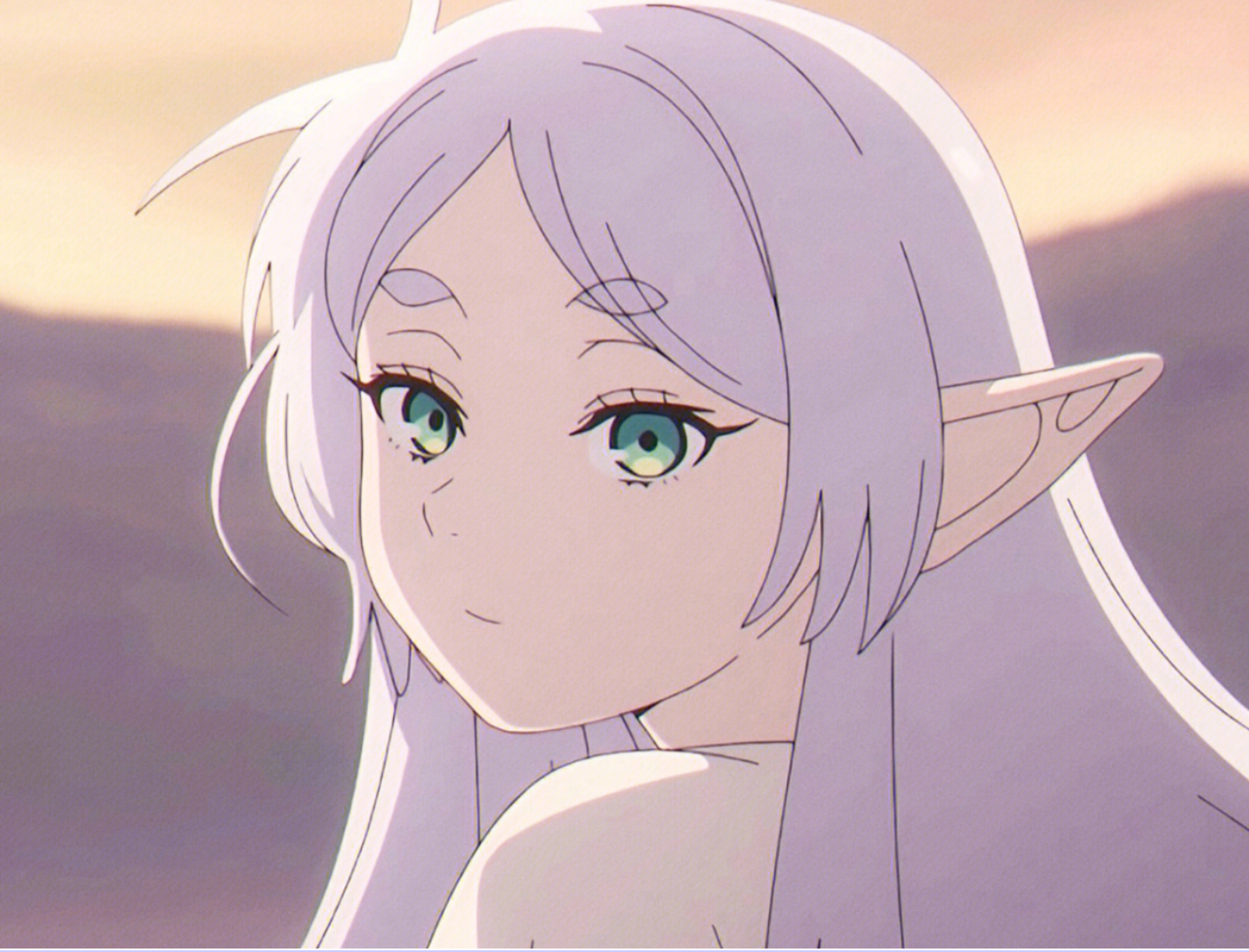

高等代数学习笔记

学习与思考

作者：杨毅涵

组织：南开大学数学科学学院

时间：June 19, 2024

我们从未理解数学，我们只是习惯它。—— 冯·诺依曼

目录

第 1 章 线性空间

1.1 未定义

命题 1.1

对任意矩阵 $A=[a_1,...,a_n]\in\mathbb{R}^{m\times n}$,我们记

$$K(A):=\{x\in\mathbb{R}^n : Ax=0\},\ R(A):=L(a_1,...,a_n)$$

(1) 设 $A\in\mathbb{R}^{n\times n}$, 证明:

$$\mathbb{R}^n=K(A)\oplus R\left(A^\top\right)$$

(2) 设 $A\in\mathbb{R}^{r\times n}, B\in\mathbb{R}^{s\times n}$, 证明:

$$K(A)\subseteq K(B)\Leftrightarrow R\left(A^\top\right)\supseteq R\left(B^\top\right)$$

命题 1.2

第 2 章 特征值

2.1 特征值与特征向量

命题 2.1

设 φ 是线性空间 V 上的线性变换, V 有一个直和分解:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$$

其中 V_i 都是 φ-不变子空间.

(1) 设 φ 限制在 V_i 上的特征多项式为 $f_i(\lambda)$,求证: φ 的特征多项式

$$f(\lambda) = f_1(\lambda) f_2(\lambda) \cdots f_m(\lambda)$$

(2) 设 λ_0 是 φ 的特征值, $V_0 = \{\mathbf{v} \in V \mid \varphi(\mathbf{v}) = \lambda_0 \mathbf{v}\}$ 为特征子空间, $V_{i,0} = V_i \cap V_0 = \{\mathbf{v} \in V_i \mid \varphi(\mathbf{v}) = \lambda_0 \mathbf{v}\}$,求证:

$$V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}$$

证明 (1) 取 V_i 的一组基,将它们拼成 V 的一组基. 记 $\mathbf{A}_i$ 是 φ 在 V_i 上的限制在 V_i 所取基下的表示矩阵,则 φ 在 V 的这组基下的表示矩阵为分块对角矩阵 $\mathbf{A} = \mathrm{diag}\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m\}$,于是

$$f(\lambda) = |\lambda \mathbf{I}_n - \mathbf{A}| = |\lambda \mathbf{I} - \mathbf{A}_1| \, |\lambda \mathbf{I} - \mathbf{A}_2| \cdots |\lambda \mathbf{I} - \mathbf{A}_m|$$

即 $f(\lambda) = f_1(\lambda) f_2(\lambda) \cdots f_m(\lambda)$.

(2) 任取 $\alpha \in V_0$,设 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m$,其中 $\alpha_i \in V_i$,则

$$\varphi(\alpha_1) + \varphi(\alpha_2) + \cdots + \varphi(\alpha_m) = \varphi(\alpha) = \lambda_0 \alpha = \lambda_0 \alpha_1 + \lambda_0 \alpha_2 + \cdots + \lambda_0 \alpha_m$$

注意到 $\varphi(\alpha_i) \in V_i$,故由直和的充要条件可得 $\varphi(\alpha_i) = \lambda_0 \alpha_i$,即 $\alpha_i \in V_{i,0}$,从而 $V_0 = V_{1,0} + V_{2,0} + \cdots + V_{m,0}$. 注意到 $V_{i,0} \subseteq V_i$,故

$$V_{i,0} \cap (V_{1,0} + \cdots + V_{i-1,0}) \subseteq V_i \cap (V_1 + \cdots + V_{i-1}) = 0, \ 2 \leq i \leq m$$

于是上述和为直和.

注 注将上面的条件和结论代数化之后,可知: 对分块对角矩阵 $\mathbf{A} = \mathrm{diag}\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m\}$ 的任一特征值 λ_0,其代数重数等于每个分块的代数重数之和,其几何重数等于每个分块的几何重数之和.

命题 2.2

设 $\mathbf{A}$ 是 n 阶整数矩阵, p, q 为互素的整数且 $q > 1$. 求证: 矩阵方程 $\mathbf{A}\mathbf{x} = \dfrac{p}{q}\mathbf{x}$ 必无非零解.

证明 用反证法. 设上述矩阵方程有非零解,则 $\dfrac{p}{q}$ 为 $\mathbf{A}$ 的特征值,即为特征多项式 $f(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$ 的根. 由于 $\mathbf{A}$ 是整数矩阵,故 $f(\lambda)$ 为整系数多项式. 由整系数多项式有有理根的必要条件可知 $q \mid 1$,从而 $q = \pm 1$,这与假设矛盾.

命题 2.3

设 n 阶矩阵 $\mathbf{A}$ 的全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n, f(x)$ 是一个多项式,求证: $f(\mathbf{A})$ 的全体特征值为 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$.

证明 因为任一 n 阶矩阵均复相似于上三角矩阵,故可设

$$\mathbf{P}^{-1}\mathbf{AP}=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

注意到上三角矩阵的和、数乘及乘方仍是上三角矩阵,经计算可得

$$\mathbf{P}^{-1}f(\mathbf{A})\mathbf{P}=f\left(\mathbf{P}^{-1}\mathbf{AP}\right)=\begin{pmatrix}f(\lambda_1) & * & \cdots & * \\ 0 & f(\lambda_2) & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(\lambda_n)\end{pmatrix}$$

因此 $f(\mathbf{A})$ 的全体特征值为 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$.

命题 2.4

求下列循环矩阵的特征值:

$$\mathbf{A}=\begin{pmatrix}a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1\end{pmatrix}$$

解 设 $\mathbf{J}=\begin{pmatrix}\mathbf{O} & \mathbf{I}_{n-1} \\ 1 & \mathbf{O}\end{pmatrix}, f(x)=a_1+a_2x+a_3x^2+\cdots+a_nx^{n-1}$,可知 $\mathbf{A}=f(\mathbf{J})$. 经简单计算可得 $|\lambda\mathbf{I}_n-\mathbf{J}|=\lambda^n-1$,于是 $\mathbf{J}$ 的特征值为

$$\omega_k=\cos\frac{2k\pi}{n}+\mathrm{i}\sin\frac{2k\pi}{n},\ 0\le k\le n-1$$

因此 $\mathbf{A}$ 的特征值为 $f(1), f(\omega_1), \cdots, f(\omega_{n-1})$.

命题 2.5

求证: n 阶矩阵 $\mathbf{A}$ 为幂零矩阵的充要条件是 $\mathbf{A}$ 的特征值全为零.

证明 若 $\mathbf{A}$ 为幂零矩阵,即存在正整数 k ,使得 $\mathbf{A}^k=\mathbf{O}$,则 $\mathbf{A}$ 的任一特征值 λ_0 也适合 x^k ,于是 $\lambda_0=0$. 反之,若 $\mathbf{A}$ 的特征值全为零,则存在可逆矩阵 $\mathbf{P}$,使得 $P^{-1}AP=B$ 为上三角矩阵且主对角元素全为零.

可知 $B^n=O$,于是 $\mathbf{A}^n=\left(\mathbf{PBP}^{-1}\right)^n=\mathbf{PB}^n\mathbf{P}^{-1}=\mathbf{O}$,即 $\mathbf{A}$ 为幂零矩阵. 也可以利用 Cayley-Hamilton 定理来证明,由于 $\mathbf{A}$ 的特征值全为零,故其特征多项式为 λ^n ,从而 $\mathbf{A}^n=O$.

2.2 对角化

2.3 极小多项式与 Clay-Hamilton 定理

2.4 特征值的估计

第 3 章 相似标准型

3.1 多项式矩阵

定义 3.1 (多项式矩阵(λ - 矩阵))

下列形式的矩阵:

$$\mathbf{A}(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix}$$

其中 $a_{ij}(\lambda)$ 是以 λ 为未定元的数域 $\mathbb{K}$ 上的多项式,称为**多项式矩阵**,或 **λ -矩阵**.

λ -矩阵的加法、数乘及乘法与数域上的矩阵运算一样,只需在运算过程中将数的运算代之以多项式运算即可.

♣

定义 3.2 (λ - 矩阵的初等行变换)

对 λ -矩阵 $\mathbf{A}(\lambda)$ 施行的下列 3 种变换称为 λ -矩阵的**初等行变换**:

(1) 将 $\mathbf{A}(\lambda)$ 的两行对换;

(2) 将 $\mathbf{A}(\lambda)$ 的第 i 行乘以 $\mathbb{K}$ 中的非零常数 c;

(3) 将 $\mathbf{A}(\lambda)$ 的第 i 行乘以 $\mathbb{K}$ 上的多项式 $f(\lambda)$ 后加到第 j 行上去.

♣

注 同理可以定义 $\lambda-$ 矩阵的初等列变换.

定义 3.3 (λ - 矩阵的相抵)

若 $\mathbf{A}(\lambda), \mathbf{B}(\lambda)$ 是同阶 λ -矩阵且 $\mathbf{A}(\lambda)$ 经过 λ -矩阵的初等变换后可变为 $\mathbf{B}(\lambda)$,则称 $\mathbf{A}(\lambda)$ 与 $\mathbf{B}(\lambda)$ 相抵.

♣

笔记 与数字矩阵一样, λ -矩阵的相抵关系也是一种**等价关系**,即

(1) $\mathbf{A}(\lambda)$ 与自身相抵;

(2) 若 $\mathbf{A}(\lambda)$ 与 $\mathbf{B}(\lambda)$ 相抵,则 $\mathbf{B}(\lambda)$ 与 $\mathbf{A}(\lambda)$ 相抵;

(3) 若 $\mathbf{A}(\lambda)$ 与 $\mathbf{B}(\lambda)$ 相抵, $\mathbf{B}(\lambda)$ 与 $\mathbf{C}(\lambda)$ 相抵,则 $\mathbf{A}(\lambda)$ 与 $\mathbf{C}(\lambda)$ 相抵.

定义 3.4 (初等 λ - 矩阵)

下列 3 种矩阵称为**初等 λ -矩阵**:

(1) 将 n 阶单位阵的第 i 行与第 j 行对换,记为 $\mathbf{P}_{ij}$;

(2) 将 n 阶单位阵的第 i 行乘以非零常数 c ,记为 $\mathbf{P}_i(c)$;

(3) 将 n 阶单位阵的第 i 行乘以多项式 $f(\lambda)$ 后加到第 j 行上去得到的矩阵，记为 $T_{ij}(f(\lambda))$.

♣

注 第一类与第二类初等 λ -矩阵与数域上的第一类与第二类初等矩阵没有什么区别. 第三类初等 λ -矩阵的形状如下:

$$
\mathbf{T}_{ij}(f(\lambda)) = \begin{pmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & & & \\
& & \vdots & \ddots & & & \\
& & f(\lambda) & \cdots & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{pmatrix}
$$

注 下列 λ -矩阵的变换不是 λ -矩阵的初等变换:

$$
\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} \lambda & \lambda \\ 0 & 1 \end{pmatrix}.
$$

这是因为前面一个矩阵的第一行乘以 λ 不是 λ -矩阵的初等变换.

同理下面的变换需第一行乘以 λ^{-1} ,因此也不是 λ -矩阵的初等变换:

$$
\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

笔记 类似数字矩阵, λ -矩阵的初等变换也对应于初等 λ -矩阵的相乘.

定理 3.1

对 λ -矩阵 $\mathbf{A}(\lambda)$ 施行第 $k\,(k=1,2,3)$ 类初等行 (列) 变换等于用第 k 类初等 λ -矩阵左 (右) 乘以 $\mathbf{A}(\lambda)$.

♡

证明 与数字矩阵的证明完全一样.

定义 3.5 (λ - 矩阵的逆矩阵)

若 $\mathbf{A}(\lambda)$, $\mathbf{B}(\lambda)$ 都是 n 阶 λ -矩阵,且

$$
\mathbf{A}(\lambda)\mathbf{B}(\lambda) = \mathbf{B}(\lambda)\mathbf{A}(\lambda) = \mathbf{I}_n
$$

则称 $\mathbf{B}(\lambda)$ 是 $\mathbf{A}(\lambda)$ 的逆 λ -矩阵.

这时称 $\mathbf{A}(\lambda)$ 为可逆 λ -矩阵,在不引起混淆的情形下, 有时简称为可逆阵.

♣

注 注意不要将数字矩阵中的一些结论随意搬到 λ -矩阵上. 比如下面的 λ -矩阵的行列式不为零,但它不是可逆 λ -矩阵:

$$
\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}
$$

这是因为下面的矩阵不是 λ -矩阵.

$$
\begin{pmatrix} \lambda^{-1} & 0 \\ 0 & 1 \end{pmatrix}
$$

命题 3.1

设 $\mathbf{M}(\lambda)$ 是一个 n 阶 λ-矩阵,则 $\mathbf{M}(\lambda)$ 可以化为如下形状:

$$\mathbf{M}(\lambda)=\mathbf{M}_m\lambda^m+\mathbf{M}_{m-1}\lambda^{m-1}+\cdots+\mathbf{M}_0$$

其中 $\mathbf{M}_i$ 为数域 $\mathbb{K}$ 上的 n 阶数字矩阵. 因此,一个多项式矩阵可以化为系数为矩阵的多项式, 反之亦然.

引理 3.1

设 $\mathbf{M}(\lambda)$ 与 $\mathbf{N}(\lambda)$ 是两个 n 阶 λ-矩阵且都不等于零. 又设 $\mathbf{B}$ 为 n 阶数字矩阵,则必存在 λ-矩阵 $\mathbf{Q}(\lambda)$ 及 $\mathbf{S}(\lambda)$ 和数字矩阵 $\mathbf{R}$ 及 $\mathbf{T}$,使下式成立:

$$\mathbf{M}(\lambda)=(\lambda\mathbf{I}-\mathbf{B})\mathbf{Q}(\lambda)+\mathbf{R} \tag{3.1}$$

$$\mathbf{N}(\lambda)=\mathbf{S}(\lambda)(\lambda\mathbf{I}-\mathbf{B})+\mathbf{T} \tag{3.2}$$

证明 将 $\mathbf{M}(\lambda)$ 写为

$$\mathbf{M}(\lambda)=\mathbf{M}_m\lambda^m+\mathbf{M}_{m-1}\lambda^{m-1}+\cdots+\mathbf{M}_0$$

其中 $\mathbf{M}_m\neq\mathbf{O}$. 可对 m 用归纳法,若 $m=0$,则已适合要求 (取 $\mathbf{Q}(\lambda)=\mathbf{O}$). 现设对小于 m 次的矩阵多项式有 (3.1) 式成立. 令

$$\mathbf{Q}_1(\lambda)=\mathbf{M}_m\lambda^{m-1}$$

则

$$\mathbf{M}(\lambda)-(\lambda\mathbf{I}-\mathbf{B})\mathbf{Q}_1(\lambda)=(\mathbf{B}\mathbf{M}_m+\mathbf{M}_{m-1})\lambda^{m-1}+\cdots+\mathbf{M}_0$$

上式是一个次数小于 m 的矩阵多项式,由归纳假设得

$$\mathbf{M}(\lambda)-(\lambda\mathbf{I}-\mathbf{B})\mathbf{Q}_1(\lambda)=(\lambda\mathbf{I}-\mathbf{B})\mathbf{Q}_2(\lambda)+\mathbf{R}$$

于是

$$\mathbf{M}(\lambda)=(\lambda\mathbf{I}-\mathbf{B})[\mathbf{Q}_1(\lambda)+\mathbf{Q}_2(\lambda)]+\mathbf{R}$$

令 $\mathbf{Q}(\lambda)=\mathbf{Q}_1(\lambda)+\mathbf{Q}_2(\lambda)$ 即得 (3.1) 式. 同理可证 (3.2) 式.

定理 3.2

设 A,B 是数域 $\mathbb{K}$ 上的矩阵,则 A 与 B 相似的充分必要条件是 λ-矩阵 $\lambda\mathbf{I}-\mathbf{A}$ 与 $\lambda\mathbf{I}-\mathbf{B}$ 相抵.

证明 若 $\mathbf{A}$ 与 $\mathbf{B}$ 相似,则存在 $\mathbb{K}$ 上的非异阵 $\mathbf{P}$,使 $\mathbf{B}=\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$,于是

$$P^{-1}(\lambda I-A)P=\lambda I-P^{-1}AP=\lambda I-B \tag{3.3}$$

把 $\mathbf{P}$ 看成是常数 λ-矩阵,上式表明 $\lambda\mathbf{I}-\mathbf{A}$ 与 $\lambda\mathbf{I}-\mathbf{B}$ 相抵.
反过来,若 $\lambda\mathbf{I}-\mathbf{A}$ 与 $\lambda\mathbf{I}-\mathbf{B}$ 相抵,即存在 $\mathbf{M}(\lambda)$ 及 $\mathbf{N}(\lambda)$,使

$$\mathbf{M}(\lambda)(\lambda\mathbf{I}-\mathbf{A})\mathbf{N}(\lambda)=\lambda\mathbf{I}-\mathbf{B} \tag{3.4}$$

其中 $\mathbf{M}(\lambda)$ 与 $\mathbf{N}(\lambda)$ 都是有限个初等矩阵之积,因而都是可逆阵. 因此可将 (3.4) 式写为

$$\mathbf{M}(\lambda)(\lambda\mathbf{I}-\mathbf{A})=(\lambda\mathbf{I}-\mathbf{B})\mathbf{N}(\lambda)^{-1} \tag{3.5}$$

由引理 3.2 可设

$$\mathbf{M}(\lambda)=(\lambda\mathbf{I}-\mathbf{B})\mathbf{Q}(\lambda)+\mathbf{R} \tag{3.6}$$

代入 (3.5) 式经整理得

$$\mathbf{R}(\lambda\mathbf{I}-\mathbf{A})=(\lambda\mathbf{I}-\mathbf{B})\left[\mathbf{N}(\lambda)^{-1}-\mathbf{Q}(\lambda)(\lambda\mathbf{I}-\mathbf{A})\right] \tag{3.7}$$

上式左边是次数小于等于 1 的矩阵多项式, 因此上式右边中括号内的矩阵多项式的次数必须小于等于零,也即必是一个常数矩阵,设为 $\mathbf{P}$. 于是

$$\mathbf{R}(\lambda\mathbf{I}-\mathbf{A})=(\lambda\mathbf{I}-\mathbf{B})\mathbf{P} \tag{3.8}$$

(3.8) 式又可整理为

$$(\mathbf{R}-\mathbf{P})\lambda=\mathbf{RA}-\mathbf{BP}$$

再次比较次数得 $\mathbf{R}=\mathbf{P},\mathbf{RA}=\mathbf{BP}$. 现只需证明 $\mathbf{P}$ 是一个非异阵即可. 由假设

$$\mathbf{P}=\mathbf{N}(\lambda)^{-1}-\mathbf{Q}(\lambda)(\lambda\mathbf{I}-\mathbf{A})$$

将上式两边右乘 $N(\lambda)$ 并移项得

$$\mathbf{PN}(\lambda)+\mathbf{Q}(\lambda)(\lambda\mathbf{I}-\mathbf{A})\mathbf{N}(\lambda)=\mathbf{I}$$

但

$$(\lambda\mathbf{I}-\mathbf{A})\mathbf{N}(\lambda)=\mathbf{M}(\lambda)^{-1}(\lambda\mathbf{I}-\mathbf{B})$$

因此

$$\mathbf{PN}(\lambda)+\mathbf{Q}(\lambda)\mathbf{M}(\lambda)^{-1}(\lambda\mathbf{I}-\mathbf{B})=\mathbf{I} \tag{3.9}$$

再由引理 3.2 可设

$$\mathbf{N}(\lambda)=\mathbf{S}(\lambda)(\lambda\mathbf{I}-\mathbf{B})+\mathbf{T}$$

代入 (3.9) 式并整理得

$$\left[\mathbf{PS}(\lambda)+\mathbf{Q}(\lambda)\mathbf{M}(\lambda)^{-1}\right](\lambda\mathbf{I}-\mathbf{B})=\mathbf{I}-\mathbf{PT}$$

上式右边是次数小于等于零的矩阵多项式,因此上式左边中括号内的矩阵多项式必须为零,从而 $\mathbf{PT}=\mathbf{I}$,即 $\mathbf{P}$ 是非异阵.

