

其中 $\alpha_1, \beta_1, \dots, \alpha_s, \beta_s, \lambda_{2s+1}, \dots, \lambda_n \in \mathbb{R}, \beta_1 \cdots \beta_s \neq 0$. 因为 \mathcal{A} 是对称算子, 所以 B 是对称矩阵(第三章第二讲命题 4.7). 设 $s > 0$. 则 $N(\alpha_1, \beta_1)$ 是对称矩阵. 由此可知, $\beta_1 = 0$. 矛盾. 故 $s = 0$.

(ii) 把 A 看成标准欧式空间中 \mathbb{R}^n 上的对称算子即可.

(iii) 因为 $A \sim_o \text{diag}(\lambda_1, \dots, \lambda_n)$, 所以 $A \sim_s \text{diag}(\lambda_1, \dots, \lambda_n)$. 因为特征多项式是相似不变量(见第二章第三讲第五页), $\chi_A = (t - \lambda_1) \cdots (t - \lambda_n)$. \square

问题.

(i) 设 $\mathcal{A} \in \mathcal{L}(V)$ 对称. 求 V 的某组单位正交基使得 \mathcal{A} 在下的矩阵是 $\text{diag}(\lambda_1, \dots, \lambda_n)$.

(ii) 设 $A \in \text{SM}_n(\mathbb{R})$. 求 $P \in O_n(\mathbb{R})$ 使得

$$P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n).$$

基本步骤.

1. 计算 \mathcal{A} 的特征根. 设互不相同的特征根是 $\lambda_1, \dots, \lambda_k$;
2. 对 $i \in \{1, 2, \dots, k\}$, 求 V^{λ_i} 的一组基;
3. 对 $i \in \{1, 2, \dots, k\}$, 利用 Gram-Schmidt 正交化求 V^{λ_i} 的一组单位正交基; $\mathbf{e}_{i,1}, \dots, \mathbf{e}_{i,d_i}$

4. 根据命题 4.17 和对角化判别法II, $\mathbf{e}_{1,1}, \dots, \mathbf{e}_{1,d_1}, \dots, \mathbf{e}_{k,1}, \dots, \mathbf{e}_{k,d_k}$ 是 V 的一组单位正交基且在该基下 \mathcal{A} 的矩阵是对角的.

对于对称矩阵, 只需把它看成标准欧式空间上的线性算子即可.

例 5.2 设 $A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix} \in \text{SM}_4(\mathbb{R})$. 计算 $P \in \text{O}_4(\mathbb{R})$ 和对角阵 B 使得 $B = P^{-1}AP$.

解. 由计算机计算得 $\chi_A(t) = (t-1)^3(t+3)$. 于是, $\lambda_1 = 1$, $\lambda_2 = -3$. 计算 V^{λ_1} 的一组基. 已知 $\dim(V^{\lambda_1}) = 3$. 于是, $\text{rank}(\lambda_1 E - A) = 1$. 我们只要考虑方程

$$x_1 - x_2 - x_3 + x_4 = 0$$

的解空间即可. 直接得出 $V^{\lambda_1} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \right\rangle$. 计算 V^{λ_2} 的一组基. 因为 $V^{\lambda_1} \perp V^{\lambda_2}$ 且 $\mathbb{R}^4 = V^{\lambda_1} \oplus V^{\lambda_2}$, 所

以 $V^{\lambda_2} = (V^{\lambda_1})^\perp$. 由此可知, $V^{\lambda_2} = \left\langle \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right\rangle$. 利用 Gram-Schmidt 正交化求 V^{λ_1} 的一组单位正交基得

$$\epsilon_1 = \frac{1}{\sqrt{2}}(1, 1, 0, 0)^t, \quad \epsilon_2 = \frac{1}{\sqrt{2}}(0, 0, 1, 1)^t, \quad \epsilon_3 = \frac{1}{2}(1, -1, 1, -1)^t.$$

利用 Gram-Schmidt 正交化求 V^{λ_2} 的一组单位正交基得:

$$\epsilon_4 = \frac{1}{2}(1, -1, -1, 1)^t. \text{ 故 } P = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}. \text{ 我们得}$$

到 $P^t A P = \text{diag}(1, 1, 1, -3)$. \square

推论 5.3 设 $A \in \text{SM}_n(\mathbb{R})$. 则 A 的正(负)惯性指数等于 A 中正(负)根的个数(在记重数的意义下). 特别地, A (半)正定当且仅当 A 的特征根都是正的(非负的).

