

0.1 实对称矩阵的正交相似标准型

回顾实对称矩阵的正交相似标准型.

0.1.1 实二次型值的估计以及实对称矩阵特征值的估计

命题 0.1

(1) 设 A 是 n 阶实对称矩阵, 其特征值为 $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, 求证: 对任意的 n 维实列向量 α , 均有

$$\lambda_1 \alpha' \alpha \leq \alpha' A \alpha \leq \lambda_n \alpha' \alpha$$

且前一个不等式等号成立的充要条件是 α 属于特征值 λ_1 的特征子空间, 后一个不等式等号成立的充要条件是 α 属于特征值 λ_n 的特征子空间.

(2) 设 A 是 n 阶 Hermite 矩阵, 其特征值为 $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, 求证: 对任意的 n 维复列向量 α , 均有

$$\lambda_1 \bar{\alpha}' \alpha \leq \bar{\alpha}' A \alpha \leq \lambda_n \bar{\alpha}' \alpha$$

且前一个不等式等号成立的充要条件是 α 属于特征值 λ_1 的特征子空间, 后一个不等式等号成立的充要条件是 α 属于特征值 λ_n 的特征子空间.

注 这个命题 0.1 是命题??的推广, 即利用实对称矩阵的正交相似标准型得到了实二次型取值的精确上下界. 命题 0.1 还可以推广到 Hermite 矩阵的情形即 (2), 其证明只要将 α' 换成 $\bar{\alpha}'$ 即可. 然而这种推广并不是平凡的, 因为当我们处理一般的实矩阵时, 不可避免地会遇到复特征值和复特征向量, 此时若把实对称矩阵当作 Hermite 矩阵来处理, 会使讨论变得更加简洁.

证明

(1) 设 P 为正交矩阵, 使得 $P'AP = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. 对任意的实列向量 α , 设 $\beta = P'\alpha = (b_1, b_2, \cdots, b_n)'$, 则

$$\begin{aligned} \alpha' A \alpha &= \beta' (P'AP) \beta = \lambda_1 b_1^2 + \lambda_2 b_2^2 + \cdots + \lambda_n b_n^2 \\ &\leq \lambda_n b_1^2 + \lambda_n b_2^2 + \cdots + \lambda_n b_n^2 = \lambda_n \beta' \beta = \lambda_n (P'\alpha)' (P'\alpha) = \lambda_n \alpha' \alpha \end{aligned}$$

等号成立的充要条件是若 $\lambda_i \neq \lambda_n$, 则 $b_i = 0$, 这也等价于 α 属于特征值 λ_n 的特征子空间. 同理可证前一个不等式及其等号成立的充要条件.

(2) 证明只要将 (1) 的证明中的 α' 换成 $\bar{\alpha}'$ 即可.

□

定义 0.1 (Rayleigh quotient(瑞丽商))

定义 Rayleigh quotient(瑞利商) 为

$$R(A, x) = \frac{x^H A x}{x^H x},$$

其中 $x \in \mathbb{R}^n$ 为非零向量, $A \in \mathbb{R}^{n \times n}$ 为实对称矩阵.

命题 0.2 (Rayleigh quotient(瑞丽商) 的基本性质)

设 A, B 是 n 阶实对称矩阵, 则有

$$\min_{x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{\min}(A), \max_{x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{\max}(A). \quad (1)$$

以及

$$\lambda_{\min}(A) + \lambda_{\min}(B) \leq \lambda(A+B) \leq \lambda_{\max}(A) + \lambda_{\max}(B). \quad (2)$$

笔记 这里 λ_{\max} 表示最大特征值, λ_{\min} 表示最小特征值, λ 表示任一特征值.

证明 设 T 是实 n 阶正交矩阵使得

$$T^T A T = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}.$$

于是对 $x \neq 0$, 就有

$$\frac{x^T A x}{x^T x} = \frac{y^T T^T A T y}{y^T T^T T y} = \frac{y^T T^T A T y}{y^T y} = \frac{\sum_{i=1}^n \lambda_i y_i^2}{\sum_{i=1}^n y_i^2} \begin{cases} \leq \frac{\sum_{i=1}^n y_i^2}{\sum_{i=1}^n y_i^2} \lambda_{\max}(A) \\ \geq \frac{\sum_{i=1}^n y_i^2}{\sum_{i=1}^n y_i^2} \lambda_{\min}(A) \end{cases}$$

这就证明了(1).

