 块的差别在于它的准对角线上有一些位置的 1 是空缺的，这些位置就是每个（除了第一个）*Jordan* 块的起始位置的上方.

注记.

$$n_1 + n_2 + \cdots + n_s = n$$

$$l_{i1} + l_{i2} + \cdots + l_{im_i} = n_i$$

注记. A 与 B 相似 $\Leftrightarrow A$ 与 B 有相同的 $Jordan$ 标准形

右推左是显然的，因为相似具有传递性

左推右大概正是我们这章要证明的东西，但是你能看到我们有一个算某个矩阵的 $Jordan$ 标准型的算法，能够看到一旦矩阵是相似的，那么它们应该就有相同的 $Jordan$ 标准形.

注记. 给定了 $Jordan$ 标准形以后，容易观察到它的最小多项式是

$$d_A(\lambda) = d_J(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_s)^{k_s}$$

其中 $k_i = \max\limits_{1 \leqslant j \leqslant m_i} l_{ij}$

这利用了上节课的习题：准对角阵的最小多项式是对角块的最小多项式的最小公倍式.

注记. A 相似于对角阵 $\Leftrightarrow l_{ij} = 1 \Leftrightarrow k_i = 1 \Leftrightarrow d_A(\lambda)$ 无重根

$l_{ij} = 1 \Leftrightarrow m_i = n_i$

注记. $d_A(\lambda) = P_A(\lambda) \Leftrightarrow m_i = 1$

例 1.1. $A = \begin{pmatrix} 3 & -2 & 1 \\ 2 & -2 & 2 \\ 3 & -6 & 5 \end{pmatrix}$, 求 A 的 $Jordan$ 标准形

先算 A 的特征多项式，看 A 的特征值以及与之与对应的代数重数，在阶数较低时，能够猜测出 $Jordan$ 标准形可能具有的形式；利用相似的矩阵的多项式也相似进而相抵进而秩相同，我们能够给矩阵减去特征值倍的单位矩阵来看它的秩进而判断它是哪个 $Jordan$ 标准形

$P_A(\lambda) = (\lambda - 2)^3$

$J_1 = \begin{pmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{pmatrix}$ $J_2 = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix}$ $J_3 = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 2 \end{pmatrix}$

$r(J - 2I) = r(A - 2I) = 1$

$\Rightarrow J = J_2$

例 1.2. $A = \begin{pmatrix} -2 & 4 & 1 & 0 & 2 \\ -4 & 6 & 1 & -1 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 \\ 0 & 0 & 1 & 0 & 3 \end{pmatrix}$, 求 A 的 $Jordan$ 标准形

该题相较上题，区别在于特征值不再只有一个，但我们立即发现这对我们并不会造成影响，作为上三角阵，它的行列式就是对角元的乘积，注意到不同 J_i 对应着不同的特征值 λ_i，这意味着用 J 减去 λ_i 倍的单位矩阵时，其他对角块 J_j 的秩永远是满的，不会发生改变.

$P_A(\lambda) = (\lambda - 2)^2 (\lambda - 3)^3$

$\lambda_1 = 2$

$J_{11} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ $J_{12} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

$\lambda_2 = 3$

$J_{21} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ $J_{22} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ $J_{23} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

$$r(J-2I)=r(A-2I)=4=rank\begin{pmatrix}J_1-2I & \\ & J_2-2I\end{pmatrix}=r(J_1-2I)+3$$
$$\therefore r(J_1-2I)=1$$
同理有 $r(J_2-3I)=1$

注记. 在上面两题中，由于 n_i 都较小，对于某个 m_i 来说，分解方式只有一种，所以我们一旦确定了 $J-\lambda_i$ 的秩，便确定了 *Jordan* 标准形.

m_i 决定了 *Jordan* 矩阵由多少个 *Jordan* 块组成. 比方说有一个 *Jordan* 矩阵是 n_i 阶的，它的次对角线有 n_i-1 个元素，其中有 m_i-1 个零，剩下的都是 1，所以说 $J_i-\lambda_iI$ 的秩一定为 n_i-m_i，然后我们要把 n_i-m_i 写成 m_i 个非负整数的和，每种能够写出来的和的形式就对应着一种 J_i 的可能的情况，我们需要继续计算 $J_i-\lambda_iI$ 的方幂的秩才能最终确定结果.

记 λ_i 对应的 p 阶 Jordan 块的个数为 $\delta_p^i, p=1,2,\cdots,n$，则应有

$$\sum_{p=1}^{n_i}\delta_p^i=m_i,\sum_{p=1}^{n_i}p\delta_p^i=n_i$$

引入

$$r_k^i=r(A-\lambda_iI)^k$$

$$\begin{aligned}r_1^i&=r(A-\lambda_iI)=r(J-\lambda_iI)\\&=n-n_i+r(J_i-\lambda_iI)\\&=n-n_i+\sum_{p=1}^{n_i}(p-1)\delta_p^i\end{aligned}$$

$$r_k^i=n-n_i+\sum_{p=k}^{n_i}(p-k)\delta_p^i$$

$$\begin{aligned}r_{k-1}^i&=n-n_i+\sum_{p=k-1}^{n_i}(p-k)\delta_p^i\\&=n-n_i+\sum_{p=k}^{n_i}(p+1-k)\delta_p^i\end{aligned}$$

$$d_k^i:=r_{k-1}^i-r_k^i=\sum_{p=k}^{n_i}\delta_p^i$$

$$d_{k+1}^i=\sum_{p=k+1}^{n_i}\delta_p^i$$

$$\Rightarrow\delta_k^i=d_k^i-d_{k+1}^i$$

注记. *你需要想清楚一个 p 阶的 Jordan 块最终在哪次消失. 在计算 $J_i-\lambda_iI$ 的 p 次方幂的秩的时候，p 阶 Jordan 块贡献了它最后令秩减少的力量，自此之后，他已经没有可以再被减的秩了. 哦... 不要忘了还有 1 阶 Jordan 块... 它在我认为的开始：从计算 $rank(J_i-\lambda_iI)$ 到计算 $rank(J_i-\lambda_iI)^2$ 时，都已经不贡献令秩减少的力量了... 它们体现在哪呢... 体现在 $rankI$ 与 $rank(J_i-\lambda_iI)$ 之间.*

例 1.3. 计算 $A^n, e^A, sinA, cosA$, 最终归结为计算 *Jordan* 标准形、*Jordan* 块的 n 次幂

例 1.4. $A^2 = A$, 证明 A 相似于,$r = r(A)$

证

A 的 *Jordan* 标准形为 $J \Rightarrow J^2 = J \Rightarrow J_{l_{ij}}^2 = J_{l_{ij}} \Rightarrow J_{l_{ij}} = (\lambda_i)$

例 1.5.

$$\frac{dx}{dt} = 3x - 2y + z$$
$$\frac{dy}{dt} = 2x - 2y + 2z$$
$$\frac{dz}{dt} = 3x - 6y + 5z$$

$$\frac{dX}{dt} = AX = PJP^{-1}X \Rightarrow \frac{d(P^{-1}X)}{dt}$$

2 矩阵相似变换

定理 2.1. $A \in \mathbb{C}^{n\times n}$, 则 A 相似于上三角阵 B, B 的对角线元素就是 A、B 共同的特征值.

定理 2.2. $A \in \mathbb{C}^{n\times n}$，$P_A(\lambda) = (\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$，则 A 相似于 $diag(A_{11},\cdots,A_{ss})$, 其中 $A_{ii} \in \mathbb{C}^{n_i\times n_i}$ 是上三角矩阵，对角元为 λ_i.

注记. 处理 A_{ii}，考虑干掉对角元 $B_{ii} = A_{ii} - \lambda_i I$, 而对于 B_{ii} 存在 k 满足 $B_{ii}^k = 0$

定义 2.1. $A \in F^{n\times n}$，若存在正整数 m 使得 $A^m = 0$, 则称 A 为幂零矩阵；使上式成立的最小的 m 称为 A 的幂零指数.

注记. 幂零指数为什么不能超过矩阵的阶数???

定理 2.3. A 为幂零矩阵 $\Leftrightarrow d_A(\lambda) = \lambda^t \Leftrightarrow A$ 的特征值全为零

定理 2.4. $A \in \mathbb{C}^{n\times n}$ 的幂零指数为 n，则 A 相似于 $J_n(0)$

证明. 用归纳法

$n = 1$ 时，显然成立

$n = 2$

归纳法，假设结论对 $n-1$ 时成立，A=$\begin{pmatrix} 0 & b \\ 0 & B \end{pmatrix}$.

下面证明 B 的幂零指数为 $n-1$，用反证法，如果 $B^{n-2} = O$, 则 $A^{n-2} = \begin{pmatrix} 0 & * \\ 0 & O \end{pmatrix} \Rightarrow A^{n-1} = O$，矛盾

□

定理 2.5. 则 A 相似于 $diag(J_{l_1}(0),\ldots,J_{l_k}(0))$, 这里

证明

用归纳法，*n=1*，显然；假设结论对于 $<n$ 成立，下面证明结论对于 n 成立

A=$\begin{pmatrix} 0 & b \\ 0 & B \end{pmatrix}$

3 根子空间分解

设 V 是 n 维复向量空间，$\mathscr{A}$ 是 V 上线性变换，$\mathscr{A}$ 的特征多项式是

$$\varphi(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}.$$

$A \in \mathbb{C}^{n\times n}$,$P_A(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}$ $\quad Ker(A - \lambda_i I)^k$ 要在 $Ker(A - \lambda_i I)^{k+1}$ 找 n_i 个线性无关向量

$W_i := Ker(A - \lambda_i I)^{k+1}$，称为根子空间

问题转化为

$$\mathbb{C}^n = W_{\lambda_1} + W_{\lambda_2} + \cdots + W_{\lambda_s}$$

定义 3.1. *$\mathscr{A}$ 是 n 维线性空间 V 上的线性变换，a 是 $\mathscr{A}$ 的特征值，如果存在正整数 k 使得 $(\mathscr{A} - a\mathscr{I})^k\beta = 0$，则称非零向量 $\beta \in V$ 为 $\mathscr{A}$ 的属于 a 的根向量. 使上式成立的最小正整数 k 称为 β 的次数. 特别地，特征向量为 1 次根向量.*

命题 3.1. 设

$$V_i = \left\{ v \in V \middle| \exists\ k \text{ s.t. } (A - \lambda_i I)^k v = 0 \right\},$$
$$W_i = \left\{ v \in V \middle| (A - \lambda_i I)^{n_i} v = 0 \right\}.$$

则 $V_i = W_i$.

定义 3.2.

定理 3.1. *$\mathscr{A}$ 是 n 维线性空间 V 上线性变换，其特征多项式为 $P_{\mathscr{A}}(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}$，则 $\mathscr{A}$ 的属于 λ_i 的根向量全体以及零向量构成 V 的 n_i 维子空间，称为根子空间，记为 W_{λ_i}.*

证明. 首先有

$$Ker(\mathscr{A} - \lambda_i \mathscr{I}) \subseteq Ker(\mathscr{A} - \lambda_i \mathscr{I})^2 \subseteq \cdots \subseteq Ker(\mathscr{A} - \lambda_i \mathscr{I})^{n_i} \subseteq \cdots$$

考虑 $\mathscr{A}$ 的矩阵表示，存在 V 的一组基 M，使得 $\mathscr{A}$ 在 M 下的矩阵 A 为上三角阵，其对角线元素为其特征值，相同的特征值连在一起

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & * & \\ & & \lambda_1 & & & & \\ & & & \lambda_2 & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & \lambda_s \end{pmatrix}$$

考虑 $(A - \lambda_i I)^k$，由于特征值互不相同，所以右下角部分的对角线元素始终不为零，所以有

$$rank(A - \lambda_i)^k = n - n_i, when \quad k \geqslant n_i$$

$$dimKer(\mathscr{A} - \lambda_i \mathscr{I})^k = n - r(A - \lambda_i)^k = n_i, when \quad k \geqslant n_i$$

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

此时我们已经证明了，

$$Ker(\mathscr{A}-\lambda_i\mathscr{I})^k \subseteq Ker(\mathscr{A}-\lambda_i\mathscr{I})^{n_i}, \forall k \geqslant 1$$

所以 $\mathscr{A}$ 的任意特征根都包含在 $Ker(\mathscr{A}-\lambda_i\mathscr{I})^{n_i}$ 中，并且其维数为 n_i. □

定理 3.2. $\mathscr{A}$ 是 n 维线性空间 V 上线性变换，其特征多项式为 $P_{\mathscr{A}}(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$，则

$$V=W_{\lambda_1}\oplus W_{\lambda_2}\oplus\cdots\oplus W_{\lambda_s}$$

定理 3.3. $\mathscr{A}$ 是 n 维线性空间 V 上线性变换,W 是 $\mathscr{A}$ 的不变子空间. 将 W 的基 $M_1=\{\alpha_1,\cdots,\alpha_m\}$ 扩充为 V 的基 $M=\{\alpha_1,\cdots,\alpha_m,\cdots,\alpha_n\}$. 则 $\mathscr{A}$ 在 M 下的矩阵为准上三角阵

$$A=\begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$$

其中 $A_{11}\in F^{m\times m}$ 是 $\mathscr{A}_W$ 在基 M_1 下的矩阵.

定理 3.4. $\mathscr{A}$ 是 n 维线性空间 V 上线性变换，其特征多项式为 $P_{\mathscr{A}}(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$，则

(1) W_{λ_i} 是 $\mathscr{A}$ 的不变子空间.

(2) $\mathscr{A}|_{W_{\lambda_i}}$ 的特征多项式为 $(\lambda-\lambda_i)^{n_i}$.

(3) 存在 V 的一组基 M，$\mathscr{A}$ 在 M 下的矩阵为 $diag(A_1,A_2,\cdots,A_s)$，其中 A_i 为上三角矩阵.

注记. $\mathscr{A}$ 的矩阵表示从**上三角阵**迈进到了**准对角阵并且对角元为上三角阵**.

定义 3.3. $\mathscr{A}$ 是 V 上的线性变换，$w\subset V$ 是子空间，如果 $\mathscr{A}W\subset W$，则称是关于 $\mathscr{A}$ 的不变子空间.

定理 3.5. 线性变换 $\mathscr{A}$ 的有限个不变子空间之和仍为不变子空间；线性变换 $\mathscr{A}$ 的任意个不变子空间之交仍为不变子空间.

定理 3.6. $\mathscr{A}$ 是 V 上线性变换，$W\subset V$ 是 $\mathscr{A}$ 的不变子空间，则存在 V 的一组基 M，$\mathscr{A}$ 在 M 下的矩阵为

$$A=\begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}, A_{11}\in\mathbb{C}^{m\times m}, m=dimW$$

反之，设 $\mathscr{A}$ 在 $M=\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵为上述 A，则 $W=<\alpha_1,\cdots,\alpha_m>$ 为 $\mathscr{A}$ 的不变子空间.

证明. 设 $\alpha_1,\cdots,\alpha_m$ 为 W 的基，扩充为 V 的一组基 $\alpha_1,\cdots,\alpha_m,\cdots,\alpha_n$

$$\mathscr{A}\alpha_1=a_{11}\alpha_1+\cdots+a_{m1}\alpha_m$$

$$\mathscr{A}\alpha_m=a_{1m}\alpha_1+\cdots+a_{mm}\alpha_m$$

$$\mathscr{A}\alpha_{m+1}=a_{1,m+1}\alpha_1+\cdots+a_{m,m+1}\alpha_m+\cdots+a_{n,m+1}\alpha_n$$

$$\mathscr{A}\alpha_n=a_{1n}\alpha_1+\cdots+a_{mn}\alpha_m+\cdots+a_{nn}\alpha_n$$

□

定理 3.7. $\mathscr{A}$ 是 V 上的线性变换，W_1, W_2 是 $\mathscr{A}$ 的不变子空间，且 $V = W_1 \oplus W_2$，设 M_i 为 W_i 的一组基，$i = 1, 2$，则 $\mathscr{A}$ 在 $M = M_1 \cup M_2$ 下的矩阵为

$$A = \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix}, A_i \in \mathbb{C}^{n_i \times n_i}, n_i = dimW_i, i = 1, 2$$

反之亦然.

推论 3.1. $\mathscr{A}$ 是 V 上线性变换, $W_1, \cdots, W_s$ 为 $\mathscr{A}$ 的不变子空间, 且 $V = W_1 \oplus \cdots \oplus W_s$, M_i 为 W_i 的基，则 $\mathscr{A}$ 在 $M = \bigcup\limits_{i=1}^{s} M_i$ 下的矩阵 $A = diag(A_1, \cdots, A_s), A_i \in \mathbb{C}^{n_i \times n_i}, n_i = dimW_i, i = 1, 2, \cdots, s.$

反之亦然. 并且 A_i 是 $\mathscr{A}|_{W_i}$ 在 M_i 下的矩阵.

推论 3.2. $\mathscr{A}$ 可对角化 $\Leftrightarrow V = W_1 \oplus \cdots \oplus W_n$，$W_i$ 为 $\mathscr{A}$ 的一维不变子空间.

例 3.1. $\mathscr{A}, \mathscr{B}$ 是 V 上线性变换，且 $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$，则 $Im\mathscr{B}, Ker\mathscr{B}$ 是 $\mathscr{A}$ 的不变子空间.

证明. $\alpha \in Im\mathscr{B}$，要证 $\mathscr{A}\alpha \in Im\mathscr{B}$

$\alpha = \mathscr{B}\beta, \beta \in V$，$\mathscr{A}\alpha = \mathscr{A}(\mathscr{B}\beta) = \mathscr{B}(\mathscr{A}\beta) \in Im\mathscr{B}$ □

4 循环子空间分解

$\mathscr{A}=\mathscr{A}|_{W_i}, P_{\mathscr{A}}(\lambda)=(\lambda-\lambda_i)^{n_i}$

先设 $d_{\mathscr{A}}(\lambda)=(\lambda-\lambda_i)^{n_i}$, 即 $(\mathscr{A}-\lambda_i\mathscr{I})^{n_i}=0, (\mathscr{A}-\lambda_i\mathscr{I})^{n_i-1}\neq 0$

定理 4.1. $\mathscr{A}$ 是 n 维线性空间 W 上的线性变换, $d_{\mathscr{A}}(\lambda)=(\lambda-a)^n$, 则存在 W 的一组基 $M=\{(\mathscr{A}-a\mathscr{I})^{n-1}\alpha,\cdots,(\mathscr{A}-a\mathscr{I})\alpha,\alpha\}, \alpha\in W$, $\mathscr{A}$ 在 M 下的矩阵为 $J_n(a).W=<(\mathscr{A}-a\mathscr{I})^{n-1}\alpha,\cdots,(\mathscr{A}-a\mathscr{I})\alpha,\alpha>=C_1$.

定义 4.1. $\mathscr{A}$ 是 W 上线性变换, $\alpha\in W$, 称由 $\alpha,\mathscr{A}\alpha,\cdots,\mathscr{A}^t\alpha,\cdots$ 生成的子空间为循环子空间.

定理 4.2. 循环子空间 $C=\{f(\mathscr{A})\alpha|f(\lambda)\in F[\lambda])\}$ 是 $\mathscr{A}$ 的包含 α 的最小不变子空间.

证明. $C=span\{\alpha,\mathscr{A}\alpha,\cdots,\mathscr{A}^t\alpha,\cdots\}, \forall\beta\in C, \beta=\sum_{i=0}^{m}c_i\mathscr{A}^i\alpha=f(\mathscr{A})\alpha$

C 是不变子空间，$\beta=\sum_{i=0}^{m}c_i\mathscr{A}^i\alpha, \mathscr{A}\beta=\sum_{i=0}^{m}c_i\mathscr{A}^{i+1}\alpha\in C$

C 包含 α

$C^{'}$ 是包含 α 的不变子空间，$\alpha\in C\to\mathscr{A}\alpha\in C^{'}\to\cdots\to\mathscr{A}^k\alpha\in C^{'}\to C\subseteq C^{'}$ □

$dimW=n$,$\alpha,\mathscr{A}\alpha,\cdots,\mathscr{A}^n$ 线性相关，存在不全为零的 c_i 使得

$$\sum_{i=0}^{n}c_i\mathscr{A}^i\alpha=0\to f(\mathscr{A})\alpha=0, f(\lambda)=\sum_{i=0}^{n}c_i\lambda^i\neq 0$$

定义 4.2. $\mathscr{A}$ 是 W 上线性变换, $\alpha\in W$, 若存在非零多项式 f 使得 $f(\mathscr{A})\alpha=0$, 则称 f 为 α 关于 $\mathscr{A}$ 的化零多项式. 次数最低的首一话零多项式称为 α 的最小多项式, 记为 $d_{\mathscr{A},\alpha}(\lambda)$

注记. $f(\mathscr{A})=0\to f(\mathscr{A})\alpha=0$

α 的最小多项式唯一.

引理 4.1. $g(\lambda)\in F[\lambda], g(\lambda)=h(\lambda)k(\lambda), (h(\lambda),k(\lambda))=1, \mathscr{A}\in End(V)$, 则有

$$Kerg(\mathscr{A})=Kerh(\mathscr{A})\oplus Kerk(\mathscr{A})$$

证明.
- 直和

因为 $(h(\lambda),k(\lambda))=1$，所以存在 $u(\lambda),v(\lambda)\in F[\lambda]$，使得

$$u(\lambda)h(\lambda)+v(\lambda)k(\lambda)=1$$

带入 $\mathscr{A}$，得到

$$u(\mathscr{A})h(\mathscr{A})+v(\mathscr{A})k(\mathscr{A})=\mathscr{I}$$

假如存在 α 满足 $\alpha\in Kerh(\mathscr{A})$ 且 $\alpha\in Kerk(\mathscr{A})$，即 $h(\mathscr{A})\alpha=0$ 且 $k(\mathscr{A})\alpha=0$

将上式作用于 α，得到

$$u(\mathscr{A})h(\mathscr{A})\alpha+v(\mathscr{A})k(\mathscr{A})\alpha=0=\alpha$$

这样就证明了和是直和

- 相等

  显然 $Kerh(\mathscr{A}) \subset Kerg(\mathscr{A}), Kerk(\mathscr{A}) \subset Kerg(\mathscr{A})$，所以显然有

  $$Kerh(\mathscr{A}) \oplus Kerk(\mathscr{A}) \subset Kerg(\mathscr{A})$$

  所以我们只需证明

  $$Kerg(\mathscr{A}) \subset Kerh(\mathscr{A}) \oplus Kerk(\mathscr{A})$$

  上面我们已经得到了

  $$u(\mathscr{A})h(\mathscr{A})\alpha + v(\mathscr{A})k(\mathscr{A})\alpha = \alpha$$

  下面只需验证这就是我们想要的分解

  $$k(\mathscr{A})u(\mathscr{A})h(\mathscr{A})\alpha = h(\mathscr{A})k(\mathscr{A})u(\mathscr{A})\alpha = g(\mathscr{A})u(\mathscr{A})\alpha = \mathscr{O}u(\mathscr{A})\alpha = 0$$

  所以 $u(\mathscr{A})h(\mathscr{A})\alpha \in Kerk(\mathscr{A})$，同理可证 $v(\mathscr{A})k(\mathscr{A})\alpha \in Kerh(\mathscr{A})$

□

定理 4.3. $\mathscr{A}$ 是 W 上线性变换，$\alpha \in W$，$d_\alpha(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda = a_0$，$C_\alpha$ 是 α 生成的循环子空间.

1. $\{\alpha, \mathscr{A}\alpha, \cdots, \mathscr{A}^{m-1}\alpha\}$ 为 $C\alpha$ 的一组基，$dimC_\alpha = m$.
2. $\mathscr{A}|_C$ 在基 $\{\alpha, \mathscr{A}\alpha, \cdots, \mathscr{A}^{m-1}\alpha\}$ 下的矩阵为 $A = \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & \\ & \ddots & 0 & -a_{m-2} \\ & & 1 & -a_{m-1} \end{pmatrix}$
3. $\mathscr{A}|_C$ 的最小多项式等于特征多项式等于 $d_\alpha(\lambda)$.

证明. 1. 若

$$c_0\alpha + c_1\mathscr{A}\alpha + \cdots + \mathscr{A}^{m-1}\alpha = 0$$

$$f(\mathscr{A})\alpha = 0, f(\lambda) = c_0 + c_1\lambda + \cdots + c^{m-1}\lambda^{m-1}$$

如果 $f \neq 0$，则 f 是化零多项式，这不可能.