例题 3.1 证明引理 3.2 中的余式 R 及 T 是唯一确定的, $Q(\lambda)$ 与 $S(\lambda)$ 也唯一确定.

证明 反证法假设存在 $Q(\lambda), Q^{\#}(\lambda), R, R^{\#}$ 使得:

$$(\lambda I-B)Q(\lambda)+R=(\lambda I-B)Q^{\#}(\lambda)+R^{\#}$$

变形得:

$$(\lambda I-B)(Q(\lambda)-Q^{\#}(\lambda))=R^{\#}-R$$

设:

$$Q(\lambda)-Q^{\#}(\lambda)=\sum_{k=0}^{n}M_k\lambda^k$$

则有

$$(\lambda I-B)(Q(\lambda)-Q^{\#}(\lambda))=(\lambda I-B)\sum_{k=0}^{n}M_k\lambda^k=M_n\lambda^{n+1}+\sum_{k=1}^{n}(M_{k-1}-BM_k)\lambda^k-BM_0$$

由$(\lambda I-B)(Q(\lambda)-Q^{\#}(\lambda))=R^{\#}-R$ 为数字矩阵知道:

$$\left\{\begin{array}{lcl}M_0 & = & 0\\ M_{n-1}-BM_n & = & 0\\ M_{n-2}-BM_{n-1} & = & 0\\ \vdots & & \\ M_0-BM_1 & = & 0\end{array}\right. \Longrightarrow \left\{\begin{array}{lcl}M_n & = & 0\\ M_{n-1} & = & 0\\ M_{n-2} & = & 0\\ \vdots & & \\ M_0 & = & 0\end{array}\right.$$

所以我们有:

$$Q(\lambda)-Q^{\#}(\lambda)=0\Longrightarrow Q(\lambda)=Q^{\#}(\lambda)\Longrightarrow R=R^{\#}$$

3.2 矩阵的法式

引理 3.2

设 $\mathbf{A}(\lambda) = (a_{ij}(\lambda))_{m\times n}$ 是任一非零 λ -矩阵,则 $\mathbf{A}(\lambda)$ 必相抵于这样的一个 λ -矩阵 $\mathbf{B}(\lambda) = (b_{ij}(\lambda))_{m\times n}$,其中 $b_{11}(\lambda) \neq 0$ 且 $b_{11}(\lambda)$ 可整除 $\mathbf{B}(\lambda)$ 中的任一元素 $b_{ij}(\lambda)$.

证明 设 $k = \min\{\deg a_{ij}(\lambda) \mid a_{ij}(\lambda) \neq 0, 1 \leq i \leq m; 1 \leq j \leq n\}$,我们对 k 用数学归纳法. 首先,经行对换及列对换可将 $\mathbf{A}(\lambda)$ 的第 $(1,1)$ 元素变成次数最低的非零多项式,因此不妨设 $a_{11}(\lambda) \neq 0$ 且 $\deg a_{11}(\lambda) = k$. 若 $k = 0$,则 $a_{11}(\lambda)$ 是一个非零常数,结论显然成立. 假设对非零元素次数的最小值小于 k 的任一 λ -矩阵,引理的结论成立,现考虑非零元素次数的最小值等于 k 的 λ -矩阵 $\mathbf{A}(\lambda)$. 若 $a_{11}(\lambda)$ 可整除所有的 $a_{ij}(\lambda)$,则结论已成立. 若否,设在第一列中有元素 $a_{i1}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除,作带余除法:

$$a_{i1}(\lambda) = a_{11}(\lambda) q(\lambda) + r(\lambda)$$

用 $-q(\lambda)$ 乘以第一行加到第 i 行上,第 $(i,1)$ 元素就变为 $r(\lambda)$. 注意到 $r(\lambda) \neq 0$ 且 $\deg r(\lambda) < \deg a_{11}(\lambda) = k$,由归纳假设即知结论成立.

同样的方法可施于第一行. 因此我们不妨设 $a_{11}(\lambda)$ 可整除第一行及第一列. 这时,设 $a_{21}(\lambda) = a_{11}(\lambda) g(\lambda)$. 将第一行乘以 $-g(\lambda)$ 加到第二行上,则第 $(2,1)$ 元素变为零. 用同样的方法可消去第一行、第一列除 $a_{11}(\lambda)$ 以外的所有元素,于是 $\mathbf{A}(\lambda)$ 经初等变换后变成下列形状:

$$\begin{pmatrix} a_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & a'_{22}(\lambda) & \cdots & a'_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & a'_{m2}(\lambda) & \cdots & a'_{mn}(\lambda) \end{pmatrix} \tag{3.10}$$

这时,若 $a_{11}(\lambda)$ 可整除所有其他元素,则结论已成立. 若否,比如 $a_{11}(\lambda)$ 不能整除 $a'_{ij}(\lambda)$,则将第 i 行加到第一行上去,这时在第一行又出现了一元素 $a'_{ij}(\lambda)$,它不能被 $a_{11}(\lambda)$ 整除. 重复上面的做法,通过归纳假设即可得到结论.

第 4 章 二次型

4.1 二次型的化简与矩阵的合同

定义 4.1 (二次型)

设 f 是数域 $\mathbb{K}$ 上的 n 元二次齐次多项式:

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n$$

$$+a_{22}x_2^2 + \cdots + 2a_{2n}x_2x_n + \cdots + a_{nn}x_n^2,$$

称 f 为数域 $\mathbb{K}$ 上的 n 元二次型,简称二次型.

♣

定义 4.2 (二次型的矩阵)

用矩阵的乘法我们可以把二次型写成矩阵相乘的形式:

$$f(x_1, x_2, \cdots, x_n) = \mathbf{x}'\mathbf{A}\mathbf{x}$$

其中

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

在矩阵 $\mathbf{A}$ 中, $a_{ij} = a_{ji}$ 对一切 i, j 成立,也就是说矩阵 $\mathbf{A}$ 是一个对称阵.

♣

笔记 由此可知,给定数域 $\mathbb{K}$ 上的一个 n 元二次型,我们就得到了 $\mathbb{K}$ 上的一个 n 阶对称阵 $\mathbf{A}$, 称为该二次型的相伴矩阵或系数矩阵. 反过来,若给定 $\mathbb{K}$ 上的一个 n 阶对称阵 $\mathbf{A}$, 则对称矩阵,我们可以得到 $\mathbb{K}$ 上的一个二次型,称为对称阵 $\mathbf{A}$ 的相伴二次型. 现在的问题是: 用这样的方法得到的对称阵是否和二次型一一对应?

回答这一点并不难. 设 $f = x'Ax = x'Bx$,我们要证明 $A = B$. 这等价于证明下面的结论:设 $\mathbf{A} = (a_{ij})$ 是 n 阶对称阵,若 $\alpha'\mathbf{A}\alpha = 0$ 对一切 α 成立,则 $\mathbf{A} = \mathbf{O}$.

令 $\alpha = \mathbf{e}_i$ 是 n 维标准单位列向量,则 $a_{ii} = \mathbf{e}_i'\mathbf{A}\mathbf{e}_i = 0$. 再令 $\alpha = \mathbf{e}_i + \mathbf{e}_j\ (i \neq j)$,则

$$0 = (e_i + e_j)' A (e_i + e_j) = e_i'Ae_i + e_j'Ae_j + e_i'Ae_j + e_j'Ae_i = a_{ij} + a_{ji},$$

因为 $a_{ij} = a_{ji}$,故 $a_{ij} = 0\,(i \neq j)$,于是 $\mathbf{A} = \mathbf{O}$. 这表明用对称阵来表示二次型时系数矩阵是唯一的.

事实上, 如果我们不限制矩阵是对称阵, 则系数矩阵将不唯一, 这样会给用矩阵方法研究二次型带来困难.

二次型理论的基本问题是要寻找一个线性变换把它变成只含平方项的标准型. 由上面我们知道, 二次型与对称阵一一对应, 而线性变换可以用矩阵来表示. 自然地, 二次型的变换与矩阵有着密切的关系, 现在我们来探讨这个关系.

设 V 是 n 维线性空间,二次型 $f(x_1, x_2, \cdots, x_n)$ 可以看成是 V 上的二次函数. 即若设 V 的一组基为 $\{e_1, e_2, \cdots, e_n\}$,向量 α 在这组基下的坐标向量为 $\mathbf{x} = (x_1, x_2, \cdots, x_n)'$,则 f 便是向量 α 的函数. 现假设 $\{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_n\}$ 是 V 的另一组基,向量 α 在 $\{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_n\}$ 下的坐标向量为 $\mathbf{y} = (y_1, y_2, \cdots, y_n)'$. 记 $C = (c_{ij})$ 是从基 $\{e_1, e_2, \cdots, e_n\}$ 到基 $\{f_1, f_2, \cdots, f_n\}$ 的过渡矩阵,则

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

或简记为

$$x = Cy$$

将上式代入 $f(x_1, x_2, \cdots, x_n) = \mathbf{x}'\mathbf{A}\mathbf{x}$,得

$$f(x_1, x_2, \cdots, x_n) = \mathbf{y}'\mathbf{C}'\mathbf{A}\mathbf{C}\mathbf{y}.$$

显然, $C'AC$ 仍是一个对称阵,故 $y'C'ACy$ 是以 $y_1, y_2, \cdots, y_n$ 为变元的二次型, 记为 $g(y_1, y_2, \cdots, y_n)$. 由此我们可看出: 若二次型 $f(x_1, x_2, \cdots, x_n)$ 所对应的对称阵为 $\mathbf{A}$,则经过变量代换之后得到的二次型 $g(y_1, y_2, \cdots, y_n)$ 所对应的对称阵为 $C'AC$.

定义 4.3 (合同关系)

设 A, B 是数域 $\mathbb{K}$ 上的 n 阶矩阵,若存在 n 阶非异阵 C ,使

$$B = C'AC$$

则称 $\mathbf{B}$ 与 $\mathbf{A}$ 是合同的,或称 $\mathbf{B}$ 与 $\mathbf{A}$ 具有合同关系. ♣

笔记 不难证明, 合同关系是一个等价关系, 即

(1) 任一矩阵 $\mathbf{A}$ 与自己合同,因为 $\mathbf{A} = \mathbf{I}'\mathbf{A}\mathbf{I}$;

(2) 若 B 与 A 合同,则 A 与 B 合同. 这是因为若 $B = C'AC$,则 $A = (C')^{-1}BC^{-1} = \left(C^{-1}\right)'BC^{-1}$;

(3) 若 B 与 A 合同, D 与 B 合同,则 D 与 A 合同. 事实上,若 $B = C'AC$, $\mathbf{D} = \mathbf{H}'\mathbf{B}\mathbf{H}$,则 $\mathbf{D} = \mathbf{H}'\mathbf{C}'\mathbf{A}\mathbf{C}\mathbf{H} = (\mathbf{C}\mathbf{H})'\mathbf{A}(\mathbf{C}\mathbf{H})$.

引理 4.1

对称阵 A 的下列变换都是合同变换:

(1) 对换 $\mathbf{A}$ 的第 i 行与第 j 行,再对换第 i 列与第 j 列;

(2) 将非零常数 k 乘以 $\mathbf{A}$ 的第 i 行,再将 k 乘以第 i 列;

(3) 将 A 的第 i 行乘以 k 加到第 j 行上,再将第 i 列乘以 k 加到第 j 列上. ♡

证明 上述变换相当于将一个初等矩阵左乘以 $\mathbf{A}$ 后再将这个初等矩阵的转置右乘之, 因此是合同变换.

引理 4.2

设 A 是数域 $\mathbb{K}$ 上的非零对称阵,则必存在非异阵 C ,使 $C'AC$ 的第 $(1,1)$ 元素不等于零.

证明 若 $a_{11}=0$,而 $a_{ii}\neq 0$,则将 $\mathbf{A}$ 的第一行与第 i 行对换,再将第一列与第 i 列对换,得到的矩阵的第 $(1,1)$ 元素不为零. 根据上述引理,这样得到的矩阵和原矩阵合同.

若所有的 $a_{ii}=0\,(i=1,2,\cdots,n)$,设 $a_{ij}\neq 0\,(i\neq j)$,将 $\mathbf{A}$ 的第 j 行加到第 i 行上,再将第 j 列加到第 i 列上. 因为 $\mathbf{A}$ 是对称阵, $a_{ij}=a_{ji}\neq 0$,于是第 (i,i) 元素是 $2a_{ij}\neq 0$,再用前面的办法使第 $(1,1)$ 元素不等于零. 根据上述引理,这样得到的矩阵和原矩阵仍合同, 这就证明了结论.

定理 4.1 (合同对角化)

设 A 是数域 $\mathbb{K}$ 上的 n 阶对称阵,则必存在 $\mathbb{K}$ 上的 n 阶非异阵 C ,使 $C'AC$ 为对角阵.

证明 由上述引理,不妨设 $\mathbf{A}=(a_{ij})$ 中 $a_{11}\neq 0$. 若 $a_{i1}\neq 0$,则可将第一行乘以 $-a_{11}^{-1}a_{i1}$ 加到第 i 行上,再将第一列乘以 $-a_{11}^{-1}a_{i1}$ 加到第 i 列上. 由于 $a_{i1}=a_{1i}$,故得到的矩阵的第 $(1,i)$ 元素及第 $(i,1)$ 元素均等于零. 由引理 4.1 可知,新得到的矩阵与 $\mathbf{A}$ 是合同的. 依次这样做下去,可把 $\mathbf{A}$ 的第一行与第一列除 a_{11} 外的元素都消去,于是 $\mathbf{A}$ 合同于下列矩阵:

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} \\ 0 & b_{32} & b_{33} & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{n2} & b_{n3} & \cdots & b_{nn} \end{pmatrix}$$

上式右下角是一个 $n-1$ 阶对称阵,记为 $\mathbf{A}_1$. 因此由归纳假设,存在 $n-1$ 阶非异阵 $\mathbf{D}$,使 $\mathbf{D}'\mathbf{A}_1\mathbf{D}$ 为对角阵,于是

$$\begin{pmatrix} 1 & O \\ O & D' \end{pmatrix}\begin{pmatrix} a_{11} & O \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & D \end{pmatrix}=\begin{pmatrix} a_{11} & O \\ O & D'A_1D \end{pmatrix}$$

是一个对角阵. 显然

$$\begin{pmatrix} 1 & O \\ O & D' \end{pmatrix}=\begin{pmatrix} 1 & O \\ O & D \end{pmatrix}'$$

于是 A 合同于对角阵.

命题 4.1

设矩阵 $\mathbf{A}$ 和 $\mathbf{B}$ 合同,求证: $\mathbf{A}^*$ 和 $\mathbf{B}^*$ 也合同.

证明 由于 $\mathbf{A}$ 与 $\mathbf{B}$ 合同,故存在非异阵 $\mathbf{C}$,使得 $\mathbf{B}=\mathbf{C}'\mathbf{A}\mathbf{C}$. 等式两边同时取伴随,可得 $B^*=(C'AC)^*=C^*A^*(C')^*=C^*A^*(C^*)'$. 注意到 C^* 也是非异阵,故 $\mathbf{A}^*$ 和 $\mathbf{B}^*$ 合同.

4.2 二次型的化简

二次型的化简有两种办法，一种是配方法，一种是初等变换法.

配方法的思想在于若没有平方项，则通过换元构造出平方项，若存在平方项，则以某一个平方项对应的文字为主元，进行配方，从而归结到 $n-1$ 个文字的情况，如此反复最终配凑成平方数的和.

下面我们介绍初等变换法，初等变换法的依据是引理 4.1.

这种方法可总结如下: 作 $n\times 2n$ 矩阵 $(\mathbf{A}!\mathbf{I}_n)$,对这个矩阵实施初等行变换, 同时施以同样的初等列变换, 将它左半边化为对角阵, 则这个对角阵就是已化简的二次型的相伴矩阵,右半边的转置便是变换矩阵 $\mathbf{C}$.

如碰到第 $(1,1)$ 元素是零的矩阵,可先设法将第 $(1,1)$ 元素化成非零,再进行上述过程.

我们用一个例子来展示这种做法:

例题 4.1 将二次型 $f\left(x_1,x_2,x_3\right)=2x_1x_2+4x_1x_3-4x_2x_3$ 化成对角型.

解 写出与 f 相伴的对称阵 $\mathbf{A}$,作 $(\mathbf{A}\quad \mathbf{I}_3)$ 并将它的第二行加到第一行上,再将第二列加到第一列上:

$$(\mathbf{A}!\mathbf{I}_3)=\begin{pmatrix}0&1&2&1&0&0\\1&0&-2&0&1&0\\2&-2&0&0&0&1\end{pmatrix}\to\begin{pmatrix}2&1&0&1&1&0\\1&0&-2&0&1&0\\0&-2&0&0&0&1\end{pmatrix}.$$

对上述矩阵进行初等变换得到

$$\begin{pmatrix}2&0&0&1&1&0\\0&-\dfrac{1}{2}&0&-\dfrac{1}{2}&\dfrac{1}{2}&0\\0&0&8&2&-2&1\end{pmatrix}$$

因此 f 化简为

$$2y_1^2-\frac{1}{2}y_2^2+8y_3^2$$

其中

$$C=\begin{pmatrix}1&-\frac{1}{2}&2\\1&\frac{1}{2}&-2\\0&0&1\end{pmatrix}$$

4.3 惯性定理

我们已经知道,任意一个实对称阵 $\mathbf{A}$ 必合同于一个对角阵:

$$\mathbf{C}'\mathbf{A}\mathbf{C}=\operatorname{diag}\left\{d_1,d_2,\cdots,d_r,0,\cdots,0\right\}$$

其中 $d_i\neq 0\,(i=1,\cdots,r)$. 注意到 $\mathbf{C}$ 是可逆阵,故 $r=\mathrm{r}\left(\mathbf{C}'\mathbf{A}\mathbf{C}\right)=\mathrm{r}\left(\mathbf{A}\right)$,即秩 r 是矩阵合同关系下的一个不变量. 如同相似标准型一样, 我们的目的是要找出实对称阵在合同关系下的全系不变量, 即找出一组足以判断两个实对称阵是否合同的合同不变量. 我们不妨设实对称阵已具有下列对角阵的形状:

$$A=\operatorname{diag}\left\{d_1,d_2,\cdots,d_r,0,\cdots,0\right\}$$

由引理 4.1 不难知道,任意调换 $\mathbf{A}$ 的主对角线上的元素得到的矩阵仍与 $\mathbf{A}$ 合同. 因此我们可把零放在一起,把正项与负项放在一起,即可设 $d_1>0,\cdots,d_p>0\,;d_{p+1}<0,\cdots,d_r<0$.$\mathbf{A}$ 所代表的二次型为

$$f\left(x_1,x_2,\cdots,x_n\right)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2 \tag{4.1}$$

令

$$\begin{cases} y_1=\sqrt{d_1}x_1,\cdots,y_p=\sqrt{d_p}x_p; \\ y_{p+1}=\sqrt{-d_{p+1}}x_{p+1},\cdots,y_r=\sqrt{-d_r}x_r; \\ y_j=x_j\,(j=r+1,\cdots,n)\,, \end{cases}$$

则 (4.1) 式变为

$$f=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2 \tag{4.2}$$

这一事实等价于说 $\mathbf{A}$ 合同于下列对角阵:

$$\operatorname{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\}$$

其中有 p 个 $1,q$ 个 $-1,n-r$ 个零.

定理 4.2 (惯性定理)

设 $f\left(x_1,x_2,\cdots,x_n\right)$ 是一个 n 元实二次型,且 f 可化为两个标准型:

$$c_1y_1^2+\cdots+c_py_p^2-c_{p+1}y_{p+1}^2-\cdots-c_ry_r^2$$

$$d_1z_1^2+\cdots+d_kz_k^2-d_{k+1}z_{k+1}^2-\cdots-d_rz_r^2$$

其中 $c_i>0,d_i>0$,则必有 $p=k$.

证明 用反证法,设 $p>k$. 由前面的说明不妨设 c_i 及 d_i 均为 1,因此

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2=z_1^2+\cdots+z_k^2-z_{k+1}^2-\cdots-z_r^2 \tag{4.3}$$

又设

$$x=By,\ x=Cz$$

其中

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\ \mathbf{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},\ \mathbf{z}=\begin{pmatrix}z_1\\z_2\\\vdots\\z_n\end{pmatrix}$$

于是 $z=C^{-1}By$. 令

$$\mathbf{C}^{-1}\mathbf{B}=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{pmatrix}$$

则

$$\begin{cases}z_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n\\z_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n\\\qquad\vdots\\z_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n\end{cases}$$

因为 $p>k$,故齐次线性方程组

$$\begin{cases}c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n&=&0\\\qquad\vdots\\c_{k1}y_1+c_{k2}y_2+\cdots+c_{kn}y_n&=&0\\\qquad y_{p+1}&=&0\\\qquad\vdots\\\qquad y_n&=&0\end{cases}$$

必有非零解 (n 个未知数, $n-(p-k)$ 个方程式). 令其中一个非零解为 $y_1=a_1,\cdots,y_p=a_p,y_{p+1}=0,\cdots,y_n=0$,把这组解代入 (4.3) 式左边得到

$$a_1^2+\cdots+a_p^2>0.$$

但这时 $z_1=\cdots=z_k=0$,故 (4.3) 式右边将小于等于零,引出了矛盾. 同理可证 $p<k$ 也不可能.