此外运用基本不等式知

$$\min_{x \neq 0} \frac{x^T A x}{x^T x} + \min_{x \neq 0} \frac{x^T B x}{x^T x} \leq \min_{x \neq 0} \frac{x^T (A+B)x}{x^T x} \leq \max_{x \neq 0} \frac{x^T (A+B)x}{x^T x} \leq \max_{x \neq 0} \frac{x^T A x}{x^T x} + \max_{x \neq 0} \frac{x^T B x}{x^T x}$$

于是由(1)知

$$\lambda_{\min}(A) + \lambda_{\min}(B) \leq \lambda_{\min}(A+B) \leq \lambda_{\max}(A+B) \leq \lambda_{\max}(A) + \lambda_{\max}(B)$$

这就证明了(2). □

命题 0.3

设 $A = (a_{ij})_{n \times n}$ 是实对称 (或 Hermite) 矩阵, 则

$$\mathbf{1}^T A \mathbf{1} = \sum_{i=1}^n \sum_{j=1}^n a_{ij},$$

其中 $\mathbf{1} = (1 \ 1 \ \dots \ 1)^T$. ▲

证明 利用矩阵乘法验证即可. □

例题 0.1 设 A, B 是 n 阶实对称矩阵, 其特征值分别为

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n, \quad \mu_1 \leq \mu_2 \leq \dots \leq \mu_n$$

求证: $A+B$ 的特征值全落在 $[\lambda_1 + \mu_1, \lambda_n + \mu_n]$ 中.

证明 由命题 0.1 可知, 对任意的 n 维实列向量 α , 有 $\alpha' A \alpha \leq \lambda_n \alpha' \alpha, \alpha' B \alpha \leq \mu_n \alpha' \alpha$. 因为 $A+B$ 仍是实对称矩阵, 故其特征值全为实数. 任取 $A+B$ 的实特征值 ν 及其特征向量 β , 则 $\beta'(A+B)\beta = \nu \beta' \beta$. 注意到

$$\beta'(A+B)\beta = \beta' A \beta + \beta' B \beta \leq \lambda_n \beta' \beta + \mu_n \beta' \beta = (\lambda_n + \mu_n) \beta' \beta$$

且 $\beta' \beta > 0$, 故 $\nu \leq \lambda_n + \mu_n$. 同理可证 $\nu \geq \lambda_1 + \mu_1$. □

例题 0.2 设 $\lambda = a + bi$ 是 n 阶实矩阵 A 的特征值, 实对称矩阵 $A+A'$ 和 Hermite 矩阵 $-i(A-A')$ 的特征值分别为

$$\mu_1 \leq \mu_2 \leq \dots \leq \mu_n, \quad \nu_1 \leq \nu_2 \leq \dots \leq \nu_n$$

求证: $\mu_1 \leq 2a \leq \mu_n, \nu_1 \leq 2b \leq \nu_n$.

证明 设 α 是 A 的属于特征值 $\lambda = a + bi$ 的特征向量, 即有 $A\alpha = \lambda\alpha$. 将此式左乘 $\bar{\alpha}'$ 可得 $\bar{\alpha}' A \alpha = \lambda \bar{\alpha}' \alpha$; 再将此式共轭转置可得 $\bar{\alpha}' A' \alpha = \bar{\lambda} \bar{\alpha}' \alpha$; 最后将上述两式相加以及相减再乘以 $-i$, 可分别得到

$$\bar{\alpha}' (A+A') \alpha = (\lambda + \bar{\lambda}) \bar{\alpha}' \alpha = 2a \bar{\alpha}' \alpha$$

$$\bar{\alpha}' (-i(A-A')) \alpha = -i(\lambda - \bar{\lambda}) \bar{\alpha}' \alpha = 2b \bar{\alpha}' \alpha$$

注意到 $\bar{\alpha}' \alpha > 0$, 故由命题 0.1(2) 即得结论. □

例题 0.3 设 A 是 n 阶实矩阵, $A'A$ 的特征值为

$$\mu_1 \leq \mu_2 \leq \dots \leq \mu_n$$

求证: 若 λ 是 A 的特征值, 则

$$\sqrt{\mu_1} \leq |\lambda| \leq \sqrt{\mu_n}$$

证明 设 α 是 A 的属于特征值 λ 的特征向量, 即有 $A\alpha = \lambda\alpha$, 将此式共轭转置可得 $\bar{\alpha}'A' = \bar{\lambda}\bar{\alpha}'$, 再将上述两式乘在一起可得 $\bar{\alpha}'A'A\alpha = |\lambda|^2\bar{\alpha}'\alpha$, 最后由命题 0.1(2) 即得结论. \square

例题 0.4 设 A_1, \dots, A_k 是 n 阶实矩阵, $A_i'A_i$ 的特征值为

$$\mu_{i1} \leq \mu_{i2} \leq \dots \leq \mu_{in}, \quad 1 \leq i \leq k$$

求证: 若 λ 是 $A_1 \cdots A_k$ 的特征值, 则

$$\sqrt{\mu_{11} \cdots \mu_{k1}} \leq |\lambda| \leq \sqrt{\mu_{1n} \cdots \mu_{kn}}$$

证明 设 α 是 $A_1 \cdots A_k$ 的属于特征值 λ 的特征向量, 即有 $A_1 \cdots A_k \alpha = \lambda \alpha$, 将此式共轭转置可得 $\bar{\alpha}'A_k' \cdots A_1' = \bar{\lambda}\bar{\alpha}'$, 再将上述两式乘在一起可得

$$\bar{\alpha}'A_k' \cdots A_1'A_1 \cdots A_k \alpha = |\lambda|^2 \bar{\alpha}'\alpha$$

