

Final

Collated by Undefined

Date: 10/1/2022 Time: 8:30 ~ 10:50

Problems

1 对称矩阵 $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ ($n \in \mathbb{N}^*$), 证明: $\text{tr}(\mathbf{A}\mathbf{B}\mathbf{A}\mathbf{B}) \leq \text{tr}(\mathbf{A}\mathbf{A}\mathbf{B}\mathbf{B})$.

2 设 n 是正整数, 矩阵 $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ 有相同的特征多项式, 有相同的极小多项式, 且 $\forall \mu \in \mathbb{C}$, $\text{rank}(\mu \mathbf{I} - \mathbf{A}) = \text{rank}(\mu \mathbf{I} - \mathbf{B})$, 其中 \mathbf{I} 是同阶单位矩阵. 问 \mathbf{A}, \mathbf{B} 是否一定相似? 如果是请证明, 否则请给出反例.

3 设 n 是正整数, 向量 $x, y \in \mathbb{C}^n$ 且满足 $x^*x = 1, x^*y = 0$, \mathbf{I} 是 n 阶单位矩阵, 证明:
 $\|\mathbf{I} - (x + y)x^*\|_2 = \|x + y\|_2$.

4 设 $n \in \mathbb{N}^*$, n 阶正规矩阵 \mathbf{A}, \mathbf{B} 满足 $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$, 证明: $\mathbf{A}\mathbf{B}$ 是正规矩阵.

5 正整数 m, n 满足 $0 < m < n$, 分块矩阵 $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$ 是 Hermite 正定矩阵, 其中 $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{A}_{11} \in \mathbb{C}^{m \times m}$. 设 $\mathbf{S} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$.
 证明: (i) \mathbf{S} 是正定矩阵; (ii) $\|\mathbf{S}\|_2 \|\mathbf{S}^{-1}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$.

6 设 $n \in \mathbb{N}^*$, 凸集 $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$, 且 $\Omega_1 \cap \Omega_2 = \emptyset$. 取 $x_1 \in \Omega_1, x_2 \in \Omega_2, \alpha \in [1, +\infty)$, 集合 $\Omega_3 = \{y + \alpha(x_2 - x_1) \mid y \in \Omega_2\}$, 证明:
 (i) Ω_3 是凸集; (ii) $\Omega_1 \cap \Omega_3 = \emptyset$.

7 设 $\mathbf{A} \in \mathbb{C}^{2 \times 2}$, $\forall t \in \mathbb{C}$, 有 $\exp(t\mathbf{A}) = e^t \begin{bmatrix} 1-t & -t \\ t & 1+t \end{bmatrix}$.

请给出一个这样的矩阵 \mathbf{A} , 或者证明为什么这样的矩阵 \mathbf{A} 不存在.

8 设 n 是正整数, 矩阵 $\mathbf{A} \in \mathbb{R}^{n \times n}$ 的所有元素非负, 其 (i, j) 元素为 a_{ij} ,
 证明: 对任何 n 维正向量 x , 有

$$\min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j \leq \rho(\mathbf{A}) \leq \max_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j.$$

Answer

1

Proof. 注意到 $(AB)^\top = B^\top A^\top = BA$, 令 $C = AB = (c_{ij})_{n \times n}$,

$$\operatorname{tr}(C^\top C) - \operatorname{tr}(C^2) = \|C\|_2^2 - \sum_{i,j} c_{ij} c_{ji} = \sum_{i=1}^{n-1} \sum_{j=i+1}^n (c_{ij} - c_{ji})^2 \geq 0;$$

$$\operatorname{tr}(ABAB) = \operatorname{tr}(C^2) \leq \operatorname{tr}(C^\top C) = \operatorname{tr}(BAAB) = \operatorname{tr}(AABB).$$