注 (4.2) 式中的二次型称为 f 的**规范标准型**.

定义 4.4

设 $f\left(x_1,x_2,\cdots,x_n\right)$ 是一个实二次型,若它能化为形如 (4.2) 式的形状,则称 r 是该二次型的秩, p 是它的正惯性指数, $q=r-p$ 是它的负惯性指数, $s=p-q$ 称为 f 的符号差.

笔记 显然,若已知秩 r 与符号差 s ,则 $p=\frac{1}{2}(r+s)$, $q=\frac{1}{2}(r-s)$. 事实上,在 p,q,r,s 中只需知道其中两个数,其余两个数也就知道了. 由于实对称阵与实二次型之间的等价关系, 我们将实二次型的秩、惯性指数及符号差也称为相应的实对称阵的秩、惯性指数及符号差.

定理 4.3

秩与符号差 (或正负惯性指数) 是实对称阵在合同关系下的全系不变量.

证明 由上面的定理知道,秩 r 与符号差 s 是实对称阵合同关系的不变量. 反之,若 n 阶实对称阵 $\mathbf{A},\mathbf{B}$ 的秩都为 r ,符号差都是 s ,则它们都合同于

$$\operatorname{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\}$$

其中有 $p=\frac{1}{2}(r+s)$ 个 1, $q=\frac{1}{2}(r-s)$ 个 -1 及 $n-r$ 个零,因此 $\mathbf{A}$ 与 $\mathbf{B}$ 合同.

对正负惯性指数的结论也同样成立.

复二次型要比实二次型简单得多.

命题 4.2

复二次型

$$f\left(x_1,x_2,\cdots,x_n\right)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2$$

必可化为

$$z_1^2+z_2^2+\cdots+z_r^2$$

其中 $z_i=\sqrt{d_i}x_i\ (i=1,2,\cdots,r)$, $z_j=x_j\ (j=r+1,\cdots,n)$.

所以复对称阵的合同关系只有一个全系不变量,那就是秩 r .

例题 4.2 把互相合同的实对称阵作为一个类,问: n 阶实对称阵共有多少个类?

解 讨论正负惯性指数即可, 共有 $\frac{(n+1)(n+2)}{2}$ 类.

例题 4.3 实对称阵 $\mathbf{A}$ 和它的伴随阵 $\mathbf{A}^*$ 一定合同吗?

解 令 $\mathbf{A}=\operatorname{diag}\{1,1,-1\}$,则 $\mathbf{A}^*=\operatorname{diag}\{-1,-1,1\}$, $\mathbf{A}$ 和 $\mathbf{A}^*$ 的正惯性指数不相同,因此不合同.

4.4 正定型与正定矩阵

定义 4.5

设 $f(x_1, x_2, \cdots, x_n) = \mathbf{x}'\mathbf{A}\mathbf{x}$ 是 n 元实二次型, $\mathbf{A}$ 是相伴矩阵.

(1) 若对任意 n 维非零列向量 α 均有 $\alpha'\mathbf{A}\alpha > 0$,则称 f 是正定二次型 (简称**正定型**),矩阵 A 称为正定矩阵 (简称**正定阵**);

(2) 若对任意 n 维非零列向量 α 均有 $\alpha'\mathbf{A}\alpha < 0$,则称 f 是负定二次型 (简称**负定型**),矩阵 A 称为负定矩阵 (简称**负定阵**);

(3) 若对任意 n 维非零列向量 α 均有 $\alpha'\mathbf{A}\alpha \geq 0$,则称 f 是半正定二次型 (简称**半正定型**),矩阵 $\mathbf{A}$ 称为半正定矩阵 (简称**半正定阵**);

(4) 若对任意 n 维非零列向量 α 均有 $\alpha'\mathbf{A}\alpha \leq 0$,则称 f 是半负定二次型 (简称**半负定型**),矩阵 $\mathbf{A}$ 称为半负定矩阵 (简称**半负定阵**);

(5) 若存在 α,使 $\alpha'\mathbf{A}\alpha > 0$; 又存在 β,使 $\beta'\mathbf{A}\beta < 0$,则称 f 是**不定型**. ♣

定理 4.4

设 $f(x_1, x_2, \cdots, x_n)$ 是 n 元实二次型,则

(1) f 是正定型的充分必要条件是 f 的正惯性指数等于 n;

(2)f 是负定型的充分必要条件是 f 的负惯性指数等于 n;

(3)f 是半正定型的充分必要条件是 f 的正惯性指数等于 f 的秩 r;

(4)f 是半负定型的充分必要条件是 f 的负惯性指数等于 f 的秩 r. ♡

证明 若 f 的正惯性指数等于 n,则 f 可化为下列标准型:

$$f = y_1^2 + y_2^2 + \cdots + y_n^2$$

显然 f 是正定型. 反之,若 f 是正定型,如果 f 的正惯性指数 $p < n$,则 f 可化为如下标准型:

$$f = y_1^2 + \cdots + y_p^2 - c_{p+1}y_{p+1}^2 - \cdots - c_n y_n^2$$

其中 $c_j \geq 0\,(j = p+1, \cdots, n)$. 这时令 $b_1 = \cdots = b_p = 0, b_{p+1} = \cdots = b_n = 1$,则 $b_1, b_2, \cdots, b_n$ 不全为零. 假设这时 $\mathbf{x} = \mathbf{C}\mathbf{y}$,其中

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

C 是非异阵,则从 $y_i = b_i\,(i = 1, \cdots, n)$ 可得 $x_i = a_i\,(i = 1, \cdots, n)$ 是一组不全为零的实数. 于是

$$f(a_1, a_2, \cdots, a_n) \leq 0$$

这与 f 是正定型矛盾. 其余结论的证明类似.

定理 4.5

n 阶实对称阵 $\mathbf{A}$ 是正定阵当且仅当它合同于单位阵 $\mathbf{I}_n$； $\mathbf{A}$ 是负定阵当且仅当它合同于 $-\mathbf{I}_n$；
$\mathbf{A}$ 是半正定阵当且仅当 $\mathbf{A}$ 合同于下列对角阵:

$$\begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$$

$\mathbf{A}$ 是半负定阵当且仅当 $\mathbf{A}$ 合同于下列对角阵:

$$\begin{pmatrix} -\mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$$

♡

定义 4.6

设 $A=(a_{ij})$ 是 n 阶矩阵, A 的 n 个子式:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} (k=1,2,\cdots,n)$$

称为 $\mathbf{A}$ 的**顺序主子式**.

♣

定理 4.6

n 阶实对称阵 $\mathbf{A}$ 是正定阵的充分必要条件是它的 n 个顺序主子式全大于零.

♡

证明 先证必要性. 设 n 阶实对称阵 $\mathbf{A}=(a_{ij})$ 为正定阵,则对应的实二次型

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

为正定型. 令

$$f_k(x_1,x_2,\cdots,x_k)=\sum_{j=1}^{k}\sum_{i=1}^{k}a_{ij}x_ix_j$$

则对任意一组不全为零的实数 $c_1,c_2,\cdots,c_k$,有

$$f_k(c_1,c_2,\cdots,c_k)=f(c_1,c_2,\cdots,c_k,0,\cdots,0)>0$$

因此 f_k 是一个正定二次型,从而它的相伴矩阵 $\mathbf{A}_k$ (由 $\mathbf{A}$ 的前 k 行及前 k 列组成) 是一个正定阵. 由于 $\mathbf{A}_k$ 合同于 $\mathbf{I}_k$,故存在 k 阶非异阵 $\mathbf{B}$,使

$$\mathbf{B}'\mathbf{A}_k\mathbf{B}=\mathbf{I}_k$$

于是

$$\det(\mathbf{B}'\mathbf{A}_k\mathbf{B})=\det(\mathbf{B})^2\det(\mathbf{A}_k)=1$$

即有 $\det(\mathbf{A}_k)>0\,(k=1,2,\cdots,n)$.

再证充分性. 对 $\mathbf{A}$ 的阶数进行归纳. 当 $n=1$ 时, $\mathbf{A}=(a)\,,a>0$,于是 $f=ax_1^2$ 是正定型,从而 A 是

正定阵. 设结论对 $n-1$ 成立,现证明对 n 阶实对称阵 $\mathbf{A}$,若它的 n 个顺序主子式全大于零,则 $\mathbf{A}$ 必是正定阵. 记 $\mathbf{A}_{n-1}$ 是 $\mathbf{A}$ 的 $n-1$ 阶顺序主子式所在的矩阵,则 $\mathbf{A}$ 可写为

$$\begin{pmatrix} \mathbf{A}_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix}$$

因为 $\mathbf{A}$ 的顺序主子式全大于零,故 $\mathbf{A}_{n-1}$ 的顺序主子式也全大于零,由归纳假设, $\mathbf{A}_{n-1}$ 是正定阵. 于是 $\mathbf{A}_{n-1}$ 合同于 $n-1$ 阶单位阵,即存在 $n-1$ 阶非异阵 $\mathbf{B}$,使

$$B'A_{n-1}B = I_{n-1}$$

令 $\mathbf{C}$ 是下列分块矩阵:

$$C = \begin{pmatrix} B & O \\ O & 1 \end{pmatrix}$$

则

$$C'AC = \begin{pmatrix} B' & O \\ O & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix} \begin{pmatrix} B & O \\ O & 1 \end{pmatrix} = \begin{pmatrix} I_{n-1} & B'\alpha \\ \alpha' B & a_{nn} \end{pmatrix}$$

这是一个实对称阵, 其形式为

$$\mathbf{C'AC} = \begin{pmatrix} 1 & \cdots & 0 & c_1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & c_{n-1} \\ c_1 & \cdots & c_{n-1} & a_{nn} \end{pmatrix}$$

用第三类初等行及列变换可将上述矩阵化为对角阵. 这相当于对 $\mathbf{C'AC}$ 右乘一个非异阵 Q 后,再左乘 Q' 得到一个对角阵,亦即 $Q'C'ACQ$ 等于

$$\mathrm{diag}\{1, \cdots, 1, c\}$$

由于 $|\mathbf{A}| > 0$,故 $c > 0$,这就证明了 $\mathbf{A}$ 是一个正定阵.

命题 4.3

若 A 是正定阵,证明:

(1) $\mathbf{A}$ 的任一 k 阶主子阵,即由 $\mathbf{A}$ 的第 $i_1, i_2, \cdots, i_k$ 行及 $\mathbf{A}$ 的第 $i_1, i_2, \cdots, i_k$ 列交点上元素组成的矩阵, 必是正定阵;

(2) $\mathbf{A}$ 的所有主子式全大于零,特别, $\mathbf{A}$ 的主对角元素全大于零;

(3) A 中绝对值最大的元素仅在主对角线上.

证明 (1) 设 $\mathbf{A}_k$ 是矩阵 $\mathbf{A}$ 的第 k 个顺序主子式所在的矩阵,则 $\mathbf{A}_k$ 是实对称阵且其顺序主子式都大于零,因此 $\mathbf{A}_k$ 是正定阵.

经过若干次行对换以及相同的列对换,我们不难将 $\mathbf{A}$ 的第 $i_1, i_2, \cdots, i_k$ 行及 $i_1, i_2, \cdots, i_k$ 列分别换成第 $1, 2, \cdots, k$ 行和第 $1, 2, \cdots, k$ 列. 利用上面的结论即知 (1) 成立.

(2) 是 (1) 的推论.

(3) 用反证法. 假设 $a_{ij}\ (i \neq j)$ 是 $\mathbf{A}$ 的绝对值最大的元素. 根据 (1),我们只需证明由第 i, j 行与第 i, j 列交点上元素组成的矩阵不是正定阵即可. 考虑矩阵 (由 $\mathbf{A}$ 的对称性不妨设 $i < j$)

$$\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}$$

注意到 $a_{ij}=a_{ji}$ 且 $|a_{ij}|\geq a_{ii},|a_{ij}|\geq a_{jj}$,上述矩阵的行列式值 $a_{ii}a_{jj}-a_{ij}^2\leq 0$, 所以这个矩阵一定不是正定阵.

例题 4.4 举例说明 $\mathbf{A}$ 的所有顺序主子式都大于等于零并不能保证 $\mathbf{A}$ 是半正定阵.

解 例如, $\mathbf{A}=\mathrm{diag}\{1,0,-1\}$ 的顺序主子式都非负,但 $\mathbf{A}$ 不是半正定阵.

定理 4.7

设 $\mathbf{A}$ 是 n 阶实对称阵,证明: $\mathbf{A}$ 是半正定阵的充分必要条件是 $\mathbf{A}$ 的所有主子式都大于等于零.

证明 这个证明[1]是我在第一次学二次型的时候做的，当时由于先学的二次型，还没有接触到特征值的概念，所以就想尝试单纯利用简单的矩阵知识去证明这一命题，于是就有了下面这个又臭又长的证明，这个证明也发在了知乎上面：半正定矩阵的充分必要条件是所有主子式非负的一个证明方法。

引理1:

若 $A=\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$,且 $\mathrm{rank}\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}=\mathrm{rank}\begin{bmatrix} A_1 & A_2 \end{bmatrix}=\mathrm{rank}\begin{bmatrix} A_1 \\ A_3 \end{bmatrix}$,

则我们有 $\mathrm{rank\,A}=\mathrm{rank\,A_1}$.

[证明]: 由 $\mathrm{rank}\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}=\mathrm{rank}\begin{bmatrix} A_1 \\ A_3 \end{bmatrix}$ ，我们有 $\exists X, s.t. \begin{bmatrix} A_1 \\ A_3 \end{bmatrix}X=\begin{bmatrix} A_2 \\ A_4 \end{bmatrix}$

进一步有 $A_1X=A_2$ ，故而有 $\mathrm{rank\,A_1}=\mathrm{rank}\begin{bmatrix} A_1 & A_2 \end{bmatrix}$

所以我们有 $\mathrm{rank\,A}=\mathrm{rank\,A_1}$. 证毕.

引理2:

设 A 是 n 阶对称矩阵，则 A 的最大的非零主子式的阶数等于 A 的秩.

[证明]：设 $\mathrm{rank\,A}=\mathrm{r}$ ，取出 A 的列向量的极大线性无关组，不失一般性，设为 $\mathbf{col}_1,\cdots,\mathbf{col}_\mathrm{r}$ 为 A 的前 r 列，由于 A 是对称阵，所以 A 的前 r 行 $\mathbf{row_1},\cdots,\mathbf{row_r}$ 也是行向量的极大线性无关组

将矩阵 A 分块为 $A=\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$, $A_1\in\mathbb{P}^{r\times r}$

我们有下面的结论： $\mathrm{rank}\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}=\mathrm{rank}\begin{bmatrix} A_1 & A_2 \end{bmatrix}=\mathrm{rank}\begin{bmatrix} A_1 \\ A_3 \end{bmatrix}=\mathrm{r}$

从而由[引理1]我们知道 $\mathrm{rank\,A_1}=\mathrm{rank\,A}=\mathrm{r}$ ，从而 A_1 满秩，故 $\det(A_1)\neq 0$. 证毕.

回到原题. 有了上面的引理，我们开始着手处理这道题目。使用数学归纳法。当 $n=2$ 时，结论平凡，容易验证。

假设 $k\leq n-1$ 时，有‘ k 阶半正定矩阵的充分必要条件是所有主子式非负’ 成立，现在我们考虑所有主子式非负的 n 阶对称矩阵 A 的情况。设 $\mathrm{rank\,A}=\mathrm{r}$. ($r=0$ 情况平凡，下面的 r 默认大于 0)

case 1: $r\leq n-1$, 不妨设 A 的列向量的极大线性无关组为 $\mathbf{col_{i1}},\cdots,\mathbf{col_{ir}}$,由于合同变换不改变矩阵的半正定性，所以我们不妨通过合同变换将这 r 列交换到前 r 列，所以不妨直接设为前 r 列，即

$$A=\begin{bmatrix} A_1 & A_2 \\ A_2^T & A_4 \end{bmatrix}, A_1\in\mathbb{R}^{r\times r}, \mathrm{rank\,A_1}=\mathrm{rank\,A}=\mathrm{r}$$

[1]可以不用看这个证明，直接看第二个证明

由于 A 的所有主子式非负，特别地，有 A_1 的所有主子式非负，包括 $\det(A_1)\geq 0$ ，则由 $k=r\leq n-1$ 的归纳假设我们知道 A_1 是半正定的，又由[引理2]知道 $\det(A_1)\neq 0$ ，所以 $\det(A_1)>0$ ，由于 A_1 是半正定的且 A_1 可逆，所以我们可以断言 A_1 是正定的，从而存在可逆矩阵 C 使得 $C^TA_1C=E_r$ ，因此我们可以对 A 进行合同变换

$$\begin{pmatrix}C^T & O\\ O & E\end{pmatrix}\begin{pmatrix}A_1 & A_2\\ A_2^T & A_4\end{pmatrix}\begin{pmatrix}C & O\\ O & E\end{pmatrix}=\begin{pmatrix}C^TA_1C & C^TA_2\\ A_2^TC & A_4\end{pmatrix}=\begin{pmatrix}E_r & C^TA_2\\ A_2^TC & A_4\end{pmatrix}$$

再进行一次合同变换，打洞处理

$$\begin{pmatrix}E_r & O\\ -A_2^TC & E\end{pmatrix}\begin{pmatrix}E_r & C^TA_2\\ A_2^TC & A_4\end{pmatrix}\begin{pmatrix}E_r & -C^TA_2\\ O & E\end{pmatrix}=\begin{pmatrix}E_r & O\\ O & A_4-A_2^TCC^TA_2\end{pmatrix}$$

设 $A_0=A_4-A_2^TCC^TA_2$ ，则我们知道 A 合同于 $\begin{pmatrix}E_r & O\\ O & A_0\end{pmatrix}$

故 $\mathrm{rank\,E_r}+\mathrm{rank\,A_0}=\mathrm{rank}\begin{pmatrix}E_r & O\\ O & A_0\end{pmatrix}=\mathrm{rank\,A}=\mathrm{r}$ ，所以 $\mathrm{rank\,A_0}=0$

所以 $A_0=0$ ，故而 A 合同于 $\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}$ ，所以 A 的负惯性指数为 0，故 A 半正定，结论对 n 成立.

case 2: $r=n$ ，即 A 是满秩矩阵.

事实上我们在这里遇到了困难，因为当 A 的秩为 n 时，我们无法像上一题一样使用归纳假设，所以需要换一条路走，而确实有一条充满巧合的小路留给了我们.

反证法: 假设此时 A 不是半正定矩阵，由于 A 满秩，所以我们可以设

$$A=C^T\begin{bmatrix}E_r & O\\ O & -E_s\end{bmatrix}C$$

其中 $r,s>0$ ，由于 $\det(A)=\overset{2}{\det}(C)(+1)^r(-1)^s>0$ ，所以我们有 $s\geq 2$ (因为 s 是偶数)

下面考虑 A 的前 $(n-1)\times(n-1)$ 项，由于 A 的主子式全部非负，

所以由归纳假设可以知道前 $(n-1)\times(n-1)$ 个元素组成的矩阵(后面记为 B)是半正定的（归纳假设中 $k=n-1$ 的情况），又由于，记 $A=\begin{pmatrix}B & b\\ b^T & c\end{pmatrix}$.

因为 B 是半正定的，设 $\mathrm{rank\,B}=\mathrm{t}$

case 2.1: $t<n-1$

设 $D^TBD=\begin{pmatrix}E_t & O\\ O & O\end{pmatrix}$ ，从而对 A 进行合同变换:

$$\begin{pmatrix}D^T & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}B & b\\ b^T & c\end{pmatrix}\begin{pmatrix}D & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}D^TBD & D^Tb\\ b^TD & c\end{pmatrix}=\begin{pmatrix}E_t & O & b_1\\ O & O & b_2\\ b_1^T & b_2^T & c\end{pmatrix}$$

$$\begin{pmatrix} E_t & O & O \\ O & E_{n-1-t} & O \\ -b_1^T & O & 1 \end{pmatrix}\begin{pmatrix} E_t & O & b_1 \\ O & O & b_2 \\ b_1^T & b_2^T & c \end{pmatrix}\begin{pmatrix} E_t & O & -b_1 \\ O & E_{n-1-t} & O \\ O & O & 1 \end{pmatrix} = \begin{pmatrix} E_t & O & O \\ O & O & b_2 \\ O & b_2^T & c \end{pmatrix}$$

故 A 合同于 $\begin{pmatrix} E_t & O & O \\ O & O & b_2 \\ O & b_2^T & c \end{pmatrix}$，有

$$\mathrm{rank\,A} = \mathrm{rank}\begin{pmatrix} E_t & O & O \\ O & O & b_2 \\ O & b_2^T & c \end{pmatrix} = \mathrm{rank\,E_t} + \mathrm{rank}\begin{pmatrix} O & b_2 \\ b_2^T & c \end{pmatrix} = \mathrm{n}$$

所以 $\begin{pmatrix} O & b_2 \\ b_2^T & c \end{pmatrix}$ 是满秩的，这要求左上角的零矩阵 O 只能为 1×1 的，故 $b_2 \in \mathbb{R}$ 是一个非零实数

但是又根据 $\det(A) > 0$ 与合同变换不改变行列式的正负，有

$$\det\begin{pmatrix} E_t & O & O \\ O & O & b_2 \\ O & b_2^T & c \end{pmatrix} = \det(E_t)\det\begin{pmatrix} 0 & b_2 \\ b_2 & c \end{pmatrix} > 0$$

但是 $\det(E_t)\det\begin{pmatrix} 0 & b_2 \\ b_2 & c \end{pmatrix} = 1 \times \left(0c - b_2^2\right) = -b_2^2 < 0$，从而导致矛盾，故 $t = n - 1$.

case 2.2: $t = n - 1$

则知道 B 半正定且可逆，故 B 正定，所以存在可逆矩阵 D 使得 $D^T BD = E_{n-1}$，所以对 A 作合同变换

$$\begin{pmatrix} D^T & O \\ O & 1 \end{pmatrix}\begin{pmatrix} B & b \\ b^T & c \end{pmatrix}\begin{pmatrix} D & O \\ O & 1 \end{pmatrix} = \begin{pmatrix} D^T BD & D^T b \\ b^T D & c \end{pmatrix} = \begin{pmatrix} E_{n-1} & D^T b \\ b^T D & c \end{pmatrix}$$

$$\begin{pmatrix} E_{n-1} & O \\ -b^T D & 1 \end{pmatrix}\begin{pmatrix} E_{n-1} & D^T b \\ b^T D & c \end{pmatrix}\begin{pmatrix} E_{n-1} & -D^T b \\ O & 1 \end{pmatrix} = \begin{pmatrix} E_{n-1} & O \\ O & c - b^T DD^T b \end{pmatrix}$$

故 A 合同于 $\begin{pmatrix} E_{n-1} & O \\ O & c - b^T DD^T b \end{pmatrix}$，但 $\begin{pmatrix} E_{n-1} & O \\ O & c - b^T DD^T b \end{pmatrix}$ 的负惯性指数 ≤ 1，这与 A 的负惯性指数 ≥ 2 矛盾，从而我们的假设错误，故 A 是一个半正定矩阵.