设 A_i 的特征值为 $a_{i1}, a_{i2}, \dots, a_{in}$. 由特征值的性质可知, A_i' 的特征值等于 A_i 的特征值. 从而 $A_i'A_i$ 的特征值为 A_i 特征值的平方, 于是

$$\mu_{il} = a_{il}^2, \quad 1 \leq l \leq n$$

考虑复数域, 则 $a_{il} = \sqrt{\mu_{il}}, 1 \leq l \leq n$. 因此 $A_1 A_2 \cdots A_k$ 的特征值为

$$\sqrt{\mu_{11}\mu_{21}\cdots\mu_{k1}} \leq \dots \leq \sqrt{\mu_{1n}\mu_{2n}\cdots\mu_{kn}}$$

最后对 $A_1 \cdots A_k$ 应用命题 0.1(2) 即得结论. \square

命题 0.4

设 n 阶复矩阵 M 的全体特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则 M 的谱半径 $\rho(M)$ 定义为 $\rho(M) = \max_{1 \leq i \leq n} |\lambda_i|$. 设 A, B, C

为 n 阶实矩阵, 使得 $\begin{pmatrix} A & B \\ B' & C \end{pmatrix}$ 为半正定实对称矩阵, 证明: $\rho(B)^2 \leq \rho(A)\rho(C)$.

证明 我们先来处理 $\begin{pmatrix} A & B \\ B' & C \end{pmatrix}$ 为正定实对称矩阵的情形. 此时, A, C 都是正定阵, 设它们的全体特征值分别为

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n, \quad \nu_1 \leq \nu_2 \leq \dots \leq \nu_n$$

则 $\rho(A) = \lambda_n, \rho(C) = \nu_n$ 且 A^{-1} 的全体特征值为 $\lambda_n^{-1} \leq \lambda_{n-1}^{-1} \leq \dots \leq \lambda_1^{-1}$. 考虑如下对称分块初等变换:

$$\begin{pmatrix} I_n & O \\ -B'A^{-1} & I_n \end{pmatrix} \begin{pmatrix} A & B \\ B' & C \end{pmatrix} \begin{pmatrix} I_n & -A^{-1}B \\ O & I_n \end{pmatrix} = \begin{pmatrix} A & O \\ O & C - B'A^{-1}B \end{pmatrix}$$

由 $\begin{pmatrix} A & B \\ B' & C \end{pmatrix}$ 是正定阵可知, $C - B'A^{-1}B$ 也是正定阵. 任取 B 的特征值 $\mu \in \mathbb{C}$ 及其特征向量 $\beta \in \mathbb{C}^n$, 即有 $B\beta = \mu\beta$ 以及 $\bar{\beta}'B' = \bar{\mu}\bar{\beta}'$. 将 $C - B'A^{-1}B$ 看成是正定 Hermite 矩阵, 则有 $\bar{\beta}'(C - B'A^{-1}B)\beta > 0$, 再由命题 0.1(2) 可得

$$\nu_n \bar{\beta}'\beta \geq \bar{\beta}'C\beta > \bar{\beta}'B'A^{-1}B\beta = |\mu|^2 \bar{\beta}'A^{-1}\beta \geq |\mu|^2 \lambda_n^{-1} \bar{\beta}'\beta$$

注意到 $\bar{\beta}'\beta > 0$, 故 $|\mu|^2 < \lambda_n \nu_n$, 即 $|\mu|^2 < \rho(A)\rho(C)$, 于是 $\rho(B)^2 < \rho(A)\rho(C)$.

我们用摄动法来处理半正定的情形. 对任意的正实数 t , $\begin{pmatrix} A+tI_n & B \\ B' & C+tI_n \end{pmatrix}$ 是正定阵, 从而由正定情形的结论可知

$$\rho(B)^2 < \rho(A+tI_n)\rho(C+tI_n) = (\rho(A)+t)(\rho(C)+t)$$

令 $t \rightarrow 0+$ 即得 $\rho(B)^2 \leq \rho(A)\rho(C)$. \square

命题 0.5

设 n 阶实对称矩阵 $A = (a_{ij})$ 为非负矩阵, 即所有的元素 $a_{ij} \geq 0$, 且 A 的全体特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 求证: 存在某个特征值 $\lambda_j = \rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$, 并可取到 λ_j 的某个特征向量 β 为非负向量, 即 β 的所有元素都大于等于零.