证明. 由定理 5.1 (ii) 和 (iii), $A \sim_o \text{diag}(\lambda_1, \dots, \lambda_n)$. 因为 $A \sim_c \text{diag}(\lambda_1, \dots, \lambda_n)$, 所以 A 的签名与 $\text{diag}(\lambda_1, \dots, \lambda_n)$ 的签名相同. \square

定理 5.4 设 $A, B \in \text{SM}_n(\mathbb{R})$ 且 A 正定. 则存在 $P \in \text{GL}_n(\mathbb{R})$ 使得 $P^t A P = E$ 和 $P^t B P$ 是对角阵.

证明. 因为 A 正定, 所以存在 $P_1 \in GL_n(\mathbb{R})$ 使得 $P_1^t A P_1 = E$ (惯性定理). 令 $C = P_1^t B P_1$. 则 C 对称. 根据第三章第三讲定理 6.1 (ii), 存在 $P_2 \in O_n(\mathbb{R})$, 使得 $D = P_2^t C P_2$ 是对角阵. 令 $P = P_1 P_2$. 则 $P^t B P = P_2^t C P_2 = D$ 且

$$P^t A P = P_2^t E P_2 = P_2^t P_2 = E (\because P_2 \in O_n(\mathbb{R})). \quad \square$$

例 5.5 设 $A, B \in SM_n(\mathbb{R})$, A 正定且 B 半正定. 证明:

$$\det(A + B) \geq \det(A) + \det(B).$$

证明. 由上述定理存在 $P \in GL_n(\mathbb{R})$ 使得 $P^t A P = E$ 和 $P^t B P = \text{diag}(\alpha_1, \dots, \alpha_n)$. 于是,

$$P^t(A + B)P = \text{diag}(1 + \alpha_1, \dots, 1 + \alpha_n).$$

两边取行列式得 $\det(P)^2 \det(A + B) = \prod_{i=1}^n (1 + \alpha_i)$. 而

$$\det(P)^2 (\det(A) + \det(B)) = \det(P^t A P) + \det(P^t B P) = 1 + \prod_{i=1}^n \alpha_i.$$

因为 B 半正定, 所以 $\alpha_1, \dots, \alpha_n$ 非负. 于是

$$\prod_{i=1}^n (1 + \alpha_i) \geq 1 + \prod_{i=1}^n \alpha_i.$$

由此可知, $\det(P)^2 \det(A + B) \geq \det(P)^2 (\det(A) + \det(B))$. 故 $\det(A + B) \geq \det(A) + \det(B)$. \square

其中 $\alpha_1, \beta_1, \dots, \alpha_s, \beta_s, \lambda_{2s+1}, \dots, \lambda_n \in \mathbb{R}, \beta_1 \cdots \beta_s \neq 0$. 因为 \mathcal{A} 是斜对称算子, 所以 B 是斜对称矩阵(第三章第三讲命题 5.7). 则 $N(\alpha_i, \beta_i)$ 是斜对称矩阵. 由此可知, $\alpha_i = 0, i = 1, 2, \dots, s$. 同理, $\lambda_{2s+1} = \dots = \lambda_n = 0$.

(ii) 把 A 看成标准欧式空间中 \mathbb{R}^n 上的斜对称算子.

(iii) 因为 $A \sim_o M$, 所以 $A \sim_s M$. 故 $\chi_A = \chi_M$. 故

$$\chi_M(t) = (t^2 + \beta_1^2) \cdots (t^2 + \beta_s^2) t^{n-2s}.$$

于是, $\text{spec}_{\mathbb{C}}(A) \subset \{0, \pm\beta_1\sqrt{-1}, \dots, \pm\beta_s\sqrt{-1}\}$. \square

例 6.2 设 $A \in \text{SSM}_n(\mathbb{R})$. 证明 $\det(E + A) \geq 1$.