下证 $\mathscr{A}^k\alpha \in span\{\alpha, \alpha, \cdots, \mathscr{A}^{m-1}\alpha\}, k = 0, 1, \cdots$

先证 $\mathscr{A}^m\alpha \in span\{alpha, \mathscr{A}\alpha, \cdots, \mathscr{A}^{m-1}\alpha\}$

$$(\mathscr{A}^m + a_{m-1}\mathscr{A}^{m-1} + \cdots + a_1\mathscr{A} + a_0\mathscr{I})\alpha = 0 \to \mathscr{A}^m\alpha = -a_{m-1}\mathscr{A}^{m-1}\alpha - \cdots - a_1\alpha - a_0\alpha$$

2.

$$\begin{aligned} \mathscr{A}(\alpha, \mathscr{A}\alpha, \cdots, \mathscr{A}^{m-1}\alpha) &= (\mathscr{A}\alpha, \mathscr{A}^2\alpha, \cdots, \mathscr{A}^m\alpha) \\ &= (\alpha, \mathscr{A}\alpha, \cdots, \mathscr{A}^{m-1}\alpha) \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & \\ & \ddots & 0 & -a_{m-2} \\ & & 1 & -a_{m-1} \end{pmatrix} \end{aligned}$$

3.

$$P_{\mathscr{A}|_{C_\alpha}}(\lambda) = P_A(\lambda) = det(\lambda I - A) = d_\alpha(\lambda)$$

$$d_{\mathscr{A}|_{C_\alpha}}|d_\alpha(\lambda)$$

下证 $d_\alpha(\lambda)|d_{\mathscr{A}|_{C_\alpha}}$

□

定理 4.4. $\mathscr{A}$ 是 W 上线性变换，$d_{\mathscr{A}} = (\lambda - a)^k, k \leqslant n, n = dimW$. 则

1. 存在 $\alpha_1, \cdots, \alpha_m \in W$ 使 $W = C_{\alpha_1} \oplus C_{\alpha_1} \oplus \cdots \oplus C_{\alpha_m}$，$\mathscr{A}|_{C_{\alpha_m}}$ 上最小多项式为 $(\lambda - a)^{l_j}, k = l_1 \geqslant l_2 \geqslant \cdots \geqslant l_m$
2. 存在 C_{α_j} 的一组基 M_j，$\mathscr{A}$ 在 $M = \cup_{j=1}^m M_j$ 下的矩阵为 $diag(J_{l_1}(\alpha), J_{l_2}(\alpha), \cdots, J_{l_m}(\alpha))$

证明.

$$\mathscr{B} = \mathscr{A} - a\mathscr{I}, d_{\mathscr{B}}(\lambda) = \lambda^k, P_{\mathscr{B}}(\lambda) = \lambda^n$$

$$C_{\alpha_1} =< \alpha_1, \mathscr{B}\alpha_1, \cdots, \mathscr{B}^{l_1-1}\alpha_1 >, \mathscr{B}^{l_1}\alpha_1 = 0, \alpha_1 \in Ker\mathscr{B}^{l_1}, \alpha_1 \notin Ker\mathscr{B}^{l_1-1}$$

$$C_{\alpha_m} =< \alpha_m, \mathscr{B}\alpha_m, \cdots, \mathscr{B}^{l_1-1}\alpha_m >, \mathscr{B}^{l_1}\alpha_m = 0, \alpha_1 \in Ker\mathscr{B}^{l_m}, \alpha_1 \notin Ker\mathscr{B}^{l_m-1}$$

$$Ker\mathscr{B} \subset Ker\mathscr{B}^2 \subset \cdots \subset Ker\mathscr{B}^{k-1} \subset Ker\mathscr{B}^k = W$$

$$r_i = rank(\mathscr{B}^i)$$

$$Ker\mathscr{B}^K = Ker\mathscr{B}^{k-1} \oplus U_k, dimU_k = r_{k-1} - rk := d_k$$

$\alpha_1, \cdots, \alpha_{d_k}$ 是 U_k 的一组基

$$C_{\alpha_1} =< \alpha_1, \mathscr{B}\alpha_1, \cdots, \mathscr{B}^{l_1-1}\alpha_1 >$$

$$C_{\alpha_{d_k}} =< \alpha_{d_k}, \mathscr{B}\alpha_{d_k}, \cdots, \mathscr{B}^{l_1-1}\alpha_{d_k} >$$

下面证明 $C_{\alpha_1} + \cdots + C_{\alpha_{d_k}}$ 为直和

设 $\beta_1 + \cdots + \beta_{d_k} = 0, \beta_i \in C_{\alpha_i}$，要证 $\beta_i = 0$

$$\beta_i = \sum_{j=0}^{k-1} \lambda_{ij}\mathscr{B}^j\alpha_i$$

$$\sum_{i=1}^{d_k} \beta_i = \sum_{i=1}^{d_k}\sum_{j=0}^{k-1} \lambda_{ij}B^j\alpha_i = 0$$

$$= \sum_{j=0}^{k-1} \mathscr{B}^j(\sum_{i=1}^{d_k} \lambda_{ij}\alpha_i) = 0$$

$$\to \mathscr{B}^{k-1}(\sum_{i=1}^{d_k} \lambda_{i0}\alpha_i) = 0 \to \sum_{i=1}^{d_k} \lambda_{i0}\alpha \in Ker\mathscr{B}^{k-1} \cap U_k = 0$$

$$\to \lambda_{i0} = 0$$

$$\to \sum_{j=1}^{k-1} \mathscr{B}^{j-1}(\sum_{i=1}^{d_k} \lambda_{ij}\alpha_i) = 0$$

$$\lambda_{ij} = 0 \to \beta_i = 0$$

再找 $\alpha \in Ker\mathscr{B}^{k-1}, \notin \mathscr{B}^{k-2}$

$$Ker\mathscr{B}^{k-1} = Ker\mathscr{B}^{k-2} \oplus V_{k-1} = Ker\mathscr{B}^{k-2} \oplus \mathscr{B}U_k \oplus U_{k-1}$$

$$dimU_{k-1} = dimKer\mathscr{B}^{k-1} - dimKer\mathscr{B}^{k-2} - dimBU_k = r_{k-2} - r_{k-1} - d_k = d_{k-1} - d_k = \delta_{k-1}$$

找 $\alpha_{\delta_k+1}, \cdots, \alpha_{\delta_k+\delta_{k-1}} \in U_{k-1}$ 为一组基

$$C_{\alpha_{\delta_k+1}} =< \alpha_{\delta_k+1}, \mathscr{B}\alpha_{\delta_k+1}, \cdots, \mathscr{B}^{l_1-1}\alpha_{\delta_k+1} >$$
$$C_{\alpha_{\delta_k+\delta_{k-1}}} =< \alpha_{\delta_k+\delta_{k-1}}, \mathscr{B}\alpha_{\delta_k+\delta_{k-1}}, \cdots, \mathscr{B}^{l_1-1}\alpha_{\delta_k+\delta_{k-1}} >$$

下面证明 $C_{\alpha_{\delta_k+1}} + \cdots + C_{\alpha_{\delta_k+\delta_{k-1}}}$ 为直和

$$Ker\mathscr{B} = \mathscr{B}^{k-1}U_k \oplus \mathscr{B}^{k-2}U_{k-1} \oplus \cdots \oplus \mathscr{B}U_2 \oplus U_1$$

□

例 4.1. *1.* 依次将 $Ker\mathscr{B}^i$ 的基扩充为 $Ker\mathscr{B}^{i+1}$ 的基，$i = 1, 2, \cdots$

$$Ker$$

2. 令 $T = \mathscr{B}^{k-1}S_k \cup \mathscr{B}^{k-2}S_{k-1} \cup \cdots \cup \mathscr{B}S_2 \cup S_1 \in Ker\mathscr{B}$ 从左到右求 T 的极大无关组 $T* = \mathscr{B}^{k-1}S_k^* \cup \mathscr{B}^{k-2}S_{k-1}^* \cup \cdots \cup \mathscr{B}S_2^* \cup S_1^*$

3. α 向量集合 $S_k^* \cup S_{k-1}^* \cdots \cup S_2^* \cup S_1^*$

5 多项式矩阵的相抵

定义 5.1. F 是数域，λ 为未定元，称矩阵

$$A(\lambda) = (a_{ij}(\lambda))_{m\times n}, a_{ij}(\lambda) \in F[\lambda]$$

为多项式矩阵或 λ 矩阵. 所有 $m \times n$ 阶多项式矩阵全体记为 $F[\lambda]^{m\times n}$.

注记. 过去矩阵的矩阵元取值于域 F，现在多项式矩阵的矩阵元取值于环 $F[\lambda]$，域与环的区别在于非零元是否可逆. 但是注意到我们对所谓伴随矩阵的定义并没有用到除法，所以对于多项式矩阵我们也可以定义行列式、定义相应的伴随矩阵，进而考虑其是否满秩、是否可逆等.

定义 5.2. 称多项式矩阵 A 的非零子式的最大阶数为它的秩，这与数域上的矩阵的定义是相同的.

注记.
- 思考:数域上的矩阵的秩的等价定义还有行向量或列向量的极大线性无关组的个数,对于多项式矩阵是否还有类似的理论？行秩是否等于列秩等于矩阵的秩？这个思考是否有价值？(感觉无)
- $A(\lambda)$ 满秩 $\Leftrightarrow det(A(\lambda)) \neq 0$

定义 5.3. 设 $A(\lambda) \in F[\lambda]^{n\times n}$，若存在 $B(\lambda) \in F[\lambda]^{n\times n}$ 使得 $A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I^{(n)}$，则称 $A(\lambda)$ 为可逆 λ-方阵，记 $B(\lambda) = A(\lambda)^{-1}$. 这与数域上的矩阵的定义是相同的.

我们可以看到上面关于多项式矩阵的秩与多项式矩阵的逆的定义都与数域上的矩阵相同，然而接下来这个定理开始展示二者的不同.

定理 5.1. $A(\lambda)$ 可逆 $\Leftrightarrow det(A(\lambda)) = C \neq 0$，$C$ 是一个常数.

证明. "$\Rightarrow$"

$A(\lambda)$ 可逆，存在 $B(\lambda)$ 使得 $A(\lambda)B(\lambda) = I$

$\Rightarrow det(A(\lambda))det(B(\lambda)) = 1$

$\Rightarrow det(A\lambda) = C \neq 0$

"$\Leftarrow$"

$det A(\lambda) = C \neq 0, A(\lambda)^{-1} = \dfrac{1}{C}A(\lambda)^* \in F[\lambda]^{n\times n}$ □

定义 5.4. 我们可以仿照数域上的矩阵的情况，定义初等变换

- 交换两行
- 乘上某个常数
- 把某一行乘上某个多项式加到另一行

注记. 容易验证初等方阵都是可逆方阵，并且其逆是同类型的初等方阵.

定义 5.5. $A(\lambda), B(\lambda) \in F[\lambda]^{m\times n}$，若存在有限个初等变换将 $A(\lambda)$ 化为 $B(\lambda)$，称 $A(\lambda)$ 与 $B(\lambda)$ 相抵，记为 $A(\lambda) \sim B(\lambda)$

注记. 容易验证相抵是等价关系，我们自然要思考最简代表元和全系不变量的问题.

定理 5.2. 若 $A(\lambda)\in F[\lambda]^{m\times n}$，则经过有限次初等变换，$A(\lambda)$ 可以化为 $\begin{pmatrix} D(\lambda) & 0 \\ 0 & 0 \end{pmatrix}$ 的形式，其中 $D(\lambda)=diag(d_1(\lambda),\cdots,d_r(\lambda)), r=rank(A(\lambda)), d_i(\lambda)|d_{i+1}(\lambda)$

证明. 若 $A(\lambda)=0$，显然；

下设 $A(\lambda)\neq 0$, 则存在 $a_{ij}(\lambda)\neq 0$，我们可以通过行列变换将该非零元挪到 $(1,1)$ 的位置，所以不妨设 $a_{11}(\lambda)\neq 0$

下证经过有限次初等变换 $A(\lambda)$ 可以化为 $\begin{pmatrix} b_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{m2}(\lambda) & \cdots & b_{mn}(\lambda) \end{pmatrix}$ 的形式,并且 $b_{11}(\lambda)\neq 0, b_{11}(\lambda)|b_{ij}(\lambda)$.

对 $d:=deg(a_{11}(\lambda))$ 进行归纳.

若 $d=0$, 那么 $a_{11}(\lambda)$ 为非零常数，显然成立.

设结论对 $d\leqslant k-1$ 都成立，下证结论对 $d=k$ 成立. 大致思路是先试图把第一行和第一列其他位置打成 0，如果有一个元素不能被整除，那做带余除法，就能得到一个次数更低的，把这个次数更低的换到左上角，由归纳假设，得证；如果成功把第一行和第一列都打成 0 了，看是否整除右下角那个块的所有元素，如果有一个不行，就把那个块所在的行直接加到第一行，此时情况转化为了前面一种情况，得证. □

推论 5.1. $A(\lambda)$ 可逆 $\Leftrightarrow A(\lambda)$ 是若干初等方阵的乘积

注记. 右推左是显然的，因为可逆方阵的乘积还是可逆方阵；而左推右我想正好揭示了为什么我们要研究初等方阵. 顺便说一句，数域上的矩阵与多项式环上的矩阵的初等变换的定义有一条存在着不同，即第二条，给某一行乘上一个常数，如果不假思索可能会认为多项式矩阵中的对应情形应该是给某一行乘上一个多项式. 之所以是乘一个数而不是乘一个多项式就是因为我们现在考虑的是环而不是域，而环中的非零元不一定可逆，也就是说“给某一行乘上一个多项式”这个变换不一定是可逆的.“但把某一行乘上一个多项式加到另一行”这个变换确实总是可逆的.

例 5.1. $\begin{pmatrix} 4 & 3 & -4 \\ -1 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix}$，将 $\lambda I-A$ 化为标准型.

$$\lambda I-A=\begin{pmatrix} \lambda-4 & -3 & -4 \\ 1 & \lambda & -2 \\ -1 & -1 & \lambda \end{pmatrix}\sim\begin{pmatrix} 1 & & \\ & 1 & \\ & & (\lambda-1)^2(\lambda-2) \end{pmatrix}$$

注记. 在后面的小节中我们将看到：

- $d_1(\lambda),\cdots,d_r(\lambda)$ 由 $A(\lambda)$ 唯一确定
- 它们是全序不变量

定义 5.6. $A(\lambda)\in F[\lambda]^{m\times n}.A(\lambda)$ 的所有 k 阶非零子式的最大公因式称为 $A(\lambda)$ 的 k 阶行列式因子，记为 $D_k(\lambda)$. 如果 $A(\lambda)$ 的所有 k 阶子式均为零，则规定 $D_k(\lambda)=0$. 虽然零阶行列式因子是不存在的，但是我们特别地规定 $D_0(\lambda)=1$.

注记. *1.* 若 $rank(A(\lambda)) = r$，则 $D_{r+1}(\lambda) = \cdots = D_s(\lambda) = 0, s = min(m, n)$.

2. $D_i(\lambda)|D_{i+1}(\lambda), i = 0, 1, \cdots, r$.

3. $A(\lambda) = \begin{pmatrix} D(\lambda) & 0 \\ 0 & 0 \end{pmatrix}, D(\lambda) = diag(d_1(\lambda), \cdots, d_r(\lambda))$

上面出现的 $D(\lambda)$ 是5.2中的 $D(\lambda)$，注意不要与新定义的行列式因子 $D_k(\lambda)$ 混淆. 区分方法：带角标的就是行列式因子，回忆，行列式因子的定义是全体该阶子式的最大公因式. 这里是直接假设 $A(\lambda)$ 具有上面的形式，而不是相似到了它. 在这种情况下，我们很容易写出 $A(\lambda)$ 的各阶行列式因子.

$D_0(\lambda) = 1,$

$D_1(\lambda) = gcd(d_1, d_2, \cdots, d_r) = d_1(\lambda)$

$D_2(\lambda) = gcd(d_1d_2, d_1d_3, \cdots, d_{r-1}d_r) = d_1(\lambda)d_2(\lambda)$

$D_i(\lambda) = d_1(\lambda) \cdots d_i(\lambda), i = 1, 2, \cdots, r$

$D_{r+1}(\lambda) = 0$

进而我们得到了递推关系：

$$d_i(\lambda) = \frac{D_i(\lambda)}{D_{i-1}(\lambda)}, i = 1, 2, \cdots, r$$

注意该递推关系现在只是我们在 $A(\lambda)$ 具有特殊形式的前提下得到的.

定理 5.3. $A(\lambda), B(\lambda) \in F[\lambda]^{m\times n}$，则

$A(\lambda) \sim B(\lambda) \Leftrightarrow A(\lambda)$ 与 $B(\lambda)$ 的各阶行列式因子相同.

证明. "$\Rightarrow$"

Cauchy-Binet 公式

"$\Leftarrow$" 思路，把 $A(\lambda)$ 和 $B(\lambda)$ 都相抵到5.2中的形式，由上面已经证明了的"$\Rightarrow$"，我们知道此时仍旧有 $D_k(\lambda) = d_1(\lambda) \cdots d_k(\lambda), 1 \leqslant k \leqslant r$，进而我们证明 $A(\lambda)$ 与 $B(\lambda)$ 相抵的这个标准形是同一个东西，由相抵的传递性，得证.

$$D_k(\lambda) = \tilde{D}_k(\lambda), k = 0, 1, \cdots, r$$

$$\Rightarrow rank(A(\lambda)) = rank(B(\lambda)) = r$$

$$A(\lambda) \sim \begin{pmatrix} D(\lambda) & 0 \\ 0 & 0 \end{pmatrix}$$

$$D_k(\lambda) = d_1(\lambda) \cdots d_k(\lambda), 1 \leqslant k \leqslant r$$

$$B(\lambda) \sim \begin{pmatrix} \tilde{D}(\lambda) & 0 \\ 0 & 0 \end{pmatrix}$$

$$\tilde{D}_k(\lambda) = \tilde{d}_1(\lambda) \cdots \tilde{d}_k(\lambda), 1 \leqslant k \leqslant r$$

已知

$$D_k(\lambda) = \tilde{D}_k(\lambda), k = 1, 2, \cdots, r$$

$$d_1(\lambda) \cdots d_k(\lambda) = \tilde{d}_1(\lambda) \cdots \tilde{d}_k(\lambda)$$

$$\Rightarrow d_i(\lambda) = \tilde{d}_i(\lambda), i = 1, 2, \cdots, r \Rightarrow D(\lambda) = \tilde{D}(\lambda)$$

$$A(\lambda) \sim \begin{pmatrix} D(\lambda) & 0 \\ 0 & 0 \end{pmatrix} \sim B(\lambda) \Rightarrow A(\lambda) \sim B(\lambda)$$

□

定义 5.7. $A(\lambda) \in F[\lambda]^{m\times n}$, $D_k(\lambda)$ 是 $A(\lambda)$ 的 k 阶行列式因子, 则称 $d_k(\lambda) = \dfrac{D_k(\lambda)}{D_{k-1}(\lambda)}$ 为 $A(\lambda)$ 的不变因子,$i = 1, 2, \cdots r$,称 $\begin{pmatrix} D(\lambda) & 0 \\ 0 & 0 \end{pmatrix}$ 为 $A(\lambda)$ 的 *Smith* 标准形,其中 $D(\lambda) = diag(d_1(\lambda), \cdots, d_r(\lambda)), d_i(\lambda)\big|d_{i+1}(\lambda)$

推论 5.2. 不变因子全体是多项式矩阵相抵的全系不变量.

证明. $A(\lambda) \sim B(\lambda) \Leftrightarrow D_k(\lambda) = \tilde{D}_k(\lambda), k = 0, \cdots, min(m, n) \Leftrightarrow d_i(\lambda) = \tilde{d}_i(\lambda), i = 1, 2, \cdots, r$ □

下面假设 $A(\lambda) \in \mathbb{C}[\lambda]^{m\times n}$，那么不变因子可以分解成一次因式的乘积

$$\begin{cases} d_1(\lambda) = (\lambda - \lambda_1)^{n_{11}}(\lambda - \lambda_2)^{n_{12}} \cdots (\lambda - \lambda_s)^{n_{1s}} \\ \cdots \\ d_r(\lambda) = (\lambda - \lambda_1)^{n_{r1}}(\lambda - \lambda_2)^{n_{r2}} \cdots (\lambda - \lambda_s)^{n_{rs}} \end{cases}$$

注记.

- 上面中的一些 n_{ij} 可能为零，只是为了整齐我们把他们补充了出来
- 显然有 $n_{1i} \leqslant n_{2i} \leqslant \cdots \leqslant n_{ri}$

定义 5.8. 称 $\{(\lambda - \lambda_j)^{n_{ij}} | n_{ij} > 0\}$ 为 $A(\lambda)$ 的初等因子组.

例 5.2. 若不变因子为 $d_1(\lambda) = \lambda(\lambda - 1)^2, d_2(\lambda) = \lambda(\lambda - 1)^3(\lambda + 1)$
那么初等因子组为 $\{\lambda, \lambda, (\lambda - 1)^2, (\lambda - 1)^3, \lambda + 1\}$

注记. 初等因子组中重复的项要重复记录.

例 5.3. 求下列矩阵的 *Smith* 标准形，行列式因子，不变因子及初等因子组

1. $A(\lambda) = \begin{pmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{pmatrix}$
2. $A(\lambda) = \lambda I - J, J = J_n(a)$
3. $A(\lambda) = \begin{pmatrix} \lambda^2(\lambda-1)^3(\lambda+1) & 0 \\ 0 & \lambda^4(\lambda-1)^2 \end{pmatrix}$
4. $A(\lambda) \in \mathbb{C}[\lambda]^{4\times 4}, r(A(\lambda)) = 3$, 初等因子组为

$$\{\lambda^2, \lambda^4, (\lambda - 1)^2, (\lambda - 1)^3, \lambda + 1\}$$

证明. 1. $A(\lambda) \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda(\lambda+1) \end{pmatrix}$

由定义立得行列式因子，不变因子及初等因子组.

2. $A(\lambda)=\begin{pmatrix}\lambda-a & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & \lambda-a\end{pmatrix}$

   观察到右上角的子式行列式不为零，且为正负 1，一大堆行列式因子就被确定下来了

3. 因为是对角阵，所以算行列式因子容易

□

定理 5.4. $A(\lambda),B(\lambda)\in\mathbb{C}[\lambda]^{m\times n}$，则 $A(\lambda)\sim B(\lambda)\Leftrightarrow r(A(\lambda))=r(B(\lambda))$ 且 $A(\lambda)$ 与 $B(\lambda)$ 有相同的初等因子组.

证明. “$\Rightarrow$”显然.

“$\Leftarrow$”$A(\lambda)$ 的不变因子由秩及初等因子组唯一决定. □

定理 5.5. $A(\lambda)=diag(f_1(\lambda),\cdots,f_r(\lambda))$，则 $\{f_i(\lambda)\}_{i=1}^r$ 初等因子组的全体就是 $A(\lambda)$ 的初等因子组.

话说这一节的证明思路都很简单，但怎么就都这么长这么繁呢……

推论 5.3. $A(\lambda)=diag(A_1(\lambda),\cdots,A_r(\lambda))$ 的初等因子组是 $A_i(\lambda)$ 的初等因子组的并.