证明 任取一个特征值 λ_k , 使得 $|\lambda_k| = \max_{1 \leq i \leq n} |\lambda_i|$, 并取 λ_k 的特征向量 $\alpha = (a_1, a_2, \dots, a_n)'$, 即有 $A\alpha = \lambda_k\alpha$, 于是 $\alpha'A\alpha = \lambda_k\alpha'\alpha$. 以下不妨设 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, 令 $\beta = (|a_1|, |a_2|, \dots, |a_n|)'$, 则 β 是非负向量且 $\beta'\beta = \sum_{i=1}^n a_i^2 = \alpha'\alpha$. 注意到 $a_{ij} \geq 0$ ($1 \leq i, j \leq n$), 故由 **命题 0.1** 可得如下不等式:

$$\begin{aligned} |\lambda_k|\alpha'\alpha &= |\lambda_k\alpha'\alpha| = |\alpha'A\alpha| = \left| \sum_{i,j=1}^n a_{ij}a_i a_j \right| \\ &\leq \sum_{i,j=1}^n a_{ij}|a_i||a_j| = \beta'A\beta \leq \lambda_n\beta'\beta = \lambda_n\alpha'\alpha \end{aligned}$$

于是 $\lambda_n \geq |\lambda_k| \geq 0$. 再由假设可知 $\lambda_n = |\lambda_k| = \max_{1 \leq i \leq n} |\lambda_i|$, 因此上述不等式取等号. 特别地, $\beta'A\beta = \lambda_n\beta'\beta$, 故由 **命题 0.1** 中不等式取等号的充要条件可知, β 就是属于特征值 λ_n 的非负特征向量. \square

0.1.2 正定阵和半正定阵的性质的研究**定义 0.2 (正定阵和半正定阵的记号)**

设 A, B 是实对称矩阵, 我们用 $A > O$ 表示 A 是正定阵, $A \geq O$ 表示 A 是半正定阵. 当 A 和 B 都是正定阵时, 用 $A > B$ 表示 $A - B$ 是正定阵; 当 A 和 B 都是半正定阵时, 用 $A \geq B$ 表示 $A - B$ 是半正定阵.

例题 0.5 求证: 若 A 是 n 阶正定实对称矩阵, 则 $A + A^{-1} \geq 2I_n$.

证明 **证法一:** 设 P 是正交矩阵, 使得 $P'AP = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 其中 $\lambda_i > 0$ 是 A 的特征值, 则

$$P'A^{-1}P = \text{diag}\{\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_n^{-1}\}$$

因为 $\lambda_i + \lambda_i^{-1} \geq 2$, 故 $P'AP + P'A^{-1}P - 2I_n$ 是半正定阵, 由此即得 $A + A^{-1} \geq 2I_n$.

证法二: 令 $B = \begin{pmatrix} A & -I_n \\ -I_n & A^{-1} \end{pmatrix}$, 注意到

$$A + A^{-1} - 2I_n = \begin{pmatrix} I_n & I_n \end{pmatrix} \begin{pmatrix} A & -I_n \\ -I_n & A^{-1} \end{pmatrix} \begin{pmatrix} I_n \\ I_n \end{pmatrix} = \begin{pmatrix} I_n & I_n \end{pmatrix} B \begin{pmatrix} I_n \\ I_n \end{pmatrix} = 2B.$$

故由 **命题 ??** 可知 $B \geq 0$, 故 $A + A^{-1} \geq 2I_n$. \square

命题 0.6

设 B 是 n 阶半正定实对称矩阵, $\mu_1, \mu_2, \dots, \mu_n$ 是 B 的全体特征值, 证明: 对任意给定的正整数 $k > 1$, 存在一个只和 $\mu_1, \mu_2, \dots, \mu_n$ 有关的实系数多项式 $f(x)$, 满足: $B = f(B^k)$.

证明 设 Q 为正交矩阵, 使得 $Q'BQ = \text{diag}\{\mu_1, \mu_2, \dots, \mu_n\}$, 其中 $\mu_i \geq 0$. 设 $\mu_{i_1}, \mu_{i_2}, \dots, \mu_{i_s}$ 是 B 的全体不同特征值, $\lambda_i = \mu_{i_1}^k$ ($1 \leq i \leq n$), 则 $\lambda_i \geq 0$ 且 $\lambda_{i_1}, \lambda_{i_2}, \dots, \lambda_{i_s}$ 两两互异. 作 Lagrange 插值多项式:

$$f(x) = \sum_{j=1}^s \mu_{i_j} \frac{(x - \lambda_{i_1}) \cdots (x - \lambda_{i_{j-1}})(x - \lambda_{i_{j+1}}) \cdots (x - \lambda_{i_s})}{(\lambda_{i_j} - \lambda_{i_1}) \cdots (\lambda_{i_j} - \lambda_{i_{j-1}})(\lambda_{i_j} - \lambda_{i_{j+1}}) \cdots (\lambda_{i_j} - \lambda_{i_s})}$$

显然 $f(\lambda_{i_j}) = \mu_{i_j}$ ($1 \leq j \leq s$), 从而 $f(\lambda_i) = \mu_i$ ($1 \leq i \leq n$), 于是

$$\text{diag}\{\mu_1, \mu_2, \dots, \mu_n\} = \text{diag}\{f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)\} = f(\text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\})$$

因此

$$\begin{aligned} B &= Q \operatorname{diag}\{\mu_1, \mu_2, \dots, \mu_n\} Q' = Q f(\operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}) Q' \\ &= f(Q \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\} Q') = f((Q \operatorname{diag}\{\mu_1, \mu_2, \dots, \mu_n\} Q')^k) = f(B^k) \end{aligned}$$

□

定义 0.3 (半正定阵的 k 次方根)

设 A 是 n 阶半正定实对称矩阵, 若存在半正定阵 B 使得 $A = B^k$, 其中 k 为正整数, 则称 B 为 A 的 k 次方根, 记为 $B = A^{\frac{1}{k}}$.

