2025 年 12 月 2 日作业 (截止: 2025 年 12 月 9 日 08:00)

1. 已知 m, n 是正整数, $A(\tau)$ 是 $m \times n$ 的矩阵, 且 $A(\tau)$ 的每个元素都是关于参数 τ 的连续函数, 其中 $0 \le \tau \le 1$. 证明: 对于任意矩阵范数 $\|\cdot\|$ 都有

$$\left\| \int_0^1 A(\tau)\,\mathrm{d}\tau \right\| \le \int_0^1 \|A(\tau)\|\,\mathrm{d}\tau.$$

2. 设 n, k 为正整数, $1 \le k \le n$. 已知 $A \in \mathbb{R}^{n\times n}$ 是对称正定矩阵. 证明:

$$\max_{\substack{X \in \mathbb{R}^{n\times k} \\ X^\top X = I_k}} \det(X^\top A X) = \prod_{i=1}^{k} \sigma_i(A).$$

3. 若 n 是正整数, $F(\cdot)$ 是 $\mathbb{R}^n$ 上的仿射变换, 证明: 由 $G(x) = F(x) - F(0)$ 定义的变换 $G(\cdot)$ 是 $\mathbb{R}^n$ 上的线性变换.

4. 已知 n 是正整数. 给定实对称半正定矩阵 $A \in \mathbb{R}^{n\times n}$, 以及 $b \in \mathbb{R}^n$, $c \in \mathbb{R}$, 证明: 集合

$$\Omega = \left\{ x \in \mathbb{R}^n : x^\top A x + b^\top x + c \le 0 \right\}$$

是凸集.

5. 若 n 是正整数, $A \in \mathbb{C}^{n\times n}$ 是正规矩阵, 证明: A 的数值域

$$w(A) = \left\{ \frac{x^* A x}{x^* x} : x \in \mathbb{C}^n \setminus \{0\} \right\}$$

是 A 的特征值的凸包 (即包含 A 的所有特征值的最小凸集).

6. (选做) 在欧氏平面 $\mathbb{R}^2$ 上, 若 I 是三角形 ABC 的内心, D, E, F 分别是 I 在 BC, CA, AB 边上的正投影. 设 $a = |BC|$, $b = |CA|$, $c = |AB|$, 证明:

$$I = \frac{aD + bE + cF}{a + b + c}.$$

7. (选做) 已知 n, k 为正整数, 且 $1 \le k \le n$. 若 $\mathcal{X}$ 和 $\mathcal{Y}$ 都是 $\mathbb{C}^n$ 的 k 维子空间, 对应的正投影算子分别为 $P_{\mathcal{X}}$ 和 $P_{\mathcal{Y}}$, 可以定义 $\mathcal{X}, \mathcal{Y}$ 之间的距离为

$$\operatorname{dist}(\mathcal{X}, \mathcal{Y}) = \|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2.$$

试建立 $\operatorname{dist}(\mathcal{X}, \mathcal{Y})$ 与 $\mathcal{X}, \mathcal{Y}$ 之间的经典角的关系.

参考答案

1

证明. 考虑将积分展开为 Riemann 和，则有

$$\int_0^1 A(\tau)\,\mathrm{d}\tau = \lim_{|P|\to 0}\sum_{i=1}^{N} A(\xi_i)\Delta\tau_i,$$

由三角不等式知

$$\left\|\sum_{i=1}^{N} A(\xi_i)\Delta\tau_i\right\| \le \sum_{i=1}^{N}\|A(\xi_i)\|\Delta\tau_i.$$

两边取极限后即得所需不等式。 □

2

证明. 考虑 A 的谱分解 $A = Q\Lambda Q^\top$，则有

$$X^\top AX = X^\top Q\Lambda Q^\top X := Y^\top\Lambda Y, \quad Y = Q^\top X.$$

容易验证 Y 仍然是正交归一的。因此，我们只需证

$$\max_{\substack{Y\in\mathbb{R}^{n\times k}\\ Y^\top Y=I_k}} \det(Y^\top\Lambda Y) = \prod_{i=1}^{k}\sigma_i(\Lambda).$$

下面我们考虑 $Y^\top\Lambda Y$ 的特征值。考虑其 Rayleigh 商

$$R(z) = \frac{z^\top Y^\top\Lambda Yz}{z^\top z} = \frac{(Yz)^\top\Lambda(Yz)}{(Yz)^\top(Yz)} = R(Yz,\Lambda).$$