证明. 根据定理 6.1, 存在 $P \in O_n(\mathbb{R})$ 使得 $P^t A P = M$, 其中 M 由定理 6.1 给出. 于是,

$$P^t(E+A)P = E+M = \text{diag}(N(1, \beta_1), \dots, N(1, \beta_s), 1, \dots, 1).$$

因为 $\det(P) = \pm 1$ 和 $\det(N(1, \beta_i)) \geq 1$, 所以 $\det(E + A) \geq 1$.

\square

7 正交算子和正交矩阵

定理 7.1 (i) 设 $\mathcal{A} \in \mathcal{L}(V)$ 正交. 则存在 $\theta_1, \dots, \theta_s \in (0, \pi) \cup (\pi, 2\pi)$ 和 $\lambda_{2s+1}, \dots, \lambda_n \in \{-1, 1\}$ 使得 \mathcal{A} 在 V 的某

$i = 1, 2, \dots, s$. 类似, λ_j 是一行一列的正交矩阵. 因为正交矩阵的行列式等于 ± 1 , 所以 $\lambda_j \in \{-1, 1\}$.

(ii) 把 A 看成标准欧式空间中 \mathbb{R}^n 上的正交算子即可.

(iii) 对任意 $\theta \in (0, \pi)$, $N(\cos(\theta), \sin(\theta))$ 的特征多项式等于 $t^2 - 2\cos(\theta)t + 1$. 其根 $\lambda = \cos(\theta) \pm \sin(\theta)\sqrt{-1}$. \square

例 7.2 设 $P \in O_n(\mathbb{R})$ 且 $\det(P) = -1$. 证明: $\det(E+P) = 0$.

证明. 根据定理 7.1, 存在正交矩阵 Q 使得 $Q^{-1}PQ = M$, 其中 M 如该定理所述. 因为 $\det(P) = -1$, 所以 $\det(M) = -1$. 于是, 存在 $j \in \{2s+1, \dots, n\}$ 使得 $\lambda_j = -1$. 故 -1 是 M 的特征根. 于是, -1 也是 P 的特征根. 我们有 $\det(E+P) = 0$.

8 欧式空间中的若干几何构造

命题 8.1 设 W 是欧式空间 V 的子空间, $\mathbf{x} \in V$, \mathbf{y} 是 \mathbf{x} 在 W 上的正交投影. 则对任意 $\mathbf{w} \in W$,

$$\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{w}\|.$$

证明. (几何法) 注意到 $\mathbf{x} - \mathbf{w} = (\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{w})$. 由 $\mathbf{y} - \mathbf{w} \in W$ 可知, $(\mathbf{x} - \mathbf{y}) \perp (\mathbf{y} - \mathbf{w})$. 再利用勾股定理得到

$$\|\mathbf{x} - \mathbf{w}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2 + \|\mathbf{y} - \mathbf{w}\|^2 \geq \|\mathbf{x} - \mathbf{y}\|^2.$$

故 $\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{w}\|$.

(代数法) 设 $\mathbf{e}_1, \dots, \mathbf{e}_d$ 是 W 的一组单位正交基. 则正交投影由映射

$$\begin{aligned} \mathcal{P}_W : V &\longrightarrow V \\ \mathbf{x} &\mapsto \sum_{i=1}^d (\mathbf{x}|\mathbf{e}_i)\mathbf{e}_i \end{aligned} \quad (1)$$

给出 (见命题 2.3 的证明). 根据定理 2.5 (ii), W^\perp 有一组单位正交基 $\mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 且 $\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 是 V 的一组基. 设 $\mathbf{x} = \sum_{j=1}^n x_j \mathbf{e}_j$. 则 $\mathbf{y} = \mathcal{P}_W(\mathbf{x}) = \sum_{i=1}^d x_i \mathbf{e}_i$. 再设 $\mathbf{u} = \sum_{i=1}^d u_i \mathbf{e}_i$. 则 $\|\mathbf{x} - \mathbf{y}\|^2 = x_{d+1}^2 + \dots + x_n^2$ 和

$$\|\mathbf{x} - \mathbf{u}\|^2 = (x_1 - u_1)^2 + \dots + (x_d - u_d)^2 + x_{d+1}^2 + \dots + x_n^2.$$

故 $\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x} - \mathbf{u}\|$. \square

根据上述命题, 我们把 $\|\mathbf{x} - \mathcal{P}_W(\mathbf{x})\|$ 称为 \mathbf{x} 到 W 的距离, 记为 $d(\mathbf{x}, W)$, 其中 \mathcal{P}_W 由 (1) 给出.