♣

命题 0.7 (半正定阵的 k 次方根的唯一性)

设 A 是 n 阶半正定实对称矩阵, 求证: 对任意的正整数 $k > 1$, 必存在唯一的 n 阶半正定实对称矩阵 B , 使得 $A = B^k$. 即: A 的任意正整数次方根都存在且唯一.

♣

证明 设 P 是正交矩阵, 使得 $P'AP = \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 其中 $\lambda_i \geq 0$ 是 A 的特征值. 令 $B = P \operatorname{diag}\{\lambda_1^{\frac{1}{k}}, \lambda_2^{\frac{1}{k}}, \dots, \lambda_n^{\frac{1}{k}}\} P'$, 则 B 为半正定阵且 $A = B^k$, 这就证明了 k 次方根的存在性.

设 B 是 A 的 k 次方根, 则对 B 的任一特征值 μ_i, μ_i^k 是 A 的特征值, 即 μ_i 是 A 的某个特征值的非负 k 次方根. 由命题 0.6 可知, 存在一个只和 A 的所有特征值的非负 k 次方根 μ_i 有关的实系数多项式 $f(x)$, 使得 $B = f(B^k) = f(A)$. 设 C 是 A 的另一个 k 次方根, 则同上讨论也有 $C = f(A)$, 从而 $B = C$, 这就证明了 k 次方根的唯一性. □

命题 0.8

若 A 是半正定实对称矩阵, B 是同阶实矩阵且 $AB = BA$, 求证: $A^{\frac{1}{2}}B = BA^{\frac{1}{2}}$.

♣

证明 由命题 0.6 可知, 存在实系数多项式 $f(x)$, 使得 $A^{\frac{1}{2}} = f(A)$, 再由 A 与 B 乘法可交换可得 $A^{\frac{1}{2}}$ 与 B 乘法可交换. □

命题 0.9

设 A 为 n 阶实对称矩阵, 求证: A 为正定阵 (半正定阵) 的充要条件是

$$c_r = \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} A \begin{pmatrix} i_1 & i_2 & \dots & i_r \\ i_1 & i_2 & \dots & i_r \end{pmatrix} > 0 \quad (\geq 0), \quad 1 \leq r \leq n$$

♣

注 若 A 为半正定实对称矩阵, 则存在实矩阵 C , 使得 $A = C'C$. 有了 k 次方根这一工具后, 通常可以取 $C = A^{\frac{1}{2}}$, 这样往往可以有效地化简问题. 这一技巧在后面一些例题中会经常用到.

证明 由正定阵 (半正定阵) 的性质可知必要性成立, 下证充分性. 由推论 ?? 可知, A 的特征多项式

$$f(\lambda) = |\lambda I_n - A| = \lambda^n - c_1 \lambda^{n-1} + \dots + (-1)^{n-1} c_{n-1} \lambda + (-1)^n c_n$$

其中所有的 $c_i > 0$ (≥ 0). 注意到 A 的特征值, 即 $f(\lambda)$ 的根全是实数, 故由命题 ??(3) 可知, $f(\lambda)$ 的根全大于零 (全大于等于零), 因此 A 是正定阵 (半正定阵). □

命题 0.10

设 A, B 都是 n 阶实对称矩阵, 证明:

- (1) 若 A 半正定或者 B 半正定, 则 AB 的特征值全是实数;
- (2) 若 A, B 都半正定, 则 AB 的特征值全是非负实数;
- (3) 若 A 正定, 则 B 正定的充要条件是 AB 的特征值全是正实数.

♣

证明 (1) 设 A 半正定, 则由定理 ?? 可知, $AB = A^{\frac{1}{2}}A^{\frac{1}{2}}B$ 与 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 有相同的特征值. 注意到 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 仍是实对称矩阵, 故由推论 ?? 可知其特征值全是实数, 于是 AB 的特征值也全是实数. 同理可证 B 为半正定阵的情形.

(2) 采用与 (1) 相同的讨论, 注意到 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 仍是半正定阵, 故其特征值全是非负实数, 于是 AB 的特征值也全

是非负实数.

(3) 采用与 (1) 相同的讨论, 注意到 $A^{\frac{1}{2}}$ 是正定阵, 故 AB 的特征值全是正实数当且仅当 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 的特征值全是正实数, 这当且仅当 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 是正定阵, 从而当且仅当 B 是正定阵. \square

命题 0.11

设 A, B 都是半正定实对称矩阵, 其特征值分别为

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n, \quad \mu_1 \leq \mu_2 \leq \cdots \leq \mu_n$$

求证: AB 的特征值全落在 $[\lambda_1\mu_1, \lambda_n\mu_n]$ 中.