综上，我们证明了命题对 n 时也成立，所以由归纳原理，我们有：对任意的 n 阶实对称矩阵，它是半正定矩阵的充分必要条件是所有的主子式非负.

证明 暂时懒得写了.

例题 4.5 判断下面命题是否正确：$A = (a_{ij})_{n\times n} \in \mathbb{R}^{n\times n}$ 是正定矩阵，则 $B = (|a_{ij}|)_{n\times n}$（即每个元素取绝对值）也正定.

解 $n = 1, 2, 3$ 容易验证正确；

$n=4$ 时取

$$A=\begin{bmatrix}10 & 3 & -2 & 1\\ 3 & 10 & 0 & 9\\ -2 & 0 & 10 & 4\\ 1 & 9 & 4 & 10\end{bmatrix}$$

容易验证为反例.

命题 4.4

$A\in\mathbb{R}^{n\times n}$ 半正定,则

$$x^{\top}Ax=0\Leftrightarrow Ax=0,\ \forall x\in\mathbb{R}^n$$

证明 设 $A=C^TC$，从而有

$$AX=0\Rightarrow C^TCX=0\Rightarrow X^TC^TCX=0\Rightarrow(CX)^T(CX)=0\Rightarrow CX=0\Rightarrow AX=0$$

命题 4.5

$A\in\mathbb{R}^{n\times n}$ 半正定,则

(1). 对角元 $a_{ii}\geq 0, i=1,2,\ldots,n$;

(2). 如果对某个 $k\in\{1,2,\ldots,n\}$,对角元 $a_{kk}=0$,则 $a_{ik}=a_{ki}=0, i=1,2,\ldots,n$.

证明 (1) A 半正定，则所有主子式非负，特别地有 $a_{ii}\geq 0$.

(2) 反证，假设对 $a_{kk}=0$，存在 i_0 使得 $a_{i_0k}=a_{ki_0}\neq 0$，从而考虑

$$X=x_{i_0}\mathbf{e}_{i_0}+\mathbf{e}_kx_k$$

从而有

$$f=X^TAX=a_{i_0i_0}x_{i_0}^2+2a_{i_0k}x_{i_0}x_k$$

从而知道当固定 i_0 时，f 是关于 x_k 的线性函数，从而存在 x_k 使得 $f<0$，从而与半正定矛盾，所以假设错误，故 $a_{i_0k}=a_{ki_0}=0$.

命题 4.6

如果 A,B 都是半正定矩阵,则

$$\operatorname{rank}(A+B)\geq\max\{\operatorname{rank}(A),\operatorname{rank}(B)\}$$

证明 有 $X^T(A+B)X=0\Leftrightarrow(A+B)X=0$，又有 $X^TAX+X^TBX=0\Rightarrow X^TAX=0, X^TBX+0$. 所以我们有

$$\mathrm{Ker}(A+B)\subseteq\mathrm{Ker}A$$

同理有：

$$\mathrm{Ker}(A+B)\subseteq\mathrm{Ker}B$$

故：

$$\operatorname{rank}(A+B)\geq\max\{\operatorname{rank}(A),\operatorname{rank}(B)\}$$

命题 4.7

设 A 是 n 阶实正定方阵, B 是 n 阶反对称方阵,则

$$|A+\lambda B|>0,\ \forall\lambda\in\mathbb{R}$$

证明 若 $|A+\lambda B|=0$，则存在 $X\neq 0$ 使得 $(A+\lambda B)X=0$，所以有：

$$X^T(A+\lambda B)X=0$$

又由于 $X^TBX=0$，所以有 $X^TAX=0$，但这与 A 的正定性矛盾，所以 $|A+\lambda B|\neq 0$.

而 $f(\lambda)=|A+\lambda B|$ 为关于 λ 的多项式，$f(0)=|A|>0$，从而由介值定理，知道 $f(\lambda)>0$，故 $|A+\lambda B|>0$.

命题 4.8

A 是反对称方阵,证明 $|A|\geq 0$.

证明 由上一题，知道零矩阵 O 为半正定矩阵，从而

$$|A|=|O+A|\geq 0$$

命题 4.9

$A,B\in\mathbb{R}^{n\times n}$ 对称,并且 A 正定。证明: 当 t 充分大时, $tA+B$ 正定.

证明

命题 4.10

设 $A=(a_{ij})\in\mathbb{R}^{n\times n}$ 对称可逆, A_{ij} 是 a_{ij} 的代数余子式. 证明: 二次型

$$f\left(x_1,\ldots,x_n\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{A_{ij}}{|A|}x_ix_j$$

与 A 有相同的正、负惯性指数.

证明

命题 4.11

设 S 是 n 阶实对称方阵. 证明 $\operatorname{rank}S=n$ 的充要条件是: 存在 n 阶实方阵 A 使得 $SA+A^{\top}S$ 正定.

命题 4.12

设 $A=(a_{ij})_{n\times n}$ 是可逆对称方阵, 计算二次型

$$\begin{vmatrix}0 & x_1 & \cdots & x_n\\ -x_1 & a_{11} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ -x_n & a_{n1} & \cdots & a_{nn}\end{vmatrix}$$

的矩阵.

命题 4.13

设 $A_i\,(1\le i\le m)$ 是同阶实对称方阵. 证明

$$A_1^2+\cdots A_m^2=0\Leftrightarrow A_1=\cdots=A_m=0$$

命题 4.14

设 A 是实对称方阵, V 是方程 $x^\top Ax=0$ 的解集. 证明:V 是线性空间 $\Leftrightarrow$ A 半正定或者半负定.

命题 4.15

A 是正定矩阵, B 是半正定矩阵, 证明

$$|A+B|\ge|A|$$

当且仅当 $B=0$ 时等号成立.

命题 4.16

设 $S=\begin{bmatrix}A & C\\ C^\top & B\end{bmatrix}$ 正定. 证明

$$|S|\le|A|\cdot|B|$$

当且仅当 $C=0$ 时等号成立.

命题 4.17

设 $A,B\in\mathbb{R}^{m\times n}$ 满足 $\operatorname{rank}(A+B)=\operatorname{rank}(A)+\operatorname{rank}(B)$. 证明: 存在可逆矩阵 P,Q ,使得

$$A=P\begin{bmatrix}E_{r\times r} & & \\ & 0_{s\times s} & \\ & & 0\end{bmatrix}Q,\ B=P\begin{bmatrix}0_{r\times r} & & \\ & E_{s\times s} & \\ & & 0\end{bmatrix}Q$$

其中, $r=\operatorname{rank}(A)$, $s=\operatorname{rank}(B)$.

命题 4.18

设

$$f_k(x):[a,b]\to\mathbb{R},\ k=1,2,\ldots,n$$

证明: $f_1(x),\ldots,f_n(x)$ 线性无关的充要条件是,在 $[a,b]$ 上存在点 $a_1,\ldots,a_n$,使得

$$D_n=\begin{vmatrix}f_1(a_1) & f_1(a_2) & \cdots & f_1(a_n)\\ f_2(a_1) & f_2(a_2) & \cdots & f_2(a_n)\\ \vdots & \vdots & & \vdots\\ f_n(a_1) & f_n(a_2) & \cdots & f_n(a_n)\end{vmatrix}\ne0$$

4.5 Hermite 型

Hermite 型看成是复数域上的二次齐次函数:

$$f(x_1, x_2, \cdots, x_n) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} \bar{x}_i x_j \tag{4.4}$$

其中 $\bar{a}_{ij} = a_{ji}$. Hermite 型虽然系数是复数且变元 x_i 是复数域上的变元,但作为函数它的值却总是实数, 这点从 Hermite 型的定义即可看出. 事实上,

$$\bar{f} = \sum_{j=1}^{n} \sum_{i=1}^{n} \bar{a}_{ij} x_i \bar{x}_j = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} \bar{x}_j x_i = f$$

因此 f 的值总是实数. 当 Hermite 型的变元 x_i 取实变元且 a_{ij} 都是实数时, f 就是实二次型. 因此, Hermite 型也可看成是实二次型的推广. 事实上, Hermite 型有许多与实二次型相同的性质.

Hermite 型 (4.4) 可写成如下的矩阵相乘的形式:

$$f(x_1, x_2, \cdots, x_n) = \overline{\mathbf{x}}' \mathbf{A} \mathbf{x}$$

其中

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

且满足 $\overline{\mathbf{A}}' = \mathbf{A}$,这样的矩阵称为 Hermite 矩阵. 与实二次型类似,Hermite 型与 Hermite 矩阵之间有着一一对应关系,即给定一个 n 变元的 Hermite 型必相伴有一个 n 阶 Hermite 矩阵,反之给定一个 n 阶 Hermite 矩阵,必有一个 n 元 Hermite 型与之对应.

设 $\mathbf{x} = \mathbf{C}\mathbf{y}$,其中 $\mathbf{C}$ 是一个非异复矩阵, $\mathbf{y} = (y_1, y_2, \cdots, y_n)'$,则

$$f = \left(\overline{Cy}\right)' A (Cy) = \bar{y}' \bar{C}' A C y.$$

矩阵 $\overline{\mathbf{C}}' \mathbf{A} \mathbf{C}$ 仍是一个 Hermite 矩阵.

定义 4.7

设 A, B 是两个 Hermite 矩阵,若存在非异复矩阵 C ,使

$$B = \bar{C}' A C$$

则称 $\mathbf{A}$ 与 $\mathbf{B}$ 是复相合的.

容易证明复相合是一个等价关系. 与实二次型类似, Hermite 型

$$a_1 \bar{y}_1 y_1 + a_2 \bar{y}_2 y_2 + \cdots + a_n \bar{y}_n y_n \tag{4.5}$$

称为标准型. Hermite 型的基本问题是如何把一个 Hermite 型化成像 (4.5) 式那样的标准型. 这个问题等价于寻找非异阵 C ,使 $\bar{C}'AC$ 成为对角阵.

定理 4.8

设 A 是一个 Hermite 矩阵,则必存在一个非异阵 C ,使 $\overline{C}'AC$ 是一个对角阵且主对角线上的元素都是实数.

证明 寻找 C 的过程类似于对称阵的情形,通过归纳法就可以做法，故从略.

现只需说明后面的结论. 事实上, $\bar{C}'AC$ 仍是 Hermite 矩阵,因此主对角线上的元素 $b_{ii}=\bar{b}_{ii}$,必是实数.

定理 4.9 (复二次型的惯性定理)

设 $f(x_1,x_2,\cdots,x_n)$ 是一个 Hermite 型,则它总可以化为如下标准型:

$$\bar{y}_1y_1+\cdots+\bar{y}_py_p-\bar{y}_{p+1}y_{p+1}-\cdots-\bar{y}_ry_r \tag{4.6}$$

且若 f 又可化为另一个标准型:

$$\bar{z}_1z_1+\cdots+\bar{z}_kz_k-\bar{z}_{k+1}z_{k+1}-\cdots-\bar{z}_rz_r$$

则 $p=k$.

我们称 (4.6) 式中的 r 为 f 的秩, p 为 f 的正惯性指数, $q=r-p$ 为 f 的负惯性指数, $p-q$ 为 f 的符号差. 同样地,秩与符号差 (或正负惯性指数) 是 Hermite 矩阵复相合的全系不变量. 上述这些结论的证明与实二次型是平行的, 我们略去其证明.

由于 Hermite 型的值总是实数, 故可定义正定型、负定型、半正定型、半负定型及不定型的概念. 相应地, 有正定 Hermite 矩阵、负定 Hermite 矩阵、半正定 Hermite 矩阵和半负定 Hermite 矩阵的概念.

定义 4.8

设 $f(x_1,x_2,\cdots,x_n)$ 是 Hermite 型,若对任一组不全为零的复数$c_1,c_2,\cdots,c_n$,均有

$$f(c_1,c_2,\cdots,c_n)>0$$

则称 f 是正定 Hermite 型,它对应的矩阵称为正定 Hermite 矩阵.

定理 4.10

n 阶 Hermite 矩阵 $\mathbf{A}$ 为正定的充分必要条件是它的 n 个顺序主子式全大于零.

证明 证明与实正定矩阵类似，故略去.

第 5 章 内积空间

5.1 内积空间的概念

定义 5.1 (实内积空间)

设 V 是实数域上的线性空间,若存在某种规则,使对 V 中任意一组有序向量 $\{\alpha, \beta\}$,都唯一地对应一个实数,记为 (α, β) ,且适合如下规则:

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

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

(3) $(c\alpha, \beta) = c(\alpha, \beta)$, c 为任一实数;

(4) $(\alpha, \alpha) \geq 0$ 且等号成立当且仅当 $\alpha = \mathbf{0}$,

则称在 V 上定义了一个内积. 实数 (α, β) 称为 α 与 β 的内积. 线性空间 V 称为实内积空间. 有限维实内积空间称为 Euclid 空间, 简称为欧氏空间.

♣

定义 5.2 (复内积空间)

设 V 是复数域上的线性空间,若存在某种规则,使对 V 中任意一组有序向量 $\{\alpha, \beta\}$, 都唯一地对应一个复数, 记为 (α, β) , 且适合如下规则:

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

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

(3) $(c\alpha, \beta) = c(\alpha, \beta)$, c 为任一复数;

(4) $(\alpha, \alpha) \geq 0$ 且等号成立当且仅当 $\alpha = \mathbf{0}$,

则称在 V 上定义了一个内积. 复数 (α, β) 称为 α 与 β 的内积. 线性空间 V 称为复内积空间. 有限维复内积空间称为酉空间.

♣

注 实内积空间的定义与复内积空间的定义是相容的. 事实上,对一个实数 a , $\bar{a} = a$. 故定义 5.1 中的 (1) 与定义 5.2 中的 (1) 是一致的. 因此,我们经常将这两种空间统称为内积空间, 在某些定理的叙述及证明中也不区别它们, 而统一作为复内积空间来处理. 但是, 需要注意的是对复内积空间, 定义 5.2 中的 (1), (3) 意味着:

$$(\alpha, c\beta) = \bar{c}(\alpha, \beta)$$

命题 5.1 (标准内积)

(1) 设 $\mathbb{R}_n$ 是 n 维实列向量空间, $\alpha = (x_1, x_2, \cdots, x_n)'$, $\beta = (y_1, y_2, \cdots, y_n)'$,定义

$$(\alpha, \beta) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

则我们定义了一个内积,这个内积称为 $\mathbb{R}_n$ 中的标准内积.

(2) 设 $\mathbb{C}_n$ 是 n 维复列向量空间, $\alpha = (x_1, x_2, \cdots, x_n)'$, $\beta = (y_1, y_2, \cdots, y_n)'$,定义

$$(\alpha, \beta) = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \cdots + x_n \bar{y}_n,$$

则在此定义下 $\mathbb{C}_n$ 成为一个酉空间,上述内积称为 $\mathbb{C}_n$ 的标准内积.

♠

对实列向量空间, 标准内积可用矩阵乘法表示

$$(\alpha, \beta) = \alpha' \beta$$

对实行向量空间, 标准内积也可表示为

$$(\alpha, \beta) = \alpha \beta'$$

对 n 维复列向量空间 U，标准内积可表示为

$$(\alpha, \beta) = \alpha' \overline{\beta}$$

当 U 是复行向量空间时,标准内积可表示为

$$(\alpha, \beta) = \alpha \overline{\beta}'$$

定义 5.3 (范数)

设 V 是内积空间 (实或复), α 是 V 中的向量,定义 α 的长度 (或范数) 为

$$\| \alpha \| = (\alpha, \alpha)^{\frac{1}{2}}$$

即实数 (α, α) 的算术根.

♣

注 注意,当 V 是复内积空间时,由于规则, (α, α) 总是实数. 从长度的定义知, $\| \alpha \| = 0$ 当且仅当 $\alpha = \mathbf{0}$. 当 $V = \mathbb{R}^n$ 且内积为标准内积时, 若 $\alpha = (x_1, x_2, \cdots, x_n)$,则

$$\| \alpha \| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

利用范数可定义内积空间中两个向量的距离. 设 $\alpha, \beta \in V$,定义 α 与 β 的距离为

$$d(\alpha, \beta) = \| \alpha - \beta \| .$$

显然 $d(\alpha, \beta) = d(\beta, \alpha)$.

定理 5.1 (范数的性质)

设 V 是实或复的内积空间, $\alpha, \beta \in V, c$ 是任一常数 (实数或复数), 则

(1) $\| c\alpha \| = |c| \| \alpha \|$;

(2) $|(\alpha, \beta)| \leq \| \alpha \| \cdot \| \beta \|$;

(3) $\| \alpha + \beta \| \leq \| \alpha \| + \| \beta \|$.

♡

证明 (1) $\| c\alpha\|^2 = (c\alpha, c\alpha) = c\bar{c} \| \alpha\|^2 = |c|^2 \| \alpha\|^2$,故 $\| c\alpha \| = |c| \| \alpha \|$.

(2) 若 $\alpha = \mathbf{0}$,则 $(\mathbf{0}, \beta) = (\mathbf{0} + \mathbf{0}, \beta) = 2(\mathbf{0}, \beta)$,故 $(\mathbf{0}, \beta) = 0$. 因此 (2) 成立. 若 $\alpha \neq \mathbf{0}$,令

$$v = \beta - \frac{(\beta, \alpha)}{\| \alpha\|^2} \alpha$$

则 $(\mathbf{v}, \alpha) = 0$,且

$$0 \leq \| \mathbf{v}\|^2 = \left(\beta - \frac{(\beta, \alpha)}{\| \alpha\|^2} \alpha, \beta - \frac{(\beta, \alpha)}{\| \alpha\|^2} \alpha \right)$$

$$= (\beta, \beta) - \frac{(\beta, \alpha)}{\| \alpha\|^2} (\alpha, \beta)$$

$$=\| \beta\|^2 - \frac{|(\alpha,\beta)|^2}{\| \alpha\|^2}.$$

由此即证.

(3) 我们有

$$\| \alpha+\beta\|^2 = (\alpha+\beta,\alpha+\beta)$$

$$=\| \alpha\|^2 + (\alpha,\beta) + (\beta,\alpha) + \| \beta\|^2$$

$$=\| \alpha\|^2 + \| \beta\|^2 + (\alpha,\beta) + \overline{(\alpha,\beta)}.$$

由 (2) 得

$$|(\alpha,\beta)| \leq \| \alpha \| \| \beta \|, \left|\overline{(\alpha,\beta)}\right| \leq \| \alpha \| \| \beta \| .$$

故

$$\| \alpha+\beta\|^2 \leq \| \alpha\|^2 + \| \beta\|^2 + 2\| \alpha \| \| \beta \| = (\| \alpha \| + \| \beta \|)^2$$

笔记 定理 5.1 中的 (2) 通常称为 Cauchy-Schwarz 不等式; (3) 通常称为三角不等式.

证明的过程中要善于使用 $(\alpha,\alpha) \geq 0$ 这一性质来构造不等关系.

定义 5.4 (向量的夹角)

当 V 是实内积空间时,定义非零向量 α,β 的夹角 θ 之余弦为

$$\cos\theta = \frac{(\alpha,\beta)}{\| \alpha \| \| \beta \|}$$

当 V 是复内积空间时,定义非零向量 α,β 的夹角 θ 之余弦为

$$\cos\theta = \frac{|(\alpha,\beta)|}{\| \alpha \| \| \beta \|}$$

内积空间中两个向量 α,β 若适合 $(\alpha,\beta)=0$,则称 α 与 β 垂直或正交,我们用记号 $\alpha\perp\beta$ 来表示. 显然,零向量和任何向量都正交; 若 α 与 β 正交,则 β 与 α 也正交; 两个非零向量 α,β 正交时夹角为 90° .

♣

笔记 注意在上式中要使 θ 有意义,必须保证 $\cos\theta \leq 1$. 而这就是定理5.1 中的 (2). 在定理5.1 (3) 的证明中我们可看出: 若 α 与 β 正交则 $(\alpha,\beta)=(\beta,\alpha)=0$,因此

$$\| \alpha+\beta\|^2 = \| \alpha\|^2 + \| \beta\|^2$$

这是平面几何中勾股定理的推广.

命题 5.2 (Cauchy-Schwarz 不等式)

设 V 是实 n 维行向量空间,内积取标准内积,从定理 5.1 的 (2) 立即可得到下列 Cauchy 不等式:

$$(x_1y_1 + x_2y_2 + \cdots + x_ny_n)^2 \leq (x_1^2 + x_2^2 + \cdots + x_n^2)(y_1^2 + y_2^2 + \cdots + y_n^2)$$

设 V 是 $[a,b]$ 上连续函数全体构成的实线性空间,定义

$$(f,g)=\int_a^b f\left(t\right)g\left(t\right)\mathrm{d}t$$

则不难验证这是一个内积,于是 V 成为内积空间，则从定理 5.1的 (2) 可得下列 Schwarz 不等式:

$$\left(\int_a^b f\left(t\right)g\left(t\right)\mathrm{d}t\right)^2\le\int_a^b f(t)^2\,\mathrm{d}t\int_a^b g(t)^2\,\mathrm{d}t$$

5.2 内积的表示与正交基

我们需要考虑，对于有限维内积空间，如果给定内积空间的一组基之后，如何用这组基的坐标向量来表示向量的内积.