即 $\operatorname{tr}(ABAB) \leq \operatorname{tr}(AABB)$, 证毕. □

2 Solution A, B 不一定相似. 作 A, B 如下:

$$A = \begin{bmatrix} 1 & 0 & & & & & & \\ & 1 & 1 & & & & & \\ & & 1 & 1 & & & & \\ & & & 1 & 0 & & & \\ & & & & 1 & 1 & & \\ & & & & & 1 & 1 & \\ & & & & & & 1 & 1 \\ & & & & & & & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & & & & & & \\ & 1 & 0 & & & & & \\ & & 1 & 1 & & & & \\ & & & 1 & 0 & & & \\ & & & & 1 & 1 & & \\ & & & & & 1 & 1 & \\ & & & & & & 1 & 1 \\ & & & & & & & 1 \end{bmatrix}$$

A, B 特征多项式均为 $(\lambda - 1)^8$, 极小多项式均为 $(\lambda - 1)^4$,

当 $\mu \neq 2$ 时, $\operatorname{rank}(\mu I - A) = \operatorname{rank}(\mu I - B) = 8$,

当 $\mu = 2$ 时, $\operatorname{rank}(\mu I - A) = \operatorname{rank}(\mu I - B) = 5$.

但 A, B 是不相似的 Jordan 标准型.

3

Proof. 令 $P = (x + y)x^*$, $P^2 = (x + y)x^*(x + y)x^* = (x + y)x^* = P$,

即 $P^2 = P$, 且 P 的有一个特征值为 $x^*(x + y) = 1$, 其余特征值均为 0.

$\forall x \in \mathbb{C}^n, x = Px + (I - P)x$, $Px \in \operatorname{Ker}(I - P)$, $(I - P)x \in \operatorname{Ker}(P)$,

设 $y \in \operatorname{Ker}(P) \cap \operatorname{Ker}(I - P)$, $y = Py + (I - P)y = 0$,

因此 $\mathbb{C}^n = \operatorname{Ker}(P) \oplus \operatorname{Ker}(I - P)$.

$\forall x \in \mathbb{C}^n \setminus \{0\}$, 设 $x = u + v$, $u \in \operatorname{Ker}(P)$, $v \in \operatorname{Ker}(I - P)$,

$$\text{构造 } z = \|v\|_2 \frac{u}{\|u\|_2} + \|u\|_2 \frac{v}{\|v\|_2}, \text{ 有 } \frac{\|Px\|_2}{\|x\|_2} = \frac{\|(I - P)z\|_2}{\|z\|_2} = \frac{\|u\|_2}{\|u + v\|_2}.$$

$$\text{由上述对应关系知 } \max_{x \neq 0} \frac{\|Px\|_2}{\|x\|_2} = \max_{z \neq 0} \frac{\|(I - P)z\|_2}{\|z\|_2}, \quad \|P\|_2 = \|I - P\|_2.$$

$$\text{又 } P \text{ 的奇异值分解为 } \|x + y\|_2 \cdot \frac{x + y}{\|x + y\|_2} x^*, \text{ 故 } \|P\|_2 = \|x + y\|_2.$$

从而 $\|I - (x + y)x^*\|_2 = \|x + y\|_2$, 证毕. □

4

Proof. 由 $\mathbf{AB} = \mathbf{BA}$, 设 \mathbf{A} 对应于特征值 λ_0 的特征子空间为 \mathcal{V}_{λ_0} ,

$\forall x \in \mathcal{V}_{\lambda_0}, \mathbf{AB}x = \mathbf{BA}x = \lambda_0 \mathbf{B}x$, 故 $\mathbf{B}x \in \mathcal{V}_{\lambda_0}$, 即 \mathcal{V}_{λ_0} 是 \mathbf{B} 的不变子空间. 将 \mathbf{B} 对应的线性变换 φ 限制在 \mathcal{V}_{λ_0} 上得 $\varphi|_{\mathcal{V}_{\lambda_0}}$, 其在 \mathcal{V}_{λ_0} 上至少存在一个特征值 μ_0 和一个特征向量 v , 因此 \mathbf{A}, \mathbf{B} 至少存在一个公共特征向量 v .