证明 采用与命题 0.10(1) 相同的讨论, 我们只要证明 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 的特征值全落在 $[\lambda_1\mu_1, \lambda_n\mu_n]$ 中即可. 任取 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 的特征值 ν 及其特征向量 α , 即有 $A^{\frac{1}{2}}BA^{\frac{1}{2}}\alpha = \nu\alpha$. 此式两边左乘 α' , 并由命题 0.1 可得

$$\nu\alpha'\alpha = (A^{\frac{1}{2}}\alpha)'B(A^{\frac{1}{2}}\alpha) \geq \mu_1(A^{\frac{1}{2}}\alpha)'(A^{\frac{1}{2}}\alpha) = \mu_1\alpha'A\alpha \geq \lambda_1\mu_1\alpha'\alpha$$

由此即得 $\nu \geq \lambda_1\mu_1$. 同理可证 $\nu \leq \lambda_n\mu_n$. \square

命题 0.12

设 A 是 n 阶正定实对称矩阵, B 是同阶实矩阵, 使得 AB 是实对称矩阵. 求证: AB 是正定阵的充要条件是 B 的特征值全是正实数.

证明 证法一: 由 A 正定可得 A^{-1} 也正定, 再由命题 0.10(3) 可知, AB 正定的充要条件是 $A^{-1}(AB) = B$ 的特征值全是正实数.

证法二: 由命题 ?? 可知, 存在可逆矩阵 C , 使得

$$C'AC = I_n, \quad C'(AB)C = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

其中 λ_i 是矩阵 $A^{-1}(AB) = B$ 的特征值. 因此 AB 是正定阵当且仅当 $C'(AB)C$ 是正定阵, 这也当且仅当 B 的特征值 λ_i 全是正实数. \square

命题 0.13

(1) 设 A, B 都是 n 阶正定实对称矩阵, 求证: AB 是正定实对称矩阵的充要条件是 $AB = BA$.

(2) 设 A, B 都是 n 阶半正定实对称矩阵, 求证: AB 是半正定实对称矩阵的充要条件是 $AB = BA$. \blacktriangle

证明 (1) 证法一: 由命题 0.10(3) 可知, AB 的特征值全大于零, 因此 AB 是正定阵当且仅当它是实对称矩阵, 即 $AB = (AB)' = BA$.

证法二: 若 AB 是正定实对称矩阵, 则 $AB = (AB)' = BA$. 反之, 若 $AB = BA$, 则 AB 是实对称矩阵. 因为 A 正定, 故 A^{-1} 也正定, 由命题 ?? 可知, 存在可逆矩阵 C , 使得

$$C'A^{-1}C = I_n, \quad C'BC = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

其中 λ_i 是矩阵 AB 的特征值. 因为 B 正定, 故 $C'BC$ 也正定, 从而 $\lambda_i > 0$, 因此 AB 是正定阵.

(2) 由命题 0.10(2) 可知, AB 的特征值全大于等于零, 因此 AB 是半正定阵当且仅当它是实对称矩阵, 即 $AB = (AB)' = BA$. \square

命题 0.14

设 A, B 都是 n 阶正定实对称矩阵, 满足 $AB = BA$, 求证: $A - B$ 是正定阵的充要条件是 $A^2 - B^2$ 是正定阵. \blacktriangle

证明 由 $AB = BA$ 可得 $A^2 - B^2 = (A + B)(A - B)$, 其中 $A + B$ 是正定阵. 由命题 0.10(3) 可知, $A - B$ 是正定阵当且仅当 $A^2 - B^2$ 的特征值全大于零, 即当且仅当 $A^2 - B^2$ 为正定阵. \square

0.1.3 利用正交相似标准型化简矩阵问题

命题 0.15

设 A, C 都是 n 阶正定实对称矩阵, 求证: 矩阵方程 $AX + XA = C$ 存在唯一解 B , 并且 B 也是正定实对称矩阵.

注 本题还可以作如下推广: 设 A 为 n 阶亚正定阵, C 为 n 阶正定 (半正定) 实对称矩阵, 则矩阵方程 $A'X + XA = C$ 存在唯一解 B , 并且 B 也是正定 (半正定) 实对称矩阵. 矩阵方程解的存在唯一性可由命题??得到, 正定 (半正定) 的证明类似于上面的讨论. 另外, 本题的逆命题并不成立, 即若 A 为正定阵, B 为正定 (半正定) 阵, 则 $AB + BA$ 不一定是正定 (半正定) 阵. 例如, $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \end{pmatrix}$, 其中 $0 \leq \varepsilon \ll 1$. 请读者自行验证具体的细节.