问题. 设 $W = \langle \mathbf{w}_1, \dots, \mathbf{w}_k \rangle$ 和 $\mathbf{x} \in V$. 求 \mathbf{x} 到 W 的正交投影和 $d(\mathbf{x}, W)$.

命题 8.2 利用上述问题中的记号, $\mathbf{y} = y_1 \mathbf{w}_1 + \dots + y_k \mathbf{w}_k$ 是 \mathbf{x} 的正交投影当且仅当

$$G(\mathbf{w}_1, \dots, \mathbf{w}_k)(y_1, \dots, y_k)^t = ((\mathbf{x}|\mathbf{w}_1), \dots, (\mathbf{x}|\mathbf{w}_k))^t, \quad (2)$$

其中 $G(\mathbf{w}_1, \dots, \mathbf{w}_k)$ 是关于 $\mathbf{w}_1, \dots, \mathbf{w}_k$ 的 Gram 矩阵.

证明. 直接推理得

$$\begin{aligned} \mathbf{y} = \mathcal{P}_W(\mathbf{x}) &\implies (\mathbf{x} - \mathbf{y}) \perp W \\ &\implies (\mathbf{x} - \mathbf{y}) \perp \mathbf{w}_j, \quad j = 1, \dots, k \\ &\implies ((\mathbf{x} - \mathbf{y}) | \mathbf{w}_j) = 0, \quad j = 1, \dots, k \\ &\implies (\mathbf{y} | \mathbf{w}_j) = (\mathbf{x} | \mathbf{w}_j), \quad j = 1, \dots, k \\ &\implies \sum_{i=1}^k (\mathbf{w}_i | \mathbf{w}_j) y_i = (\mathbf{x} | \mathbf{w}_j), \quad j = 1, \dots, k \\ &\implies (2) \text{ 成立. } \square \end{aligned}$$

设 $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ 和 $\mathbf{x} = (x_1, \dots, x_n)$. 考虑线性方程组,

$$A\mathbf{x} = \mathbf{b}. \quad (3)$$

设 \mathbf{b} 在 $V_c(A)$ 中的正交投影是 $\mathbf{v} = \alpha_1 \vec{A}^{(1)} + \dots + \alpha_n \vec{A}^{(n)}$. 则 $(\alpha_1, \dots, \alpha_n)^t$. 称为 (3) 的一个最小平方解.

注意到 (3) 有解当且仅当 $\mathbf{b} \in V_c(A)$. 此时 $\mathbf{v} = \mathbf{b}$. 故每个 (3) 的最小平方解都是 (3) 的解, 反之亦然. 当 A 列满秩时, (3) 的最小平方解是唯一的.

例 8.3 设 x 代表某种杂质, y 代表产品的成品率. 已知

$$y = a x\% + b.$$

根据下列实验数据

$x\%$	3.6	3.7	3.8	4.0	4.1	4.2
y	1.0	0.9	0.9	0.6	0.56	0.35

求 a, b .

解. a, b 满足的方程组是

$$\begin{cases} 3.6a + b = 1.0 \\ 3.7a + b = 0.9 \\ 3.8a + b = 0.9 \\ 4.0a + b = 0.6 \\ 4.1a + b = 0.56 \\ 4.2a + b = 0.35 \end{cases}$$

该方程组无解. 计算其最小平方解得到 $a = -1.05, b = 4.81$.

故 $y = -1.05x\% + 4.81$.