推论 5.4. $J=diag(J_1,\cdots,J_s),J_i=diag(J_{l_{i1}}(\lambda_i),\cdots,J_{l_{im_i}}(\lambda_i))$，则 $\lambda I-J$ 的初等因子组为 $\{(\lambda-\lambda_i)^{l_{im_i}}|1\leqslant i\leqslant s\}$

6 多项式矩阵相抵与矩阵的相似标准形

定理 6.1. *$\mathscr{A}$ 是 V 上线性变换，$\mathscr{A}$ 在 V 的一组基下的矩阵为 A，设 $\lambda I-A$ 的 $Smith$ 标准形为 $S(\lambda)=diag(1,\cdots,1,f_{s+1}(\lambda),\cdots,f_n(\lambda))$，$f_i(\lambda)$ 是次数为 d_i 的首一多项式，则*

1. *存在 $\beta_i\in V$，$i=s+1,\cdots,n$，使 $V=F[\mathscr{A}]\beta_{s+1}\oplus\cdots\oplus F[\mathscr{A}]\beta_n$*

   且 $d_{\beta_i(\lambda)}=f_i(\lambda), dimF[\mathscr{A}]\beta_{s+1}=d_i$

2. *$B_i=\{\beta_i,\mathscr{A}\beta_i,\cdots,\mathscr{A}^{d_i-1}\beta_i\}$ 是 C_{β_i} 的一组基，$\mathscr{A}$ 在*

$$B=B_{s+1}\cup\cdots\cup B_n$$

   下的矩阵为 $B=diag(B_{s+1},\cdots,B_n), B_i=\begin{pmatrix}0 & & & -a_{i0}\\ 1 & & & \vdots\\ & \ddots & 0 & -a_{i,i-2}\\ & & 1 & -a_{i,i-1}\end{pmatrix}$

证明. 1. 设 $M=\{\alpha_1,\cdots,\alpha_n\}$ 为 V 的一组基，$\mathscr{A}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)A$

$$(\alpha_1,\cdots,\alpha_i)(\mathscr{A}I-A)=0$$

存在可逆方阵

□

定义 6.1. *$A\in F^{n\times n}$，称 $\lambda I-A$ 为 A 的特征方阵，B 称为 A 的有理相似标准形.*

定理 6.2. *$A,B\in F^{n\times n}$，则 A 与 B 在 F 上相似 $\Leftrightarrow\lambda I-A$ 与 $\lambda I-B$ 在 F 上相抵.*

定理 6.3. *$A,B\in F^{n\times n}$，$F\subset K$，K 为数域，则 A 与 B 在 F 上相似 $\Leftrightarrow$ 在 K 上相似*

定理 6.4. *$A\in\mathbb{C}^{n\times n}$，$\lambda I-A$ 的初等因子组为，则*

例 6.1. *$\mathscr{A}$ 是 V 上线性变换，$\mathscr{A}$ 在 V 的一组基 M 下的矩阵为 A，证明：*

1. *$P_{\mathscr{A}}(\lambda)$ 等于 $\lambda I-A$ 的所有不变因子的乘积，$d_{\mathscr{A}}=d_n(\lambda)$*
2. *$d_{\mathscr{A}}(\lambda)=P_{\mathscr{A}}(\lambda)\Leftrightarrow\mathscr{A}$ 为循环变换（$V=F[\mathscr{A}]\beta,\beta\in V$）*

证明. 1. $P(\lambda)(\lambda I-A)Q(\lambda)=diag(d_1(\lambda),\cdots,d_n(\lambda))$ 两边取行列式

□

定义 6.2. *$\mathscr{A}$ 是 V 上线性变换，若存在 $\beta\in V$ 使得 $V=F[\alpha]\beta$，称 $\mathscr{A}$ 为循环变换*
A 为方阵，若 $d_A(\lambda)=P_A(\lambda)$，则称 A 为单纯方阵
$\mathscr{A}$ 为循环变换 $\Leftrightarrow\mathscr{A}$ 在 V 的一组基下的矩阵为单纯方阵

7 实方阵相似

引理 7.1. 设 $A \in \mathbb{R}^{n\times n}$，则 A 的虚特征值对应的初等因子成对共轭出现

引理 7.2. $\tau = a + bi, a, b \in \mathbb{R}, b \neq 0$，则 $A = \begin{pmatrix} \tau & \\ & \bar{\tau} \end{pmatrix}$ 相似于 $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$

证明. $A = \begin{pmatrix} a+bi & \\ & a-bi \end{pmatrix} = aI + b\begin{pmatrix} i & \\ 0 & -i \end{pmatrix}$

只要证 $\begin{pmatrix} i & \\ & -i \end{pmatrix} \sim \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ □

8 Jordan 标准形的应用

例 8.1. A 为单纯方阵且 $AB=BA$，证明：存在多项式 f 使得 $B=f(A)$

几何方法. $V=\mathbb{C}^n, \mathscr{A}: X\to AX, X\in\mathbb{C}^n, \mathscr{B}: X\to BX, X\in\mathbb{C}^n$

A 为单纯方阵 $\Leftrightarrow \mathscr{A}$ 是循环变换，$V=F[\mathscr{A}]\beta, \beta\in V$

β 是存在的还是任意的都可以???

$\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$

要证 $\mathscr{B}=f(\mathscr{A})$

上面是转化成几何的语言

要说明两个变换相等，就要说明作用在每个元素上相等

$\mathscr{B}v=f(\mathscr{A})v, \forall v\in V$

$v=g(\mathscr{A})\beta$

我们需要知道 $\mathscr{B}\beta$ 是什么东西

因为 $\mathscr{B}\beta$ 也是 V 中的元素，那么 $\mathscr{B}\beta=f(\mathscr{A})\beta$

$\mathscr{B}v=\mathscr{B}g(\mathscr{A})\beta=g(\mathscr{A})\mathscr{B}\beta=g(\mathscr{A})f(\mathscr{A})\beta=f(\mathscr{A})g(\mathscr{A})\beta=f(\mathscr{A})v$

$\Rightarrow \mathscr{B}=f(\mathscr{A})$ □

代数方法. A 为单纯方阵，$A=PJP^{-1}, J=diag(J_{m1}(\lambda_1),\cdots,J_{ms}(\lambda_s))$，$\lambda$ 互不相等.

$AB=BA\Rightarrow PJP^{-1}B=BPJP^{-1}\Rightarrow JP^{-1}P=P^{-1}BPJ$

$\Rightarrow J\tilde{B}=\tilde{B}J, \tilde{B}=(\tilde{B_i}j_{s\times s})$

把问题化为其中一个是标准形的问题

$\Rightarrow J_i\tilde{B}_{ij}=\tilde{B}_{ij}J_j, 1\leqslant i,j\leqslant s$

当 $i\neq j$，J_i 与 J_j 特征值不相同，所以上面的方程只有零解，所以 B 是准对角阵

当 $i=j$，$J_i\tilde{B}_{ii}=\tilde{B}_{ii}J_i\Rightarrow(J_i-\lambda_iI)\tilde{B}_{ii}=\tilde{B}_{ii}(J_i-\lambda_iI)$

$\tilde{B}_{ii}=b_{11}I+b_{12}N+b_{13}N^2+\cdots+b_{1n}N^{n-1}$

$\Rightarrow \tilde{B}_{ii}=\tilde{f}_i(N)=\tilde{f}_i(J_i-\lambda_iI)=f_i(J_i)$

$\Rightarrow \tilde{B}=diag(f_1(J_1),\cdots,f_s(J_s))$

找一个 f 使得 $f(J_i)=f_i(J_i), \forall i$

关键：J_i 有化零多项式！

$g_i(\lambda)=(\lambda-\lambda_i)^{m_i}, f_i(J_i)=0$

如果 $f=q_ig_i+f_i, i=1,\cdots,s\Leftrightarrow f\equiv f_i mod g_i, i=1,\cdots,s$

$\tilde{B}=diag(f(J_1),\cdots,f(J_s))=f(diag(J_1,\cdots,J_s))$

$B=P\tilde{B}P^{-1}=f(Pdiag(J_1,\cdots,J_s)P^{-1})=f(A)$ □

例 8.2. $A\in\mathbb{C}^{n\times n}$，则 A 相似于上三角阵 T，T 的非对角元为任意小的正数或零

证明. $A=PJP^{-1}, J=diag(\cdots J_{m_i}(\lambda),\cdots)$

只需要对 $J=J_m(\lambda)$ 进行证明

取 $\begin{pmatrix}\epsilon & & \\ & \epsilon^2 & \\ & & \\ & & \epsilon^m\end{pmatrix}, \epsilon>0,$

则 $P^{-1}JP$ 满足要求 □

例 8.3. 相似于 $diag(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, 1, \cdots, 1)$ 的方阵为反射方阵，若 $A^2 = I(A \neq \pm I)$，则 A 是有限个反射方阵之积.

证明. $d_A(\lambda) = \lambda^2 - 1 \Rightarrow A = P \begin{pmatrix} -I_r & \\ & I_{n-r} \end{pmatrix} P^{-1}$

$\begin{pmatrix} -I_r & \\ & I_{n-r} \end{pmatrix} = R_1 R_2 \cdots R_r, R_i = diag(1, \cdots, 1, -1, 1, \cdots, 1)$

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

例 8.4. $A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{n \times n}, C \in \mathbb{C}^{m \times m}$

设 $AB = CA$ 且 B C 无公共特征值，则 $A = 0$

证明. $B = PJP^{-1}, C = \tilde{P}\tilde{J}\tilde{P}^{-1}, J = diag(J_1, \cdots, J_s), \tilde{J} = (\tilde{J}_1, cdots, \tilde{J}_t)$

$AB = CA \Rightarrow APJP^{-1} = \tilde{P}\tilde{J}\tilde{P}^{-1}A$

$\Rightarrow \tilde{A}J = \tilde{J}\tilde{A}$

$\Rightarrow \tilde{J}_i \tilde{A}_{ij} = \tilde{A}_{ij} J_j$

相当于把原来的任意的矩阵，转化到对于 $Jordan$ 块来验证 □

例 8.5. $A, B \in \mathbb{F}^{n \times n}$，且 $diag(A, A)$ 与 $diag(B, B)$ 相似，证明 A 与 B 相似

证明. 设 A 的初等因子组为 $\{(\lambda - \lambda^{l_{ij}} \big| 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant m_i)\}$

所以 $diag(A, A)$ 的初等因子组为 $\{(\lambda - \lambda^{l_{ij}}, (\lambda - \lambda^{l_{ij}} \big| 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant m_i)\}$

所以 $diag(B, B)$ 的初等因子组为 $\{(\lambda - \lambda^{l_{ij}}, (\lambda - \lambda^{l_{ij}} \big| 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant m_i)\}$

所以 B 的初等因子组为 $\{(\lambda - \lambda^{l_{ij}} \big| 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant m_i)\}$ □

例 8.6. $2n$ 阶实方阵 A 满足 $A^2 + I = 0$，则 A 实相似于 $\begin{pmatrix} 0 & I^{(n)} \\ -I^{(n)} & 0 \end{pmatrix}$

证明. $d_A(\lambda) = \lambda^2 + 1, A \sim J = \begin{pmatrix} iI^{(n)} & \\ & -iI^{(n)} \end{pmatrix} \sim \begin{pmatrix} 0 & I^{(n)} \\ -I^{(n)} & 0 \end{pmatrix}$

为什么数目相等??? □

Chapter 3

内积空间

为什么要研究内积空间？线性代数的主要内容其实已经完结了，但为什么还要讲一个空间？线性空间是一个代数结构，从实际应用的角度来说还有一些距离，为什么呢？因为线性空间里面没有度量，你不能度量向量的长度，不能度量向量的夹角. 所以要想在实际应用中有用，光有代数结构是不够用的. 内积空间就是在线性空间的基础上加上度量. 为什么叫做内积空间？因为这里我们的度量是内积诱导出来的. 在实际应用中，我们就不考虑一般的数域了，我们只讲两种数域上的度量空间，一个是 $\mathbb{R}$ 上的 Euclid 空间，一个是 $\mathbb{C}$ 上的

1 Euclid 空间

- 实线性空间上的内积，对称正定双线性函数
- 选定一组基后内积的度量矩阵，全体内积与全体对称正定矩阵之间的一一对应
- 在基 $\alpha_1,\cdots,\alpha_n$ 下 $(\alpha,\beta)=X^TAY$，若 $(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)P$，那么在基 $\beta_1,\cdots,\beta_n$ 下 $(\alpha,\beta)=(PX')^TA(PY')=X'P^TAPY'$，由此引出矩阵的相合
- 同一内积在不同基下的度量矩阵相合，选定一组基后得到全体内积与全体对称正定矩阵之间有一一对应，这告诉我们 $\sigma: A\mapsto P^TAP$ 是从对称正定矩阵到对称正定矩阵的一个可逆映射
- 对称正定矩阵在相合关系下只有一个等价类，所有对称正定矩阵都相合于单位阵，这在几何上是说对任意的内积都存在标准正交基

例 1.1. $V=C[a,b],(f,g)=\int_a^b f(x)g(x)dx$

定理 1.1. 内积有下列性质：

(1) $(\alpha,0)=(0,\beta)=0$

(2) $(\sum_{i=1}^n\lambda_i\alpha_i,\sum_{j=1}^m\mu_j\beta_j)=\sum_{i=1}^n\sum_{j=1}^m\lambda_i\mu_j(\alpha_i,\beta_j)$

(3) *(Cauchy-Schwarz 不等式)*$(\alpha,\beta)^2\leqslant(\alpha,\alpha)(\beta,\beta)$，等号成立当且仅当 α,β 线性相关.

证明. 考虑 $(t\alpha+\beta,t\alpha+\beta)=(\alpha,\alpha)t^2+2(\alpha,\beta)t+(\beta,\beta)\geqslant 0$，对 $\forall t\in\mathbb{R}$ 成立.

- 当 $(\alpha,\alpha)=0$ 时，
- 当 $(\alpha,\alpha)\neq 0$ 时，

$\Longrightarrow 4(\alpha,\beta)^2-4(\alpha,\alpha)(\beta,\beta)\leqslant 0$，等号成立当且仅当 $\exists t_0$ 使得 $(t_0\alpha+\beta,t_0\alpha+\beta)=0\Rightarrow t_0\alpha+\beta=0$ □

- $\mathbb{R}^n,(x,y)=x_1y_1+\cdots+x_ny_n\Rightarrow(x_1y_1+\cdots+x_ny_n)^2\leqslant(x_1^2+\cdots+x_n^2)(y_1^2+\cdots+y_n^2)$
- $C[a,b],(f,g)=\int_a^b f(x)g(x)dx\Rightarrow(\int_a^b f(x)g(x))^2\leqslant\int_a^b f^2(x)dx\int_a^b g^2(x)dx$

定义 1.1. V 为 *Euclid* 空间，$\alpha\in V$ 的模长 $|\alpha|=\sqrt{(\alpha,\alpha)}$，$\alpha,\beta\in V$ 的夹角 $\theta=\arccos\dfrac{(\alpha,\beta)}{|\alpha||\beta|}$.

注记. *Cauchy-Schwarz* 公式可记忆为

$$(\alpha,\beta)=|\alpha||\beta|\cos\theta\leqslant|\alpha||\beta|$$

定理 1.2 (三角不等式)**.** $|\alpha+\beta|\leqslant|\alpha|+|\beta|$

证明. $|\alpha+\beta|^2=(\alpha+\beta,\alpha+\beta)=(\alpha,\alpha)+2(\alpha,\beta)+(\beta,\beta)\leqslant|\alpha|^2+2|\alpha||\beta|+|\beta|^2=(|\alpha|+|\beta|)^2$ □

例 1.2. $V=\mathbb{R}^{n\times n}$，可以定义内积 $(A,B)=Tr(AB^T)$

1. $(A,B)=Tr(AB^T)=Tr(BA^T)=(B,A)$

2. $(\lambda A_1+\mu A_2, B) = Tr((\lambda A_1+\mu A_2)B^T) = Tr(\lambda A_1 B^T+\mu A_2 B^T) = \lambda Tr(A_1B^T)+\mu Tr(A_2B^T) = \lambda(A_1, B) + \mu(A_2, B)$

3. $(A, A) = Tr(AA^T) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 \geqslant 0$

此时的 *Cauchy-Schwarz* 公式为 $(Tr(AB^T))^2 \leqslant Tr(AA^T)Tr(BB^T)$

$$A^TA = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_n^T \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}$$

2 标准正交基

- 正交的概念是由内积定义的，并且高度依赖于内积的选取. 比如我们选定一组基后，全体内积便与全体正定矩阵形成了一一对应，容易看出在一些内积下这组基成为正交基，甚至是标准正交基，在一些内积下就不是.
- 我们一般不会将内积换来换去，讨论一个问题都是选定一个内积. 内积之间有没有好坏之分？两个内积之间相差一个什么？
- 选定一个内积后，我们便能将一组基对应到一个正定矩阵. 而我们知道全体基与全体可逆矩阵之间有着一一对应，因此就建立了全体可逆矩阵到全体正定矩阵的一个满射. 特殊的例子可以帮助我们理解这个满射，全体标准正交基对应的度量矩阵都是单位阵，而两个标准正交基之间相差一个正交方阵，这促使我们猜测两个相差一个正交方阵的两个可逆矩阵可以对应到同一个正定方阵，换句话说，可逆矩阵可以分解为正交方阵与正定方阵的乘积.

定义 2.1. 称一组基是标准正交基，如果它满足下列等价叙述：

- 两两正交，模长为 1；
- $(\alpha_i, \alpha_j) = \delta_{ij}$；
- 对应的度量矩阵 $G = I$.

实内积空间 V 中向量之间的关系，从加法和数量乘法运算的角度看，有线性相关与线性无关之区分；从度量的角度看，有正交与不正交之区分. 内积作为增添在线性空间上的结构，必定受到已有的线性结构的约束，

命题 2.1. 若非零向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性相关，则它们不可能两两正交.

证明. 存在不全为零的 $\lambda_1, \lambda_2, \cdots, \lambda_m$，使得

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_m\alpha_m = 0.$$

假如它们两两正交，令上式左右同时与 α_1 作内积，得到

$$\lambda_1(\alpha_1, \alpha_1) = 0,$$

由于 $\alpha_1 \neq 0$，因此 $\lambda_1 = 0$，同理可得其他 $\lambda_i = 0, i = 2, \cdots, m$. 矛盾！ □

定理 2.1. n 维 Euclid 空间中存在标准正交基. 代数上是说，任何正定方阵都相合于单位阵.

设内积 $(\cdot,\cdot)$ 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的度量矩阵为 G，$(\cdot,\cdot)$ 在 $\{\beta_1, \cdots, \beta_n\}$ 下的度量矩阵为 I，那么存在上三角阵 U 使得

$$U^T G U = I \Rightarrow G = (U^{-1})^T U^{-1} = P^T P.$$

例 2.1. 设三维 Euclid 空间 V 上的某内积在基 $\alpha_1, \alpha_2, \alpha_3$ 下的度量矩阵为

$$G = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}.$$

求 V 的一组标准正交基，用 $\alpha_1, \alpha_2, \alpha_3$ 表示.

证明.

$$
\begin{aligned}
\beta_1 &= \alpha_1 \\
(\beta_1,\beta_1) &= (\alpha_1,\alpha_1) = 1 \\
\beta_2 &= \alpha_2 + \lambda_{21}\beta_1 \\
0 &= (\alpha_2,\beta_1) + \lambda_{21}(\beta_1,\beta_1) \\
&= (\alpha_2,\alpha_1) + \lambda_{21}(\alpha_1,\alpha_1) \\
&= 1 + \lambda_{21} \\
\beta_2 &= \alpha_2 - \alpha_1 \\
(\beta_2,\beta_2) &= (\alpha_2-\alpha_1,\alpha_2-\alpha_1) = 1 \\
\beta_3 &= \alpha_3 + \lambda_{31}\beta_1 + \lambda_{32}\beta_2 \\
0 &= (\alpha_3,\beta_1) + \lambda_{31}(\beta_1,\beta_1) \\
&= (\alpha_3,\alpha_1) + \lambda_{31} \\
&= 1 + \lambda_{31} \\
0 &= (\alpha_3,\beta_2) + \lambda_{32}(\beta_2,\beta_2) \\
&= (\alpha_3,\alpha_2-\alpha_1) + \lambda_{32} \\
&= 2 + \lambda_{32} \\
\beta_3 &= \alpha_3 - \beta_1 - 2\beta_2 \\
&= \alpha_3 - \alpha_1 - 2(\alpha_2-\alpha_1) \\
&= \alpha_3 - 2\alpha_2 + \alpha_1 \\
(\beta_3,\beta_3) &= (\alpha_3-2\alpha_2+\alpha_1,\alpha_3-2\alpha_2+\alpha_1) \\
&= 1-4+2+8-12+6 = 1
\end{aligned}
$$

□

定义 2.2. 设 $\{\alpha_1,\cdots,\alpha_n\}$ 与 $\{\beta_1,\cdots,\beta_n\}$ 均为标准正交基，并且

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

因为内积 $(\cdot,\cdot)$ 在这两组基下的度量矩阵都为 I，所以有

$$P^TIP = P^TP = I.$$

我们将满足 $P^TP = I$ 的矩阵称为正交矩阵.

注记. 有趣的是，虽然我们借助内积在两组基下的度量矩阵是 I 这一事实推出了 $P^TP = I$，但 P 作为两组基之间的过渡矩阵，它只反映了线性的结构，$P^TP = I$ 这一结果并不依赖于那两组基到底是不是标准正交基，也就是并不依赖于我们到底选择哪种内积. 也就是说，根据是否相差一个正交矩阵，我们可以将基划分成若干个等价类，

存在多少种内积？有多少个正定矩阵就有多少个内积

例 2.2. $P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 为正交方阵.

命题 2.2. P 为正交方阵 $\Leftrightarrow P$ 的列向量构成 $\mathbb{R}^n$ 的标准正交基 $\Leftrightarrow P$ 的行向量构成 $\mathbb{R}^n$ 的标准正交基

证明. $P=(\alpha_1,\cdots,\alpha_n), P^TP=\begin{pmatrix}(\alpha_1,\alpha_1) & \cdots & (\alpha_1,\alpha_n)\\ \vdots & & \vdots\\ (\alpha_n,\alpha_1) & \cdots & (\alpha_n,\alpha_n)\end{pmatrix}=I$

$\Leftrightarrow (\alpha_i,\alpha_j)=\delta_{ij}$ □

命题 2.3. 正交方阵 P 有以下性质

(1) $\det P=\pm 1$

(2) P 的特征值的模长为 1

(3) P 正交 $\Rightarrow P^T=P^{-1}$ 正交

(4) P,Q 正交 $\Rightarrow PQ$ 正交

证明.

(1) 对 $P^TP=I$ 两边同时取行列式

(2) 设 λ 是特征值，X 是与之对应的一个特征向量.

$PX=\lambda X$，对该式做共轭转置（基本技巧），得到

$$\begin{aligned}&\bar{X}^TP^T=\bar{\lambda}\bar{X}^T\\ \Rightarrow&\bar{X}^TP^TPX=\bar{\lambda}\bar{X}^T\lambda X\\ \Rightarrow&\bar{X}^TX=|\lambda|^2\bar{X}^TX\\ &|\lambda|=1\end{aligned}$$

□

命题 2.4. *Euclid* 空间中标准正交基之间的过渡方阵为正交方阵，反之，若从标准正交基 M 到另一组基 M' 的过渡矩阵为正交方阵，则 M' 也为标准正交基.

例 2.3. A 为可逆实方阵，证明 $A=QR$，Q 为正交方阵，R 为上三角阵.

证明. 令 $A=(\alpha_1,\cdots,\alpha_n)$，$Q=(\beta_1,\cdots,\beta_n)$ 实际上这是 Gram-Schmidt 正交化证明过的东西了.

把代数问题理解为几何问题. □

定义 2.3. U 与 V 是 *Euclid* 空间，$\sigma:U\to V$ 是双射，满足

1. $\sigma(\lambda\alpha+\mu\beta)=\lambda\sigma(\alpha)+\mu\sigma(\beta)$

2. $(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta)$

则称 U 与 V 同构.

定理 2.2. *Euclid* 空间 U 与 V 同构 $\Leftrightarrow$ 线性空间 U 与 V 同构

证明. "$\Rightarrow$"，显然

"$\Leftarrow$"，因线性空间 U 与 V 同构，则有 $dimU=dimV$.

设 $(\alpha_1,\cdots,\alpha_n)$ 为 U 的标准正交基，$(\beta_1,\cdots,\beta_n)$ 为 V 的标准正交基.

定义线性映射 σ，$\sigma(\alpha_i)=\beta_i$ □

3 正交补，实内积空间的保距同构

定义 3.1. 正交补

定理 3.1.

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

4 正交变换

定义 4.1. $\mathscr{A}$ 是 *Euclid* 空间 V 上的线性变换，如果

$$(\mathscr{A}\alpha, \mathscr{A}\beta) = (\alpha, \beta)$$

对 $\forall \alpha, \beta \in V$ 成立，则称 $\mathscr{A}$ 为正交变换（*orthogonal transformation*）.