定理 5.2 (Gram 矩阵)

V 是欧氏空间 (酉空间), $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ 是 V 的一组基. 如果 $(\mathbf{v}_i, \mathbf{v}_j) = g_{ij}(i, j = 1, 2, \cdots, n), \alpha = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n, \beta = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n$

当 V 是欧氏空间时,

$$(\alpha, \beta) = \left(\sum_{i=1}^{n} a_i\mathbf{v}_i, \sum_{j=1}^{n} b_j\mathbf{v}_j\right) = \sum_{i,j=1}^{n} a_i g_{ij} b_j.$$

我们把上述结论写成矩阵形式:

$$(\alpha, \beta) = (a_1, a_2, \cdots, a_n)\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

其中矩阵

$$G = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} = \begin{pmatrix} (\mathbf{v}_1, \mathbf{v}_1) & (\mathbf{v}_1, \mathbf{v}_2) & \cdots & (\mathbf{v}_1, \mathbf{v}_n) \\ (\mathbf{v}_2, \mathbf{v}_1) & (\mathbf{v}_2, \mathbf{v}_2) & \cdots & (\mathbf{v}_2, \mathbf{v}_n) \\ \vdots & \vdots & & \vdots \\ (\mathbf{v}_n, \mathbf{v}_1) & (\mathbf{v}_n, \mathbf{v}_2) & \cdots & (\mathbf{v}_n, \mathbf{v}_n) \end{pmatrix}$$

称为基向量 $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ 的 Gram (格列姆) 矩阵或内积空间 V 在给定基下的度量矩阵.

于是, 我们得到了内积在给定基下的表示:

$$(\alpha, \beta) = \mathbf{x}'\mathbf{G}\mathbf{y}$$

其中 $\mathbf{x}, \mathbf{y}$ 分别是向量 α, β 在给定基下的坐标向量.

再来看矩阵 $\mathbf{G}$. 因为 $(\mathbf{v}_i, \mathbf{v}_j) = (\mathbf{v}_j, \mathbf{v}_i)$,所以 $\mathbf{G}$ 是实对称阵. 又因为对任意的非零向量 α ,总有 $(\alpha, \alpha) > 0$,所以 $\mathbf{x}'\mathbf{G}\mathbf{x} > 0$ 对一切 n 维非零实列向量 $\mathbf{x}$ 成立. 这表明 $\mathbf{G}$ 是一个正定阵. 反之,给定 n 阶正定实对称阵, 我们也不难定义 V 上的内积. 由此我们可以看出,若给定了 n 维欧氏空间的一组基,则欧氏空间上的内积和 n 阶正定实对称阵之间存在着一个一一对应.

♡

命题 5.3

在 n 维欧氏空间 V 中,两两夹角大于直角的向量个数至多是 $n+1$ 个.

♠

证明 用反证法证明. 假设存在 $n+2$ 个两两夹角大于直角的向量 $\alpha_1, \alpha_2, \cdots, \alpha_{n+1}, \alpha_{n+2} \in V$,则由 $\dim V = n$ 可知, $\alpha_1, \alpha_2, \cdots, \alpha_{n+1}$ 必线性相关,即存在不全为零的实数 $c_1, c_2, \cdots, c_{n+1}$,使得 $c_1\alpha_1 +$

$c_2\alpha_2+\cdots+c_{n+1}\alpha_{n+1}=\mathbf{0}$. 将此式按照系数正负整理为如下形式:

$$\sum_{c_i>0} c_i\alpha_i=\sum_{c_j<0}\left(-c_j\right)\alpha_j \tag{5.1}$$

由 $c_1,c_2,\cdots,c_{n+1}$ 不全为零不妨设存在某个 $c_i>0$. 若 (5.1) 式两边都等于零,则有

$$0=\left(\sum_{c_i>0} c_i\alpha_i,\alpha_{n+2}\right)=\sum_{c_i>0} c_i\left(\alpha_i,\alpha_{n+2}\right)<0,$$

矛盾. 因此 (5.1) 式两边都非零,从而也存在某个 $c_j<0$,于是

$$0<\left(\sum_{c_i>0} c_i\alpha_i,\sum_{c_i>0} c_i\alpha_i\right)=\left(\sum_{c_i>0} c_i\alpha_i,\sum_{c_j<0}\left(-c_j\right)\alpha_j\right)=\sum_{c_i>0}\sum_{c_j<0} c_i\left(-c_j\right)\left(\alpha_i,\alpha_j\right)<0$$

矛盾. 例如,不妨设 $V=\mathbb{R}^n$ (取标准内积),则向量 $\alpha_1=(n,-1,\cdots,-1)',\alpha_2=(-1,n,\cdots,-1)',\alpha_n=(-1,-1,\cdots,n)',\alpha_{n+1}=(-1,-1,\cdots,-1)'$ 就满足两两夹角大于直角. 因此, $n+1$ 就是两两夹角大于直角的向量个数的最佳上界,结论得证.

笔记 对于这类题目要善于利用正负分离的思想.

注 进一步，还能证明：

(1) $\alpha_1,\alpha_2,\cdots,\alpha_{n+1}$ 中任意 n 个向量必线性无关;

(2) $\alpha_1,\alpha_2,\cdots,\alpha_{n+1}$ 中任一向量必为其余向量的负系数线性组合.

命题 5.4

设 $\mathbf{A}$ 是 n 阶半正定实对称矩阵,求证: 对任意的 n 维实列向量 $\mathbf{x},\mathbf{y}$,有

$$\left(\mathbf{x}'\mathbf{A}\mathbf{y}\right)^2\leq\left(\mathbf{x}'\mathbf{A}\mathbf{x}\right)\left(\mathbf{y}'\mathbf{A}\mathbf{y}\right)$$

证明 对任意正实数 $t,\mathbf{A}+t\mathbf{I}_n$ 都是正定阵,这决定了 n 维列向量空间 $\mathbb{R}^n$ 上的一个内积,故由 Cauchy-Schwarz 不等式可得

$$\left(\mathbf{x}'\left(\mathbf{A}+t\mathbf{I}_n\right)\mathbf{y}\right)^2\leq\left(\mathbf{x}'\left(\mathbf{A}+t\mathbf{I}_n\right)\mathbf{x}\right)\left(\mathbf{y}'\left(\mathbf{A}+t\mathbf{I}_n\right)\mathbf{y}\right)$$

注意到上式两边都是关于 t 的连续函数,同时取极限,令 $t\to 0+$,即得结论.

笔记 对于半正定阵的问题，可以考虑利用摄动法变为正定阵再进行处理.

命题 5.5

设 U_1,U_2,U 是 n 维内积空间 V 的子空间,求证:

(1) $\left(U^{\perp}\right)^{\perp}=U$;

(2) $\left(U_1+U_2\right)^{\perp}=U_1^{\perp}\cap U_2^{\perp}$;

(3) $\left(U_1\cap U_2\right)^{\perp}=U_1^{\perp}+U_2^{\perp}$;

(4) $V^{\perp}=0,0^{\perp}=V$.

证明 (1) 因为 $V=U^{\perp}\oplus\left(U^{\perp}\right)^{\perp}$,故 $\dim\left(U^{\perp}\right)^{\perp}=n-\dim U^{\perp}=\dim U$. 另一方面,显然有 $U\subseteq\left(U^{\perp}\right)^{\perp}$,因此 $\left(U^{\perp}\right)^{\perp}=U$.

(2) 显然 $\left(U_1+U_2\right)^{\perp}\subseteq U_1^{\perp},\left(U_1+U_2\right)^{\perp}\subseteq U_2^{\perp}$,于是 $\left(U_1+U_2\right)^{\perp}\subseteq U_1^{\perp}\cap U_2^{\perp}$. 反之,对任一 $\alpha\in U_1^{\perp}\cap U_2^{\perp},\beta\in U_1+U_2$,记 $\beta=\beta_1+\beta_2$,其中 $\beta_1\in U_1,\beta_2\in U_2$,则

$$(\alpha,\beta)=\left(\alpha,\beta_1+\beta_2\right)=\left(\alpha,\beta_1\right)+\left(\alpha,\beta_2\right)=0,$$

故 $\alpha \in (U_1+U_2)^{\perp}$,于是 $U_1^{\perp} \cap U_2^{\perp} \subseteq (U_1+U_2)^{\perp}$. 因此 $(U_1+U_2)^{\perp} = U_1^{\perp} \cap U_2^{\perp}$.

(3) 由 (1) 及 (2),有 $\left(U_1^{\perp}+U_2^{\perp}\right)^{\perp} = \left(U_1^{\perp}\right)^{\perp} \cap \left(U_2^{\perp}\right)^{\perp} = U_1 \cap U_2$.

(4) 显然成立.

命题 5.6

设 $\mathbf{A}$ 为 $m \times n$ 实矩阵,齐次线性方程组 $\mathbf{Ax}=\mathbf{0}$ 的解空间为 U ,求 $U^{\perp}$ 适合的线性方程组.

解 设 $\mathbf{A}$ 的秩为 r ,则解空间 U 是 $\mathbb{R}^n$ (取标准内积) 的 $n-r$ 维子空间. 取 U 的一组基 $\eta_1,\cdots,\eta_{n-r}$,令 $\mathbf{B}=(\eta_1,\cdots,\eta_{n-r})$ 为 $n \times (n-r)$ 实矩阵,则可得 $U^{\perp}=\{\eta_1,\cdots,\eta_{n-r}\}^{\perp}$,因此 $U^{\perp}$ 适合的线性方程组为 $\mathbf{B}'\mathbf{x}=\mathbf{0}$.

命题 5.7

设 $\mathbf{A}$ 为 $m \times n$ 实矩阵,求证: 非齐次线性方程组 $\mathbf{Ax}=\beta$ 有解的充要条件是向量 β 属于齐次线性方程组 $\mathbf{A}'\mathbf{y}=\mathbf{0}$ 解空间的正交补空间.

证明 设 $\mathbf{A}=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 为列分块, $U=L(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 为 $\mathbb{R}^m$ (取标准内积) 的子空间,则 $\mathbf{Ax}=\beta$ 有解当且仅当 $\beta \in U$. 另一方面, $\mathbf{A}'\mathbf{y}=\mathbf{0}$ 的解空间即为 $\{\mathbf{y} \in \mathbb{R}^m \mid (\alpha_i,\mathbf{y})=0, 1 \le i \le n\} = U^{\perp}$,注意到 $U=\left(U^{\perp}\right)^{\perp}$,故结论得证.

5.3 伴随

定理 5.3 (伴随的矩阵)

设 V 是 n 维内积空间, $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ 是 V 的一组**标准正交基**. 若 V 上的线性算子 φ 在这组基下的表示矩阵为 $\mathbf{A} = (a_{ij})$,则如果 V 是酉空间,那么 φ^* 在同一组基下的表示矩阵为 $\overline{\mathbf{A}}' = (\bar{a}_{ij})'$,即 $\mathbf{A}$ 的共轭转置; 如果 V 是欧氏空间,那么 φ^* 的表示矩阵为 $\mathbf{A}'$,即 $\mathbf{A}$ 的转置.

命题 5.8 (伴随的几何性质)

设 V 是 n 维内积空间, φ 是 V 上的线性算子.
(1) 若 U 是 φ 的不变子空间,则 $U^\perp$ 是 φ^* 的不变子空间;
(2) 若 φ 的全体特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$,则 φ^* 的全体特征值为 $\bar{\lambda}_1, \bar{\lambda}_2, \cdots, \bar{\lambda}_n$.

命题 5.9

设 φ 是有限维内积空间 V 上的线性变换, φ 的极小多项式为 $g(x)$,证明: φ^* 的极小多项式为 $\bar{g}(x)$,这里 $\bar{g}(x)$ 的系数等于 $g(x)$ 系数的共轭.

证明 取 V 的一组标准正交基,设 $\mathbf{A}$ 是 φ 的表示矩阵,则无论 V 是酉空间还是欧氏空间, φ^* 的表示矩阵总可写为 $\overline{\mathbf{A}}'$. 注意到 $g(\mathbf{A}) = \mathbf{O}$ 当且仅当 $\bar{g}\left(\overline{\mathbf{A}}'\right) = \mathbf{O}$, 故结论成立.

命题 5.10

设 φ 是内积空间 V 上的线性变换,若 U 是 φ 的不变子空间,求证: $U^\perp$ 是 φ^* 的不变子空间.

证明 任取 $\alpha \in U, \beta \in U^\perp$,由 $(\alpha, \varphi^*(\beta)) = (\varphi(\alpha), \beta) = 0$ 即得结论.

笔记 这个命题为内积空间中的相关问题提供了归纳的基础.

命题 5.11

设 V 是 n 维内积空间, φ 是 V 上的线性变换,求证: $\operatorname{Im}\varphi^* = (\operatorname{Ker}\varphi)^\perp$.

证明 由例 5.5 可知,只要证明 $\operatorname{Ker}\varphi = (\operatorname{Im}\varphi^*)^\perp$ 即可. 一方面,任取 $\alpha \in \operatorname{Ker}\varphi$,则对任一 $\beta \in V$ 有 $(\alpha, \varphi^*(\beta)) = (\varphi(\alpha), \beta) = (\mathbf{0}, \beta) = 0$,即 $\alpha \in (\operatorname{Im}\varphi^*)^\perp$, 于是 $\operatorname{Ker}\varphi \subseteq (\operatorname{Im}\varphi^*)^\perp$. 另一方面,任取 $\alpha \in (\operatorname{Im}\varphi^*)^\perp$,则对任一 $\beta \in V$ 有 $0 = (\alpha, \varphi^*(\beta)) = (\varphi(\alpha), \beta)$,令 $\beta = \varphi(\alpha)$ 即得 $\varphi(\alpha) = \mathbf{0}$,即 $\alpha \in \operatorname{Ker}\varphi$,于是 $(\operatorname{Im}\varphi^*)^\perp \subseteq \operatorname{Ker}\varphi$,因此结论得证.

5.4 内积空间的同构，正交变换和酉变换

定理 5.4 (保积同构的性质)

设 V 与 U 都是 n 维内积空间 (同为实空间或同为复空间),若 φ 是 $V \to U$ 的线性映射,则下列命题等价:

(1) φ 保持内积;

(2) φ 是保积同构;

(3) φ 将 V 的任一组标准正交基变成 U 的一组标准正交基;

(4) φ 将 V 的某一组标准正交基变成 U 的一组标准正交基.

♡

定义 5.5 (正交变换与酉变换)

设 V 是欧氏空间,若 φ 是 V 上保持内积的线性变换,则称 φ 为 V 上的正交变换或正交算子. 若 U 是酉空间,则 U 上保持内积的线性变换称为酉变换或酉算子.

♣

笔记 显然正交变换及酉变换都是可逆线性变换. 由定理 5.4, 正交变换可定义为把欧氏空间中一组标准正交基变成标准正交基的线性变换. 酉变换也类似.

定理 5.5

设 φ 是欧氏空间或酉空间上的线性变换,则 φ 是正交变换或酉变换的充分必要条件是 φ 非异,且

$$\varphi^{-1} = \varphi^*$$

♡

定义 5.6 (正交矩阵与酉矩阵)

设 $\mathbf{A}$ 是 n 阶实方阵,若 $\mathbf{A}' = \mathbf{A}^{-1}$,则称 $\mathbf{A}$ 是正交矩阵. 又若 $\mathbf{C}$ 是 n 阶复方阵且 $\overline{\mathbf{C}}' = \mathbf{C}^{-1}$,则称 $\mathbf{C}$ 是酉矩阵.

♣

定理 5.6 (正交变换与酉变换的表示矩阵)

设 φ 是欧氏空间 (酉空间) V 上的正交变换 (酉变换),则在 V 的任一组标准正交基下,φ 的表示矩阵是正交矩阵 (酉阵).

♡

定理 5.7 (正交矩阵的充要条件)

$\mathbf{A}$ 为正交矩阵的充分必要条件是它的 n 个行向量是 n 维实行向量空间组成的欧氏空间 (取标准内积) 的标准正交基; 或它的 n 个列向量是 n 维实列向量空间组成的欧氏空间 (取标准内积) 的标准正交基.

♡

定理 5.8 (正交矩阵的特征值)

若 n 阶实矩阵 $\mathbf{A}$ 是正交矩阵,则

(1) $\mathbf{A}$ 的行列式值等于 1 或 -1 ;

(2) $\mathbf{A}$ 的特征值的绝对值 (模长) 等于 1 .

♡

定理 5.9 (QR 分解)

设 A 是 n 阶实 (复) 矩阵,则 A 可分解为

$$\mathbf{A}=\mathbf{QR}$$

其中 $\mathbf{Q}$ 是正交 (酉) 矩阵, $\mathbf{R}$ 是一个上三角阵且主对角线上的元素均大于等于零, 并且若 $\mathbf{A}$ 是非异阵,则这样的分解必唯一.

命题 5.12

设 V 是 n 维欧氏空间, $\alpha_1,\alpha_2,\cdots,\alpha_n,\beta_1,\beta_2,\cdots,\beta_n\in V$. 证明: 若存在非零向量 $\alpha\in V$,使得 $\sum\limits_{i=1}^{n}(\alpha,\alpha_i)\beta_i=\mathbf{0}$,则必存在非零向量 $\beta\in V$,使得 $\sum\limits_{i=1}^{n}(\beta,\beta_i)\alpha_i=\mathbf{0}$.

证明 取 V 的一组标准正交基 $\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n$,设 α,β 的坐标向量分别为 $\mathbf{x},\mathbf{y}$;α_i 的坐标向量为 $\mathbf{x}_i\,(1\le i\le n)$; β_i 的坐标向量为 $\mathbf{y}_i\,(1\le i\le n)$; n 阶实矩阵 $\mathbf{A}=(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n)$, $\mathbf{B}=(\mathbf{y}_1,\mathbf{y}_2,\cdots,\mathbf{y}_n)$,则由抽象向量映射到坐标向量的保积同构 $\varphi:V\to\mathbb{R}^n$,可把本题化为如下矩阵问题: 若存在非零列向量 $\mathbf{x}$,使得

$$\sum_{i=1}^{n}(\mathbf{x}'\mathbf{x}_i)\,\mathbf{y}_i=\mathbf{BA}'\mathbf{x}=\mathbf{0}\tag{5.2}$$

则必存在非零列向量 $\mathbf{y}$,使得

$$\sum_{i=1}^{n}(\mathbf{y}'\mathbf{y}_i)\,\mathbf{x}_i=\mathbf{AB}'\mathbf{y}=\mathbf{0}\tag{5.3}$$

事实上,由齐次线性方程组 (5.2) 有非零解可得 $\mathrm{r}\,(\mathbf{BA}')<n$,注意到 $\mathbf{AB}'=(\mathbf{BA}')'$. 故 $\mathrm{r}\,(\mathbf{AB}')<n$,于是齐次线性方程组 (5.3) 也有非零解,结论得证.

命题 5.13

设 V 是 n 维欧氏空间, $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 是一组向量, $\mathbf{G}=\mathbf{G}\,(\alpha_1,\alpha_2\,,\cdots,\alpha_m)$ 是其 Gram 矩阵,求证: $\mathrm{r}\,(\alpha_1,\alpha_2,\cdots,\alpha_m)=\mathrm{r}\,(\mathbf{G})$.

证明 取 V 的一组标准正交基 $\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n$,设 α_i 的坐标向量为 $\mathbf{x}_i(1\le i\le m)$, $\mathbf{A}=(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_m)$ 为 $n\times m$ 实矩阵,则由抽象向量映射到坐标向量的保积同构 $\varphi:V\to\mathbb{R}^n$ 可知 $\mathbf{G}=\mathbf{A}'\mathbf{A}$,于是只要证明 $\mathrm{r}\,(\mathbf{A})=\mathrm{r}\,(\mathbf{A}'\mathbf{A})$ 成立即可.

命题 5.14

设 V,U 都是 n 维欧氏空间, $\{\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n\}$ 和 $\{\mathbf{f}_1,\mathbf{f}_2,\cdots,\mathbf{f}_n\}$ 分别是 V 和 U 的一组基 (不一定是标准正交基),线性映射 $\varphi:V\to U$ 满足 $\varphi\,(\mathbf{e}_i)=f_i\,(1\le i\le n)$. 求证: φ 是保积同构的充要条件是这两组基的 Gram 矩阵相等,即

$$\mathbf{G}\,(\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n)=\mathbf{G}\,(\mathbf{f}_1,\mathbf{f}_2,\cdots,\mathbf{f}_n)$$

证明 φ 把 V 的一组基映为 U 的一组基保证了 φ 是线性同构. 若 φ 保持内积,则 $(\mathbf{e}_i,\mathbf{e}_j)=(\varphi\,(\mathbf{e}_i)\,,\varphi\,(\mathbf{e}_j))=(\mathbf{f}_i,\mathbf{f}_j)$,从而它们的 Gram 矩阵相同. 反之,若它们的 Gram 矩阵相同,任取 $\alpha,\beta\in V$,设它们在基 $\{\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n\}$

下的坐标向量分别为 $\mathbf{x},\mathbf{y}$,则 $\varphi(\alpha),\varphi(\beta)$ 在基 $\{\mathbf{f}_1,\mathbf{f}_2,\cdots,\mathbf{f}_n\}$ 下的坐标向量也分别为 $\mathbf{x},\mathbf{y}$,于是

$$(\varphi(\alpha),\varphi(\beta))=x'G(f_1,f_2,\cdots,f_n)y=x'G(e_1,e_2,\cdots,e_n)y=(\alpha,\beta)$$

故 $\varphi:V\to U$ 是保积同构.

命题 5.15

设 V,U 都是 n 维欧氏空间, $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 和 $\{\beta_1,\beta_2,\cdots,\beta_m\}$ 分别是 V 和 U 中的向量组. 证明: 存在保积同构 $\varphi:V\to U$,使得

$$\varphi(\alpha_i)=\beta_i\,(1\le i\le m)$$

成立的充要条件是这两组向量的 Gram 矩阵相等.

证明 必要性类似于命题5.14的必要性的证明,下证充分性. 设向量组 $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 和 $\{\beta_1,\beta_2,\cdots,\beta_m\}$ 有相同的 Gram 矩阵, $V_1=L(\alpha_1,\alpha_2,\cdots,\alpha_m),U_1=L(\beta_1,\beta_2,\cdots,\beta_m)$.

设 $\{\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\}$ 是向量组 $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 的极大无关组,若设 $c_1\beta_{i_1}+c_2\beta_{i_2}+\cdots+c_r\beta_{i_r}=\mathbf{0}$,则可得 $c_1\alpha_{i_1}+c_2\alpha_{i_2}+\cdots+c_r\alpha_{i_r}=\mathbf{0}$,从而 $c_1=c_2=\cdots=c_r=0$,即 $\beta_{i_1},\beta_{i_2},\cdots,\beta_{i_r}$ 线性无关.