命题 8.4 设 $\mathbf{x} \in V$, W 是 V 的子空间, $\mathbf{w}_1, \dots, \mathbf{w}_k$ 是 W 的一组基. 则

$$d(\mathbf{x}, W)^2 = \frac{\det(G(\mathbf{x}, \mathbf{w}_1, \dots, \mathbf{w}_k))}{\det(G(\mathbf{w}_1, \dots, \mathbf{w}_k))}.$$

证明. 设 \mathbf{x} 在 W 中的正交投影是 $\mathbf{y} = \sum_{i=1}^k y_i \mathbf{w}_i$,

$$M = G(\mathbf{x}, \mathbf{w}_1, \dots, \mathbf{w}_k) \quad \text{和} \quad N = G(\mathbf{w}_1, \dots, \mathbf{w}_k).$$

则 M 可逆 (命题 1.7). 直接计算:

$$\begin{aligned}
 \frac{\det(N)}{\det(M)} &= \det(M^{-1}) \det(N) \\
 &= \det \begin{pmatrix} 1 & O \\ O & M^{-1} \end{pmatrix} \det(N) \\
 &= \det \left(\begin{pmatrix} 1 & O \\ O & M^{-1} \end{pmatrix} \begin{pmatrix} (\mathbf{x}|\mathbf{x}) & (\mathbf{x}|\mathbf{w}_1) & \dots, & (\mathbf{x}|\mathbf{w}_k) \\ (\mathbf{x}|\mathbf{w}_1) & & & \\ \vdots & & M & \\ (\mathbf{x}|\mathbf{w}_k) & & & \end{pmatrix} \right) \\
 &= \det \begin{pmatrix} (\mathbf{x}|\mathbf{x}) & (\mathbf{x}|\mathbf{w}_1) & \dots, & (\mathbf{x}|\mathbf{w}_k) \\ y_1 & & & \\ \vdots & & E_n & \\ y_k & & & \end{pmatrix} \quad (\because (2)) \\
 &= (\mathbf{x}|\mathbf{x}) - (\mathbf{x}|\mathbf{y}) = d(\mathbf{x}, W)^2. \quad \square
 \end{aligned}$$

利用上述命题中的符号, $\det(G(\mathbf{w}_1)) = \|\mathbf{w}_1\|^2$. 故

$$\det G(\mathbf{w}_1, \mathbf{w}_2) = d(\mathbf{w}_1, \langle \mathbf{w}_2 \rangle)^2 \det(G(\mathbf{w}_2)) = \underbrace{d(\mathbf{w}_1, \langle \mathbf{w}_2 \rangle)}_{\text{高}} \underbrace{\|\mathbf{w}_2\|}_{\text{底}}^2.$$

故 $\det G(\mathbf{w}_1, \mathbf{w}_2)$ 代表由 \mathbf{w}_1 和 \mathbf{w}_2 张成的平行四边形的面积的平方. 通过归纳法, 我们可以利用上述定理看成 $\det G(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k)$ 是由 $\mathbf{w}_1, \dots, \mathbf{w}_k$ 张成的 $2k$ 面体的体积的平方.

设 $A \in GL_n(\mathbb{R})$. 则把其列向量看成标准欧式空间 \mathbb{R}^n 中向量, 我们有 $G(\vec{A}^{(1)}, \dots, \vec{A}^{(n)}) = A^t A$. 故

$$\det G(\vec{A}^{(1)}, \dots, \vec{A}^{(n)}) = \det(A)^2.$$

于是, $|\det(A)|$ 是由 A 的列向量张成的 $2n$ 面体的体积.

例 8.5 在正交变换下确定 $3x^2 + 2xy + 3y^2 = 1$ 在标准欧式空间 \mathbb{R}^2 中的轨迹.

解. $q(x, y) = 3x^2 + 2xy + 3y^2 = (x, y) \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. 设

$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. 则 $\chi_A(t) = (t - 4)(t - 2)$. 设 $\lambda_1 = 4$ 和

$\lambda_2 = 2$. 则 $V^{\lambda_1} = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$ 和 $V^{\lambda_2} = \langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rangle$. 令

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

则 P 正交且 $P^t A P = \text{diag}(4, 2)$. 即 $3x^2 + 2xy + 3y^2 = 1$ 在单位正交基 $\vec{P}^{(1)}$ 和 $\vec{P}^{(2)}$ 下的形式是

$$4X^2 + 2Y^2 = 1, \quad \text{其中} \quad \begin{pmatrix} X \\ Y \end{pmatrix} = P^t \begin{pmatrix} x \\ y \end{pmatrix}.$$