证明 设 P 为正交矩阵, 使得 $P'AP = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 其中 $\lambda_i > 0$. 注意到问题的条件和结论在同时正交相似变换 $A \mapsto P'AP, X \mapsto P'XP, C \mapsto P'CP$ 下不改变, 故不妨从一开始就假设 $A = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 为正交相似标准型. 设 $X = (x_{ij}), C = (c_{ij})$, 则矩阵方程 $AX + XA = C$ 等价于方程组 $(\lambda_i + \lambda_j)x_{ij} = c_{ij}$, 由此可唯一地解出 $x_{ij} = \frac{c_{ij}}{\lambda_i + \lambda_j} (1 \leq i, j \leq n)$, 从而矩阵方程有唯一解 $B = (\frac{c_{ij}}{\lambda_i + \lambda_j})$. 显然 B 是实对称矩阵, 任取 B 的特征值 λ 及其特征向量 α , 将等式 $AB + BA = C$ 左乘 α' , 右乘 α 可得

$$\alpha' A(B\alpha) + (B\alpha)' A\alpha = \alpha' C\alpha$$

即有 $2\lambda\alpha' A\alpha = \alpha' C\alpha$, 于是 $\lambda = \frac{\alpha' C\alpha}{2\alpha' A\alpha} > 0$, 因此 B 为正定阵. □

命题 0.16

设 A, B 是 n 阶实对称矩阵, 满足 $AB + BA = O$, 证明: 若 A 半正定, 则存在正交矩阵 P , 使得 $P'AP = \text{diag}\{\lambda_1, \dots, \lambda_r, 0, \dots, 0\}, P'BP = \text{diag}\{0, \dots, 0, \mu_{r+1}, \dots, \mu_n\}$.

证明 由于 A 半正定, 故存在正交矩阵 Q , 使得 $Q'AQ = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 其中 $\lambda_i > 0 (1 \leq i \leq r), \lambda_{r+1} = \dots = \lambda_n = 0$. 注意到问题的条件和结论在同时正交相似变换 $A \mapsto Q'AQ, B \mapsto Q'BQ$ 下不改变, 故不妨从一开始就假设 A 为正交相似标准型 $\text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. 设 $B = (b_{ij})$, 则由 $AB + BA = O$ 可得 $(\lambda_i + \lambda_j)b_{ij} = 0$. 当 i, j 至少有一个落在 $[1, r]$ 中时, 有 $\lambda_i + \lambda_j > 0$, 从而 $b_{ij} = 0$, 于是 $B = \text{diag}\{O, B_{n-r}\}$, 其中 B_{n-r} 是 B 右下角的 $n-r$ 阶主子阵. 由于 B_{n-r} 是一个实对称矩阵, 故存在 $n-r$ 阶正交矩阵 R , 使得 $R'B_{n-r}R = \text{diag}\{\mu_{r+1}, \dots, \mu_n\}$. 令 $P = \text{diag}\{I_r, R\}$, 则 P 是 n 阶正交矩阵, 使得 $P'AP = \text{diag}\{\lambda_1, \dots, \lambda_r, 0, \dots, 0\}, P'BP = \text{diag}\{0, \dots, 0, \mu_{r+1}, \dots, \mu_n\}$. □

命题 0.17

设 A 为 n 阶半正定实对称矩阵, S 为 n 阶实反对称矩阵, 满足 $AS + SA = O$. 证明: $|A + S| > 0$ 的充要条件是 $r(A) + r(S) = n$.

证明 由于 A 半正定, 故存在正交矩阵 Q , 使得 $Q'AQ = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 其中 $\lambda_i > 0 (1 \leq i \leq r), \lambda_{r+1} = \dots = \lambda_n = 0$. 注意到问题的条件和结论在同时正交相似变换 $A \mapsto Q'AQ, S \mapsto Q'SQ$ 下不改变, 故不妨从一开始就假设 A 为正交相似标准型 $\text{diag}\{\Lambda, O\}$, 其中 $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_r\}$. 设 $S = (b_{ij})$, 则由 $AS + SA = O$ 可得 $(\lambda_i + \lambda_j)b_{ij} = 0$. 当 i, j 至少有一个落在 $[1, r]$ 中时, 有 $\lambda_i + \lambda_j > 0$, 从而 $b_{ij} = 0$, 于是 $S = \text{diag}\{O, S_{n-r}\}$, 其中 S_{n-r} 是 S 右下角的 $n-r$ 阶主子阵. 注意到 S_{n-r} 是一个实反对称矩阵, 故由命题??可知 $|S_{n-r}| \geq 0$, 从而

$$|A + S| = |\text{diag}\{\Lambda, S_{n-r}\}| = |\Lambda| \cdot |S_{n-r}| \geq 0$$

因此, $|A + S| > 0$ 当且仅当 $|S_{n-r}| > 0$, 即当且仅当 $r(S_{n-r}) = n - r$, 这也当且仅当 $r(A) + r(S) = r(\Lambda) + r(S_{n-r}) = r + (n - r) = n$. □

0.1.4 可对角化判定准则 7: 相似于实对称矩阵

命题 0.18

设 A 是 n 阶实矩阵, B 是 n 阶正定实对称矩阵, 满足 $A'B = BA$, 证明: A 可对角化.

注 若 B 只是半正定阵, 则命题??的结论一般并不成立. 例如, $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, 则 $A'B = BA = O$, 但 A 不可对角化.