由于向量的长度和向量间的夹角都是利用内积来定义的，自然有正交变换保持长度和角度不变，反过来，我们也有

定理 4.1. 如果 $|\mathscr{A}\alpha| = |\alpha|$ 对 $\forall \alpha \in V$ 成立，那么 $\mathscr{A}$ 是正交变换.

证明.

$$|\mathscr{A}\alpha| = |\alpha|, \forall \alpha \in V$$

$$(\mathscr{A}\alpha, \mathscr{A}\alpha) = (\alpha, \alpha), \forall \alpha \in V$$

$$(\mathscr{A}(\alpha+\beta), \mathscr{A}(\alpha+\beta)) = (\alpha+\beta, \alpha+\beta), \forall \alpha, \beta \in V$$

$$\Rightarrow (\mathscr{A}\alpha, \mathscr{A}\beta) = (\alpha, \beta), \forall \alpha, \beta \in V$$

□

定理 4.2. 下列命题等价：

1. $\mathscr{A}$ 为正交变换
2. $\mathscr{A}$ 将标准正交基变为标准正交基
3. $\mathscr{A}$ 在任意一组标准正交基下的矩阵为正交方阵

证明.

- $(1) \Rightarrow (2)$

  设 $(\alpha_1, \cdots, \alpha_n)$ 为标准正交基

  $(\mathscr{A}\alpha_i, \mathscr{A}\alpha_j) = (\alpha_i, \alpha_j) = \delta_{ij}$

  因此 $(\mathscr{A}\alpha_1, \cdots, \mathscr{A}\alpha_n)$ 也为标准正交基.

- $(2) \Rightarrow (3)$

  设 $(\alpha_1, \cdots, \alpha_n)$ 为标准正交基，$\mathscr{A}$ 在该组基下的矩阵为 A，即

  $$\mathscr{A}(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A$$

  由 (2)，$\mathscr{A}$ 将标准正交基变为标准正交基，则

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

  也是标准正交基.

  综合以上两式，有

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

  那么 A 是两组标准正交基之间的过渡矩阵，由命题 2.4 知，A 为正交方阵.

- $(3) \Rightarrow (1)$

  设 $(\alpha_1, \cdots, \alpha_n)$ 为标准正交基，$\mathscr{A}(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A$，$A$ 为正交矩阵.

  设 $\alpha = (\alpha_1, \cdots, \alpha_n)x, \beta = (\alpha_1, \cdots, \alpha_n)y$,

  则有 $\mathscr{A}\alpha = (\alpha_1, \cdots, \alpha_n)Ax, \mathscr{A}\beta = (\alpha_1, \cdots, \alpha_n)Ay$

  $(\mathscr{A}\alpha, \mathscr{A}\beta) = (Ax)^T Ay = x^T A^T Ay = x^T y = (\alpha, \beta)$

□

推论 4.1. 正交变换的特征值的模长为 1.

命题 4.1. 若 $\mathscr{A}, \mathscr{B}$ 均是 V 上的正交变换，则 $\mathscr{A}\mathscr{B}, \mathscr{A}^{-1}$ 也是 V 上的正交变换.

命题 4.2. 若 ± 1 是正交变换的特征值，则 $V_1 \perp V_{-1}$

证明. 设 $\alpha \in V_1, \beta \in V_{-1}$

则 $\mathscr{A}\alpha = \alpha, \mathscr{A}\beta = -\beta$

$(\alpha, \beta) = (\mathscr{A}\alpha, \mathscr{A}\beta) = (\alpha, -\beta) = -(\alpha, \beta) \Rightarrow (\alpha, \beta) = 0$ □

例 4.1. 在集合平面上任取一点 O 作为原点，将平面上每个点 P 与向量 $\vec{OP}$ 对应起来，将平面看成 *2* 维欧氏空间 V. 设 $\mathscr{A}$ 是 V 上的正交变换，则：

1. 当 $det\mathscr{A} = 1$，$\mathscr{A}$ 是绕原点的旋转，在 V 的任意一组标准正交基下的矩阵 $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$

2. 当 $det\mathscr{A} = -1$，$\mathscr{A}$ 是关于过原点的某条直线 l 的轴对称，在 V 的某一组标准正交基下的矩阵为 $diag(1, -1)$.

例 4.2. 设 $\mathbb{R}^3$ 建立了直角坐标系的几何空间，$\mathscr{A}$ 是 $\mathbb{R}^3$ 上的正交变换，且 $det\mathscr{A} = 1$，求证：$\mathscr{A}$ 是绕过某点的直线的旋转.

证明. $\mathscr{A} \to A, \lambda_1, \lambda_2, \lambda_3$

$|A| = \lambda_1\lambda_2\lambda_3$

$|\lambda I - A| = \lambda^3 + \cdots \Rightarrow$

1. 至少有一实根

2. 若有复根，必共轭出现

□

设 $\mathscr{A}$ 为 Euclid 空间 V 上的正交变换，在标准正交基 $M_1 = \{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵表示为 A，在另一组标准正交基 $M_2 = \{\beta_1, \cdots, \beta_2\}$ 下的矩阵表示为 B，转移矩阵为 $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)P$，其中 P 为正交方阵，那么我们有 $B = P^{-1}AP$.

定义 4.2. 设 A, B 是同阶实方阵，如果存在正交方阵 P 使 $B = P^{-1}AP$，则称 A 与 B 正交相似.

注记. 同一个变换在不同基下的矩阵相似，同一个变换在不同标准正交基下的矩阵正交相似.

引理 4.1. 设 $\mathscr{A}$ 是 *Euclid* 空间 V 上的正交变换，W 是 $\mathscr{A}$ 的不变子空间，则 $W^{\perp}$ 也是 $\mathscr{A}$ 的不变子空间.

证明. 先证 $\mathscr{A}|_W$ 是可逆变换. □

注记. 我们将在后面看到，这对于范围更广的一类变换——规范变换，仍然是对的.

5 对称变换

- 线性变换本身只与线性结构有关，但它是不是对称变换却依赖于内积的选取，但是它本身的与内积无关的性质并不依赖于它是不是一个对称变换，比如特征值均为实数，不同特征值对应的特征向量彼此正交. 因此，只要是在某组基下的矩阵表示是对称矩阵的线性变换，都在某个内积下是对称变换

定义 5.1. $\mathscr{A}$ 是 Euclid 空间 V 上的线性变换，若 $\mathscr{A}$ 满足

$$(\mathscr{A}\alpha,\beta)=(\alpha,\mathscr{A}\beta)$$

对 $\forall\ \alpha,\beta\in V$ 成立，则称 $\mathscr{A}$ 为对称变换.

定理 5.1. $\mathscr{A}$ 为对称变换当且仅当 $\mathscr{A}$ 在 V 的任何一组标准正交基 M 下的矩阵 A 为对称方阵.

证明. $M=\{\alpha_1,\cdots,\alpha_n\}$，$\mathscr{A}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)A$

设 $\alpha=(\alpha_1,\cdots,\alpha_n)x,\beta=(\alpha_1,\cdots,\alpha_n)y$

则 $\mathscr{A}\alpha=(\alpha_1,\cdots,\alpha_n)Ax,\mathscr{A}\beta=(\alpha_1,\cdots,\alpha_n)Ay$

$(\mathscr{A}\alpha,\beta)=(Ax)^Ty=x^TA^Ty,(\alpha,\mathscr{A}\beta)=x^T(Ay)=x^TAy$

$(\mathscr{A}\alpha,\beta)=(\alpha,\mathscr{A}\beta)\Leftrightarrow x^TA^Ty=x^TAy,\forall x,y\in\mathbb{R}^n\Leftrightarrow A^T=A$ □

注记. 若基是一组普通的基，那么有 $A^TG=GA$，其中 G 是内积在这组基下的度量矩阵.

定理 5.2. 设 $\mathscr{A}$ 是 Euclid 空间 V 上的对称变换，那么 $\mathscr{A}$ 的特征值均为实数.

证明. 设 M 是 V 的标准正交基，$\mathscr{A}$ 在 M 下的矩阵为 A.

设 λ 是 A 的特征值，X 是从属于 λ 的一个特征向量.

对 $AX=\lambda X$ 做共轭转置，得到 $\bar{X}^TA=\bar{\lambda}\bar{X}^T$

$$\Rightarrow\begin{cases}\bar{X}^TAX=\lambda\bar{X}^TX\\ \bar{X}^TAX=\bar{\lambda}\bar{X}^TX\end{cases}$$

$\Rightarrow\lambda=\bar{\lambda}$ □

定理 5.3. 对称变换 $\mathscr{A}$ 的不同特征值对应的特征向量彼此正交.

证明. $\mathscr{A}\alpha_1=\lambda_1\alpha_1,\mathscr{A}\alpha_2=\lambda_2\alpha_2,\lambda_1\neq\lambda_2$

$(\mathscr{A}\alpha_1,\alpha_2)=(\alpha_1,\mathscr{A}\alpha_2)\Rightarrow\lambda_1(\alpha_1,\alpha_2)=\lambda_2(\alpha_1,\alpha_2)\Rightarrow(\alpha_1,\alpha_2)=0$ □

注记. 而在从属于同一个特征值的全体特征向量中，即在该特征值对应的特征子空间中，我们总能运用 *Gram-Schmidt* 正交化方法找到一组正交基. 想到这一点，有下面的定理便是自然的.

定理 5.4. $\mathscr{A}$ 是 *Euclid* 空间 V 上的对称变换，则存在 V 的一组标准正交基 M，使得 $\mathscr{A}$ 在该组基下的矩阵是对角阵.

证明. 这等价于要找到一组标准正交基由特征向量组成.

我们证明对任意特征值不存在二次根向量，即要证明，如若 $(A-\lambda I)^2x=0$，那么一定有 $(A-\lambda I)x=0$.

对于对称阵 A，显然有 $A-\lambda I$ 仍是对称阵，所以我们只需证明 $\lambda=0$ 的情况.

考虑 Ax 与自身的内积，

$$(Ax)^T Ax = x^T A^T Ax = x^T AAx = x^T(A^2 x) = 0$$

因此有 $Ax = 0$.

我们已经知道 V 可以被分解为根子空间的直和，上面证明了，根子空间就是特征子空间. 对于单个的特征子空间，我们总能对它的一组基进行 Gram-Schimdt 正交化而得到一组标准正交基，这些基拼起来显然构成 V 的一组基. 这组基为什么还是正交的呢？因为我们知道有对称方阵的不同特征值对应的特征向量彼此正交. □

定理 5.4′. 实对称方阵正交相似于对角阵，对角元就是它的全部特征值.

证明. 用归纳法，$n = 1$，显然

设该结论对 $n-1$ 阶方阵成立，下证结论对 n 阶方阵成立.

设 λ_1 为 A 的特征值，X_1 为对应的单位特征向量

将 X_1 扩充为 $\mathbb{R}^n$ 的标准正交基 $X_1, X_2, \cdots, X_n$

$A(X_1, \cdots, X_n) = (X_1, \cdots, X_n)\begin{pmatrix} \lambda_1 & C \\ 0 & A_1 \end{pmatrix}, C \in \mathbb{R}^{n-1}, A_1 \in \mathbb{R}^{n-1\times n-1}$

令 $P_1 = (X_1, \cdots, X_n)$，$P_1^{-1}AP_1 = P_1^T AP_1 = \begin{pmatrix} \lambda_1 & C \\ 0 & A_1 \end{pmatrix}$ 实对称，所以 $C = 0, A_1$ 也是对称阵.

$\Rightarrow P_1^T AP_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix}$

由归纳假设，存在正交阵 P_2，使得 $P_2^T A_1 P_2 = diag\{\lambda_2, \cdots, \lambda_n\}$

$\left(P_1 \begin{pmatrix} 1 & \\ & P_2 \end{pmatrix}\right)^T A \left(P_1 \begin{pmatrix} 1 & \\ & P_2 \end{pmatrix}\right) = diag\{\lambda_1, \cdots, \lambda_n\}$ □

例 5.1. $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$，求正交方阵 P 使得 $P^{-1}AP$ 为对角阵.

解. $P_A(\lambda) = (\lambda - 5)(\lambda + 1)^2$

$\lambda_1 = 5$ 对应的特征向量 $X_1 = (1, 1, 1)^T$

$\lambda_2 = \lambda_2$ 对应的特征向量 $X_2 = (-1, 1, 0)^T, X_3 = (-1, 0, 1)^T$

将 X_1 单位化，将 X_2, X_3 单位正交化. □

6 Euclid 空间上线性函数

V 是线性空间,V 上的线性函数 $f: V \to F$,$f(\lambda\alpha+\mu\beta) = \lambda f(\alpha)+\mu f(\beta)$. $V^* = \{f|f\text{是}V\text{上的线性函数}\}$,$V^*$ 构成线性空间 V 的对偶空间.$M = \{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一组基，$\sigma: V^* \to V, f \mapsto \sum_{i=1}^{n} f(\alpha_i)x_i$，同构. 令 $f_i := \alpha = x_1\alpha_1 + \cdots + x_n\alpha_n \mapsto x_i, f_i \in V^*, \{f_1, \cdots, f_n\}$ 是 V^* 的一组基，$f_i(\alpha_j) = \delta_{ij}$，称 $\{f_1, \cdots, f_n\}$ 为 $\{\alpha_1, \cdots, \alpha_n\}$ 的对偶基.

V 是 Euclid 空间，固定一个 α，$f_\alpha: \beta \mapsto (\alpha, \beta), f_\alpha \in V^*$，固定 β，$f_\beta: \alpha \mapsto (\alpha, \beta), f_\beta \in V^*$

定理 6.1. *V 是 n 维 Euclid 空间，则*

1. *$f_\alpha: \beta \mapsto (\alpha, \beta)$ 是 V 上的线性函数*
2. *$\sigma: \alpha \mapsto f_\alpha$ 是 V 到 V^* 的同构映射.*
3. *$\forall f \in V^*, \exists \alpha \in V$ 使 $f_\alpha = f$*
4. *设 $\alpha_1, \cdots, \alpha_n$ 为 V 的标准正交基，则 $f_{\alpha_1}, \cdots, f_{\alpha_n}$ 是 $\{\alpha, \cdots, \alpha\}$ 的对偶基.*

证明. 1. 略

2. $\sigma: \alpha \mapsto f_\alpha, V \to V^*$ 同构.

   注记. *这个同构是线性空间之间的同构，而不是 Euclid 空间之间的同构，因为我们没有在 V^* 中定义内积.*

   •
   $$f_{C_1\alpha_1+C_2\alpha_2}(\beta) = (C_1\alpha_1 + C_2\alpha_2, \beta) = C_1(\alpha_1, \beta) + C_2(\alpha_2, \beta) = C_1 f_{\alpha_1}(\beta) + C_2 f_{\alpha_2}(\beta)$$
   $$\Rightarrow f_{C_1\alpha_1+C_2\alpha_2} = C_1 f_{\alpha_1} + C_2 f_{\alpha_2}$$

   •
   $$\alpha \in Ker\sigma \Leftrightarrow f_\alpha = 0 \Leftrightarrow \forall \beta \in V, f_\alpha(\beta) = (\alpha, \beta) = 0 \Leftrightarrow \alpha = 0$$

3. 由 (2) 显然

4. $f_{\alpha_i}(\alpha_j) = (\alpha_i, \alpha_j) = \delta_{ij}$

□

推论 6.1. *$\mathscr{A}$ 是 Euclid 空间 V 上的线性变换，则存在唯一的变换 $\mathscr{A}^*$ 使得 $(\mathscr{A}\alpha, \beta) = (\alpha, \mathscr{A}^*\beta)$ 对 $\forall \alpha, \beta \in V$ 成立.*

证明.
$$f_\beta: \alpha \mapsto (\mathscr{A}\alpha, \beta), f_\beta \in V^*$$
$\exists \tilde{\beta}$ 使 $f_\beta(\alpha) = (\alpha, \tilde{\beta})$
$$\beta \mapsto \tilde{\beta} = \mathscr{A}^*\beta$$

□

7 酉空间

- 复线性空间上的内积
-

定义 7.1. 设 V 是 $\mathbb{C}$ 上线性空间，在 V 上定义内积 $(\cdot,\cdot)$ 满足

1. $(\alpha,\beta)=\overline{(\beta,\alpha)}$
2. 共轭线性性
3. 正定性：$(\alpha,\alpha)\geqslant 0$，等号成立当且仅当 $\alpha=0$

称定义了上述内积的空间 V 为酉空间.

注记. 保持了正定性，损失了一定程度的线性性和对称性

例 7.1. $\mathbb{C}^n:(x,y)=\bar{x}_1y_1+\cdots+\bar{x}_ny_n$

例 7.2. $C[a,b]:$

$$(f(x),g(x))=\int_a^b\overline{f(x)}g(x)dx$$

命题 7.1. 内积的性质：

1. $(\alpha,0)=(0,\beta)=0$
2. $(\sum_{i=1}^m\lambda_i\alpha_i,\sum_{j=1}^n\mu_j\beta_j)=\sum_{i=1}^m\sum_{j=1}^n\bar{\lambda}_i\mu_j(\alpha_i,\beta_j)$
3. (Cauchy-Schwarz 不等式)$|(\alpha,\beta)|^2\leqslant(\alpha,\alpha)(\beta,\beta)$，等号成立 $\Leftrightarrow\alpha,\beta$ 线性相关.
4. $|\alpha+\beta|\leqslant|\alpha|+|\beta|$

定义 7.2. $|\alpha|=\sqrt{(\alpha,\alpha)}$，如果 $|\alpha|=1$，称 α 为单位向量.
如果 $(\alpha,\beta)=0$，称 α 与 β 正交.

注记. 没有定义夹角！定义不了夹角！

设 $\{\alpha,\cdots,\alpha_n\}$ 为 V 的一组基，$\alpha=(\alpha_1,\alpha_n)x,\beta=(\alpha,\cdots,\alpha_n)y$

$$(\alpha,\beta)=\sum_{i=1}^n\sum_{j=1}^n\bar{x}_iy_j(\alpha_i,\alpha_j)=\bar{x}^TGy,G=$$

注记.
1. $\bar{G}^T=G$，即 $G^*=G$，Hermite 方阵.
2. $\forall x\in\mathbb{C}^n,x^*Gx\geqslant 0$，等号成立 $\Leftrightarrow x=0$
   G 称为正定方阵.

定理 7.1. V 是酉空间，内积在两组基 $\{\alpha_1,\cdots,\alpha_n\}$ 及 $\{\beta_1,\cdots,\beta_n\}$ 下的度量矩阵分别是 G 与 G'，$(\beta_1,\cdots.\beta_n)=(\alpha,\cdots,\alpha_n)P$，则 $G'=P^*GP$

证明. 略. □

定义 7.3. $A,B\in\mathbb{C}^{n\times n}$，如存在可逆方阵 P 使 $B=P^*AP$，则称 A 与 B 是共轭相合.

定义 7.4. V 是 n 维酉空间 $\{\alpha_1,\cdots,\alpha_n\}$ 称为 V 的标准正交基，如果 $(\alpha_i,\alpha_j)=\delta_{ij}$

定理 7.2. n 维酉空间存在标准正交基.

证明. 略. □

$\{\beta_1,\cdots,\beta_n\}=\{\alpha_1,\cdots,\alpha_n\}P$，$\{\alpha_1,\cdots,\alpha_n\},\{\beta_1,\cdots,\beta_n\}$ 均为标准正交基，$G'=G=I,I=P^*IP\Rightarrow P^*P=I$，正交方阵.

定理 7.3. 两组标准正交基之间的过渡方阵为酉方阵，反之，从一组标准正交基到另一组基的过渡矩阵为酉方阵，则另一组基也为标准正交基.

注记. *1.* U_1,U_2 为酉方阵，则 U_1U_2 是酉方阵.

2. U 是酉方阵，则 $U^{-1}=U^*$ 也是酉方阵.

3. U 的行（列）均构成 $\mathbb{C}^n$ 的标准正交基.

例 7.3. $A\in\mathbb{C}^n$，A 可逆，则 A 可分解为 $A=UT$，U 为酉方阵，T 为上三角，并且对角元为正.

定理 7.4. W 是酉空间 V 的子空间，则 $V=W\oplus W^{\perp}$

8 规范变换

- 伴随变换
  - $(\mathscr{A}\alpha, \beta) = (\alpha, \mathscr{A}^*\beta)$
  - 有限维复（实）内积空间伴随变换一定存在，无限维不一定存在
  - 存在则唯一
  - 如果 $\mathscr{A}$ 在一组标准正交基下的矩阵是 A，那么 $\mathscr{A}^*$ 在该组基下的矩阵是 A^*
  - 实内积空间中，A^* 自动退化为 A^T，因此可统一记忆为 A^*
  - 如果 W 是 $\mathscr{A}$ 的不变子空间，那么 $W^\perp$ 是 $\mathscr{A}^*$ 的不变子空间
  - 有一个和上一条长得很像的命题：当 $\mathscr{A}$ 是**规范变换**时，如果 W 是 $\mathscr{A}$ 的不变子空间，那么 $W^\perp$ 也是 $\mathscr{A}$ 的不变子空间
  - 如果 $\mathscr{A}$ 可逆，并且 $\mathscr{A}^{-1}$ 也有伴随变换，那么 $\mathscr{A}^*$ 也可逆，且 $(\mathscr{A}^*)^{-1} = (\mathscr{A}^{-1})^*$
  - 如果 $\mathscr{A}$ 有伴随变换，那么 $\mathscr{A}$ 可以唯一地表示成 $\mathscr{A} = \mathscr{A}_1 + \mathrm{i}\mathscr{A}_2$，其中 $\mathscr{A}_1, \mathscr{A}_2$ 都是 Hermite 变换
- 规范变换
  - 存在伴随变换并且 $\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A}$
  - 复内积空间：存在标准正交基使得 $\mathscr{A}$ 在此基下的矩阵为对角矩阵，并且对角元是特征值
  - $A = U^* \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) U$

上面关于规范变换的所有性质中最重要的一条毫无疑问是存在一组标准正交基使得 $\mathscr{A}$ 在此基下的矩阵为对角阵并且对角元是特征值. 我们将用三种不同的方法来证明这一结论，在不同的证明方法中逐条揭露规范变换的其他重要性质.

8.1 规范矩阵酉相似于对角阵的三种证明

证法一

引理 8.1. 设 $\mathscr{A}$ 是复（实）内积空间 V 上的规范变换，则

$$\|\mathscr{A}\alpha\| = \|\mathscr{A}^*\alpha\|, \quad \forall \alpha \in V.$$

证明.

$$\|\mathscr{A}\alpha\|^2 = (\mathscr{A}\alpha, \mathscr{A}\alpha) = (\alpha, \mathscr{A}^*\mathscr{A}\alpha) = (\alpha, \mathscr{A}\mathscr{A}^*\alpha)$$
$$\|\mathscr{A}^*\alpha\|^2 = (\mathscr{A}^*\alpha, \mathscr{A}^*\alpha) = (\alpha, (\mathscr{A}^*)^*\mathscr{A}^*\alpha) = (\alpha, \mathscr{A}\mathscr{A}^*\alpha)$$

□

事实上（虽然在我们的证明中没有用，但为了理论的完整还是在这里给出），我们还有：

命题 8.1. 若 $\|\mathscr{A}\alpha\| = \|\mathscr{A}^*\alpha\|$ 对任意 $\alpha \in V$ 成立，那么 $\mathscr{A}$ 是规范变换.

证明. 此时我们无法从

$$(\alpha, \mathscr{A}^*\mathscr{A}\alpha) = (\alpha, \mathscr{A}\mathscr{A}^*\alpha), \quad \forall \alpha \in V$$

直接推出 $\mathscr{A}^*\mathscr{A} = \mathscr{A}\mathscr{A}^*$. 但按照我们之前的证明的经验，如果有

$$(\alpha, \mathscr{A}^*\mathscr{A}\beta) = (\alpha, \mathscr{A}\mathscr{A}^*\beta), \quad \forall \alpha, \beta \in V$$

成立就好了.

$$(\mathscr{A}(\alpha+\beta), \mathscr{A}(\alpha+\beta)) = (\mathscr{A}^*(\alpha+\beta), \mathscr{A}^*(\alpha+\beta))$$
$$(\mathscr{A}\alpha, \mathscr{A}\beta) = (\mathscr{A}^*\alpha, \mathscr{A}^*\beta)$$
$$(\alpha, \mathscr{A}^*\mathscr{A}\beta) = (\alpha, \mathscr{A}\mathscr{A}^*\beta)$$

□

注记. $\|\mathscr{A}\alpha\| = \|\mathscr{A}^*\alpha\|$ 最好用的是从 $\mathscr{A}\alpha = 0$ 推出 $\mathscr{A}^*\alpha = 0$. 我们还得到了一个推论：$\operatorname{Ker}\mathscr{A} = \operatorname{Ker}\mathscr{A}^*$.