取 \mathbf{A}, \mathbf{B} 的一个公共特征向量 q_1 ($\|q_1\|_2 = 1$), 并将其扩展为全空间上一组标准正交基, 记作酉矩阵 $\mathbf{Q}_1 = [q_1, \tilde{q}_2, \dots, \tilde{q}_n]$.

$$\mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 = \begin{bmatrix} \lambda_1 & \alpha^* \\ 0 & \mathbf{A}_1 \end{bmatrix}, \quad \mathbf{Q}_1^* \mathbf{A}^* \mathbf{Q}_1 = \begin{bmatrix} \bar{\lambda}_1 & 0 \\ \alpha & \mathbf{A}_1^* \end{bmatrix}$$

由于 \mathbf{A} 是正规矩阵, 知 $\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$,

$$\begin{bmatrix} |\lambda_1|^2 & \bar{\lambda}_1 \alpha^* \\ \alpha \lambda_1 & \mathbf{A}_1^* \mathbf{A}_1 + \alpha \alpha^* \end{bmatrix} = \begin{bmatrix} |\lambda_1|^2 + \alpha^* \alpha & \alpha^* \mathbf{A}_1^* \\ \mathbf{A}_1 \alpha & \mathbf{A}_1 \mathbf{A}_1^* \end{bmatrix}$$

比较可知 $\alpha = 0$, $\mathbf{A}_1^* \mathbf{A}_1 = \mathbf{A}_1 \mathbf{A}_1^*$, $\mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \mathbf{A}_1 \end{bmatrix}$.

同理 $\mathbf{Q}_1^* \mathbf{B} \mathbf{Q}_1 = \begin{bmatrix} \mu_1 & 0 \\ 0 & \mathbf{B}_1 \end{bmatrix}$, 且 \mathbf{B}_1 是正规矩阵, $\mathbf{A}_1 \mathbf{B}_1 = \mathbf{B}_1 \mathbf{A}_1$.

对 $\mathbf{A}_1, \mathbf{B}_1$ 继续进行上述操作, 最后可以找到酉矩阵 $\mathbf{Q}_1, \dots, \mathbf{Q}_{n-1}$ 满足

$$\mathbf{Q}_{n-1}^* \cdots \mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 \cdots \mathbf{Q}_{n-1} = \text{diag}\{\lambda_1, \dots, \lambda_n\} = \mathbf{\Lambda},$$

$$\mathbf{Q}_{n-1}^* \cdots \mathbf{Q}_1^* \mathbf{B} \mathbf{Q}_1 \cdots \mathbf{Q}_{n-1} = \text{diag}\{\mu_1, \dots, \mu_n\} = \mathbf{M}.$$

令酉阵 $\mathbf{Q} = \mathbf{Q}_1 \cdots \mathbf{Q}_{n-1}$, 有 $\mathbf{Q}^* \mathbf{A} \mathbf{B} \mathbf{Q} = \mathbf{\Lambda} \mathbf{M}$, 故 $\mathbf{A} \mathbf{B}$ 可酉对角化,

因此 $\mathbf{A} \mathbf{B}$ 是正规矩阵, 证毕. □

5

Proof. 对矩阵 \mathbf{A} 作分块合同变换:

$$\begin{bmatrix} \mathbf{I} & 0 \\ -\mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \\ 0 & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & 0 \\ 0 & \mathbf{S} \end{bmatrix},$$

正定矩阵的任意主子阵正定, 且合同变换不改变正定性, 可知 \mathbf{A}_{11} , \mathbf{A}_{22} , \mathbf{S} 均为正定矩阵. 同时 \mathbf{A}_{11}^{-1} 也是正定矩阵, $\mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}$ 是半正定矩阵.

用 $\lambda_1, \lambda_2, \dots$ 表示一个矩阵按升序排列的特征值.

$$\lambda_{n-m}(\mathbf{S}) = \max_{\|x\|_2=1} x^* (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) x \leq \max_{\|x\|_2=1} x^* \mathbf{A}_{22} x = \lambda_{n-m}(\mathbf{A}_{22});$$

再由Cauchy交错定理得 $\|\mathbf{S}\|_2 = \lambda_{n-m}(\mathbf{S}) \leq \lambda_{n-m}(\mathbf{A}_{22}) \leq \lambda_n(\mathbf{A}) = \|\mathbf{A}\|_2$.