证明 **证法一:** 注意到 B 正定, 故由 $A'B = BA$ 可得

$$B^{\frac{1}{2}}AB^{-\frac{1}{2}} = B^{-\frac{1}{2}}A'B^{\frac{1}{2}} = (B^{\frac{1}{2}}AB^{-\frac{1}{2}})',$$

即 $B^{\frac{1}{2}}AB^{-\frac{1}{2}}$ 是实对称矩阵, 又 A 相似于 $B^{\frac{1}{2}}AB^{-\frac{1}{2}}$, 故 A 可对角化.

证法二: 设 $V = \mathbb{R}^n$, 取由正定阵 B 定义的内积, φ 为由矩阵 A 的乘法定义的线性变换. 由条件 $A'B = BA$ 经过简单的计算不难验证 $(\varphi(x), y) = (x, \varphi(y))$ 对任意的 $x, y \in V$ 成立, 因此 φ 是 V 上的自伴随算子, 从而可对角化, 于是 A 也可对角化. \square

命题 0.19

设 A, B 都是 n 阶半正定实对称矩阵, 证明: AB 可对角化.

注 由命题 0.10(2)或下述证明中 B_{11} 的半正定性可知, AB 相似于主对角元全大于等于零的对角矩阵. 另外, 若 A 是正定阵, B 是实对称矩阵, 则 AB 也可对角化. 事实上, AB 相似于 $A^{-\frac{1}{2}}(AB)A^{\frac{1}{2}} = A^{\frac{1}{2}}BA^{\frac{1}{2}}$, 这是一个实对称矩阵, 从而 AB 可对角化. 又若 A 是半正定阵, B 是实对称矩阵, 则 AB 一般不可对角化. 例如, $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, 则

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ 不可对角化.}$$

证明 设 C 为非异实矩阵, 使得 $C'AC = \text{diag}\{I_r, O\}$, 则 AB 相似于

$$C'AB(C')^{-1} = (C'AC)(C^{-1}B(C^{-1})').$$

注意到 $C^{-1}B(C^{-1})'$ 仍然是半正定阵, 故不妨从一开始就假设 A 是合同标准型 $\text{diag}\{I_r, O\}$. 设 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ 为对应的分块, 则 $AB = \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}$. 因为 B 半正定, 故由命题??可得 $r(B_{11} \ B_{12}) = r(B_{11})$, 于是存在实矩阵 M , 使得 $B_{12} = B_{11}M$. 考虑如下相似变换:

$$\begin{pmatrix} I & M \\ O & I \end{pmatrix} AB \begin{pmatrix} I & -M \\ O & I \end{pmatrix} = \begin{pmatrix} I & M \\ O & I \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix} \begin{pmatrix} I & -M \\ O & I \end{pmatrix} = \begin{pmatrix} B_{11} & O \\ O & O \end{pmatrix},$$

于是 AB 相似于 $\text{diag}\{B_{11}, O\}$, 这是一个实对称矩阵, 从而 AB 可对角化. \square

命题 0.20

设 n 阶实矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & a_{n-1,n-1} & a_{n-1,n} \\ & & & a_{n,n-1} & a_{nn} \end{pmatrix}$$

求证: 若 $a_{i,i+1}a_{i+1,i} \geq 0 (1 \leq i \leq n-1)$, 则 A 的特征值全为实数;

若 $a_{i,i+1}a_{i+1,i} > 0 (1 \leq i \leq n-1)$, 则 A 在实数域上可对角化.

证明 在三对角矩阵 A 中, 若存在某个 $a_{i,i+1} = 0$ 或 $a_{i+1,i} = 0$, 则 $|\lambda I - A| = |\lambda I - A_1| |\lambda I - A_2|$, 其中 A_1, A_2 是满足相同条件的低阶三对角矩阵. 不断这样做下去, 故我们只要证明若 $a_{i,i+1}a_{i+1,i} > 0 (1 \leq i \leq n-1)$, 则 A 在实数域上可对角化即可. 考虑如下相似变换: 将 A 的第二行乘以 $\sqrt{\frac{a_{12}}{a_{21}}}$, 再将第二列乘以 $\sqrt{\frac{a_{21}}{a_{12}}}$, 这样第 $(1, 2)$ 元素和第 $(2, 1)$ 元素都变成了 $\sqrt{a_{12}a_{21}}$; 第 $(1, 1)$ 元素和第 $(2, 2)$ 元素保持不变; 第 $(2, 3)$ 元素变为 $a_{23}\sqrt{\frac{a_{12}}{a_{21}}}$, 第 $(3, 2)$ 元素变为 $a_{32}\sqrt{\frac{a_{21}}{a_{12}}}$. 一般地, 令 $d_{i+1} = \sqrt{\frac{a_{12}a_{23} \cdots a_{i,i+1}}{a_{21}a_{32} \cdots a_{i+1,i}}}$, 依次将第 $i+1$ 行乘以 d_{i+1} , 再将第 $i+1$ 列乘以 $d_{i+1}^{-1} (1 \leq i \leq n-1)$, 最后可得到 A 实相似于一个实对称矩阵, 从而 A 在实数域上可对角化. \square