引理 8.2. 设 $\mathscr{A}$ 是复（实）内积空间 V 上的正规变换，c 是任一复（实）数，则 $c\mathscr{I} - \mathscr{A}$ 也是 V 上的正规变换.

证明. 注意到 $(c\mathscr{I} - \mathscr{A})^* = \bar{c}\mathscr{I} - \mathscr{A}^*$，直接验证即可. □

定理 8.1. 设 $\mathscr{A}$ 是复（实）内积空间 V 上的正规变换，则

- λ 是 $\mathscr{A}$ 的一个特征值当且仅当 $\bar{\lambda}$ 是 $\mathscr{A}^*$ 的一个特征值；
- α 是 $\mathscr{A}$ 的属于特征值 λ 的一个特征向量当且仅当 α 是 $\mathscr{A}^*$ 的属于特征值 $\bar{\lambda}$ 的一个特征向量.

同样地，为了理论的完整，虽然在我们的证明中没用，我们依旧给出：

命题 8.2. 设 $\mathscr{A}$ 是 n 维酉空间上的线性变换，若 $\mathscr{A}$ 的每个特征向量也是 $\mathscr{A}^*$ 的特征向量，则 $\mathscr{A}$ 是规范变换.

证明一. 注意到 $\mathscr{A}\mathscr{A}^H$ 和 $\mathscr{A}^H\mathscr{A}$ 都是 Hermite 阵, 故它们半单, 即可酉相似对角化. 如果 $\mathscr{A}v=\lambda v$, 由题意, 存在 $\exists\mu\in\mathbb{C}$, 使 $\mathscr{A}^Hv=\mu v$. 因此,$v^H\lambda v=v^H\mathscr{A}v=\overline{v^H\mathscr{A}^Hv}=\overline{v^H\mu v}$, 即 $\lambda|v|^2=\overline{\mu}|v|^2$, 因此 $\lambda=\overline{\mu}$. 如果 $\exists v\neq 0$,$(\mathscr{A}-\lambda\mathscr{I})^{k-1}v\neq 0$,$(\mathscr{A}-\lambda\mathscr{I})^kv=0$,$k\geqslant 2[/imath]$, 则 $(\mathscr{A}-\lambda\mathscr{I})^H(\mathscr{A}-\lambda\mathscr{I})^{k-1}v=0$, 故 $|(\mathscr{A}-\lambda\mathscr{I})^{k-1}v|^2=v^H(\mathscr{A}^H-\overline{\lambda}\mathscr{I})^{k-1}(\mathscr{A}-\lambda\mathscr{I})^{k-1}v=0$, 矛盾! 故 $\mathscr{A}$ 只有单特征值, 即 $\mathscr{A}$ 可酉相似对角化. 因此,$\mathscr{A}$ 的所有特征向量全体能够张成 $\mathbb{C}^n$. 对于 $\mathscr{A}v=\lambda v$,$(\mathscr{A}\mathscr{A}^H-\mathscr{A}^H\mathscr{A})v=\mathscr{A}\overline{\lambda}v-\lambda\mathscr{A}^Hv=|\lambda|^2v-|\lambda|^2v=0$, 故 $\mathscr{A}\mathscr{A}^H=\mathscr{A}^H\mathscr{A}$. □ 核心想法: 只要证明 $\mathscr{A}$ 半单即可, 这样能够利用题目所给的特征向量条件, 说明 $\mathscr{A}\mathscr{A}^H-\mathscr{A}^H\mathscr{A}$ 作用到所有特征向量上为 0, 进而得证. □

注记. 要特别感谢吴天学长给出的上面的证明，正是模仿这个证明，我才完成了规范方阵可酉相似于对角矩阵的第二种证明.

定理 8.2. 设 $\mathscr{A}$ 是复（实）内积空间 V 上的正规变换，如果 W 是 $\mathscr{A}$ 的不变子空间，则 $W^\perp$ 也是 $\mathscr{A}$ 的不变子空间.

定理 8.3. 设 $\mathscr{A}$ 是有限维酉空间 V 上的正规变换，则 V 中存在一个标准正交基，使得 $\mathscr{A}$ 在此基下的矩阵是对角矩阵.

证明. □

推论 8.1. 设 $\mathscr{A}$ 是 n 维酉空间 V 上的正规变换，那么：

(1) $\mathscr{A}$ 是酉变换当且仅当 $\mathscr{A}$ 的特征值的模为 1;

(2) $\mathscr{A}$ 是 *Hermite* 变换当且仅当 $\mathscr{A}$ 的特征值都是实数;

(3) $\mathscr{A}$ 是斜 *Hermite* 变换当且仅当 $\mathscr{A}$ 的特征值是 0 或纯虚数.

证明. 由于 $\mathscr{A}$ 是 V 上的正规变换，因此 V 中存在一个标准正交基，使得 $\mathscr{A}$ 在此基下的矩阵 A 为对角矩阵 $\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\mathscr{A}$ 的全部特征值.

(1)

$$\begin{aligned}\mathscr{A}\text{是酉变换} &\iff A\text{是酉矩阵}\\ &\iff A^*=A^{-1}\\ &\iff \bar{\lambda}_i=\lambda_i^{-1}\\ &\iff |\lambda_i|=1\end{aligned}$$

(2)

$$\begin{aligned}\mathscr{A}\text{是 Hermite 变换} &\iff A\text{是 Hermite 矩阵}\\ &\iff A^*=A\\ &\iff \bar{\lambda}_i=\lambda_i\\ &\iff \lambda_i\text{是实数}\end{aligned}$$

(3)

$$\mathscr{A}\text{是斜 Hermite 变换} \iff A\text{是斜 Hermite 矩阵}$$

$$\begin{aligned}&\Longleftrightarrow A^* = -A\\&\Longleftrightarrow \bar{\lambda}_i = -\lambda_i\\&\Longleftrightarrow \lambda_i\text{是 0 或纯虚数}\end{aligned}$$

□

证法二

命题 8.3. 如果 $\mathscr{A}$ 是复（实）内积空间 V 上的规范变换，那么 $\mathscr{A}$ 的属于不同特征值的特征向量彼此正交.

证明.

$$\lambda_1(\xi_1,\xi_2)=(\lambda_1\xi_1,\xi_2)=(\mathscr{A}\xi_1,\xi_2)=(\xi_1,\mathscr{A}^*\xi_2)=(\xi_1,\bar{\lambda}_2\xi_2)=\lambda_2(\xi_1,\xi_2)$$

□

命题 8.4. 内容...

证法三

定理 8.4. 设 $\mathscr{A}$ 是 n 维酉空间 V 上的线性变换，那么 V 中存在一组标准正交基，使得 $\mathscr{A}$ 在此基下的矩阵 A 是上三角矩阵.

定理 8.5. 任意复方阵酉相似于上三角阵.

注记. 过去我们曾经证明过任意复方阵相似于上三角阵，这里我们得到的结论更强. 而事实上两者的证明完全类似，过去我们是找特征向量，扩充为一组基，现在我们是找单位特征向量，扩充为一组标准正交基.

引理 8.3. 与规范方阵酉相似的方阵也是规范方阵.

证明. A 是规范的，$B=U^{-1}AU$，其中 U 为酉方阵.

$$BB^*=U^*AUU^*A^*U=U^*AA^*U=U^*A^*AU=U^*A^*UU^*AU=B^*B$$

□

引理 8.4. 设复方阵 T 为上三角阵，则 T 为规范方阵当且仅当 T 为对角阵.

证明. $\Leftarrow$ 显然.

$\Rightarrow$

$$T=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ 0 & 0 & \lambda_n\end{pmatrix},T^*=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ 0 & 0 & \lambda_n\end{pmatrix},TT^*=\begin{pmatrix}|\lambda_1|^2+\sum\limits_{j=2}^{n}|t_{1j}|^2 & & * \\ 0 & \ddots & \\ & & |\lambda_n|^2\end{pmatrix}$$

$$T^*T=\begin{pmatrix}|\lambda_1|^2+\sum\limits_{j=2}^{n}|t_{1j}|^2 & & * \\ 0 & \ddots & \\ & & |\lambda_n|^2\end{pmatrix}$$

$$t_{1j}=0,$$

□

定理 8.6. $A\in\mathbb{C}^{n\times n}$，则 A 为规范方阵当且仅当 A 酉相似于对角阵.

证明. $\Leftarrow$ 对角阵为规范方阵，由引理8.3，A 规范.

$\Rightarrow$ 由定理8.5，A 可酉相似于上三角阵 T，由引理8.3，T 是规范方阵，再由引理8.4，T 为上三角阵.

□

8.2 实规范矩阵的正交相似

引理 8.5. 与实规范方阵正交相似的方阵也是实规范方阵.

证明. $B = O^TAO, B^T = O^TA^TO$

$B^TB = O^TA^TOO^TAO = O^TA^TAO$

$BB^T = O^TAOO^TA^TO = O^TAA^TO$ □

引理 8.6. 实方阵 $\begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$ 是规范的 $\Leftrightarrow A_2 = 0$，且 A_1, A_3 规范.

引理 8.7.

$$A^TA = \begin{pmatrix} A_1^T & 0 \\ A_2^T & A_3^T \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} = \begin{pmatrix} A_1^TA_1 & A_1^TA_2 \\ A_2^TA_1 & A_2^TA_2 + A_3^TA_3 \end{pmatrix}$$

$$AA^T = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \begin{pmatrix} A_1^T & 0 \\ A_2^T & A_3^T \end{pmatrix} = \begin{pmatrix} A_1A_1^T + A_2A_2^T & A_2A_3^T \\ A_3A_2^T & A_3A_3^T \end{pmatrix}$$

$$\begin{aligned} & A_1^TA_1 = A_1A_1^T + A_2A_2^T \\ & \Rightarrow Tr(A_1^TA_1) = Tr(A_1A_1^T) + Tr(A_2A_2^T) \\ & \Rightarrow Tr(A_2A_2^T) = 0 \\ & \Rightarrow A_2 = 0 \\ & \Rightarrow A_1^TA_1 = A_1A_1^T, A_3^TA_3 = A_3A_3^T \end{aligned}$$

推论 8.2. $\mathscr{A}$ 是 *Euclid* 空间 V 上的规范变换，W 是 $\mathscr{A}$ 的不变子空间，则 $W^\perp$ 也是 $\mathscr{A}$ 的不变子空间.

引理 8.8. 二阶实规范方阵正交相似于对角阵 $\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$ 或 $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, 其中 $b \neq 0$.

证明. $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，$AA^T = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}$，$A^TA = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix}$

$ac + bd = ab + cd \Rightarrow a(c-b) = d(c-b)$

1. $b = c$

   $$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \sim \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$$

2. $b \neq c$

   $$\Rightarrow a = d, b = -c \Rightarrow A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

□

定理 8.7. 设实规范方阵 A 的特征值为 $a_1 \pm b_1 i, \cdots, a_s \pm b_s i (b_j > 0, a_j \in \mathbb{R})$ 及 $\lambda_{2s+1}, \cdots, \lambda_n (\lambda_j \in \mathbb{R})$，则 A 正交相似于

$$D = diag\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{2s+1}, \cdots, \lambda_n\right)$$

证明. 归纳法，$n=1$ 时，成立

$n=2$ 时，由引理 3，成立

设结论对小于 n 阶规范方阵成立，下证结论对 n 阶规范方阵成立

1. A 有实特征值 λ_n，对应单位特征向量 X_n,$AX_n=\lambda_n X_n$，将 X_n 扩充为一组标准正交基 $M=\{X_1,\cdots,X_n\}$.
$$A(X_1\cdots X_n)=(X_1\cdots X_n)\begin{pmatrix}A_1 & 0\\ C & \lambda_n\end{pmatrix}$$
$P_1^TAP_1=\begin{pmatrix}A_1 & 0\\ C & \lambda_n\end{pmatrix}$ 规范 $\Rightarrow C=0,A_1$ 规范，由归纳假设，存在正交方阵 P_2 使得 $P_2^TA_1P_2$ 成为所要求的形状，得证.

2. A 无实特征值，设 $a_1\pm b_i$ 为 A 的特征值 $(b_1>0)$，$X_1\pm iX_2$ 是对应特征向量. 将 X_1,X_2 正交化为 β_1,β_2，将 β_1,β_2 扩充为 $\mathbb{R}^n$ 的标准正交基 $\beta_1,\cdots,\beta_n$.
$A(\beta_1\cdots\beta_n)=(\beta_1\cdots\beta_n)\begin{pmatrix}A_1 & A_2\\ 0 & A_3\end{pmatrix}$，$A_1$ 为 2 阶，特征值 $a_1\pm b_1i$
$P_1^TAP_1=\begin{pmatrix}A_1 & A_2\\ 0 & A_3\end{pmatrix}$ 规范 $\Rightarrow A_2=0,A_1,A_3$ 规范,存在正交方阵 P_2 使得 $P_2^TA_1P_2=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}$,
由归纳假设，存在正交方阵 P_3 使得 $P_3^TA_3P_3$ 为标准形, 得证

□

推论 8.3. $\mathscr{A}$ 是 V 上规范变换，$V=W_1\oplus\cdots\oplus W_s\oplus W_{2s+1}\oplus\cdots\oplus W_n$，$W_j$ 是 $\mathscr{A}$ 的不变子空间，$dimW_j=1$ 或 2.

例 8.1. 证明：正交方阵可以分解为两个对称的正交方阵之积.

证明. $\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}-\sin\theta & \cos\theta\\ \cos\theta & \sin\theta\end{pmatrix}$ □

例 8.2. $\mathscr{A}$ 是 V 上的规范变换，证明 $(Im\mathscr{A})^\perp=Ker\mathscr{A}$

证明. 设 $\beta\in(Im\mathscr{A})^\perp$，那么 $\forall\alpha\in V$，$(\mathscr{A}\alpha,\beta)=0\Leftrightarrow\forall\alpha\in V,(\alpha,\mathscr{A}^*\beta)=0$
$\Rightarrow\mathscr{A}^*\beta=0\Leftrightarrow\mathscr{A}\beta=0\Leftrightarrow\beta\in Ker\mathscr{A}$. □

定理 8.8 (实方阵的正交相似). $A\in\mathbb{R}^{n\times n}$，$A$ 的特征值为 $a_1\pm b_1i,\cdots,a_s\pm b_si,\lambda_{2s+1},\cdots,\lambda_n$，则

A 正交相似于准上三角阵 $\begin{pmatrix}A_1 & * & * & * & * & *\\ & \ddots & * & * & * & *\\ & & A_s & * & * & *\\ & & & \lambda_{2s+1} & * & *\\ & & & & \ddots & *\\ & & & & & \lambda_n\end{pmatrix}$

其中 $A_j\sim\begin{pmatrix}a_j & b_j\\ -b_j & a_j\end{pmatrix}$

证明. 归纳法

$n=1$ 时，显然.

$n=2$ 时，$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

1. A 有两个实特征根 λ_1,λ_2，对应 λ_1 的单位特征向量为 X_1，将 X_1 扩充为 $\mathbb{R}^2$ 的一组标准正交基 X_1,X_2，$A(X_1,X_2)=(X_1,X_2)\begin{pmatrix} \lambda_1 & * \\ 0 & \lambda_2 \end{pmatrix}$

   注记. 关于右下角为什么是 λ_2 的解释：

   - $A\sim\begin{pmatrix} \lambda_1 & * \\ 0 & * \end{pmatrix}$，而相似矩阵有相同的特征值，这就要求右下角的元素必须是 λ_2.
   - 我们还可以这样看这个问题，设 λ_2 对应的特征向量为 X_2，则 $\{X_1,X_2\}$ 是一组基，对这组基运用 *Gram-Schimdt* 正交化方法，得到

     $$\gamma_2=X_2-(X_2,X_1)X_1$$

     将 A 作用于上式的两侧，得到

     $$\begin{aligned} A\gamma_2 &= AX_2-(X_2,X_1)AX_1 \\ &= \lambda_2X_2-\lambda_1(X_2,X_1)X_1 \end{aligned}$$

     显然若是要将右侧写为 γ_2 与 X_1 的线性组合，γ_2 前面的系数必须为 λ_2.

2. A 有一对共轭虚特征值 $a\pm bi$，由实方阵的实相似理论得到

$$A\sim\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

下面假设定理对阶数小于 n 阶的方阵成立. 下面证明定理对 n 成立.

1. A 有实特征值 λ_n，X_n 为对应单位特征向量，将 X_n 扩充为 $\mathbb{R}^n$ 的标准正交基 $\{X_1,\cdots,X_n\}$，

$$A(X_1,\cdots,X_n)=(X_1,\cdots,X_n)\begin{pmatrix} \tilde{A}_1 & * \\ 0 & \lambda_n \end{pmatrix}$$

由归纳假设，存在正交阵 P_2，

$$P_2^T\tilde{A}_1P_2=\begin{pmatrix} A_1 & & & & & \\ & \ddots & & * & & \\ & & A_s & & & \\ & & & \lambda_{2s+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_{n-1} \end{pmatrix}$$

$$
\begin{pmatrix} P_2 & \\ & 1 \end{pmatrix}^T P_1^T A P_1 \begin{pmatrix} P_2 & \\ & 1 \end{pmatrix} = \begin{pmatrix} A_1 & & & & & \\ & \ddots & & * & & \\ & & A_s & & & \\ & & & \lambda_{2s+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_n \end{pmatrix}
$$

2. A 没有实特征值，设 $a_1 \pm b_1$ 为 A 的特征值，$X_1 \pm iX_2$ 为对应的特征向量，将 $\{X_1, X_2\}$ 正交化为 Y_1, Y_2，将它扩充为 $\mathbb{R}^n$ 的标准正交基 $\{Y_1, \cdots, Y_n\}$，

$$
A(Y_1, \cdots, Y_n) = (Y_1, \cdots, Y_n) \begin{pmatrix} A_1 & * \\ O & \tilde{A} \end{pmatrix}
$$

感觉这里和下面这个题目多少有点关系，抄录于此. □

例 8.3. 设 $X_1 + \mathrm{i}X_2$ 是 n 阶实方阵 A 的属于虚特征值 $a + b\mathrm{i}$ 的特征向量，其中 $X_1, X_2 \in \mathbb{R}^n$. 证明 X_1, X_2 生成的子空间 W 是 $\mathbb{R}^n$ 的线性变换 $\mathscr{A} : X \mapsto AX$ 的 2 维不变子空间，并求出 $\mathscr{A}|_W$ 在基 $\{X_1, X_2\}$ 下的矩阵.

定理 8.9 (Schur 定理)**.** $A \in \mathbb{R}^{n\times n}$，$\lambda_1, \cdots, \lambda_n$ 是 A 的特征值，则

$$
Tr(AA^T) \leqslant \sum_{j=1}^{n} |\lambda_j|^2
$$

且等式成立当且仅当 A 为规范方阵.

证明. 存在正交阵 P 使得 $A = P^T DP$，

$$
Tr(AA^T) = Tr(P^T DPP^T D^T P) = Tr(P^T DD^T P) = Tr(PP^T DD^T) = Tr(DD^T)
$$

$$
\begin{aligned}
Tr(DD^T) &= 2\sum_{j=1}^{s}(a_j^2 + b_j^2) + \sum_{j=2s+1}^{n} \lambda_j^2 + \sigma \\
&= \sum_{j=1}^{n} |\lambda_j|^2 + \sigma \\
&\geqslant \sum_{j=1}^{n} |\lambda_j|^2
\end{aligned}
$$

等号成立 $\Leftrightarrow \sigma = 0 \Leftrightarrow$ 非准对角元为零 $\Leftrightarrow A$ 为规范方阵 □

例 8.4. 设 A, B 为 n 阶实规范方阵，且 AB 也为规范方阵，证明：BA 也为规范方阵.

解.

$$
\begin{aligned}
Tr(BA(BA)^T) &= Tr(AB(AB)^T) \\
&= \sum_{j=1}^{n} |\lambda_j(AB)|^2
\end{aligned}
$$

$$= \sum_{j=1}^{n} |\lambda_j(BA)|^2$$

$\Rightarrow BA$是规范的

□

8.3 正规矩阵酉相似于对角阵的应用

定理 8.10 (Schur 不等式)**.** 设 $A \in \mathbb{C}^{n\times n}$，$\lambda_1, \lambda_2, \cdots \lambda_n$ 是 A 的全部特征值，证明：

$$\mathrm{Tr}(A^*A) = \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 \geqslant \sum_{i=1}^{n}|\lambda_i|^2,$$

其中等号成立当且仅当 A 为规范方阵.

证明. $A = U^{-1}\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_n\end{pmatrix}U$，$U$ 为酉方阵

$$AA^* = U^{-1}\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_n\end{pmatrix}UU^*\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_n\end{pmatrix}U = U^{-1}\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_n\end{pmatrix}U$$

$$Tr(AA^*) = \sum_{i=1}^{n}|\lambda_i|^2 + \sigma \geqslant \sum_{i=1}^{n}|\lambda_i|^2$$

σ 是上三角阵所有非对角元模长平方和. □

例 8.5. 证明：规范方阵不同特征值对应的特征向量彼此正交.

证明. $\lambda_1, \cdots, \lambda_s$ 为 A 的不同特征值，重数为 $n_1, \cdots, n_s$，

$$A = U^{-1}\begin{pmatrix}\lambda_1 I^{(n_1)} & & \\ & \ddots & \\ & & \lambda_s I^{(n_s)}\end{pmatrix}$$

$U = (\alpha_1, \cdots, \alpha_n)$ 为酉方阵

$$A(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)\begin{pmatrix}\lambda_1 I^{(n_1)} & & \\ & \ddots & \\ & & \lambda_s I^{(n_s)}\end{pmatrix}$$

□

例 8.6. 给定 Hermite 阵 $H = \begin{pmatrix}0 & -\mathrm{i} & 1 \\ \mathrm{i} & 0 & \mathrm{i} \\ 1 & -\mathrm{i} & 0\end{pmatrix}$，求酉方阵 U 使得 $U^{-1}HU$ 为对角阵.

证明. 算特征多项式，找特征向量，分别正交化. □

例 8.7. A 为规范方阵 $\Leftrightarrow$ 存在多项式 $f(\lambda) \in \mathbb{C}[\lambda]$ 使得 $A^* = f(A)$.

证明. A 的不同特征值为 $\lambda_1, \cdots, \lambda_s$，重数为 $n_1, \cdots, n_s$，则存在酉方阵 U 使得

$$A = U^{-1}\big)\big(A^* = U^{-1}\big)\big(U$$

$$f(A) = U^{-1}\Big(\ \ldots\Big)U$$

□

9 习题

9.1 丘维声习题 8.3 正交补，实内积空间的保距同构

11. 设 V 是一个实内积空间，W 是 V 的一个子空间. 设 $\alpha \in V$，证明：

(1) $\beta \in W$ 是 α 在 W 上的最佳逼近元当且仅当 $\alpha - \beta \in W^{\perp}$；

(2) 若 α 在 W 上的最佳逼近元存在，则它是唯一的.

证明. (1)

$\Rightarrow$

$$d(\alpha,\beta) \leqslant d(\alpha,\gamma)$$
$$|\alpha-\beta|^2 \leqslant |\alpha-\gamma|^2$$
$$|\alpha-\beta|^2 \leqslant |\alpha-\beta+\beta-\gamma|^2$$

由于 $\beta,\gamma \in W$，因此 $\beta-\gamma \in W$，再由 γ 的任意性，我们不妨将 $\beta-\gamma$ 重记为 γ，即

$$|\alpha-\beta|^2 \leqslant |\alpha-\beta+\gamma|^2$$

$$(\alpha-\beta,\alpha-\beta) \leqslant (\alpha-\beta+\gamma,\alpha-\beta+\gamma)$$
$$(\gamma,\gamma)+2(\alpha-\beta,\gamma) \geqslant 0$$

由 γ 的任意性，该式子对 $k\gamma$ 也对，代入得

$$k^2(\gamma,\gamma)+2k(\alpha-\beta,\gamma) \geqslant 0.$$

当 $\gamma \neq 0$ 时，我们得到一个开口向上二次函数，它恒大于等于零要求判别式小于等于零，即

$$4(\alpha-\beta,\gamma)^2 \leqslant 0,$$

这就迫使

$$(\alpha-\beta,\gamma)=0.$$

由 γ 的任意性，$\alpha-\beta \in W^{\perp}$.