令 $\mathbf{X} = \begin{bmatrix} -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{I} \end{bmatrix}$, 有 $\mathbf{X}^*\mathbf{A}\mathbf{X} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} = \mathbf{S}$.

$\forall x \in \mathbb{C}^{n-m}$, 有 $\mathbf{X}x = \begin{bmatrix} -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}x \\ x \end{bmatrix}$, 即 $\|\mathbf{X}x\|_2 \geq \|x\|_2$.

又知 $\frac{y^*\mathbf{A}y}{y^*y}$ 在 $\text{Range}(\mathbf{X}) \setminus \{0\}$ 上的局部最小值不小于 $\mathbb{C}^n \setminus \{0\}$ 上的全局最小值.

$$\text{故 } \lambda_1(\mathbf{S}) = \min_{x \neq 0} \frac{x^*\mathbf{X}^*\mathbf{A}\mathbf{X}x}{x^*x} \geq \min_{x \neq 0} \frac{(\mathbf{X}x)^*\mathbf{A}\mathbf{X}x}{(\mathbf{X}x)^*\mathbf{X}x} \geq \min_{y \neq 0} \frac{y^*\mathbf{A}y}{y^*y} = \lambda_1(\mathbf{A}),$$

$$\|\mathbf{S}^{-1}\|_2 = \frac{1}{\lambda_1(\mathbf{S})} \leq \frac{1}{\lambda_1(\mathbf{A})} = \|\mathbf{A}^{-1}\|_2.$$

故 $\|\mathbf{S}\|_2\|\mathbf{S}^{-1}\|_2 \leq \|\mathbf{A}\|_2\|\mathbf{A}^{-1}\|_2$, 证毕.

□

6

Proof. Ω_1, Ω_2 是凸集, $\Omega_3 = \{y + \alpha(x_2 - x_1) \mid y \in \Omega_2\}$, 则

$\forall z_1, z_2 \in \Omega_3, \exists y_1, y_2 \in \Omega_2$ s.t. $z_1 = y_1 + \alpha(x_2 - x_1), z_2 = y_2 + \alpha(x_2 - x_1)$;

$\forall \theta \in [0, 1], \theta z_1 + (1 - \theta)z_2 = \theta y_1 + (1 - \theta)y_2 + \alpha(x_2 - x_1)$, 由于 Ω_2 是凸集, 故 $\theta y_1 + (1 - \theta)y_2 \in \Omega_2$, $\theta z_1 + (1 - \theta)z_2 \in \Omega_3$. 因此 Ω_3 是凸集.

由凸集分离定理, $\Omega_1 \cap \Omega_2 = \emptyset$, 故存在 $w \in \mathbb{R}^n \setminus \{0\}, \gamma \in \mathbb{R}$, 使得

$\forall u \in \Omega_1, v \in \Omega_2, w^\top u \leq \gamma \leq w^\top v$.

$\forall z \in \Omega_3$, 令 $\Omega_2 \ni v = z - \alpha(x_2 - x_1)$,

$w^\top z = w^\top v + \alpha w^\top (x_2 - x_1) \geq \gamma + \alpha w^\top (x_2 - x_1) \geq \gamma$;

设 Ω_1, Ω_3 存在公共元素 z_* , 可知 $w^\top z_* = \gamma$,

z_* 落在超平面 $H = \{x \in \mathbb{R}^n \mid w^\top x = \gamma\}$ 上, 且 $v_* = z_* - \alpha(x_2 - x_1) \in \Omega_2$;

作 $s = \frac{1}{1 + \alpha}v_* + \frac{\alpha}{1 + \alpha}x_2, s \in \Omega_2$, 但同时 $s = \frac{1}{1 + \alpha}z_* + \frac{\alpha}{1 + \alpha}x_1, s \in \Omega_1$,

与 $\Omega_1 \cap \Omega_2 = \emptyset$ 矛盾, 故 Ω_1, Ω_3 在超平面 H 上无公共元素.