又对任意的 $i\ne i_1,i_2,\cdots,i_r$,若设 $\alpha_i=a_1\alpha_{i_1}+a_2\alpha_{i_2}+\cdots+a_r\alpha_{i_r}$,则可得 $\beta_i=a_1\beta_{i_1}+a_2\beta_{i_2}+\cdots+a_r\beta_{i_r}$,于是 $\{\beta_{i_1},\beta_{i_2},\cdots,\beta_{i_r}\}$ 也是向量组 $\{\beta_1,\beta_2,\cdots,\beta_m\}$ 的极大无关组,从而 $\{\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\}$ 和 $\{\beta_{i_1},\beta_{i_2},\cdots,\beta_{i_r}\}$ 分别是 V_1,U_1 的一组基.

定义线性映射 $\varphi_1:V_1\to U_1$ 为 $\varphi_1(\alpha_{i_k})=\beta_{i_k}\,(1\le k\le r)$,则由命题5.14的充分性可知, $\varphi_1:V_1\to U_1$ 是保积同构. 对任意的 $i\ne i_1,i_2,\cdots,i_r$,

$$\varphi_1(\alpha_i)=\varphi_1\left(\sum_{k=1}^r a_k\alpha_{i_k}\right)=\sum_{k=1}^r a_k\varphi_1(\alpha_{i_k})=\sum_{k=1}^r a_k\beta_{i_k}=\beta_i$$

从而 $\varphi_1(\alpha_i)=\beta_i\,(1\le i\le m)$. 注意到 $V=V_1\perp V_1^{\perp},U=U_1\perp U_1^{\perp}$,故可取 $V_1^{\perp}$ 的一组标准正交基 $\gamma_{r+1},\cdots,\gamma_n,U_1^{\perp}$ 的一组标准正交基 $\delta_{r+1},\cdots,\delta_n$,定义线性映射 $\varphi_2:V_1^{\perp}\to U_1^{\perp}$ 为 $\varphi_2(\gamma_j)=\delta_j\,(r+1\le j\le n)$,则 $\varphi_2:V_1^{\perp}\to U_1^{\perp}$ 也是保积同构.

下面定义线性映射 $\varphi:V\to U$,对任一 $v=\alpha+\gamma\in V$,其中 $\alpha\in V_1,\gamma\in V_1^{\perp}$,定义 $\varphi(\mathbf{v})=\varphi_1(\alpha)+\varphi_2(\gamma)$,容易验证 $\varphi:V\to U$ 是线性同构. 我们还有

$$\begin{aligned}(\varphi(\mathbf{v}),\varphi(\mathbf{v}))&=(\varphi_1(\alpha)+\varphi_2(\gamma),\varphi_1(\alpha)+\varphi_2(\gamma))=(\varphi_1(\alpha),\varphi_1(\alpha))+(\varphi_2(\gamma),\varphi_2(\gamma))\\&=(\alpha,\alpha)+(\gamma,\gamma)=(\alpha+\gamma,\alpha+\gamma)=(\mathbf{v},\mathbf{v})\end{aligned}$$

故 $\varphi:V\to U$ 保持范数,从而是满足题目条件的保积同构.

命题 5.16

设 A,B 是 $m\times n$ 实矩阵,求证: $A'A=B'B$ 的充要条件是存在 m 阶正交矩阵 $\mathbf{Q}$,使得 $\mathbf{A}=\mathbf{QB}$.

证明 充分性显然成立,下证必要性. 取 $V=\mathbb{R}^m$ 上的标准内积,设 $\mathbf{A}=(\alpha_1,\alpha_2,\cdots,\alpha_n),\mathbf{B}=(\beta_1,\beta_2,\cdots,\beta_n)$ 为列分块,则由 $\mathbf{A'A}=\mathbf{B'B}$ 可得 $\mathbf{G}(\alpha_1,\alpha_2,\cdots,\alpha_n)=\mathbf{G}(\beta_1,\beta_2,\cdots,\beta_n)$,再由命题5.15可知,存在 V 上的正交变换 φ ,使得 $\varphi(\beta_i)=\alpha_i\,(1\le i\le n)$. 设 φ 在 V 的标准单位列向量构成的标准正交基下的表示矩阵为 $\mathbf{Q}$,则 $\mathbf{Q}$ 为正交矩阵且 $\mathbf{Q}\beta_i=\alpha_i\,(1\le i\le n)$,因此

$$\mathbf{QB}=(\mathbf{Q}\beta_1,\mathbf{Q}\beta_2,\cdots,\mathbf{Q}\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)=\mathbf{A}$$

命题 5.17

设 $\alpha \in \mathbb{R}^n$，$|\alpha| = 1$，证明：存在一个对称的正交矩阵 A，使得 A 的第一列为 α.

证明 令 $e_i = \mathrm{col}_i I_n$，$\mathbb{R}^n$ 的镜面反射 $R_{\alpha-e_1}$ 满足：

$$R_{\alpha-e_1} \in O(\mathbb{R}^n), R_{\alpha-e_1}^2 = id, R_{\alpha-e_1}(e_1) = \alpha$$

设 $R_{\alpha-e_1}$ 在 $e_1, e_2, \cdots, e_n$ 下的矩阵为 A，则有 $A'A = I_n, A^2 = I_n, \mathrm{col}_1 A = \alpha$，从而知道结论得证.

5.5 自伴随算子

引理 5.1

欧氏空间中两组标准正交基之间的过渡矩阵是正交矩阵，酉空间中两组标准正交基之间的过渡矩阵是酉矩阵.

定义 5.7 (正交相似与酉相似)

设 $\mathbf{A}, \mathbf{B}$ 是 n 阶实矩阵,若存在正交矩阵 $\mathbf{P}$,使 $\mathbf{B} = \mathbf{P}'\mathbf{A}\mathbf{P}$ 成立,则称 $\mathbf{B}$ 和 $\mathbf{A}$ 正交相似.
设 $\mathbf{A}, \mathbf{B}$ 是 n 阶复矩阵,若存在酉矩阵 $\mathbf{P}$, 使 $\mathbf{B} = \overline{\mathbf{P}}'\mathbf{A}\mathbf{P}$,则称 $\mathbf{B}$ 和 $\mathbf{A}$ 酉相似.

定义 5.8 (自伴随算子)

设 φ 是内积空间 V 上的线性变换, φ^* 是 φ 的伴随,若 $\varphi^* = \varphi$, 则称 φ 是自伴随算子. 在 V 是欧氏空间的情形, φ 称为对称算子或对称变换, 在 V 是酉空间的情形, φ 称为 Hermite 算子或 Hermite 变换.

定理 5.10

设 V 是 n 维酉空间, φ 是 V 上的自伴随算子,则 φ 的特征值全是实数且属于不同特征值的特征向量互相正交.

推论 5.1

Hermite 矩阵的特征值全是实数, 实对称阵的特征值也全是实数. 这两种矩阵属于不同特征值的特征向量互相正交.

定理 5.11

设 V 是 n 维内积空间, φ 是 V 上的自伴随算子,则存在 V 的一组标准正交基,使得 φ 在这组基下的表示矩阵为实对角阵,且这组基恰为 φ 的 n 个线性无关的特征向量.

推论 5.2

设 $\mathbf{A}$ 是 n 阶 Hermite 矩阵,则存在酉矩阵 $\mathbf{P}$,使得 $\overline{\mathbf{P}}'\mathbf{A}\mathbf{P}$ 为实对角阵,即 Hermite 矩阵酉相似于实对角阵. 又若 $\mathbf{A}$ 是 n 阶实对称阵,则存在正交矩阵 $\mathbf{P}$,使得 $\mathbf{P}'\mathbf{A}\mathbf{P}$ 为对角阵,即实对称阵正交相似于对角阵. 上述正交矩阵或酉矩阵 $\mathbf{P}$ 的 n 个列向量恰为矩阵 $\mathbf{A}$ 的 n 个两两正交且长度等于 1 的特征向量.

定理 5.12

实对称 (Hermite) 矩阵的特征值是实对称 (Hermite) 矩阵正交 (酉) 相似的全系不变量.

定理 5.13

设 $f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ 是 n 变元实二次型, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是矩阵 $\mathbf{A}$ 的特征值,则 f 经正交变换可以化为下列标准型:

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

因此, f 的正惯性指数等于 $\mathbf{A}$ 的正特征值的个数,负惯性指数等于 $\mathbf{A}$ 的负特征值的个数. f 的秩等于 $\mathbf{A}$ 的非零特征值的个数.

♡

定理 5.14

设 $f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ 是 n 变元实二次型,则 f 是正定型当且仅当矩阵 $\mathbf{A}$ 的特征值全是正数, f 是负定型当且仅当矩阵 $\mathbf{A}$ 的特征值全是负数, f 是半正定型当且仅当 $\mathbf{A}$ 的特征值全非负, f 是半负定型当且仅当 $\mathbf{A}$ 的特征值全非正.

♡

5.6 复正规算子

定义 5.9

设 φ 是内积空间 V 上的线性变换, φ^* 是其伴随,若 $\varphi\varphi^* = \varphi^*\varphi$, 则称 φ 是 V 上的正规算子. 为了不引起混淆,我们也称酉空间 (欧氏空间) V 上的正规算子 φ 为复正规算子 (实正规算子). 一个复矩阵 A 若适合 $\overline{A}'A = A\overline{A}'$, 则称其为复正规矩阵. 若 $\mathbf{A}$ 是实矩阵且 $\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}'$,则称其为实正规矩阵. ♣

注 (1) 容易验证酉算子 (酉矩阵)、Hermite 算子 (Hermite 矩阵) 都是复正规算子 (矩阵). 正交变换 (矩阵)、对称变换 (矩阵) 都是实正规算子 (矩阵). 我们将在下一节中详细讨论实正规算子和实正规矩阵.

(2) 容易证明复 (实) 正规算子在任一组标准正交基下的表示矩阵都是复 (实) 正规矩阵. 反之, 选定一组标准正交基, 一个复 (实) 正规矩阵就代表了一个复 (实) 正规算子. 因此复 (实) 矩阵的正规性在酉 (正交) 相似下是不变的.

引理 5.2 (正规算子的性质)

设 φ 是内积空间 V 上的正规算子,则对任意的 $\alpha \in V$,成立

$$\| \varphi(\alpha) \| = \| \varphi^*(\alpha) \|$$

♡

命题 5.18

设 V 是 n 维酉空间, φ 是 V 上的正规算子.

(1) 向量 $\mathbf{u}$ 是 φ 属于特征值 λ 的特征向量的充分必要条件为 $\mathbf{u}$ 是 φ^* 属于特征值 $\bar{\lambda}$ 的特征向量;

(2) 属于 φ 不同特征值的特征向量必正交. ♠

笔记 需要注意到如果 φ 是正规算子, 那么 $\lambda I - \varphi$ 也是正规算子.

引理 5.3

设 V 是 n 维酉空间, φ 是 V 上的线性变换,又 $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ 是 V 的一组标准正交基. 设 φ 在这组基下的表示矩阵 $\mathbf{A}$ 是一个上三角阵,则 φ 是正规算子的充分必要条件是 A 为对角阵. ♡

定理 5.15 (Schur (舒尔) 定理)

设 V 是 n 维酉空间, φ 是 V 上的线性算子,则存在 V 的一组标准正交基,使 φ 在这组基下的表示矩阵为上三角阵. ♡

推论 5.3 (Schur 定理)

任一 n 阶复矩阵均酉相似于一个上三角阵. ♡

定理 5.16

设 V 是 n 维酉空间, φ 是 V 上的正规算子,则 V 有一组标准正交基,在这组标准正交基下, φ 的表示矩阵是对角阵,且这组基向量恰为 φ 的 n 个线性无关的特征向量. ♡

定理 5.17

复矩阵 $\mathbf{A}$ 酉相似于对角阵的充分必要条件是 $\mathbf{A}$ 为复正规矩阵.

命题 5.19

设 φ 是 n 维酉空间 V 上的线性算子,其所有不同的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_k$,则 φ 是正规算子的充分必要条件是

$$V = V_1 \perp V_2 \perp \cdots \perp V_k$$

其中 $V_i\,(i = 1, 2, \cdots, k)$ 是属于特征值 λ_i 的特征子空间.

定理 5.18

任一 n 阶酉矩阵必酉相似于下列对角阵:

$$\operatorname{diag}\{c_1, c_2, \cdots, c_n\}$$

其中 c_i 为模长等于 1 的复数.

5.7 实正规算子

由上一节我们知道,实矩阵 $\mathbf{A}$ 称为正规矩阵若 $\mathbf{A}\mathbf{A}' = \mathbf{A}'\mathbf{A}$. 实正规矩阵的正交相似标准型比复正规矩阵的酉相似标准型要复杂一些, 这是因为任一复矩阵总有特征值与非零特征向量, 而实正规矩阵可能没有实特征值及实特征向量. 为了求得实正规矩阵的正交相似标准型, 我们采用"几何" 方法.

首先利用欧氏空间 V 上的正规算子 φ 的极小多项式的不可约分解将 V 分解成为若干个不变子空间的正交直和,这时要求 φ 限制在每个不变子空间上的极小多项式不超过二次. 这样, 就把问题的研究归结为极小多项式次数不超过二次的正规算子. 因为极小多项式为一次的线性变换就是数量变换, 因此接下去就对极小多项式为二次不可约多项式的正规算子进行讨论. 这样就可以得到实正规矩阵的正交相似标准型.

引理 5.4

设 V 是 n 维欧氏空间, $f(x)$ 是一个实多项式,若 φ 是 V 上的正规算子,则 $f(\varphi)$ 也是 V 上的正规算子.

证明 设

$$f(x) = a_0 + a_1 x + \cdots + a_m x^m$$

则

$$f(\varphi) = a_0\mathbf{I} + a_1\varphi + \cdots + a_m\varphi^m,$$

$$f(\varphi)^* = a_0 I + a_1\varphi^* + \cdots + a_m(\varphi^*)^m = f(\varphi^*)$$

由 $\varphi\varphi^* = \varphi^*\varphi$,不难验证

$$f(\varphi) f(\varphi)^* = f(\varphi)^* f(\varphi)$$

引理 5.5

设 φ 是欧氏空间 V 上的正规算子, $f(x), g(x)$ 是互素的实多项式. 假定 $\mathbf{u} \in \operatorname{Ker} f(\varphi), \mathbf{v} \in \operatorname{Ker} g(\varphi)$,则

$$(\mathbf{u}, \mathbf{v}) = 0$$

证明 因为 $f(x), g(x)$ 互素,故存在实多项式 $s(x), t(x)$,使

$$f(x)s(x) + g(x)t(x) = 1.$$

于是

$$f(\varphi)s(\varphi) + g(\varphi)t(\varphi) = \mathbf{I}.$$

因此 $\mathbf{u} = g(\varphi) t(\varphi)(\mathbf{u})$,

$$(\mathbf{u}, \mathbf{v}) = (g(\varphi) t(\varphi)(\mathbf{u}), \mathbf{v}) = \left(t(\varphi)(\mathbf{u}), g(\varphi)^*(\mathbf{v})\right).$$

由上面的引理 $g(\varphi)$ 是正规算子且 $g(\varphi)(\mathbf{v}) = \mathbf{0}$,由引理 5.2 可得 $g(\varphi)^*(\mathbf{v}) = \mathbf{0}$, 因此 $(\mathbf{u}, \mathbf{v}) = 0$.

定理 5.19

设 V 是 n 维欧氏空间, φ 是 V 上的正规算子. 令 $g(x)$ 是 φ 的极小多项式,且 $g_1(x), \cdots, g_k(x)$ 为 $g(x)$ 的所有互不相同的首一不可约因子, 则 $\deg g_i(x) \le 2$ 且

$$g(x) = g_1(x) \cdots g_k(x).$$

又若 $W_i = \operatorname{Ker} g_i(\varphi)$,则

(1) $W_i \perp W_j \ (i \neq j)$;

(2) $V = W_1 \perp \cdots \perp W_k$;

(3) $W_i\ (i = 1, \cdots, k)$ 是 φ 的不变子空间,且若 φ_i 表示 φ 在 W_i 上的限制则 $g_i(x)$ 是 φ_i 的极小多项式且 φ_i 是 W_i 上的正规算子.

♡

证明 设 φ 在 V 的一组标准正交基下的表示矩阵为正规矩阵 $\mathbf{A}$. 将 $\mathbf{A}$ 看成复矩阵,则 $\mathbf{A}$ 酉相似于对角阵. 因此 $\mathbf{A}$ 的极小多项式 $g(x)$ 无重根. $g(x)$ 是实系数多项式, 其不可约因子或为一次式, 或为二次式, 故上式成立.

令 $f_i(x) = g(x) / g_i(x)$,则 $f_1(x), \cdots, f_k(x)$ 互素, 知

$$f_1(x) h_1(x) + \cdots + f_k(x) h_k(x) = 1.$$

对任意的 $\mathbf{v} \in V$,

$$\mathbf{v} = [f_1(\varphi) h_1(\varphi) + \cdots + f_k(\varphi) h_k(\varphi)](\mathbf{v})$$

注意到 $g_i(\varphi) f_i(\varphi) = g(\varphi) = 0$,故对任一 i, $f_i(\varphi) h_i(\varphi)(v) \in \operatorname{Ker} g_i(\varphi) = W_i$. 当 $i \neq j$ 时, $g_i(x)$ 与 $g_j(x)$ 互素,由引理 5.5 得到 $W_i \perp W_j$. 由上式可知

$$V = W_1 + \cdots + W_k$$

但 W_i 两两正交,故

$$V = W_1 \perp \cdots \perp W_k$$

若 $\mathbf{u} \in W_i$,则

$$g_i(\varphi) \varphi(\mathbf{u}) = \varphi g_i(\varphi)(\mathbf{u}) = \mathbf{0}$$

因此 $\varphi(\mathbf{u}) \in W_i$,即 W_i 是 φ 的不变子空间. 又从 $g_i(\varphi)(\mathbf{u}) = \mathbf{0}$ 对一切 $\mathbf{u} \in W_i$ 成立且 $g_i(x)$ 不可约知道, φ_i 的极小多项式为 $g_i(x)$.

最后因为 $\varphi\varphi^* = \varphi^*\varphi$,所以对任意的 $\alpha \in W_i$,

$$g_i(\varphi)\varphi^*(\alpha)=\varphi^* g_i(\varphi)(\alpha)=\mathbf{0}$$

这说明 W_i 也是 φ^* 的不变子空间. 容易验证 φ_i^* 等于 φ^* 在 W_i 上的限制,于是 φ_i 是 W_i 上的正规算子.

引理 5.6

设 V 是 n 维欧氏空间, φ 是 V 上的正规算子且 φ 适合多项式 $g(x)=x^2+1$. 设 $\mathbf{v}\in V,\mathbf{u}=\varphi(\mathbf{v})$,则

$$\varphi^*(\mathbf{v})=-\mathbf{u},\varphi^*(\mathbf{u})=\mathbf{v},$$

且 $\|\ \mathbf{u}\ \|=\|\ \mathbf{v}\ \|,\mathbf{u}\perp\mathbf{v}$.

证明 由 $\varphi(\mathbf{u})=\varphi^2(\mathbf{v})=-\mathbf{v}$,得

$$0=\|\ \varphi(\mathbf{v})-\mathbf{u}\|^2+\|\ \varphi(\mathbf{u})+\mathbf{v}\|^2$$

$$=\|\ \varphi(\mathbf{v})\ \|^2-2(\varphi(\mathbf{v}),\mathbf{u})+\|\ \mathbf{u}\|^2+\|\ \varphi(\mathbf{u})\ \|^2+2(\varphi(\mathbf{u}),\mathbf{v})+\|\ \mathbf{v}\|^2.$$

因为 φ 是正规算子,由引理 9.6.1 知 $\|\ \varphi(v)\ \|^2=\left\|\varphi^*(v)\right\|^2,\|\ \varphi(u)\ \|^2=\left\|\varphi^*(u)\right\|^2$. 故

$$0=\left\|\varphi^*(\mathbf{v})\right\|^2+2(\varphi^*(\mathbf{v}),\mathbf{u})+\|\ \mathbf{u}\|^2+\left\|\varphi^*(\mathbf{u})\right\|^2-2(\varphi^*(\mathbf{u}),\mathbf{v})+\|\ \mathbf{v}\|^2$$

$$=\left\|\varphi^*(\mathbf{v})+\mathbf{u}\right\|^2+\left\|\varphi^*(\mathbf{u})-\mathbf{v}\right\|^2.$$

于是 $\varphi^*(\mathbf{v})=-\mathbf{u},\varphi^*(\mathbf{u})=\mathbf{v}$. 又

$$(\mathbf{v},\mathbf{u})=(\varphi^*(\mathbf{u}),\mathbf{u})=(\mathbf{u},\varphi(\mathbf{u}))=(\mathbf{u},-\mathbf{v})=-(\mathbf{v},\mathbf{u}),$$

因此 $(\mathbf{v},\mathbf{u})=0,\mathbf{v}\perp\mathbf{u}$. 最后

$$\|\ \mathbf{v}\|^2=(\varphi^*(\mathbf{u}),\mathbf{v})=(\mathbf{u},\varphi(\mathbf{v}))=(\mathbf{u},\mathbf{u})=\|\ \mathbf{u}\|^2.$$

引理 5.7

设 V 是 n 维欧氏空间, φ 是 V 上的正规算子且 φ 适合多项式 $g(x)=(x-a)^2+b^2$,其中 a,b 都是实数且 $b\neq 0$. 设 $\mathbf{v}\in V,\mathbf{u}=b^{-1}(\varphi-a\mathbf{I})(\mathbf{v})$ 则 $\|\ \mathbf{u}\ \|=\|\ \mathbf{v}\ \|,\mathbf{u}\perp\mathbf{v}$,且

$$\varphi(\mathbf{v})=a\mathbf{v}+b\mathbf{u},\varphi(\mathbf{u})=-b\mathbf{v}+a\mathbf{u}$$

$$\varphi^*(\mathbf{v})=a\mathbf{v}-b\mathbf{u},\varphi^*(\mathbf{u})=b\mathbf{v}+a\mathbf{u}.$$

证明 令 $\psi=b^{-1}(\varphi-aI)$,则 ψ 适合多项式 x^2+1 ,且 $\mathbf{u}=\psi(\mathbf{v})$. 又 $\psi^*=b^{-1}(\varphi^*-a\mathbf{I}),\psi$ 是 V 上的正规算子. 由引理 5.7 即可得到所需结论.

笔记 注意到此时 φ 实际上是在 $L(\mathbf{u},\mathbf{v})$ 上的一个仿射变换.