$\Leftarrow$

$$|\alpha-\gamma|^2 = |\alpha-\beta+\beta-\gamma|^2 = |\alpha-\beta|^2+|\beta-\gamma|^2 \geqslant |\alpha-\beta|^2$$

(2) 若 β_1 和 β_2 都是 α 在 W 上的最佳逼近元，那么由第 (1) 问知 $\alpha-\beta_1, \alpha-\beta_2 \in W^{\perp}$，因此 $\beta_1-\beta_2 \in W^{\perp}$，但 $\beta_1-\beta_2 \in W$，$W \cap W^{\perp}=\{0\}$，因此 $\beta_1=\beta_2$.

□

12. 设 W 是实内积空间 V 的一个子空间，证明：V 中每个向量 α 都有在 W 上的最佳逼近元当且仅当 $V = W \oplus W^{\perp}$.

证明.

$\Rightarrow$ 先证明 $V = W + W^{\perp}$.

对任意 $\alpha \in V$，设它的最佳逼近元为 $\beta \in W$，由第 11 题知，$\alpha - \beta \in W^{\perp}$，即存在 $\gamma \in W^{\perp}$ 使得 $\alpha - \beta = \gamma$，即 $\alpha = \beta + \gamma$.

又 $W \cap W^{\perp} = \{0\}$，因此 $V = W \oplus W^{\perp}$.

$\Leftarrow$ 显然. □

13. 设 W 是实内积空间 V 的一个子空间.

(1) 证明：若 V 在 W 上的正交投影 $\mathscr{P}$ 存在，则 $\mathscr{P}$ 是 V 上的一个线性变换，且是幂等的，还有

$$\operatorname{Ker} \mathscr{P} = W^{\perp}, \operatorname{Im} \mathscr{P} = W.$$

(2) 证明：若 V 在 W 上的正交投影存在，则 V 在 $W^{\perp}$ 上的正交投影也存在，它等于 $\mathscr{I} - \mathscr{P}$.

证明. (1)

□

9.2 丘维声习题 8.4 正交变换

1. 设 $\mathscr{A}$ 是实内积空间 V 上的一个正交变换，证明：如果 $\mathscr{A}$ 有特征值，那么 $\mathscr{A}$ 的特征值为 1 或 -1.

证明.

$$(\alpha, \alpha) = (\mathscr{A}\alpha, \mathscr{A}\alpha) = \lambda^2(\alpha, \alpha)$$

□

2. 设 $\mathscr{A}$ 是 n 维 Euclid 空间 V 上的一个正交变换，证明：$\mathscr{A}$ 的特征多项式的复根为 ± 1，或 $\cos\theta \pm \mathrm{i}\sin\theta$，其中 $0 < \theta < \pi$.

3. 证明：n 级正交矩阵 A 如果有两个不同的特征值，那么 A 的属于不同特征值的特征向量在 $\mathbb{R}^n$ 中是正交的.

9.3 丘维声习题 8.6 酉空间

10. 设 W 是酉空间 V 的一个子空间，则 V 在 W 上的正交投影存在当且仅当 $V = W \oplus W^{\perp}$.

9.4 丘维声习题 8.8 线性变换的伴随变换，正规变换

1. 证明：酉空间 V 上的线性变换 $\mathscr{A}$ 是酉变换当且仅当 $\mathscr{A}^* = \mathscr{A}^{-1}$.

证明.

$\Rightarrow$

$$(\mathscr{A}\alpha, \beta) = (\mathscr{A}\alpha, \mathscr{A}\mathscr{A}^{-1}\beta) = (\alpha, \mathscr{A}^{-1}\beta) \Rightarrow \mathscr{A}^* = \mathscr{A}^{-1}$$

$\Leftarrow$

$$(\mathscr{A}\alpha, \mathscr{A}\beta) = (\alpha, \mathscr{A}^{-1}\mathscr{A}\beta) = (\alpha, \beta)$$

□

2. 证明：酉空间 V 上的线性变换 $\mathscr{A}$ 是 Hermite 变换当且仅当 $\mathscr{A}^* = \mathscr{A}$.

证明.

$\Rightarrow$ 这是定义.

$\Leftarrow$ 这也是定义. □

3. 证明：酉空间 V 上的线性变换 $\mathscr{A}$ 如果满足下列 3 个条件中的任意 2 个，那么它满足第 3 个条件：

(1) $\mathscr{A}$ 是酉变换；

(2) $\mathscr{A}$ 是 Hermite 变换；

(3) $\mathscr{A}$ 是对合变换（即 $\mathscr{A}^2 = \mathscr{I}$）.

证明. 从 $a = b$、$b = c$ 和 $a = c$ 中的任意两个推出第三个. □

4. 设 $\mathscr{A}, \mathscr{B}$ 是酉空间 V 上的两个 Hermite 变换，证明：$\mathscr{A}\mathscr{B}$ 是 Hermite 变换当且仅当 $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$.

证明. $\mathscr{A}\mathscr{B}$ 是 Hermite 变换

$\Leftrightarrow \mathscr{A}\mathscr{B} = (\mathscr{A}\mathscr{B})^* = \mathscr{B}^*\mathscr{A}^* = \mathscr{B}\mathscr{A}$ □

5. 设 $\mathscr{A}, \mathscr{B}$ 是酉空间上的两个 Hermite 变换，证明：$\mathscr{A}\mathscr{B} + \mathscr{B}\mathscr{A}$ 与 $\mathrm{i}(\mathscr{A}\mathscr{B} - \mathscr{B}\mathscr{A})$ 都是 Hermite 变换.

证明. 只证第二个.

$$(\mathrm{i}(\mathscr{A}\mathscr{B} - \mathscr{B}\mathscr{A}))^* = \bar{\mathrm{i}}(\mathscr{B}^*\mathscr{A}^* - \mathscr{A}^*\mathscr{B}^*) = \mathrm{i}(\mathscr{A}\mathscr{B} - \mathscr{B}\mathscr{A})$$

□

6. 证明：酉空间 V 上的线性变换 $\mathscr{A}$ 如果有伴随变换 $\mathscr{A}^*$，那么 $\mathscr{A}$ 可以唯一地表示成

$$\mathscr{A} = \mathscr{A}_1 + \mathrm{i}\mathscr{A}_2,$$

其中 $\mathscr{A}_1, \mathscr{A}_2$ 都是 Hermite 变换.

证明. 设 $\mathscr{A}_1=\dfrac{\mathscr{A}+\mathscr{A}^*}{2},\mathscr{A}_2=\dfrac{(\mathscr{A}-\mathscr{A}^*)}{2\mathrm{i}}$, 容易验证满足条件.

假设 $\mathscr{B}_1+\mathrm{i}\mathscr{B}_2$ 也满足条件，那么

$$\mathscr{A}_1-\mathscr{B}_1=\mathrm{i}(\mathscr{A}_2-\mathscr{B}_2),$$

取伴随，得

$$\mathscr{A}_1-\mathscr{B}_1=-\mathrm{i}(\mathscr{A}_2-\mathscr{B}_2).$$

因此有 $\mathscr{A}_1=\mathscr{B}_1,\mathscr{A}_2=\mathscr{B}_2$. □

注记. 让人联想到欧拉公式，让人联想到复数.

15. 证明：酉空间 V 上正规变换 $\mathscr{A}$ 的属于不同特征值的特征向量一定正交. 19. 设 $\mathscr{A},\mathscr{B}$ 是 n 维酉空间 V 上的正规变换，证明：若 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$，则 V 中存在一个标准正交基，使得 $\mathscr{A},\mathscr{B}$ 在此基下的矩阵都是对角矩阵.

Chapter 4

双线性函数，二次型

1 双线性函数

- 线性空间 V 上双线性函数的定义
- 有限维线性空间上双线性函数在固定一组基下的矩阵表示 X^TAY
- 固定一组基后，全体双线性函数与全体 n 维矩阵之间有一一对应
- 对同一个双线性函数，在不同基下的度量矩阵是合同的

定义 1.1. 设 V 是域 $\mathbb{F}$ 上的线性空间，$f: V\times V\to\mathbb{F}$ 满足

$$f(\lambda_1\alpha_1+\lambda_2\alpha_2,\beta)=\lambda_1 f(\alpha_1,\beta)+\lambda_2 f(\alpha_2,\beta)$$
$$f(\alpha,\mu_1\beta_1+\mu_2\beta_2)=\mu_1 f(\alpha_1,\beta_1)+\mu_2 f(\alpha,\beta_2)$$

则称 $f(\alpha,\beta)$ 为 V 上的双线性函数.

设 V 是域 $\mathbb{F}$ 上的 n 维线性空间，f 是 V 上的一个双线性函数，我们来探索 f 的表达式. 设 $M=\{\xi_1,\xi_2,\cdots,\xi_n\}$ 是 V 的一组基，向量 α 与 β 在这组基下的坐标分别是 $X=(x_1,x_2,\cdots,x_n)^T$ 与 $Y=(y_1,y_2,\cdots,y_n)^T$，即

$$\alpha=\sum_{i=1}^n x_i\xi_i,\beta=\sum_{j=1}^n y_j\xi_j.$$

$$f(\alpha,\beta)=f\left(\sum_{i=1}^n x_i\xi_i,\sum_{j=1}^n y_j\xi_j\right)=\sum_{i=1}^n\sum_{j=1}^n x_iy_jf(\xi_i,\xi_j)$$

记 n 阶方阵 A 为

$$A=(f(\xi_i,\xi_j))=\begin{pmatrix} f(\xi_1,\xi_1) & f(\xi_1,\xi_2) & \cdots & f(\xi_1,\xi_n)\\ f(\xi_2,\xi_1) & f(\xi_2,\xi_2) & \cdots & f(\xi_2,\xi_n)\\ \vdots & \vdots & \ddots & \vdots\\ f(\xi_n,\xi_1) & f(\xi_n,\xi_2) & \cdots & f(\xi_n,\xi_n)\end{pmatrix}$$

则

$$f(\alpha,\beta)=X^TAY.$$

称 A 是双线性函数 $f(\alpha,\beta)$ 在 V 的基 $\{\xi_1,\xi_2,\cdots,\xi_n\}$ 下的度量矩阵，它是由 f 及 V 的基 $\{\xi_1,\xi_2,\cdots,\xi_n\}$ 唯一决定的.

以矩阵 A^T 为度量矩阵的是双线性函数 $g(x,y):=f(y,x)$.

数域 $\mathbb{F}$ 上的 n 维线性空间 V 的所有双线性函数集合记为 $L(V,V,\mathbb{F})$. 定义 $L(V,V,\mathbb{F})$ 上的加法和数乘为

$$(f+g)(\alpha,\beta)=f(\alpha,\beta)+g(\alpha,\beta)$$
$$(\lambda f)(\alpha,\beta)=\lambda f(\alpha,\beta)$$

可以验证 $L(V,V,\mathbb{F})$ 在如此的加法与数乘下构成线性空间.

定理 1.1.

$$L(V,V,\mathbb{F})\cong\mathbb{F}^{n\times n}$$

现在考虑一个双线性函数在不同基下的方阵表示之间的联系. 为此重述一下方阵相合的概念. 设 A 与 B 是数域 $\mathbb{F}$ 上的 n 阶方阵. 如果存在数域 $\mathbb{F}$ 上的 n 阶可逆方阵 P，使得 $B=P^TAP$，则称方阵 A 与 B 是相合的.

定理 1.2. 设数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数 f 在 V 的基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 与 $\{\beta_1,\beta_2,\cdots,\beta_n\}$ 下的方阵分别是 A 与 B，前者到后者的过渡矩阵为

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)P,$$

其中 P 是数域 $\mathbb{F}$ 上的 n 阶可逆方阵. 则

$$B=P^TAP.$$

证明. 任取 $\alpha,\beta\in V$，设

$$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)X=(\beta_1,\beta_2,\cdots,\beta_n)\tilde{X},$$

$$\beta=(\alpha_1,\alpha_2,\cdots,\alpha_n)Y=(\beta_1,\beta_2,\cdots,\beta_n)\tilde{Y}.$$

由于基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵是 P，因此

$$X=P\tilde{X},Y=P\tilde{Y}.$$

分别在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 下计算 $f(\alpha,\beta)$，得

$$f(\alpha,\beta)=X^TAY,f(\alpha,\beta)=\tilde{X}^TB\tilde{Y}$$

$$\tilde{X}^TB\tilde{Y}=X^TAY=(P\tilde{X})^TA(P\tilde{Y})=\tilde{X}^T(P^TAP)\tilde{Y}.$$

从而

$$B=P^TAP.$$

□

由于方阵的秩是方阵在相合下的不变量，因此表明双线性函数 $f(\alpha,\beta)$ 在 V 的某组基下的方阵的秩并不依赖于基的选取，而是由双线性函数 $f(\alpha,\beta)$ 自身所确定的.

定义 1.2. f 是线性空间 V 上的双线性函数,A 是 f 在一组基 M 下的矩阵,定义 $rank(f) = rank(A)$. 若 A 非奇异，则称 f 非奇异，否则称 f 奇异（退化）.

向量关于内积 $f(\alpha, \beta)$ 的正交性可以推广到一般双线性函数.

定义 1.3. 设 f 是域 $\mathbb{F}$ 上线性空间 V 上的一个双线性函数，$\alpha, \beta \in V$，如果 $f(\alpha, \beta) = 0$，那么称 α（关于 f）左正交于 β，记作 $\alpha \perp_L \beta$，称 β（关于 f）右正交于 α，记作 $\beta \perp_R \alpha$.

一般地说，向量关于双线性函数 $f(\alpha, \beta)$ 的正交性并不是对称的，也就是说，向量 α 关于 f 左正交于向量 β 并不意味着向量 β 关于 f 也左正交于向量 α.

例 1.1. $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, V = \mathbb{F}^2$, 定义 $f(x, y) = x^T A y$，则 $e_1 \perp_L e_2$，但 $e_2 \perp_L e_1$ 不成立.

设 f 是域 $\mathbb{F}$ 上线性空间 V 上的一个双线性函数. 考虑 V 中具有下述性质的向量 α 组成的子集，使得线性函数 α_L 等于零函数，即 V 的下述子集：

$$\{\alpha \in V | f(\alpha, \beta) = 0, \forall\, \beta \in V\}.$$

称 V 的这个子集是 f 在 V 中的左根，记作 $rad_L V$. 类似地，可以考虑 f 在 V 中的右根.

命题 1.1. $rad_L V$ 和 $rad_R V$ 都是 V 的子空间.

命题 1.2. f 限制在 $rad_L V$ 和 $rad_R V$ 上是零函数.

注记. 若 f 限制在 V 的某个子空间上是零函数，则该子空间要么是 $rad_L V$ 的子空间要么是 $rad_L R$ 的子空间.

向量关于双线性函数的正交性可以推广到子空间情形.

定义 1.4. 设 $f(\alpha, \beta)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数，W 是 V 的子空间，$\alpha \in V$. 如果对任意向量 $\beta \in W$，均有 $f(\alpha, \beta) = 0(\text{或} f(\beta, \alpha) = 0)$，则向量 α 称为关于 f 左（或右）正交于子空间 W，记为 $\alpha \perp_L W(\text{或}\alpha \perp_R W)$.

命题 1.3. $S_1 \perp_L S_2 \Leftrightarrow V(S_1) \perp V(S_2)$.

定义 1.5. $S \subset V$,定义 $S^{\perp_L} = \{\alpha \in V | f(\alpha, \beta) = 0, \forall \beta \in S\}, S^{\perp_R} = \{\beta \in V | f(\alpha, \beta) = 0, \forall \alpha \in S\}. V^{\perp_L}$ 称为 V 在 f 下的左根基，$V^{\perp_R}$ 称为 V 在 f 下的右根基.

例 1.2. $V = \mathbb{F}^3$，$f(x, y) = x^T A y$，$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \\ -1 & -2 & 0 \end{pmatrix}$，$S = \{x_1, x_2\}$，$x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$，$X_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$，$W = V(S)$. 求 $W^{\perp_L}, W^{\perp_R}, V^{\perp_L}, V^{\perp_R}$

定理 1.3. f 是线性空间 V 上的双线性函数，A 是 f 的一组基 M 下的矩阵，W 是 V 的子空间，W 的一组基在 M 下的矩阵为 B，则

$$W^{\perp_L} = \left\{(\alpha_1, \cdots, \alpha_n) | x \in \mathbb{F}^n, B^T A^T x = 0\right\}$$

命题 1.4. $W = V(S) \subset V$，则

1. $S^{\perp_L} = W^{\perp_L}, S^{\perp_R} = W^{\perp_R}$

2. $W \subseteq (W^{\perp_L})^{\perp_R}$

推论 1.1. $\dim W^{\perp_L} = n - r(AB), \dim W^{\perp R} = n - r(A^T B)$

定理 1.4. f 非退化当且仅当下列条件之一成立.

1. $V^{\perp_L} = 0$

2. $V^{\perp R} = 0$

3. 令 $f_\alpha : \beta \to f(\alpha, \beta)$，则 $\alpha \to f_\alpha$ 是 V 到 V^* 的同构

4. $\forall \varphi$，存在 $\alpha \in V$ 使 $f_\alpha = \varphi$

命题 1.5. W 关于 f 非退化，则 $V = W \oplus W^{\perp_L} = W \oplus W^{\perp_R}$

证明. 只要证 $V = W + W^{\perp_L}$

固定 $\alpha \in V$，$\varphi_\alpha : \beta \to f(\alpha, \beta)$，$\varphi_\alpha \in W^*$，$f|W$ 非退化，存在 $\alpha_1 \in W$，使得 $f_{\alpha_1} \in W$. $f(\alpha_1, \beta) = f(\alpha, \beta), \forall \alpha \in W \Leftrightarrow f(\alpha - \alpha_1, \beta) = 0, \forall \beta \in W \Rightarrow \alpha - \alpha_1 \in W^{\perp_L}$. □

定理 1.5. f 关于正交是对称的当且仅当 f 对称或 f 反对称.

命题 1.6. f 在 V 的一组基 M 下的矩阵为 A，则

1. f 对称当且仅当 A 对称

2. f 反对称当且仅当 A 反对称

例 1.3. H 为 Hermite 阵，证明：

1. $I \pm \mathrm{i}H$ 可逆

2. $A = (I = \mathrm{i}H)(I - \mathrm{i}H)$ 是酉方阵

证明. 1. $H = U^{-1} \Lambda U$，

□

例 1.4. H 为 Hermite 阵，$r(H) = r$，证明 $(\mathrm{tr} H)^2 \leqslant r \mathrm{tr}(H^2)$.

命题 1.7. 设 f 是域 $\mathbb{F}$ 上 n 维线性空间 V 上的一个双线性函数，其在基 $M = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 下的矩阵表示为 A，则

(1) L_f 和 R_f 都是 V 到 V^* 的线性映射，矩阵表示分别为 A^T 和 A.

(2) $\operatorname{Ker} L_f = \operatorname{rad}_L V, \operatorname{Ker} R_f = \operatorname{rad}_R V$.

(3) $\operatorname{rank} L_f = \operatorname{rank}_m f = \operatorname{rank} R_f$.

(4) f 是非退化的当且仅当 L_f 是线性空间 V 到 V^* 的一个同构映射.

(5) f 是非退化的当且仅当 R_f 是线性空间 V 到 V^* 的一个同构映射.

2 对称和斜对称双线性函数

定理 2.1. 设 f 是特征不为 2 的域 $\mathbb{F}$ 上 n 维线性空间 V 上的一个对称双线性函数，则 V 中存在一个基，使得 f 在此基下的度量矩阵为对角矩阵.

证明.

注记. 要时刻牢记 A 的 (i,j) 元的意义是 f 作用在第 i 个基与第 j 个基上.

对线性空间的维数 n 作数学归纳法.

$n=1$ 时，$V=\langle\alpha_1\rangle$.f 在基 α_1 下的度量矩阵为 $\Big(f(\alpha_1,\alpha_1)\Big)$，这是对角矩阵. 因此当 $n=1$ 时命题为真.

假设当维数为 $n-1$ 时命题为真，现在来看 n 维线性空间 V 上的对称双线性函数 f.

若 $f=0$，则 $A=0$，是对角矩阵，成立.

下面设 $f\neq 0$，断言存在 $\alpha\in V$ 使得 $f(\alpha,\alpha)\neq 0$.

假如 $\forall\ \alpha\in V$ 都有 $f(\alpha,\alpha)=0$，则对任意 $\alpha,\beta\in V$，有

$$0=f(\alpha+\beta,\alpha+\beta)=f(\alpha,\alpha)+f(\alpha,\beta)+f(\beta,\alpha)+f(\beta,\beta)=2f(\alpha,\beta).$$

由于 $\mathrm{char}\mathbb{F}\neq 2$，因此从 $2f(\alpha,\beta)=0$ 得 $f(\alpha,\beta)=0$ 即 $f=0$，矛盾！

因此存在 $\alpha_1\neq 0$ 使得 $f(\alpha_1,\alpha_1)\neq 0$.

把 α_1 扩充成 V 的一个基 $\alpha_1,\beta_1,\cdots,\beta_{n-1}$. 令

$$\tilde{\beta}_i=\beta_i-\frac{f(\beta_i,\alpha_1)}{f(\alpha_1,\alpha_1)}\alpha_1, i=1,2,\cdots,n-1$$

则

$$f(\alpha_1,\tilde{\beta}_i)=f(\alpha_1,\beta_i)-\frac{f(\beta_i,\alpha_1)}{f(\alpha_1,\alpha_1)}f(\alpha_1,\alpha_1)=f(\alpha_1,\beta_i)-f(\beta_i,\alpha_1)=0.$$

而 $\alpha_1,\tilde{\beta}_1,\cdots,\tilde{\beta}_{n-1}$ 也是 V 的一个基.

令

$$W=\left\langle\tilde{\beta}_1,\cdots,\tilde{\beta}_{n-1}\right\rangle,$$

则

$$V=\langle\alpha_1\rangle\oplus W.$$

注记. 直和并不是重要的，很多个子空间都能够与 $\langle\alpha_1\rangle$ 形成直和. 重要的是这个子空间是与 $\langle\alpha_1\rangle$ 具有某种正交性.

$f|W$ 是 W 上的对称双线性函数，根据归纳假设，W 中存在一个基 $\alpha_2,\cdots,\alpha_n$，使得 $f|W$ 在此基下的度量矩阵为对角矩阵：

$$\begin{pmatrix} f(\alpha_2,\alpha_2) & & \\ & \ddots & \\ & & f(\alpha_n,\alpha_n) \end{pmatrix}$$

而显然 $f(\alpha_1,\alpha_i)=0, i=2,\cdots,n$.

由于 $V=\langle\alpha_1\rangle\oplus W$，因此 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基. 于是 f 在 V 的这个基下的度量矩阵为

$$\begin{pmatrix} f(\alpha_1,\alpha_1) & & & \\ & f(\alpha_2,\alpha_2) & & \\ & & \ddots & \\ & & & f(\alpha_n,\alpha_n) \end{pmatrix}.$$

□

定理 2.2. 对称方阵相合于对角阵.

注记. 这个定理应该结合之前的定理 5.4 理解. 之前我们知道，任何对称方阵正交相似于对角阵，而这里我们放宽了条件，只要求相合，那么我们就可以让标准形变得更加简单，之前对角阵的对角元是该特征方阵的特征值，而这里我们可以把对角元搞成正负一.

证明. 对矩阵的阶数 n 进行归纳.

当 $n=1$ 时，显然成立.

设结论对 $n-1$ 阶对称方阵成立. 设 $A=\begin{pmatrix} a_{11} & C^T \\ C & A_1 \end{pmatrix}$.