又 $\forall u \in \Omega_1, z \in \Omega_3, w^\top u \leq \gamma \leq w^\top z$, 故 $\Omega_1 \cap \Omega_3 = \emptyset$, 证毕. □

7 Solution. 首先设 A 可对角化, 即 $P^{-1}AP = \text{diag}\{\lambda_1, \lambda_2\}$, 则

$$\exp(tA) = P \begin{bmatrix} e^{t\lambda_1} & \\ & e^{t\lambda_2} \end{bmatrix} P^{-1}, \text{ 为将幂因子提出, 只能 } \lambda_1 = \lambda_2 = 1,$$

此时有 $\exp(tA) = e^{t\lambda_1}I$, 矛盾, 因此 A 不可对角化.

$$\text{设 } P^{-1}JP = A, J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, tJ = \begin{bmatrix} t & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t\lambda & 1 \\ 0 & t\lambda \end{bmatrix} \begin{bmatrix} 1/t & -1/t \\ 0 & 1 \end{bmatrix};$$

$$\exp(tA) = P \begin{bmatrix} t & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^{t\lambda} & e^{t\lambda} \\ 0 & e^{t\lambda} \end{bmatrix} \begin{bmatrix} 1/t & -1/t \\ 0 & 1 \end{bmatrix} P^{-1} = P \begin{bmatrix} e^{t\lambda} & te^{t\lambda} \\ 0 & e^{t\lambda} \end{bmatrix} P^{-1},$$

$$\text{为将幂因子提出, 只能 } \lambda = 1, \text{ 从而 } P \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} 1 - t & -t \\ t & 1 + t \end{bmatrix};$$

$$\text{解得一个 } P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}, A = PJP^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}.$$

$$\text{因此存在满足题意的 } A, \text{ 例如 } A = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}.$$

8

Proof. 设 $r = \min_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}$;

作矩阵 B , 其 (i, j) 元素 $b_{ij} = \frac{ra_{ij}}{\sum_{k=1}^n a_{ik}}$, 知 B 也是非负矩阵且 $B \leq A$:

进一步有 $\mathbf{B}^k \leq \mathbf{A}^k$, $\|\mathbf{B}^k\|_\infty \leq \|\mathbf{A}^k\|_\infty$ ($k \in \mathbb{N}$), 由 Gelfand 谱半径公式,

$$\rho(\mathbf{B}) = \lim_{k \rightarrow \infty} \|\mathbf{B}^k\|_\infty^{1/k} \leq \lim_{k \rightarrow \infty} \|\mathbf{A}^k\|_\infty^{1/k} = \rho(\mathbf{A});$$

r 既是 \mathbf{B} 的一个特征值, 也是 \mathbf{B} 的行和范数, 故 $r \leq \rho(\mathbf{B}) \leq \|\mathbf{B}\|_\infty \leq r$,

从而 $\rho(\mathbf{A}) \geq \rho(\mathbf{B}) = r = \min_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}$;

由 Gershgorin 第一圆盘定理, $\rho(\mathbf{A}) \leq \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}$;

故 $\min_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} \leq \rho(\mathbf{A}) \leq \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}$;

对任意 n 维正向量, 设其第 i 分量为 x_i ($x_i > 0$, $i = 1, 2, \dots, n$),

作 $\mathbf{X} = \text{diag}\{x_1, \dots, x_n\}$, 再作 $\mathbf{C} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$, 知 \mathbf{C} 仍为非负矩阵.

其 (i, j) 元素 $c_{ij} = x_i^{-1}a_{ij}x_j$, 对 \mathbf{C} 使用上述结论即得

$$\min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j \leq \rho(\mathbf{C}) \leq \max_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j;$$

注意到 $\rho(\mathbf{C}) = \rho(\mathbf{A})$, 因此

$$\min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j \leq \rho(\mathbf{A}) \leq \max_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j;$$

证毕. □