定理 5.20

设 φ 是 n 维欧氏空间 V 上的正规算子, φ 的极小多项式为 $g(x)=(x-a)^2+b^2$,其中 a,b 是实数且 $b\neq 0$,则存在 s ,使 $g(x)^s$ 是 φ 的特征多项式且存在 V 的 s 个二维子空间 $V_1,\cdots,V_s$,使

$$V=V_1\perp\cdots\perp V_s$$

每个 V_i 有标准正交基 $\{\mathbf{u}_i,\mathbf{v}_i\}$,且

$$\varphi(\mathbf{u}_i)=a\mathbf{u}_i-b\mathbf{v}_i,\varphi(\mathbf{v}_i)=b\mathbf{u}_i+a\mathbf{v}_i.$$

证明 对 V 的维数 n 进行归纳.

当 $n=0$ 时,结论是平凡的.

当 $n=1$ 时,任取 V 中的非零向量 $\mathbf{v}$,设 $\varphi(\mathbf{v})=c\mathbf{v}$,其中 c 是实数,则 c 是 φ 的特征值. 因此 c 必须适合 φ 的极小多项式 $g(x)$,但 $g(c)=(c-a)^2+b^2>0$,这个矛盾说明 $n=1$ 的情形不可能发生. 设维数小于 n 时结论已成立,现证明 n 维欧氏空间的情形.

任取 V 中长度等于 1 的向量 $\mathbf{v}_1$,令 $\mathbf{u}_1=b^{-1}(\varphi-a\mathbf{I})(\mathbf{v}_1)$,则 $\mathbf{v}_1,\mathbf{u}_1$ 是两个长度等于 1 的正交向量. 令 V_1 是由 $\mathbf{v}_1,\mathbf{u}_1$ 张成的子空间,则

$$\varphi(\mathbf{u}_1)=a\mathbf{u}_1-b\mathbf{v}_1,\varphi(\mathbf{v}_1)=b\mathbf{u}_1+a\mathbf{v}_1$$

$$\varphi^*(\mathbf{u}_1)=a\mathbf{u}_1+b\mathbf{v}_1,\varphi^*(\mathbf{v}_1)=-b\mathbf{u}_1+a\mathbf{v}_1.$$

因此 V_1 是 φ 和 φ^* 的不变子空间. 令 $W=V_1^{\perp}$,则由命题 5.8 知 W 是 φ 和 φ^* 的不变子空间. 因为 $\dim W=n-2$,由归纳假设存在 $s-1$ 个二维子空间 $V_2,\cdots,V_s$,使

$$W=V_2\perp\cdots\perp V_s$$

且每个 $V_i\,(i=2,\cdots,s)$ 有标准正交基 $\{\mathbf{u}_i,\mathbf{v}_i\}$ 满足条件. 因此

$$V=V_1\perp V_2\perp\cdots\perp V_s$$

且 φ 在标准正交基 $\{\mathbf{u}_1,\mathbf{v}_1,\mathbf{u}_2,\mathbf{v}_2,\cdots,\mathbf{u}_s,\mathbf{v}_s\}$ 下的表示矩阵为分块对角阵:

$$\mathbf{A}=\operatorname{diag}\{\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_s\}$$

其中

$$\mathbf{A}_i=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

显然 $\mathbf{A}$ 的特征多项式为 $g(x)^s$.

笔记 利用本定理的结论可以将复杂的实正规算子进行分析，变成简单的情况.

定理 5.21

设 V 是 n 维欧氏空间, φ 是 V 上的正规算子,则存在一组标准正交基,使 φ 在这组基下的表示矩阵为下列分块对角阵:

$$\operatorname{diag}\left\{\mathbf{A}_1,\cdots,\mathbf{A}_r,c_{2r+1},\cdots,c_n\right\}$$

其中 $c_j\,(j=2r+1,\cdots,n)$ 是实数, $\mathbf{A}_i$ 为形如

$$\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$$

的二阶实矩阵.

证明 由定理 5.19 可知

$$V=W_1\perp W_2\perp\cdots\perp W_k$$

其中 $W_i=\operatorname{Ker} g_i(\varphi)$, $g_i(x)$ 是次数不超过 2 的多项式,并且 φ 在 W_i 上的限制 φ , 是 W_i 上的正规算子,其极小多项式是 $g_i(x)$. 对每个 W_i ,若 $g_i(x)$ 是二次多项式则由定理 5.20 知 W_i 可分解为若干个二维子空间的正交直和. 若 $g_i(x)=x-c_i$, 则 $\varphi_i-c_i\mathbf{I}=\mathbf{0}$,即 $\varphi_i=c_i\mathbf{I}$. 由此即可得到所需结论.

笔记 定理5.21 中的分块对角阵就是实正规矩阵的正交相似标准型. 在不计对角线上块的次序的意义下, 实正规矩阵的正交相似标准型是唯一确定的. 因此, 实正规矩阵的特征值是正交相似关系的全系不变量.

定理 5.22 (正交矩阵的正交相似标准型)

设 $\mathbf{A}$ 是 n 阶正交矩阵,则 $\mathbf{A}$ 正交相似于下列分块对角阵:

$$\operatorname{diag}\left\{\mathbf{A}_1,\cdots,\mathbf{A}_r;1,\cdots,1;-1,\cdots,-1\right\}$$

其中

$$\mathbf{A}_i=\begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix}, i=1,\cdots,r$$

证明 由于正交矩阵是正规矩阵,由定理知道 $\mathbf{A}$ 必正交相似于

$$\operatorname{diag}\left\{\mathbf{A}_1,\cdots,\mathbf{A}_r;c_{2r+1},\cdots,c_n\right\}.$$

又由正交矩阵的性质知 $|c_i|=1\,(i=2r+1,\cdots,n)$. 另一方面,设

$$A_i=\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$$

则对应的极小多项式为 $g(x)=(x-a_i)^2+b_i^2=x^2-2a_ix+a_i^2+b_i^2$, 由正交矩阵的特征值模长为 1, 应用韦达定理得$a_i^2+b_i^2=\lambda\bar{\lambda}=1$,故可设 $a_i=\cos\theta_i, b_i=\sin\theta_i$.

定义 5.10 (反对称阵)

设 $\mathbf{A}$ 是 n 阶实方阵,若 $\mathbf{A}$ 适合下列条件:

$$\mathbf{A}' = -\mathbf{A}$$

则称 $\mathbf{A}$ 是实反对称阵或实斜对称阵.

由于 $\mathbf{A}\mathbf{A}' = -\mathbf{A}^2 = \mathbf{A}'\mathbf{A}$,故实反对称阵是正规矩阵. 若 $\mathbf{P}$ 是正交矩阵,则

$$(\mathbf{P}'\mathbf{A}\mathbf{P})' = \mathbf{P}'\mathbf{A}'\mathbf{P} = -\mathbf{P}'\mathbf{A}\mathbf{P}.$$

因此, 反对称性在正交相似下保持不变. 下面是实反对称阵的正交相似标准型定理.

定理 5.23 (反对称矩阵的正交相似标准型)

设 $\mathbf{A}$ 是实反对称阵,则 $\mathbf{A}$ 正交相似于下列分块对角阵:

$$\operatorname{diag}\{\mathbf{B}_1, \cdots, \mathbf{B}_r; 0, \cdots, 0\}$$

其中

$$B_i = \begin{pmatrix} 0 & b_i \\ -b_i & 0 \end{pmatrix}, i = 1, \cdots, r$$

证明 设 $\mathbf{A}$ 正交相似于

$$\operatorname{diag}\{\mathbf{B}_1, \cdots, \mathbf{B}_r; c_{2r+1}, \cdots, c_n\}.$$

由 $\mathbf{B}' = -\mathbf{B}$ 即得 $c_i = 0\,(i = 2r+1, \cdots, n)$. 再由 $\mathbf{B}_i' = -\mathbf{B}_i$ 知道, $\mathbf{B}_i$ 必具有定理所需之形状.

推论 5.4

实反对称阵的秩必是偶数, 且其实特征值必为 0 , 虚特征值为纯虚数.

5.8 谱

设 V 是 n 维酉空间, φ 是 V 上的正规算子, φ 全体不同的特征值设为 $\lambda_1, \lambda_2, \cdots, \lambda_k$. 由命题 5.19 我们知道,存在 V 的一组标准正交基,使得 φ在这组基下的表示矩阵为对角阵

$$\mathbf{A} = \operatorname{diag}\{\lambda_1, \cdots, \lambda_1; \lambda_2, \cdots, \lambda_2; \cdots; \lambda_k, \cdots, \lambda_k\}.$$

假定特征值 λ_i 的重数等于 r_i ,则在分解式

$$V = W_1 \perp W_2 \perp \cdots \perp W_k$$

中, W_i 是 φ 的特征子空间且维数等于 r_i . 若记 $\mathbf{D_i}$ 为如下对角阵

$$\mathbf{D}_i = \operatorname{diag}\{0, \cdots, 0; \cdots; 1, \cdots, 1; \cdots; 0, \cdots, 0\},$$

即主对角元有 r_i 个 1,其余都是 0 的对角阵,则

$$\mathbf{A} = \lambda_1 \mathbf{D}_1 + \lambda_2 \mathbf{D}_2 + \cdots + \lambda_k \mathbf{D}_k$$

诸 $\mathbf{D}_i$ 适合条件 $\mathbf{D}_i^2 = \mathbf{D}_i, \mathbf{D}_i \mathbf{D}_j = \mathbf{O}\,(i \neq j), \mathbf{D}_1 + \mathbf{D}_2 + \cdots + \mathbf{D}_k = \mathbf{I}_n$.

对实对称阵也有类似的结论. 如果把上面的结论“翻译” 成几何语言就是下面的谱分解定理. 我们将给出一个用“几何” 方法的证明.

定理 5.24 (谱分解定理)

设 V 是有限维内积空间, φ 是 V 上的线性算子,当 V 是酉空间时 φ 为正规算子; 当 V 是欧氏空间时 φ 为自伴随算子. $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 φ 全体不同的特征值, W_i 为 φ 属于 λ_i 的特征子空间,则 V 是 $W_i\,(i = 1, 2, \cdots, k)$ 的正交直和. 这时若设 $\mathbf{E}_i$ 是 V 到 W_i 上的正交投影,则 φ 有下列分解式:

$$\varphi = \lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \cdots + \lambda_k \mathbf{E}_k$$

♡

证明 由命题 5.19 知道

$$V = W_1 \perp W_2 \perp \cdots \perp W_k$$

又因为 $\mathbf{E}_i$ 是 $V \to W_i$ 的正交投影,故

$$\mathbf{I} = \mathbf{E}_1 + \mathbf{E}_2 + \cdots + \mathbf{E}_k$$

注意 $\varphi \mathbf{E}_i = \lambda_i \mathbf{E}_i$,于是

$$\begin{aligned} \varphi &= \varphi \mathbf{E}_1 + \varphi \mathbf{E}_2 + \cdots + \varphi \mathbf{E}_k \\ &= \lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \cdots + \lambda_k \mathbf{E}_k \end{aligned}$$

引理 5.8

令 $f_j(x)=\prod\limits_{i\neq j}\dfrac{x-\lambda_i}{\lambda_j-\lambda_i}$,则 $\mathbf{E}_j=f_j(\varphi)$.

证明 当 $i\neq j$ 时 $\mathbf{E}_i\mathbf{E}_j=\mathbf{0}$,故

$$\varphi^2=\lambda_1^2\mathbf{E}_1+\lambda_2^2\mathbf{E}_2+\cdots+\lambda_k^2\mathbf{E}_k$$

同理不难证明

$$\varphi^n=\lambda_1^n\mathbf{E}_1+\lambda_2^n\mathbf{E}_2+\cdots+\lambda_k^n\mathbf{E}_k$$

对一切正整数 n 成立. 若设

$$f(x)=a_0+a_1x+\cdots+a_nx^n,$$

则

$$f(\varphi)=a_0\mathbf{I}+a_1\varphi+\cdots+a_n\varphi^n$$

$$=a_0\left(\sum_{i=1}^{k}\mathbf{E}_i\right)+a_1\left(\sum_{i=1}^{k}\lambda_i\mathbf{E}_i\right)+\cdots+a_n\left(\sum_{i=1}^{k}\lambda_i^n\mathbf{E}_i\right)$$

$$=\sum_{i=1}^{k}f(\lambda_i)\mathbf{E}_i$$

由 $f_j(\lambda_j)=1, f_j(\lambda_i)=0\,(j\neq i)$ 即得 $f_j(\varphi)=\mathbf{E}_j$.

推论 5.5

设 φ 是酉空间 V 上的线性算子,则 φ 是正规算子的充分必要条件是存在复系数多项式 $f(x)$,使 $\varphi^*=f(\varphi)$.

证明 若存在复系数多项式 $f(x)$,使 $\varphi^*=f(\varphi)$,则 $\varphi\varphi^*=\varphi^*\varphi$,即 φ 是正规算子.

若 φ 是正规算子,由定理 5.24 知 φ 存在谱分解:

$$\varphi=\lambda_1\mathbf{E}_1+\lambda_2\mathbf{E}_2+\cdots+\lambda_k\mathbf{E}_k$$

注意到 $\mathbf{E}_i$ 是自伴随算子,故

$$\varphi^*=\bar{\lambda}_1\mathbf{E}_1+\bar{\lambda}_2\mathbf{E}_2+\cdots+\bar{\lambda}_k\mathbf{E}_k$$

采用与引理 5.8 相同的记号,令 $f(x)=\sum\limits_{j=1}^{k}\bar{\lambda}_jf_j(x)$,则

$$f(\varphi)=\sum_{j=1}^{k}\bar{\lambda}_jf_j(\varphi)=\sum_{j=1}^{k}\bar{\lambda}_j\mathbf{E}_j=\varphi^*$$

定义 5.11

设 φ 是内积空间 V 上的自伴随算子,若对任意的非零向量 $\alpha \in V$,总有 $(\varphi(\alpha),\alpha) > 0\,((\varphi(\alpha),\alpha) \geq 0)$,则称 φ 为正定 (半正定) 自伴随算子.

笔记 容易证明 φ 是正定自伴随算子当且仅当 φ 在 V 的一组标准正交基下的表示矩阵是正定 Hermite 矩阵 (酉空间时) 或正定实对称阵 (欧氏空间时); φ 是半正定自伴随算子当且仅当 φ 在 V 的一组标准正交基下的表示矩阵是半正定 Hermite 矩阵 (酉空间时) 或半正定实对称阵 (欧氏空间时).

定理 5.25

设 φ 是酉空间 V 上的正规算子. 若 φ 的特征值全是实数,则 φ 是自伴随算子; 若 φ 的特征值全是非负实数,则 φ 是半正定自伴随算子; 若 φ 的特征值全是正实数,则 φ 是正定自伴随算子; 若 φ 的特征值的模长等于 1,则 φ 是酉算子.

证明 设 φ 的谱分解为

$$\varphi = \lambda_1\mathbf{E}_1 + \lambda_2\mathbf{E}_2 + \cdots + \lambda_k\mathbf{E}_k$$

则

$$\varphi^* = \bar{\lambda}_1\mathbf{E}_1 + \bar{\lambda}_2\mathbf{E}_2 + \cdots + \bar{\lambda}_k\mathbf{E}_k.$$

若 φ 的特征值全是实数,则 $\varphi^* = \varphi$,即 φ 是自伴随算子.
若 λ_i 全是非负实数,任取非零向量 $\alpha \in V$,有

$$\alpha = \mathbf{E}_1(\alpha) + \mathbf{E}_2(\alpha) + \cdots + \mathbf{E}_k(\alpha),$$

$$\varphi(\alpha) = \lambda_1\mathbf{E}_1(\alpha) + \lambda_2\mathbf{E}_2(\alpha) + \cdots + \lambda_k\mathbf{E}_k(\alpha),$$

从而

$$(\varphi(\alpha),\alpha) = \lambda_1\left\|\mathbf{E}_1(\alpha)\right\|^2 + \lambda_2\left\|\mathbf{E}_2(\alpha)\right\|^2 + \cdots + \lambda_k\left\|\mathbf{E}_k(\alpha)\right\|^2 \geq 0.$$

同理,若特征值全是正实数,则 φ 是正定自伴随算子. 最后若 $|\lambda_i| = 1$,则

$$\begin{aligned}
\varphi\varphi^* &= \lambda_1\bar{\lambda}_1\mathbf{E}_1 + \lambda_2\bar{\lambda}_2\mathbf{E}_2 + \cdots + \lambda_k\bar{\lambda}_k\mathbf{E}_k \\
&= |\lambda_1|^2\mathbf{E}_1 + |\lambda_2|^2\mathbf{E}_2 + \cdots + |\lambda_k|^2\mathbf{E}_k \\
&= \mathbf{E}_1 + \mathbf{E}_2 + \cdots + \mathbf{E}_k \\
&= I
\end{aligned}$$

也即 φ 是酉算子.

定理 5.26

设 V 是有限维内积空间, φ 是 V 上的半正定自伴随算子,则存在 V 上唯一的半正定自伴随算子 ψ ,使 $\psi^2 = \varphi$.

证明 设 φ 的谱分解式为

$$\varphi = \lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \cdots + \lambda_k \mathbf{E}_k$$

令 $d_i = \sqrt{\lambda_i}\,(i = 1, 2, \cdots, k)$,则

$$\psi = d_1 \mathbf{E}_1 + d_2 \mathbf{E}_2 + \cdots + d_k \mathbf{E}_k$$

适合 $\psi^2 = \varphi$ 且 ψ 也是半正定自伴随算子.

现设 θ 是 V 上的半正定自伴随算子且 $\theta^2 = \varphi$,我们要证明 $\theta = \psi$,令

$$\theta = b_1 \mathbf{F}_1 + b_2 \mathbf{F}_2 + \cdots + b_r \mathbf{F}_r$$

是 θ 的谱分解,其中 $\mathbf{F}_i$ 是正交投影算子且 b_i 为非负实数. 由 $\theta^2 = \varphi$ 得

$$\varphi = b_1^2 \mathbf{F}_1 + b_2^2 \mathbf{F}_2 + \cdots + b_r^2 \mathbf{F}_r$$

故 b_i^2 是 φ 的特征值而 b_i 互不相同,因此 $r = k, b_i = d_i$ (这里允许差一个次序). 注意到 $\mathbf{E}_i(V)$ 及 $\mathbf{F}_i(V)$ 都是 φ 的关于特征值 λ_i 的特征子空间,因此 $\mathbf{F}_i = \mathbf{E}_i$,这就证明了 $\theta = \psi$.

推论 5.6

设 $\mathbf{A}$ 是半正定实对称 (Hermite) 矩阵,则必存在唯一的半正定实对称 (Hermite) 矩阵 $\mathbf{B}$,使 $\mathbf{A} = \mathbf{B}^2$. ♡

定理 5.27

设 V 是 n 维酉空间 (欧氏空间), φ 是 V 上的任一线性算子,则存在 V 上的酉算子 (正交算子) ω 以及 V 上的半正定自伴随算子 ψ ,使 $\varphi = \omega\psi$, 其中 ψ 是唯一的,并且若 φ 是非异线性算子,则 ω 也唯一. ♡

证明 若已有 $\varphi = \omega\psi$,其中 ω 为酉算子, ψ 为半正定自伴随算子,则

$$\varphi^* = \psi^* \omega^* = \psi \omega^*$$

$$\varphi^* \varphi = \psi \omega^* \omega \psi = \psi^2$$

由定义容易验证 $\varphi^*\varphi$ 是半正定自伴随算子,故由定理 5.26 知, ψ 被 φ 唯一确定.

现来证明存在性. 令 ψ 是上面定理中的 ψ 且使 $\psi^2 = \varphi^*\varphi$. 若 φ 是非异线性变换,则 ψ 也是非异的. 事实上,这时

$$(\psi(\mathbf{v}), \psi(\mathbf{v})) = \left(\psi^2(\mathbf{v}), \mathbf{v}\right) = (\varphi^*\varphi(\mathbf{v}), \mathbf{v}) = (\varphi(\mathbf{v}), \varphi(\mathbf{v}))$$

对一切 $\mathbf{v} \in V$ 成立. 显然从 φ 非异即可推出 ψ 也非异. 这时可令 $\omega = \varphi\psi^{-1}$,只需证明 ω 是酉算子 (正交算子) 即可. 注意到

$$\omega^* = \left(\varphi\psi^{-1}\right)^* = \left(\psi^{-1}\right)^* \varphi^* = (\psi^*)^{-1} \varphi^* = \psi^{-1} \varphi^*$$

$$\omega^*\omega=\psi^{-1}\varphi^*\varphi\psi^{-1}=\psi^{-1}\psi^2\psi^{-1}=\mathbf{I}$$

因此 ω 是酉算子 (正交算子).

若 φ 不是非异线性变换,现来定义 ω . 设 $W=\operatorname{Im}\psi, W^{\perp}$ 是其正交补空间. 定义 $W=\operatorname{Im}\psi\to\operatorname{Im}\varphi$ 的映射 η 如下: 若 $\psi(\mathbf{u})\in W$,则

$$\eta(\psi(\mathbf{u}))=\varphi(\mathbf{u})$$

这时我们必须验证 η 定义的合理性. 即若 $\psi(\mathbf{u})=\psi(\mathbf{v})$,必须 $\varphi(\mathbf{u})=\varphi(\mathbf{v})$. 由上式可知对任意的 $\alpha\in V,\|\psi(\alpha)\|^2=\|\varphi(\alpha)\|^2$,因此 $\psi(\mathbf{u}-\mathbf{v})=\mathbf{0}$ 当且仅当 $\varphi(\mathbf{u}-\mathbf{v})=\mathbf{0}$. 这表明 η 是一个合理定义的映射. 又 η 显然是线性的.