1. $a_{11}\neq 0$

   设 $P_1=\begin{pmatrix} 1 & -a_{11}^{-1}C^T \\ 0 & I^{(n-1)} \end{pmatrix}$，则 $P_1^T=\begin{pmatrix} 1 & 0 \\ -a_{11}^{-1}C & I^{(n-1)} \end{pmatrix}$.

   $$A'=P_1^TAP_1=\begin{pmatrix} 1 & 0 \\ -a_{11}^{-1}C & I^{(n-1)} \end{pmatrix}\begin{pmatrix} a_{11} & C^T \\ C & A_1 \end{pmatrix}\begin{pmatrix} 1 & -a_{11}^{-1}C^T \\ 0 & I^{(n-1)} \end{pmatrix}=\begin{pmatrix} a_{11} & 0 \\ 0 & A_1-a_{11}^{-1}CC^T \end{pmatrix}$$

   由归纳，存在可逆阵 P_2，使得

   $$P_2^T\left(A_1-a_{11}^{-1}CC^T\right)P_2=\operatorname{diag}(\tilde{a}_{22},\cdots,\tilde{a}_{nn}),$$

   即有

   $$\begin{pmatrix} 1 & 0 \\ 0 & P_2^T \end{pmatrix}\begin{pmatrix} a_{11} & 0 \\ 0 & A_1-a_{11}^{-1}CC^T \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & P_2 \end{pmatrix}=\operatorname{diag}(a_{11},\tilde{a}_{22},\cdots,\tilde{a}_{nn}).$$

2. $a_{11}=0$，但某个 $a_{ii}\neq 0, i\neq 1$.

   $$P_{1i}^TAP_{1i}$$

   归为情形 1.

3. 所有 $a_{ii}=0$

   如果 $a_{1j}=0$，由归纳假设，成立

   如果 $a_{1j}\neq 0$，

   $$T_{1j}(1)^TAT_{ij}(1)$$

   归为情形 2.

□

3 双线性函数空间，Witt 消去定理

设 V 是域 $\mathbb{F}$ 上的线性空间，记 V 上双线性函数的全体为 $T_2(V)$.

命题 3.1. 若 $\operatorname{char}\mathbb{F} \neq 2$，则
$$T_2(V) = S_2(V) \oplus \varLambda_2(V).$$

命题 3.2. 若 $\dim V = n$，则
$$T_2(V) \cong M_n(\mathbb{F}) \cong \operatorname{Hom}(V, V).$$

接下来直接建立 $T_2(V)$ 与 $\operatorname{Hom}(V, V)$ 的一个同构映射.

命题 3.3. 设 V 是域 $\mathbb{F}$ 上的 n 维线性空间，σ 是 V^* 到 V 的一个同构映射. 对于任意的 $g \in T_2(V)$，指派 $\mathscr{G} \in \operatorname{End}(V)$ 为
$$\mathscr{G} : V \xrightarrow{L_g} V^* \xrightarrow{\sigma} V,$$
则 $\tau : g \to \mathscr{G}$ 是同构映射.

证明.

(1) 验证 $\tau(g_1 + g_2) = \tau(g_1) + \tau(g_2)$，即要验证它们作用到每个 $\alpha \in V$ 上的象是相同的，考虑到 $\mathscr{G} = \sigma \circ L_g$ 且 σ 是双射，因此 $\mathscr{G}$ 作用下的象相同当且仅当 L_g 作用下的象相同. 即要验证 $L_{g_1+g_2} = L_{g_1} + L_{g_2}$，这是显然的.

(2) 同理可验证 $\tau(kg) = k\tau(g)$.

(3) 验证它是单射，即要从 $\mathscr{G}_1 = \mathscr{G}_2$ 推出 $g_1 = g_2$.

$\mathscr{G}_1(\alpha) = \mathscr{G}_2(\alpha), \forall\, \alpha \in V \Rightarrow g_1(\alpha, \beta) = g_2(\alpha, \beta), \forall\, \alpha, \beta \in V \Rightarrow g_1 = g_2.$

□

推论 3.1. 设 V 是域 $\mathbb{F}$ 上的 n 维线性空间，f 是 V 上的一个非退化双线性函数，则对于任意的 $g \in T_2(V)$，存在 V 上的唯一一个线性变换 $\mathscr{G}$，使得
$$g(\alpha, \beta) = f(\mathscr{G}(\alpha), \beta), \forall\, \alpha, \beta \in V.$$

例 3.1. 设 V 是特征不为 2 的域 $\mathbb{F}$ 上的 n 维线性空间，f 和 g 是 V 上的对称双线性函数，其中 f 是非退化的. 设 $\mathscr{G}$ 是 V 上唯一的一个线性变换，使得
$$g(\alpha, \beta) = f(\mathscr{G}(\alpha), \beta), \forall \alpha, \beta \in V$$

证明. $\Rightarrow$

设 $\mathscr{G}$ 可对角化，则
$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s},$$
其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $\mathscr{G}$ 的全部不同的特征值.

当 $i \neq j$ 时，对于 $\eta_i \in V_{\lambda_i}, \eta_j \in V_{\lambda_j}$，有
$$g(\eta_i, \eta_j) = f(\mathscr{G}(\eta_i), \eta_j) = f(\lambda_i \eta_i, \eta_j) = \lambda_i f(\eta_i, \eta_j),$$

$$g(\eta_j,\eta_i)=f(\mathscr{G}(\eta_j),\eta_i)=f(\lambda_j\eta_j,\eta_i)=\lambda_j f(\eta_j,\eta_i).$$

由于 f,g 都是对称双线性函数，于是

$$(\lambda_i-\lambda_j)f(\eta_i,\eta_j)=0.$$

由于 $\lambda_i\neq\lambda_j$，因此 $f(\eta_i,\eta_j)=0$，从而 $g(\eta_i,\eta_j)=0$.

下面考虑 f 限制在某个特征子空间 V_{λ_i} 上.

由于 $f|V_{\lambda_i}$ 是 V_{λ_i} 上的一个对称双线性函数，且 $\mathrm{char}\,\mathbb{F}\neq 2$，因此在 V_{λ_i} 中存在一组基 $M_i=\{\alpha_{i1},\cdots,\alpha_{ir_i}\}$，使得 f 在此基下的度量矩阵 A_i 为对角矩阵. 于是当 $k\neq j$ 时，有 $f(\alpha_{ik},\alpha_{ij})=0$.

此时也有

$$g(\alpha_{ik},\alpha_{ij})=f(\mathscr{G}\alpha_{ik},\alpha_{ij})=f(\lambda_i\alpha_{ik},\alpha_{ij})=\lambda_i f(\alpha_{ik},\alpha_{ij})=0,$$

从而 $g|V_{\lambda_i}$ 在基 M_i 下的度量矩阵 B_i 也是对角矩阵.

把 $M_1,M_2,\cdots,M_s$ 合起来成为 V 的一个基，f 和 g 在此基下的度量矩阵都是对角矩阵.

$\Leftarrow$

设 f 和 g 在 V 的一组基 $M=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 下的度量矩阵都是对角矩阵，则当 $i\neq j$ 时，有 $f(\alpha_i,\alpha_j)=0,g(\alpha_i,\alpha_j)=0$.

从 $f(\alpha_i,\alpha_j)=0$ 得出，

$$\alpha_i\in<\alpha_1,\cdots,\alpha_{i-1},\alpha_{i+1},\cdots,\alpha_n>.$$

从 $g(\alpha_i,\alpha_j)=f(\mathscr{G}(\alpha_i),\alpha_j)=0$ 得出，

$$\mathscr{G}(\alpha_i)\in<\alpha_1,\cdots,\alpha_{i-1},\alpha_{i+1},\cdots,\alpha_n>.$$

由于 f 是非退化的，

$$\dim<\alpha_1,\cdots,\alpha_{i-1},\alpha_{i+1},\cdots,\alpha_n>^{\perp}=\dim V-\dim<\alpha_1,\cdots,\alpha_{i-1},\alpha_{i+1},\cdots,\alpha_n>=1.$$

从而存在 $\lambda_i\in\mathbb{F}$，使得 $\mathscr{G}(\alpha_i)=\lambda_i\alpha_i$. 这表明 α_i 是 $\mathscr{G}$ 的属于特征值 λ_i 的一个特征向量. 从而 $\mathscr{G}$ 有 n 个线性无关的特征向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 因此 $\mathscr{G}$ 可对角化. □

4 二次型

当我们将对称双线性函数中的 y 取成 x，我们就得到了二次型. 因此二次型具有许多与双线性函数相同的性质. 域 $\mathbb{F}$ 上的 n 元二次型可以用连加号写成

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j,$$

其中 $a_{ij} = a_{ji}, 1 \leqslant i, j \leqslant n$.

二次型还可以写成

$$f(x_1, x_2, \cdots, x_n) = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = X^TAX.$$

命题 4.1. 记域 $\mathbb{F}$ 上所有 n 元二次型组成的集合为 $\mathbb{F}[x_1, x_2, \cdots, x_n]_q$，那么在自然的加法和标量乘法下，$\mathbb{F}[x_1, x_2, \cdots, x_n]_q$ 成为域 $\mathbb{F}$ 上的线性空间，并且

$$\mathbb{F}[x_1, x_2, \cdots, x_n]_q \cong S_n(\mathbb{F}).$$

例 4.1.

$$f(x) = f(x^0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^0)(x_i - x_i^0) + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\xi)(x_i - x_i^0)(x_j - x_j^0)$$

几何中二次曲面的类型的判别问题，以及其他数学分支和自然科学以及工程技术中的问题，都要求研究域 $\mathbb{F}$ 上的一个 n 元二次型能否通过非退化坐标变换化成一个只含平方项的二次型，也就是二次型的标准形.

命题 4.2. 域 $\mathbb{F}$ 上任一 n 元二次型 $f(x_1, x_2, \cdots, x_n) = X^TAX$，经过一个非退化坐标变换 $X = PY$，变成一个 n 元二次型

$$g(y_1, y_2, \cdots, y_n) = (PY)^TAPY = Y^TP^TAPY,$$

它的矩阵是 P^TAP.

受上面的命题的启发，我们引出下述概念：

定义 4.1. 域 $\mathbb{F}$ 上两个 n 元二次型 X^TAX 与 Y^TBY，如果存在一个非退化坐标变换 $X = PY$，使得 X^TAX 变成 Y^TBY，那么称二次型 X^TAX 与 Y^TBY 等价.

从命题4.2和定义4.1立即得到下述结论：

命题 4.3. 域 $\mathbb{F}$ 上两个 n 元二次型 X^TAX 与 Y^TBY 等价当且仅当它们的矩阵 A 与 B 合同.

- 域 $\mathbb{F}$ 上的一个 n 元二次型 X^TAX 能否通过非退化坐标变换化成一个只含平方项的二次型
- 在 X^TAX 的等价类力能否找到一个只含平方项的二次型
- 域 $\mathbb{F}$ 上的 n 级对称矩阵能否合同于对角阵

4.1 配方法求标准形

- 配方法
- 注意是 $X = PY$

例 4.2. 作非退化线性替换把数域 K 上的下述二次型化成标准形，并且写出所作的非退化线性替换.

(1)

(2)

定理 4.1. 任何二次型 $Q(x) = x^T Ax$ 均可以通过配方法化为不含交叉项的形式——标准形.

证明. 用归纳法

$n = 1, Q(x_1) = a_{11}x_1^2$，已经成立.

设结论对 $n-1$ 个变量成立，下面考虑 n 个变量.

$$Q(x_1, \cdots, x_n) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + 2\sum_{1 \leqslant i < j \leqslant n} a_{ij}x_i x_j$$

1. $a_{11} \neq 0$

$$\begin{aligned} Q(x_1, \cdots, x_n) &= a_{11}\left(x_1^2 + 2\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j x_1\right) + R(x_2, \cdots, x_n) \\ &= a_{11}\left(x_1 + \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\right)^2 \underbrace{- a_{11}\left(\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\right)^2 + R(x_2, \cdots, x_n)}_{S(x_2, \cdots, x_n)} \end{aligned}$$

令

$$\begin{aligned} x_1' &= x_1 + \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j \\ x_j' &= x_j, j = 2, \cdots, n \end{aligned}$$

则

$$Q(x_1, \cdots, x_n) = \widetilde{Q}(x_1', \cdots, x_2') = a_{11}x_1'^2 + S(x_2', \cdots, x_n')$$

由归纳假设，$S(x_2', \cdots, x_n')$ 可以通过配方法化为只含平方项的二次型，故 $Q(x_1, \cdots, x_n)$ 也通过配方法化为只含平方项的二次型.

2. $a_{11} = 0$，某个 $a_{ii} \neq 0$，以 x_i 为主变元进行配方，同情形一.

3. 所有平方项为零，即 $a_{ii} = 0, i = 1, \cdots, n$

   设存在 $a_{ij} \neq 0 (i \neq j)$，否则，该二次型是零，已经成立.

   令

$$\begin{aligned} x_i &= x_i' + x_j' \\ x_j &= x_i' - x_j' \end{aligned}$$

   出现 $x_i'^2$ 与 $x_j'^2$，化为了前面的情形.

□

例 4.3. 将二次型 $Q(x,y,z)=x^2+y^2+z^2+4xy-6xz-4yz$ 化为标准型.

解.

$$\begin{aligned}Q(x,y,z)&=x^2+y^2+z^2+4xy-6xz-4yz\\&=(x^2+4xy-6xz)+y^2-4yz+z^2\\&=(x+2y-3z)^2-(2y-3z)^2+y^2-4yz+z^2\\&=(x+2y-3z)^2-3y^2+8yz-8z^2\\&=(x+2y-3z)^2-3(y-\frac{4}{3}z)^2-\frac{8}{3}z^2\end{aligned}$$

令 $x'=x+2y-3z, y'=y-\frac{4}{3}z, z'=z$，得到

$$Q(x,y,z)=\widetilde{Q}(x',y',z')=x'^2-3y'^2-\frac{8}{3}z'^2$$

□

例 4.4. 实二元多项式 $f(x,y)=x^2+2xy+2y^2-3x-4y+1$ 是否恒大于零.

证明. 观察，令

$$x=\frac{x'}{z'}$$

$$y=\frac{y'}{z'}$$

$$=\frac{1}{z'^2}R(x',y',z')$$

但实际上直接配方就好了

$$=(x+y-\frac{3}{2})^2+(y+?)^2+?$$

□

定理 4.2. 实二次型可以通过可逆线性变换化为规范型

$$\widetilde{Q}(y_1,\cdots,y_n)=y_1^2+\cdots+y_r^2-y_{r+1}^2-\cdots-y_{r+s}^2$$

规范型唯一，即 r,s 唯一由 A 决定，称为 A 的正负惯性指数.

证明. 由定理4.1，$Q(x)$ 可以化为标准形 $b_1x_1^2+\cdots,b_nx_n^2$ 不妨设 $b_1,\cdots,b_r>0,b_{r+1},\cdots,b_{r+s}<0,b_{r+s+1}=\cdots,b_n=0$

$$Q(x_1,\cdots,x_n)=(\sqrt{b_1}x_1)^2+\cdots(\sqrt{b_r}x_r)^2-(\sqrt{-b_{r+1}}x_{r+1})^2-\cdots-(\sqrt{-b_{r+s}}x_{r+s})^2$$

下证唯一性，用反证法，设存在可逆变换 $x=P_1y,x=P_2z$

$$Q(x)|_{x=P_1y}=y_1^2+\cdots+y_r^2-y_{r+1}^2-\cdots-y_{r+s}^2$$

$$Q(x)|_{x=P_2z}=z_1^2+\cdots+z_p^2-y_{p+1}^2-\cdots-y_{p+q}^2$$

要证 $r=p, s=q$，首先证 $r+s=p+q$

$$P_1^T A P_1 = \begin{pmatrix} I^r & & \\ & -I^s & \\ & & 0 \end{pmatrix}$$

$$P_1^T A P_1 = \begin{pmatrix} I^p & & \\ & -I^q & \\ & & 0 \end{pmatrix}$$

不妨设 $r<p$. 要找某一个 $x\neq 0$ 使得对应的

$$y=(0,\cdots,0,y_{r+1},\cdots,y_{r+s},0,\cdots,0)$$
$$z=(z_1,\cdots,z_p,0,\cdots,0,0,\cdots,0)$$

$Q(x)|_{x=P_1y}<0, Q(x)|_{x=P_2z}=z_1^2+\cdots+z_p^2>0$

$$y_1=b_{11}x_1+\cdots+b_{1n}x_n=0$$
$$y_r=b_{r1}x_1+\cdots+b_{rn}x_n=0$$
$$z_{p+1}=c_{p+1,1}x_1+\cdots+c_{p+1,n}x_n=0$$
$$z_{p+q}=c_{p+q,1}x_1+\cdots+c_{p+q,n}x_n=0$$

而方程的个数 $r+q<p+q\leqslant n$，所以方程一定有非零解，即存在 $x\neq 0$ 使得对应的 y,z 是需要的形式. □

定理 4.3. 复系数二次型通过可逆线性变换可以化为规范型

$$Q(x)\big|_{x=Py}=y_1^2+\cdots+y_r^2$$

4.2 行列变换求标准形

注记.

$$\begin{pmatrix} A \\ I \end{pmatrix} \longrightarrow \begin{pmatrix} P^TAP \\ P \end{pmatrix}$$

用 I 来记录列变换.

例 4.5. 将 $Q(x,y,z)=$ 化为标准形.

证明.

$$A=\begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & -2 \\ -3 & -2 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & -2 \\ -3 & -2 & 1 \\ 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \xrightarrow[-2C_1+C_2]{3C_1+C_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 4 \\ 0 & 4 & -8 \\ 1 & -2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[-2C_1\to C_2]{3C_1\to C_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -\frac{8}{3} \\ 1 & -2 & \frac{1}{3} \\ 0 & 1 & \frac{4}{3} \\ 0 & 0 & 1 \end{pmatrix}$$

令

$$P=\begin{pmatrix}1 & -2 & \frac{1}{3}\\ 0 & 1 & \frac{4}{3}\\ 0 & 0 & 1\end{pmatrix},\begin{pmatrix}x\\ y\\ z\end{pmatrix}=P\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}$$

$$Q(x,y,z)=\begin{pmatrix}x' & y' & z'\end{pmatrix}\begin{pmatrix}1 & & \\ & -3 & \\ & & -\frac{8}{3}\end{pmatrix}\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}=x'^2-3y'^2-\frac{8}{3}z'^2$$

□

定理 4.4. 实对称方阵相合于规范形

$$\begin{pmatrix}I_r & & \\ & -I_s & \\ & & 0\end{pmatrix}$$

其中 r,s 由 A 唯一决定. 两个同阶实对称方阵相合当且仅当它们的正负惯性指数相同.

定理 4.5. 复对称方阵相合于规范形

$$\begin{pmatrix}I_r & \\ & 0\end{pmatrix}$$

4.3 二次曲线与曲面分类问题

二次曲线

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2b_1x+2b_2y+c=0$$

改写成

$$\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}a_{11} & a_{12}\\ a_{12} & a_{22}\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}+2\begin{pmatrix}b_1 & b_2\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}+c=0$$

令

$$\begin{pmatrix}x\\ y\end{pmatrix}=P\begin{pmatrix}x'\\ y'\end{pmatrix}$$

得到

$$\lambda_1x'^2+\lambda_2y'^2+2b_1'x+2b_2'y+c=0$$

1. 椭圆型，$\lambda_1\lambda_2=\det(A)=a_{11}a_{22}-a_{12}^2>0$

$$\tilde{x}=x'+\frac{b_1'}{\lambda_1},\tilde{y}=y'+\frac{b_2'}{\lambda_2}$$

$$\lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2+\tilde{C}=0$$

2. 双曲型，$\lambda_1\lambda_2<0$

$$\lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2+\tilde{C}=0$$

3. 抛物型，$\lambda_1\lambda_2=0$，不妨设 $\lambda_2=0,\lambda_1\neq 0$

$$\lambda_1\tilde{x}^2+2\tilde{b}\tilde{y}+\tilde{C}=0$$

二次曲面

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz+2b_1x+2b_2y+2b_3z+c=0$$

改写成

$$\begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} + 2\begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + C$$

求方阵 P 使得

$$P^TAP=\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}$$

$$\lambda_1x'^2+\lambda_2y'^2+\lambda_3z'^2+2b_1'x'+2b_2'y'+2b_3'z'+c=0$$

1. 椭球形，$\lambda_1,\lambda_2,\lambda_3$ 同号.
$$\lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2+\lambda_3\tilde{z}^2+\tilde{C}=0$$
2. 双曲型，$\lambda_1,\lambda_2,\lambda_3$ 不同号，不妨设 $\lambda_1,\lambda_2>0,\lambda_3<0$.
$$\lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2+\lambda_3\tilde{z}^2+\tilde{C}=0$$
3. 抛物型，$\lambda_1,\lambda_2,\lambda_3$ 中至少一个为零.
   (a) $\lambda_3=0,\lambda_1\lambda_2\neq 0$
   $$\lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2+2\tilde{b}_3\tilde{z}+\tilde{c}=0$$
      i. $\tilde{b}\neq 0$
      ii. $\tilde{b}=0$

例 4.6. 判断下列二次曲面的类型，形状及位置

1.
$$x^2+4y^2+z^2-4xy-8xz-4yz-1=0$$

证明.

$$\begin{pmatrix} x & y & z \end{pmatrix}\begin{pmatrix} 1 & -2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=1$$

$$P_A(\lambda)=(\lambda-5)^2(\lambda+4),\lambda_1=\lambda_2=5,\lambda_3=-4$$

$$P=\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{4}{15}\sqrt{5} & \frac{2}{3} \\ -\frac{2}{\sqrt{5}} & \frac{2}{\sqrt{15}}\sqrt{5} & \frac{1}{3} \\ 0 & -\frac{1}{3}\sqrt{5} & \frac{2}{3} \end{pmatrix}$$

□

5 Hermite 型

定义 5.1. A 为 n 阶 Hermite 阵，称 $H(x) = x^*Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\bar{x}_i x_j$ 为 Hermite 型.

注记. 二次型 A 的数域也可以取成复的，二次型与 Hermite 型的关键的区别是 Hermite 型中是共轭转置，矩阵是 Hermite 方阵，而二次型的矩阵是对称阵. 比如 $\begin{pmatrix} 1 & i \\ i & 2 \end{pmatrix}$ 是对称阵，但不是 Hermite 阵，而 $\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$ 是 Hermite 阵. 所以二者当然有不同的性质.

对 $\forall x \in \mathbb{C}^n$，$H(x)$ 是实数.

例 5.1.

$$H(x) = H(x_1, x_2, x_3) = \bar{x_1}x_2 - \bar{x_2}x_2 + 2\bar{x_3}x_3 + (2-i)\bar{x_1}x_2 + (2+i)x_1\bar{x_2}$$

找可逆线性变换 $x = Py$，使 $H(x)\big|_{x=Py} = (Py)^*A(Py) = y^*(P^*AP)y = \tilde{H}(y)$ 尽量简单.

定义 5.2. $A, B \in \mathbb{C}^{n\times n}$，若存在可逆复方阵 P 使 $B = P^*AP$，称 A 与 B 是共轭相合.

定理 5.1. Hermite 阵共轭相合于对角阵.

证明. 略 □

例 5.2. 将 Hermite 阵 $\begin{pmatrix} 0 & -i & 1 \\ i & 0 & i \\ 1 & -i & 0 \end{pmatrix}$ 做共轭相合运算化为对角阵.

解.

$$\begin{pmatrix} 0 & -i & 1 \\ i & 0 & i \\ 1 & -i & 0 \\ 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \xrightarrow[R_3\to R_1]{C_3\to C_1} \begin{pmatrix} 2 & -2i & 1 \\ 2i & 0 & i \\ 1 & -i & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \xrightarrow[C_1\cdot -\frac{1}{2}\to C_3]{C_1\cdot i\to C_2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 1 & i & -\frac{1}{2} \\ 0 & 1 & 0 \\ 1 & i & \frac{1}{2} \end{pmatrix}$$

□

定理 5.2. Hermite 阵 A 共轭相合于规范型 $\begin{pmatrix} I_r & \\ & 0 \end{pmatrix}$

证明. 略 □

6 二次型的正定性

- 正定方阵的全体似乎是结构很糟的东西

  1. 首先，它不是线性空间. 虽然两个正定方阵相加得到的仍是正定方阵，正定方阵数乘上一个正的实数得到的仍是正定方阵，但是乘上一个负的实数就没办法了. 哪怕同时考虑正定方阵和负定方阵的全体，还需要补充上一个零方阵.
  2. 其次，它不是群. 虽然正定方阵的逆是正定方阵，但两个正定方阵的乘积不一定是正定方阵.

命题 6.1. *$A>0$ 当且仅当对于任意同阶可逆方阵 P，$P^TAP>0$.*

定理 6.1. *设 A 为 n 阶实对称方阵，则下列命题等价：*

(1) $A>0$；

(2) A 的特征值全为正；

(3) A 相合于单位阵；

(4) 存在可逆方阵 P 使得 $A=P^TP$；

(5) 存在 $B>0$ 使得 $A=B^2$；

(6) A 的所有主子式为正；

(7) A 的所有顺序主子式为正.

证明.

- (1) $\Rightarrow$ (2)

  A 正交相似于对角阵，其对角元是 A 的特征值. 正交相似是特殊的相合，而正定性在相合下保持不变，因此该对角阵也正定，这显然要求各对角元都为正.

  (2) $\Rightarrow$ (1) 显然.

- (1) $\Rightarrow$ (3)

  A 是正定方阵，正定方阵相合于单位阵，这是早在定理2.1就已经知道的结果.

  (3) $\Rightarrow$ (1) 显然.

  (3) $\Leftrightarrow$ (4) 显然.

- (1) $\Rightarrow$ (5)

  $A=O^T\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O$，其中 O 是正交阵，$\lambda_1,\cdots,\lambda_n>0$.

  $\mathrm{diag}(\lambda_1,\cdots,\lambda_n)=\mathrm{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})^2$

  $A=\underbrace{O^T\mathrm{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})O}_{B}\underbrace{O^T\mathrm{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})O}_{B}$

  (5) $\Rightarrow$ (4) 显然.

- (1) $\Rightarrow$ (6)

  显然有 $A > 0 \Rightarrow \det(A) > 0$.

  我们要从 $A > 0 \Rightarrow A\begin{pmatrix} i_1 & \cdots & i_r \\ i_1 & \cdots & i_r \end{pmatrix} > 0$.

  A 对应着二次型 $Q(x) = x^T Ax$，只需要把第 $i_1, \cdots, i_r$ 个之外的 x 的分量取成 0，我们就得到了子矩阵对应的二次型，而这个子矩阵的二次型显然也是正定的，得证.

  (6) $\Rightarrow$ (7) 显然.

  (7) $\Rightarrow$ (1) 用数学归纳法.

  当 $n = 1$ 时，显然成立.

  假设结论对 $n-1$ 阶方阵实对称方阵成立，现在来看 n 阶实对称方阵

$$A = \begin{pmatrix} A_1 & C \\ C^T & a_{nn} \end{pmatrix}, C \in \mathbb{R}^{n\times 1}.$$

  显然 A_1 是 $n-1$ 阶实对称方阵，并且由条件，A_1 的各阶顺序主子式大于零，从而 A_1 是可逆的，因此我们可以对 A 进行相合运算从而把 C 和 C^T 的位置打成零.

$$A = \begin{pmatrix} A_1 & C \\ C^T & a_{nn} \end{pmatrix} \xrightarrow{(-C^T A_1^{-1})\cdot(1)+(2)} \begin{pmatrix} A_1 & C \\ 0 & a_{nn} - C^T A_1^{-1} C \end{pmatrix}$$
$$\xrightarrow{(1)\cdot(-A_1^{-1}C)+(2)} \begin{pmatrix} A_1 & 0 \\ 0 & a_{nn} - C^T A_1^{-1} C \end{pmatrix}$$

  写成矩阵运算的形式就是

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

  两边取行列式易得 $a_{nn} - C^T A_1^{-1} C > 0$.

  由归纳假设，A_1 正定，从而相合于单位阵. 因此

$$A \simeq \begin{pmatrix} I_{n-1} & 0 \\ 0 & a_{nn} - C^T A_1^{-1} C \end{pmatrix}.$$

  从而 A 是正定的.

□

例 6.1. $\operatorname{diag}(A, B) > 0$ 当且仅当 $A > 0, B > 0$.

证明. 充分性

存在可逆阵 P_1, P_2，使得 $P_1^T AP_1 = I, P_2^T BP_2 = I$

$$\begin{pmatrix} P_1 & \\ & P_2 \end{pmatrix}^T \begin{pmatrix} A & \\ & B \end{pmatrix} \begin{pmatrix} P_1 & \\ & P_2 \end{pmatrix} = \begin{pmatrix} P_1^T AP_1 & \\ & P_2^T BP_2 \end{pmatrix} = \begin{pmatrix} I & \\ & I \end{pmatrix} = I$$

必要性

主子式. □

例 6.2. $A>0$，证明 $A^k>0, k\in\mathbb{Z}$.

例 6.3. $A>0$，证明存在上三角阵 P 使得 $P^TAP=I$.

证明. $A>0 \Rightarrow A=B^2, B>0$

可以证明 B 是唯一的. 那么 $B=\sqrt{A}$ □

定理 6.2. $A>0$，则存在唯一 $B>0$，使得 $A=B^2$，且若 $AC=CA$，则 $BC=CB$

证明. 假设 $A=B_1^2=B_2^2, B_1>0, B_2>0$，要证 $B_1=B_2$.

$A=O^T\Lambda O$

B_1, B_2 的特征值为 $\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}$

$B_1=O_1^T\sqrt{\Lambda}O_1, B_2=O_2^T\sqrt{\Lambda}O_2$

$B_1^2=B_2^2 \Rightarrow O_1^T\Lambda O_1=O_2^T\Lambda O_2$

$\underbrace{O_2O_1^T}_{P}\Lambda=\Lambda O_2O_1^T$

$P_{ij}\lambda_j=\lambda_i p_{ij}\begin{cases}\lambda_i=\lambda_j & p_{ij}\sqrt{\lambda_j}=\sqrt{\lambda_i}p_{ij}\\ \lambda_i\neq\lambda_j & p_{ij}\sqrt{\lambda_j}=\sqrt{\lambda_i}p_{ij}\end{cases}$

$\Rightarrow p_{ij}\sqrt{\lambda_j}=\sqrt{\lambda_i}p_{ij} \Rightarrow O_2O_1^T\lambda_1=\lambda_1O_2O_1^T$

$\Rightarrow O_1^T\Lambda_1O_1=O_2\Lambda_1O_2 \Rightarrow B_1=B_2$

$AC=CA \Rightarrow O^T\Lambda OC=CO^T\Lambda O$

$\Rightarrow \Lambda\underbrace{OCO^T}_{\tilde{C}}=\underbrace{OCO^T}_{\tilde{C}}\Lambda$

$\Lambda\tilde{C}=\tilde{C}\Lambda$ □

定理 6.3. A 为 n 阶实对称方阵，则下列命题等价.

1. $A\geqslant 0$.
2. A 的特征值非负.
3. A 相合于 $\begin{pmatrix} I_r & \\ & 0\end{pmatrix}$
4. $A=P^TP$，其中 P 为任意矩阵.
5. 存在 $B\geqslant$ 使得 $A=B^2$.
6. A 的所有主子式非负.
7. A 的所有 k 阶主子式之和非负.

证明. (7) $\Rightarrow$ (2) 用 S_k 表示 A 的 k 阶主子式之和.

$$P_A(\lambda)=\det(\lambda I-A)=\lambda^n-s_1\lambda^{n-1}+s_2\lambda^{n-2}+\cdots+(-1)^nS_n$$

没有负实根. □

例 6.4. 实二次型 $Q(x,y,z)=\lambda(x^2+y^2+z^2)-2xy-2yz-2xz$

1. λ 为何值时 $Q(x)$ 正定.

2. λ 为何值时 $Q(x)=(ax+by+cz)^2$.

解. 1. $A=\begin{pmatrix}\lambda & -1 & -1\\ -1 & \lambda & -1\\ -1 & -1 & \lambda\end{pmatrix}$

各阶顺序主子式大于零，最后推出 $\lambda>1$.

2. $A\geqslant 0, r(A)=1$，那么行列式等于零，解得 λ，带回去验证一下.

$\square$

7 正交相抵标准形

定义 7.1. $A, B \in \mathbb{R}^{m\times n}$，如果存在正交方阵 $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$ 使得 $B = UAV$，则称 A 与 B 正交相抵.

定理 7.1. $A \in \mathbb{R}^{m\times n}$，则存在正交方阵 U, V 使得 $A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V, \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r), \sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0. \sigma_1, \sigma_2, \cdots, \sigma_r$ 为方阵 AA^T 的非零特征值的平方根，称 $\sigma_1, \sigma_2, \cdots, \sigma_r$ 为 A 的奇异值（*singular value*）. 称 $A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V$ 为 A 的奇异值分解（*singular value decomposition,SVD*）. 且 $\sigma_1, \sigma_2, \cdots, \sigma_r$ 为正交相抵的全系不变量.

注记. 若 A 是对称方阵，那么它的奇异值是 A 的非零特征值的绝对值.

证明.

$$AA^T = O\mathrm{diag}(\sigma_1^2, \cdots, \sigma_r^2, \underbrace{0, \cdots, 0}_{m-r\text{个}})O^T，O \in \mathbb{R}^{m\times m} \text{ 是正交矩阵，} \sigma_1 \geqslant \cdots \geqslant \sigma_r > 0$$

$$U^T AA^T U = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_r^2, 0, \cdots 0)$$

$$\text{令 } B = U^T A = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}, \beta_i \in \mathbb{R}^n$$

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \begin{pmatrix} \beta_1^T & \cdots & \beta_m^T \end{pmatrix} = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_r^2, 0, \cdots, 0)$$

$$(\beta_i, \beta_j) = \begin{cases} 0 & i \neq j \\ \sigma_i^2 & i = j = 1, \cdots, r \\ 0 & i = j = r+1, \cdots, m \end{cases}$$

$\Rightarrow \beta_j = 0, j = r+1, \cdots, m$，$\beta_1, \cdots, \beta_r$ 彼此正交.

令 $\gamma_i = \sigma_i^{-1}\beta_i, i = 1, \cdots, r$，$\gamma_1, \cdots, \gamma_r$ 是彼此正交的单位向量.

$$B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \begin{pmatrix} \sigma_1\gamma_1 \\ \vdots \\ \sigma_r\gamma_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

将 $\gamma_1, \cdots, \gamma_r$ 扩充为 $\mathbb{R}^m$ 的一组标准正交基，易知仍有

$$B = \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_r \\ \gamma_{r+1} \\ \vdots \\ \gamma_n \end{pmatrix} =: \mathrm{diag}(\sigma_1, \cdots, \sigma_r, 0, \cdots, 0)V$$

$$A = U\operatorname{diag}(\sigma_1, \cdots, \sigma_r, 0, \cdots, 0)V$$

下面证明 $\sigma_1, \cdots, \sigma_r$ 为正交相抵的全系不变量.

设 $B = UAV, BB^T = UAVV^TA^TU^T = UAA^TU^T$

BB^T 与 AA^T 特征值相同，A 与 B 奇异值相同.

反之，若 A 与 B 奇异值相同，则 A 与 B 正交相抵标准形相同，因此 A 与 B 正交相抵. □

定理 7.2 (矩阵极分解)**.** $A \in \mathbb{R}^{n\times n}$, 则 $A = SO$ (或 OS), 其中 $S \geqslant 0$,O 正交, 则 S 由 A 唯一决定.

证明. 假设 $A = S_1O_1 = S_2O_2$

$AA^T = S_1^2 = S_2^2 \geqslant 0.$ □

8 Hermite 型的正定性

定义 8.1. *Hermite* 型 $H(x) = x^*AX$ 称为正定、负定、半正定、半负定，如果 $H(x) > 0, < 0, \geqslant 0, \leqslant$，称 *Hermite* 方阵 A 正定、负定、半正定、半负定，如果 $H(x)$ 正定、负定、半正定、半负定.

定理 8.1. H 为 *Hermite* 阵，则下列命题等价.

1. $H > 0$
2. H 的特征值全为正.
3. H 共轭相合于 I.
4. $H = P^*P$，P 可逆.
5. $H = \tilde{H}^2, \tilde{H} > 0$
6. H 的所有主子式为正
7. H 的顺序主子式为正

定理 8.2. H 为 *Hermite* 阵，则下列命题等价.

1. $H \geqslant 0$.
2. H 特征值非负.
3. H 共轭相合于 $\begin{pmatrix} I_r & \\ & 0 \end{pmatrix}$
4. $H = P^*P$
5. $H = \tilde{H}^2, \tilde{H} \geqslant 0$
6. H 的所有主子式非负.
7. H 的 k 阶主子式之和非负.

定理 8.3. $H > 0$，则存在唯一 $\tilde{H} \geqslant 0$ 使得 $H = \tilde{H}^2$，且若 $HC = CH$，则 $\tilde{H}C = C\tilde{H}$.

定义 8.2. $A, B \in \mathbb{C}^{m\times n}$，若存在酉方阵 U, V 使得 $B = UAV$，则称 A 与 B 酉相抵.

定理 8.4. $A \in \mathbb{C}^{m\times n}$ 酉相抵于标准形 $\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}, \Sigma = \operatorname{diag}(\sigma_1, \cdots, \sigma_r)$，$\sigma_1, \sigma_r$ 为 AA^* 的非零特征值的平方根，称为 A 的奇异值，$A = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$ 称为奇异值分解. 且 $\sigma_1, \cdots, \sigma_r$ 为酉相抵的全系不变量.

定理 8.5. $A \in \mathbb{C}^{n\times n}$，则 $A = HU, H \geqslant 0$，U 为酉方阵.

例 8.1 (Moore-Penrose 广义逆). $\begin{cases} AXA = A \\ XAX = X \\ (AX)^* = AX \\ (XA)^* = XA \end{cases}$

例 8.2. $A \in \mathbb{R}^{n\times n}$，证明 $Tr(A) \leqslant Tr(\sqrt{AA^T})$，等号成立当且仅当 $A \geqslant 0$

例 8.3. A, B 均为 n 阶实对称方阵且 $A > 0$，证明 A 与 B 同时相合于对角阵.

证明. 分析: $A = I$ 时，找 P 使得 P^TIP 及 P^TBP 均为对角阵. 我们知道存在正交阵 P 使得 P^TBP 为对角阵，此时 $P^IP = I$.

存在可逆阵 P_1 使得 $P_1^TAP_1 = I = \tilde{A}$，$P_1^TBP_1 = \tilde{B}$，仍为实对称阵.

存在正交阵 P_2 使得 $P_2^T\tilde{A}P_2 = I$，$P_2^T\tilde{B}P_2 = \Lambda$.

$P_2^TP_1^TP_1P_2 = I, P_2^TP_1^TBP_1P_2 = \Lambda$ □

例 8.4. A, B 为同阶实对称方阵，$A > 0$，定义 $f(\lambda) = \det(\lambda A - B)$，证明 $f(\lambda) = 0$ 的根均为实数.

证明. $A = P^TP$，P 可逆，$f(\lambda) = \det(\lambda P^TP - B) = \det(P^T(\lambda I - (P^{-1})^TBP^{-1})P)$

$= \det(P)^2\det(\lambda I - \tilde{B})$，$\tilde{B} = (P^{-1})^TBP^{-1}$

$\tilde{B}$ 的特征值为实，因此 □

注记. $Bx = \lambda Ax$，广义特征值问题.

例 8.5. $A > 0, B > 0$，证明 AB 的特征值均为正. 如果加上 $AB = BA$ 的条件，那么 AB 正定.

证明. $A = P^TP$，P 可逆，$AB = P^TPB$ 相似于 $(P^T)^{-1}(P^TPB)P^T = PBP^T > 0$

虽然 AB 不是正定方阵，但是它相似于正定方阵.

$$AB \sim A^{-\frac{1}{2}}(AB)A^{\frac{1}{2}} = A^{\frac{1}{2}}BA^{\frac{1}{2}}$$

□

证明. $ABx = \lambda x \Rightarrow x^TBABx = \lambda x^TBx$

$\lambda = \dfrac{(Bx)^TA(Bx)}{x^TBx} > 0$ □

例 8.6. 设 $A \geqslant B > 0$，证明 $\det A \geqslant B > 0$.

证明. B 是正定的，A 是正定的，$A - B$ 是半正定的.

$A - I \geqslant 0$ 证 $\det A \geqslant 1$

$A = O^T\mathrm{diag}(\lambda_1, \cdots, \lambda_n)O$

一般地，$B = P^TP$，P 可逆，$A - B \geqslant 0$，$\Rightarrow A - P^TP \geqslant 0 (P^{-1})^TAP^{-1} - I \geqslant 0$

$\tilde{A} - I \geqslant \det(\tilde{A}) \geqslant 1$ □

例 8.7. 设 $A > 0$，K 为实反对称方阵，证明 $\det(A + K) \geqslant \det(A)$

例 8.8. $A = (a_{ij})_{n\times n} > 0$，证明：$\det(A) \leqslant a_{11}\cdots a_{nn}$

例 8.9 (Hadamard 不等式)**.** $A \in \mathbb{R}^{n\times n}$，证明 $|\det A| \leqslant \prod_{i=1}^{n}\sqrt{\sum_{k=1}^{n}a_{ki}^2} = \prod_{i=1}^{n}\|\alpha_i\|$

证明. 考虑 A^TA，然后用上一题的结论. □

例 8.10. 证明：$\forall 0 \neq x \in \mathbb{R}^n$，均有

$$f(x) = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n-1} x_i x_{i+1} > 0$$

并求 $f(x)$ 在 $x_n = 1$ 处的最小值.

证明. $f(x) = x^T A x$，要证 $A > 0$. □

例 8.11. $A \in \mathbb{R}^{n\times n}$，$A > 0$，证明，对任意 $x, y \in \mathbb{R}^{n\times n}$ 均有

$$x^T A x + y^T A^{-1} y \geqslant 2x^T y$$

证明.

$$\begin{pmatrix} x & y \end{pmatrix}^T \begin{pmatrix} A & I \\ I & A^{-1} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \geqslant 0$$

□

Chapter 5

张量积

1　双线性映射和张量积

有几个方法定义张量积，这里是利用“要求存在有某种性质的唯一态射”这个条件来定义张量积. 这个条件是关于双线性映射的.

固定域 F，我们熟知 F 向量空间线性映射 $\phi : V \to Z$. 自然地我们把只有一个变元的映射 ϕ 推广为两个变元的映射，加上适当的线性条件后便是以下定义的双线性映射.

定义 1.1. 设 V, W, Z 都是域 F 上的向量空间. 称 $f : V \times W \to Z$ 为 F 双线性映射，如果对任意 $a \in F$,$v_1, v_2, v \in C, w, w_1, w_2 \in W$ 有

$$f(av_1 + v_2, w) = af(v_1, w) + f(v_2, w),$$

$$f(v, aw_1 + w_2) = af(v, w_1) + f(v, w_2).$$

由所有从 $V \times W$ 到 Z 的双线性映射组成的集合记为 $\mathscr{L}(V, W; Z)$. 不难证明这是 F 向量空间.

注记. 设 V, W, Z 的维数分别为 n, m, l，则 $\mathscr{L}(V \times W, Z)$ 同构于 $F^{(m+n)\times l}$.

只需要分别选定 V, W 的一组基，**任意**指定基向量的像，便可得到从 $V \times W$ 到 Z 的线性映射.

但是，对于 $V \times W$ 到 Z 的双线性映射，基的像不可被随意指派. 比如，考虑 $\mathbb{R} \times \mathbb{R}$ 到 $\mathbb{R}$ 的双线性映射 f，$f(1, 0) = f(1, 0 + 0) = 2f(1, 0)$，所以 $f(1, 0) = 0$，同理 $f(0, 1) = 0$.

常规说，变元的个数越少，映射就越简单. 可以问：给出两个 F 向量空间 V, W，可否从 V, W 得出 F 向量空间 U 使得 U 到 Z 的线性映射可以取代 $V \times W$ 到 Z 的双线性映射？以下定义说清楚这个问题.

定义 1.2. 设 V, W, Z 都是 F 向量空间. 称 F 双线性映射 $f : V \times W \to U$ 为 V 与 W 在 F 上的张量积，如果对任意的 F 双线性映射 $g : V \times W \to Z$，存在唯一的 F 双线性映射 $h : U \to Z$，使得

$$h \circ f = g.$$

也就是说，下图是可交换的

$$\begin{tikzcd} V \times W \arrow[r, "f"] \arrow[dr, "g"'] & U \arrow[d, dotted, "h"] \\ & Z \end{tikzcd}$$

注意到上述定义表明 f 在 $V \times W$ 上的 F 双线性映射中具有万有性或说为“泛性”.

注记. 容易验证 (U, f) 在同构的意义下是唯一的.

这就诱导了映射

$$\theta : \mathscr{L}(V, W; Z) \to \mathscr{L}(U; Z), \quad g \mapsto h,$$

它是线性映射吗？它是单射吗？它是满射吗？

1.1 张量积的记号

正如我们所知，V 与 W 在域 F 上的张量积

$$f : V \times W \to U$$

在同构的意义下是唯一的，我们用 $V \otimes_F W$ 表示 U，或者在域 F 明确的情况下，简记为 $F \otimes W$，并且用 $v \otimes w$ 表示 $f(v, w)$.

2 张量积的存在性

2.1 定义一

设 T 是由有序对 $\{(v,w)\big|v\in V,w\in W\}$ 生成的域 F 上的向量空间（不是 $V\times W$!）. 向量空间 T 中的任一个元素可表为一种如下形式上的有限和

$$\sum_{i=1}^{n} a_i(v_i,w_i),\quad a_i\in F.$$

设 S 是由 T 中具有如下形式的元素生成的 T 的一个子空间：

$$(av_1+v_2,w)-a(v_1,w)-(v_2,w),$$

$$(v,aw_1+w_2)-a(v,w_1)-(v,w_2),$$

其中 $a\in F, v,v_1,v_2\in V, w,w_1,w_2\in W$.

令 $U=T/S$ 并定义映射 f 如下

$$\begin{aligned} f: V\times W &\to U,\\ (v,w)&\mapsto (v,w)+S. \end{aligned}$$

由 U 的构造方式可得

$$f(av_1+v_2,w)-af(v_1,w)-f(v_2,w)=0\in U,$$

$$f(v,aw_1+w_2)-af(v,w_1)-f(v,w_2)=0\in U.$$

可验证按如上方式构造的 (U,f) 是 V 与 W 在域 F 上的张量积.

设 $g:V\times W\to Z$ 是双线性映射，定义

$$\begin{gathered} h: U\to Z\\ (v,w)+S\mapsto g(v,w) \end{gathered}$$

则容易验证 h 是良定的，并且 $h\circ f=g$，而 h 的唯一性是显然的.

我们是从映射的观点来定义张量积，没有用坐标，比较简洁，但是以上张量积的构造是会帮助计算的.

如果我们把元素 $(v,w)+S$ 记为 $v\otimes w$，则张量积 U 的任意元素可表达为有限和 $\sum_i v_i\otimes w_i$.

如果 $\{e_i\}$ 生成 V，$\{f_j\}$ 生成 W，则 $\{e_i\otimes f_j\}$ 生成张量积 $V\otimes W$.

命题 2.1. 设 V,W 是域 F 上的向量空间. 取 $v_i\in V$ 和 $W_i\in W$. 设 $\{v_1,\cdots,v_n\}$ 线性无关. 如果 $\sum_{i=1}^{n} v_i\otimes w_i=0$，则 $w_1=\cdots=w_n=0$.

证明. □

2.2 定义二

为了叙述清楚，我们先说明对偶空间 V^* 和 W^* 的张量积.

1. 设 $v^* \in V^*, w^* \in W^*$，定义 $v^* \otimes w^* \in \mathscr{L}(V, W; F)$ 为

$$v^* \otimes w^*(v, w) := v^*(v) \cdot w^*(w).$$

2.
$$\begin{aligned} \bigotimes : V^* \times W^* &\longrightarrow \mathscr{L}(V, W; F) \\ (v^*, w^*) &\longmapsto v^* \otimes w^* \end{aligned}$$

   是双线性映射.

3. V^* 和 W^* 的张量积 $V^* \otimes W^*$ 定义为形如 $v^* \otimes w^*$ 的元素生成的 $\mathscr{L}(V, W; F)$ 的子空间.

   要指出的是，张量积 $V^* \otimes W^*$ 的元素是形如 $v^* \otimes w^*$ 的元素的有限线性组合，一般不能写成单项式.

4. 事实上，$V^* \otimes W^* = \mathscr{L}(V, W; F)$.

3 线性映射的张量积

4 张量积的另一种构造方式