再定义 $W^{\perp}\to(\operatorname{Im}\varphi)^{\perp}$ 的映射 ξ: 由 $(\psi(\mathbf{v}),\psi(\mathbf{v}))=(\varphi(\mathbf{v}),\varphi(\mathbf{v}))$ 式可知 φ 与 ψ 的核空间相同, 故像空间的维数相等. 于是 $W^{\perp}$ 与 $(\operatorname{Im}\varphi)^{\perp}$ 的维数相同,故必存在一个保持内积的同构,这个同构记为 ξ . 由于 $V=W\oplus W^{\perp}$,因此 V 中任一向量 $\mathbf{v}$ 均可唯一地表示为

$$\mathbf{v}=\mathbf{w}+\mathbf{w}'$$

其中 $\mathbf{w}\in W,\mathbf{w}'\in W^{\perp}$. 令

$$\omega(\mathbf{v})=\eta(\mathbf{w})+\xi(\mathbf{w}')$$

不难看出 ω 是 V 上的线性变换且若设 $\mathbf{w}=\psi(\mathbf{u})$,则

$$\begin{aligned}(\omega(\mathbf{v}),\omega(\mathbf{v})) &= (\eta(\mathbf{w})+\xi(\mathbf{w}'),\eta(\mathbf{w})+\xi(\mathbf{w}'))\\ &= (\varphi(\mathbf{u})+\xi(\mathbf{w}'),\varphi(\mathbf{u})+\xi(\mathbf{w}'))\\ &= (\varphi(\mathbf{u}),\varphi(\mathbf{u}))+(\xi(\mathbf{w}'),\xi(\mathbf{w}'))\\ &= (\psi(\mathbf{u}),\psi(\mathbf{u}))+(\mathbf{w}',\mathbf{w}')\\ &= (\mathbf{w},\mathbf{w})+(\mathbf{w}',\mathbf{w}')\\ &= (\mathbf{v},\mathbf{v})\end{aligned}$$

因此 ω 是酉算子 (正交算子). 显然这时 $\varphi=\omega\psi$.

定义 5.12 (极分解)

我们称 $\varphi=\omega\psi$ 是一个极分解.

φ 也可作这样的分解: $\varphi=\psi_1\omega_1$,这里 ω_1 为酉算子 (正交算子), ψ_1 为半正定自伴随算子. 这个式子也称为极分解.

证明只需先对 φ^* 作如上述定理之极分解,再求 φ^* 的伴随即 φ 就可以了.

♣

推论 5.7

设 $\mathbf{A}$ 是 n 阶实矩阵,则存在 n 阶正交矩阵 $\mathbf{Q}$ 以及 n 阶半正定实对称阵 $\mathbf{S}$,使 $\mathbf{A}=\mathbf{QS}$. 又设 $\mathbf{B}$ 是 n 阶复矩阵,则存在 n 阶酉矩阵 $\mathbf{U}$ 以及 n 阶半正定 Hermite 矩阵 $\mathbf{H}$,使 $\mathbf{B}=\mathbf{UH}$. 上述分解式当 $\mathbf{A},\mathbf{B}$ 为非异阵时被唯一确定.

♡

5.9 奇异值分解

定义 5.13 (奇异值及奇异向量)

设 $\mathbf{A}$ 是 $m\times n$ 实矩阵,如果存在非负实数 σ 以及 n 维非零实列向量 α, m 维非零实列向量 β ,使

$$\mathbf{A}\alpha=\sigma\beta,\ \ \mathbf{A}'\beta=\sigma\alpha,$$

则称 σ 是 $\mathbf{A}$ 的奇异值, α,β 分别称为 $\mathbf{A}$ 关于 σ 的右奇异向量与左奇异向量. ♣

为了从几何上描述奇异值问题, 我们引入线性映射的伴随概念, 它可以看成是内积空间上线性变换的伴随概念的推广.

定义 5.14 (线性映射的伴随)

设 V,U 分别是 n 维, m 维欧氏空间, φ 是 $V\to U$ 的线性映射. 若存在 $U\to V$ 的线性映射 φ^* ,使对任意的 $\mathbf{v}\in V,\mathbf{u}\in U$,都有

$$(\varphi(\mathbf{v}),\mathbf{u})=(\mathbf{v},\varphi^*(\mathbf{u}))$$

成立,则称 φ^* 是 φ 的伴随. ♣

定理 5.28 (线性映射伴随的存在性与唯一性)

设 V,U 分别是 n 维, m 维欧氏空间, φ 是 $V\to U$ 的线性映射, 则 φ 的伴随 φ^* 存在且唯一. ♡

证明

不妨设:

$$\begin{array}{rccc}\varphi: & V & \to & U\\ & \mathbf{v} & \mapsto & A\mathbf{v}\end{array}$$

我们设:

$$\begin{array}{rccc}\psi: & U & \to & V\\ & \mathbf{u} & \mapsto & A'\mathbf{u}\end{array}$$

从而对任意的 $\mathbf{v}\in V,\mathbf{u}\in U$，我们有：

$$(\varphi(\mathbf{v}),\mathbf{u})=(A\mathbf{v})'\mathbf{u}=\mathbf{v}'(A'\mathbf{u})=(\mathbf{v},\psi(\mathbf{u}))$$

故 ψ 为 φ 的伴随，从而存在性得证，下面证明唯一性：

假设存在 φ^* 与 $\varphi^\#$ 均为 φ 的伴随，从而我们知道：

$$(\alpha,\varphi^*(\beta))=(\alpha,\varphi^\#(\beta))$$

也即对任意的 $\alpha\in V,\beta\in U$：

$$(\alpha,(\varphi^*-\varphi^\#)(\beta))=0$$

从而可以知道 $\varphi^*-\varphi^\#=0$，故 $\varphi^*=\varphi^\#$，唯一性得证.

从伴随的定义我们不难发现,若取定 V 的一组标准正交基 $\{e_1,e_2,\cdots,e_n\}$, U 的一组标准正交基

$\{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_m\}$,设 φ 在这两组基下的表示矩阵为 $\mathbf{A}$,则 φ^* 在这两组基下的表示矩阵为 A' ,证明也和线性变换的情形相同. 因此,奇异值与奇异向量的几何定义即为下列等式成立:

$$\varphi(\mathbf{v}) = \sigma\mathbf{u},\ \varphi^*(\mathbf{u}) = \sigma\mathbf{v},$$

其中 $\sigma \geq 0, v \in V, u \in U$ 都是非零向量. 不难验证 $\varphi^*\varphi$ 是 V 上的半正定自伴随算子, $\varphi\varphi^*$ 是 U 上的半正定自伴随算子. 又

$$\varphi^*\varphi(\mathbf{v}) = \varphi^*(\sigma\mathbf{u}) = \sigma\varphi^*(\mathbf{u}) = \sigma^2\mathbf{v},$$

因此, σ^2 是 $\varphi^*\varphi$ 的特征值, v 是 $\varphi^*\varphi$ 的属于 σ^2 的特征向量. 同理, σ^2 也是 $\varphi\varphi^*$ 的特征值, $\mathbf{u}$ 是 $\varphi\varphi^*$ 的属于 σ^2 的特征向量.

定理 5.29

设 V, U 分别是 n 维, m 维欧氏空间, φ 是 $V \to U$ 的线性映射, 则存在 V 和 U 的标准正交基,使 φ 在这两组基下的表示矩阵为

$$\begin{pmatrix} S & O \\ O & O \end{pmatrix}$$

其中

$$S = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}$$

是一个 r 阶对角阵, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ 是 φ 的非零奇异值.

证明 因为 $\varphi^*\varphi$ 是 V 上的半正定自伴随算子,故存在 V 的一组标准正交基 $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$,使 $\varphi^*\varphi$ 在这组基下的表示矩阵为 n 阶对角阵 $\mathrm{diag}\{\lambda_1, \cdots, \lambda_r, 0, \cdots, 0\}$,其中 $r = \mathrm{r}(\varphi^*\varphi) = \mathrm{r}(\varphi)$ 且 $\lambda_1 \geq \cdots \geq \lambda_r > 0$ 为 $\varphi^*\varphi$ 的正特征值,从而有

$$\varphi^*\varphi(e_i) = \lambda_i e_i,\ 1 \leq i \leq r;\ \varphi^*\varphi(e_j) = \mathbf{0},\ r+1 \leq j \leq n.$$

令 $\sigma_i = \sqrt{\lambda_i}\,(i = 1, 2, \cdots, r)$ 为算术平方根. 注意到对任意的 $1 \leq i \leq r$,

$$\|\varphi(\mathbf{e}_i)\|^2 = (\varphi(\mathbf{e}_i), \varphi(\mathbf{e}_i)) = (\varphi^*\varphi(\mathbf{e}_i), \mathbf{e}_i) = \lambda_i(\mathbf{e}_i, \mathbf{e}_i) = \sigma_i^2\|\mathbf{e}_i\|^2 = \sigma_i^2,$$

即 $\left\|\varphi(\mathbf{e}_i)\right\| = \sigma_i$; 对任意的 $1 \leq i \neq j \leq r$,

$$(\varphi(\mathbf{e}_i), \varphi(\mathbf{e}_j)) = (\varphi^*\varphi(\mathbf{e}_i), \mathbf{e}_j) = \lambda_i(\mathbf{e}_i, \mathbf{e}_j) = 0$$

又对任意的 $r+1 \leq j \leq n$,

$$\left\|\varphi\left(\mathbf{e}_j\right)\right\|^2=\left(\varphi\left(\mathbf{e}_j\right),\varphi\left(\mathbf{e}_j\right)\right)=\left(\varphi^*\varphi\left(\mathbf{e}_j\right),\mathbf{e}_j\right)=0,$$

即 $\varphi\left(e_j\right)=0$. 令

$$\mathbf{f}_i=\frac{1}{\sigma_i}\varphi\left(\mathbf{e}_i\right),i=1,2,\cdots,r$$

则 $\mathbf{f}_1,\mathbf{f}_2,\cdots,\mathbf{f}_r$ 是 U 中一组两两正交的单位向量, 可将它们扩张为 U 的一组标准正交基 $\{\mathbf{f}_1,\cdots,\mathbf{f}_r,\mathbf{f}_{r+1},\cdots,\mathbf{f}_m$. 于是在 V 的标准正交基 $\{\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n\}$ 和 U 的标准正交基 $\{\mathbf{f}_1,\mathbf{f}_2,\cdots,\mathbf{f}_m\}$ 下, φ 满足:

$$\varphi\left(\mathbf{e}_i\right)=\sigma_i\mathbf{f}_i,1\leq i\leq r;\ \varphi\left(\mathbf{e}_j\right)=\mathbf{0},r+1\leq j\leq n.$$

由 φ^* 在上述两组标准正交基下的表示矩阵是 φ 的表示矩阵的转置可得

$$\varphi^*\left(\mathbf{f}_i\right)=\sigma_i\mathbf{e}_i,1\leq i\leq r;\ \varphi^*\left(\mathbf{f}_j\right)=\mathbf{0},r+1\leq j\leq m,$$

这就得到了要证的结论.

推论 5.8

设 $\mathbf{A}$ 是 $m\times n$ 实矩阵,则存在 m 阶正交矩阵 $\mathbf{P}$ 以及 n 阶正交矩阵 $\mathbf{Q}$,使

$$P'AQ=\begin{pmatrix} S & O \\ O & O \end{pmatrix}$$

其中

$$S=\begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}$$

是一个 r 阶对角阵, $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>0$ 是 $\mathbf{A}$ 的非零奇异值.

♡

笔记 $P'AQ=\begin{pmatrix} S & O \\ O & O \end{pmatrix}$ 称为矩阵 A 的正交相抵标准型,而 $A=P\begin{pmatrix} S & O \\ O & O \end{pmatrix}Q'$ 则称为矩阵 $\mathbf{A}$ 的奇异值分解. 矩阵的奇异值分解在信息理论、控制理论和大数据科学等领域有着重要的应用.

笔记 从 n 阶矩阵 $\mathbf{A}$ 的奇异值分解很容易得到 $\mathbf{A}$ 的极分解. 事实上,由奇异值分解的式子可得

$$A=\left(PQ'\right)Q\begin{pmatrix} S & O \\ O & O \end{pmatrix}Q'$$

其中 $\mathbf{R}=\mathbf{PQ}'$ 是 n 阶正交矩阵, $\mathbf{B}=\mathbf{Q}\,\mathrm{diag}\{\mathbf{S},\mathbf{O}\}\mathbf{Q}'$ 是 n 阶半正定实对称阵, 从而 $A=RB$ 即为 A 的极分解. 通过奇异值分解来求极分解,在处理奇异阵时很有用.

例题 5.1 设 V,U 分别是 n 维, m 维欧氏空间, φ,ψ 是 $V\to U$ 的线性映射, φ^* 和 ψ^* 分别是 φ 和 ψ 的伴随. 若 $\varphi^*\varphi=\psi^*\psi$,证明: 存在 U 上的正交变换 ω ,使 $\varphi=\omega\psi$.

证明 我们将沿用定理 5.29 证明中的记号.

设在 V 的标准正交基 $\{e_1, e_2, \cdots, e_n\}$ 下, $\varphi^*\varphi = \psi^*\psi$ 的表示矩阵为 n 阶对角阵 $\operatorname{diag}\{\lambda_1, \cdots, \lambda_r, 0, \cdots, 0\}$, 则由定理 5.29 的证明知存在 U 的标准正交基 $\{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_m\}$,使 φ 在这两组基下的表示矩阵为分块对角阵 $\operatorname{diag}\{S, O\}$.

同理,存在 U 的另一组标准正交基 $\{g_1, g_2, \cdots, g_m\}$,使 ψ 在 $\{e_1, e_2, \cdots, e_n\}$ 和 $\{g_1, g_2, \cdots, g_m\}$ 下的表示矩阵也是 $\operatorname{diag}\{S, O\}$. 现定义 U 上的线性变换 ω 为 $\omega(g_i) = f_i\,(i = 1, 2, \cdots, m)$,则 ω 是 U 上的正交变换且

$$\omega\psi(e_i) = \omega(\sigma_i g_i) = \sigma_i \omega(g_i) = \sigma_i f_i = \varphi(e_i), i = 1, \cdots, r$$

$$\omega\psi(\mathbf{e}_j) = \omega(\mathbf{0}) = \mathbf{0} = \varphi(\mathbf{e}_j), j = r+1, \cdots, n,$$

故 $\varphi = \omega\psi$.

5.10 最小二乘解

设 V 是一个内积空间, W 是 V 的子空间 $(W \neq V)$, $\mathbf{v}$ 是 V 中的一个向量,现要在 W 中找一个向量 $\mathbf{u}$,使 $\| \mathbf{v}-\mathbf{u} \|$ 最小. 如果这样的 $\mathbf{u}$ 存在,则称 $\mathbf{u}$ 是 $\mathbf{v}$ 在 W 中的正交投影向量,长度 $\| \mathbf{v}-\mathbf{u} \|$ 被称为 $\mathbf{v}$ 到 W 的距离.

定理 5.30

设 W 是有限维内积空间 V 的子空间, $\mathbf{v} \in V$,则

(1) 在 W 中存在唯一的向量 $\mathbf{u}$,使 $\| \mathbf{v}-\mathbf{u} \|$ 最小且这时 $(\mathbf{v}-\mathbf{u}) \perp W$;

(2) 若 $\{e_1, \cdots, e_m\}$ 是 W 的标准正交基,又 $\{e_{m+1}, \cdots, e_n\}$ 是 $W^{\perp}$ 的标准正交基,这样 $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ 就成为 V 的一组标准正交基,则

$$\mathbf{u}=(\mathbf{v}, \mathbf{e}_1)\, \mathbf{e}_1+(\mathbf{v}, \mathbf{e}_2)\, \mathbf{e}_2+\cdots+(\mathbf{v}, \mathbf{e}_m)\, \mathbf{e}_m$$

$$\mathbf{v}-\mathbf{u}=(\mathbf{v}, \mathbf{e}_{m+1})\, \mathbf{e}_{m+1}+\cdots+(\mathbf{v}, \mathbf{e}_n)\, \mathbf{e}_n,$$

$$\| \mathbf{v}-\mathbf{u} \|=\left(|(\mathbf{v}, \mathbf{e}_{m+1})|^2+\cdots+|(\mathbf{v}, \mathbf{e}_n)|^2\right)^{\frac{1}{2}}.$$

♡

证明 设 $\mathbf{w}$ 是 W 中任一向量,要证明 $\| \mathbf{v}-\mathbf{u} \| \leq \| \mathbf{v}-\mathbf{w} \|$. 注意到这时 $\mathbf{v}-\mathbf{w}=(\mathbf{v}-\mathbf{u})+(\mathbf{u}-\mathbf{w})$,其中 $(\mathbf{v}-\mathbf{u}) \perp W, \mathbf{u}-\mathbf{w} \in W$. 因此,由勾股定理有

$$\| \mathbf{v}-\mathbf{w}\|^2=\| \mathbf{v}-\mathbf{u}\|^2+\| \mathbf{u}-\mathbf{w}\|^2.$$

若 $\mathbf{w} \neq \mathbf{u}$,则

$$\| \mathbf{v}-\mathbf{w}\|^2>\| \mathbf{v}-\mathbf{u}\|^2.$$

这样我们不仅证明了 $\| \mathbf{v}-\mathbf{u} \|$ 的最小性,也证明了 $\mathbf{u}$ 的唯一性.

现在我们举例来说明定理 5.30 在求最小二乘解方面的应用. 在许多实际问题中我们经常会碰到所谓的“矛盾线性方程组” 的求解问题. 为说得更清楚一些, 举一个简单的例子.

年	1	2	3
u	1.6	1.7	2.0
v	1.2	1.4	1.8

表 5.1: u, v关系

在经济学中, 个人的收入与消费之间存在着密切的关系. 收入越多, 消费水平也越高. 收入较少, 消费水平也较低. 从一个社会整体来看, 个人的平均收入与平均消费之间大致呈线性关系. 若 u 表示收入, v 表示支出,则 u, v 适合

$$u=a+bv$$

其中 a, b 是两个常数,需要根据具体的统计数据来确定. 假设现在有一组统计数据表示 3 年中每年的

收入与消费情况 (表 5.1),现要根据这一组统计数据求出 a, b .

将 u, v 的值代入得到一个两个未知数 3 个方程式的线性方程组:

$$\begin{cases} a+1.2b=1.6 \\ a+1.4b=1.7 \\ a+1.8b=2.0 \end{cases}$$

从第一、第二个方程式可求出 $a=1, b=0.5$,代入第三个方程式:

$$1+1.8\times 0.5=1.9\neq 2.0,$$

这说明上面的线性方程组无解. 那么这样一来是不是说我们的问题就没有意义了呢? 当然不是! 事实上, 收入与消费的关系通常极为复杂, 我们把它当成线性关系只是一种近似的假设. 另外, 统计数据本身不可避免地会产生误差, 也就是说统计表只是实际情况的近似反映. 我们的目的是求出 a, b 的值以供理论分析之用. 既然统计数据有误差,就不可能也没有必要求出 a, b 的精确解. 此矛盾的出现并不意味着我们因此而束手无策了. 我们可以对 a, b 提出这样的要求: 求出 a 与 b ,使得到的关系式 $u=a+bv$ 能尽可能好地符合实际情形. 用数学语言来说就是求 a, b , 使平方偏差

$$((a+1.2b)-1.6)^2+((a+1.4b)-1.7)^2+((a+1.8b)-2.0)^2$$

取最小值. 这就是所谓的最小二乘问题. 一般来说, 为了使理论关系式更符合实际, 通常要求统计数据多一点, 即方程式的个数多一点.

如何求“最小二乘解”? 让我们来作一些理论上的分析. 假设有下列矛盾线性方程组:

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases}$$

其中 $m>n$,它的系数矩阵记为 $\mathbf{A}$. 为方便,将上述线性方程组写为矩阵形式:

$$Ax=\beta$$

其中

$$\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},\ \beta=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

现要求出 $\mathbf{x}$,使

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}x_j-b_i\right)^2$$

取最小值. 记

$$\alpha_1=\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix},\ \alpha_2=\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix},\ \cdots,\ \alpha_n=\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix},$$

考虑向量

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n-\beta$$

这是一个 m 维列向量,第 i 个坐标为 $\sum\limits_{j=1}^{n}a_{ij}x_j-b_i$. 因此,若记 $V=\mathbb{R}_m$,内积取标准内积, 则

$$\left\|(x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n)-\beta\right\|^2=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}x_j-b_i\right)^2.$$

而

$$\left\|(x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n)-\beta\right\|=\|\mathbf{Ax}-\beta\|=\|\beta-\mathbf{Ax}\|.$$

当 $x_1,x_2,\cdots,x_n$ 变动时, $x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$ 就是由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 张成的 V 的子空间,记为 W . 要使 $\|\beta-\mathbf{Ax}\|$ 取最小值,实际上就是要求 β 到 W 的距离. 由定理知道,只需取 β 在 W 上的正交投影向量 γ 就可以了. 于是求方程组的最小二乘解又归结为求下列线性方程组的解:

$$Ax=\gamma$$

由定理 5.30 知道这样的解总是存在的. 我们可以按照定理 5.30 中的办法先求 W 的标准正交基,再扩展为 V 的标准正交基,从而求出 γ ,最后解线性方程组便可求出 $\mathbf{x}$. 但是这样做比较麻烦,我们希望能有一个比较简便的方法求出 x 来.

首先我们注意到在实际问题中,矩阵 $\mathbf{A}$ 的秩通常都等于 n . 因此,可设 $\mathbf{A}$ 是列满秩阵,即 $\mathrm{r}(\mathbf{A})=n$. 这也就是说,向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关, W 是由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 张成的 n 维子空间. 设

$$\gamma=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$$

由定理 5.30 知道 $(\beta-\gamma)\perp W$,因此 $(\beta-\gamma)\perp\alpha_i\,(i=1,2,\cdots,n)$,即 $((\beta-\gamma),\alpha_i)=0$,或者 $(\gamma,\alpha_i)=(\beta,\alpha_i)$. 这样便有下列线性方程组:

$$\begin{cases}(\alpha_1,\alpha_1)\,x_1+(\alpha_2,\alpha_1)\,x_2+\cdots+(\alpha_n,\alpha_1)\,x_n=(\beta,\alpha_1)\,,\\(\alpha_1,\alpha_2)\,x_1+(\alpha_2,\alpha_2)\,x_2+\cdots+(\alpha_n,\alpha_2)\,x_n=(\beta,\alpha_2)\,,\\\qquad\qquad\cdots\cdots\cdots\cdots\\(\alpha_1,\alpha_n)\,x_1+(\alpha_2,\alpha_n)\,x_2+\cdots+(\alpha_n,\alpha_n)\,x_n=(\beta,\alpha_n)\,.\end{cases}$$

若记 $\mathbf{A}'$ 为 $\mathbf{A}$ 的转置,则上面的方程组可写为矩阵形式:

$$A'Ax = A'\beta.$$

由于 $\mathbf{A}$ 的秩为 n ,故 $\mathbf{A}'\mathbf{A}$ 的秩也是 n ,即 $\mathbf{A}'\mathbf{A}$ 为 n 阶非异阵,于是

$$\mathbf{x} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\beta$$

这就是线性方程组的最小二乘解.