而由于 $Y\in\mathbb{R}^{n\times k}$，最多能够张成 k 维空间，最多能抓到 Λ 的前 k 个最大特征值。具体地，Poincaré 分离定理告诉我们，$Y^\top\Lambda Y$ 的特征值最多为 $\sigma_1(\Lambda),\sigma_2(\Lambda),\ldots,\sigma_k(\Lambda)$，从而

$$\det(Y^\top\Lambda Y) \le \prod_{i=1}^{k}\sigma_i(\Lambda).$$

另一方面，取 $Y = [q_1,\ldots,q_k]$，即 $X = [I_k, 0]^\top$，其中 q_i 是 Λ 的第 i 个标准正交特征向量，则有

$$\det(Y^\top\Lambda Y) = \prod_{i=1}^{k}\sigma_i(\Lambda).$$

因此

$$\max_{\substack{X\in\mathbb{R}^{n\times k}\\ X^\top X=I_k}} \det(X^\top AX) = \prod_{i=1}^{k}\sigma_i(A).$$

□

3

证明. 我们考虑 F 在数乘下会表现出怎样的行为。任意 $x \in \mathbb{R}^n$，$\alpha \in \mathbb{R}$，我们有

$$F(x\alpha) = F(x\alpha + 0 \cdot (1-\alpha)) = F(x)\alpha + F(0)(1-\alpha).$$

因此

$$G(x\alpha) = F(x\alpha) - F(0) = (F(x) - F(0))\alpha = G(x)\alpha.$$

亦即 G 保持数乘。

下面考虑 G 在加法下的行为。任取 $x, y \in \mathbb{R}^n$，有

$$\begin{aligned} G(x+y) &= F\left(\frac{1}{2}(2x) + \frac{1}{2}(2y)\right) - F(0) \\ &= \frac{1}{2}F(2x) + \frac{1}{2}F(2y) - F(0) \\ &= \frac{1}{2}(G(2x) + F(0)) + \frac{1}{2}(G(2y) + F(0)) - F(0) \\ &= G(x) + G(y) + F(0) - F(0) \\ &= G(x) + G(y). \end{aligned}$$

综上所述，G 同时保持加法与数乘，因此 G 是线性变换。 □

4

证明. 任取 $x, y \in \Omega$ 以及 $\alpha, \beta \in [0,1]$，其中 $\alpha + \beta = 1$. 设 $f(x) = x^\top Ax + b^\top x + c$，则

$$\begin{aligned} \alpha f(x) + \beta f(y) - f(\alpha x + \beta y) &= \alpha x^\top Ax + \beta y^\top Ay - \alpha^2 x^\top Ax - 2\alpha\beta x^\top Ay \\ &= \alpha\beta(x-y)^\top A(x-y) \geq 0, \end{aligned}$$

因此

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y) \leq 0,$$

即 $\alpha x + \beta y \in \Omega$，从而 Ω 是凸集。 □

当然，还可以直接求二阶导，证明 f 是凸函数，从而得到 Ω 是凸集。

5

证明. 考虑谱分解 $A = Q\Lambda Q^*$，A 的数值域等价于 Λ 的数值域。不妨设 $\xi = x^*\Lambda x, \eta = y^*\Lambda y$，其中 $\xi \neq \eta$ 且 x 和 y 是单位向量，则 x 与 y 线性无关。不妨设 $\xi < \eta$，任取

$\alpha, \beta \in [0,1]$ 且 $\alpha + \beta = 1$，则有

$$\begin{aligned} \xi\alpha + \eta\beta &= x^*\Lambda x\alpha + y^*\Lambda y\beta \\ &= \sum_{i=1}^{n} \lambda_i(\alpha|x_i|^2 + \beta|y_i|^2) \\ &= \sum_{i=1}^{n} \lambda_i|z_i|^2, \quad z_i = \sqrt{|x_i|^2\alpha + |y_i|^2\beta} \\ &= z^*\Lambda z, \quad z = (z_1, z_2, \ldots, z_n)^\top. \end{aligned}$$

因此，A 的数值域包含于 A 的特征值的凸包。另一方面，任取 A 的任一特征值 λ_j，都有

$$e_j^*\Lambda e_j = \lambda_j,$$

其中 e_j 是第 j 个标准正交基向量。因此，A 的特征值的凸包包含于 A 的数值域。综上所述，A 的数值域即为 A 的特征值的凸包。 □

6

证明. 由于

$$\begin{aligned} S_{\triangle EIF} &= k\sin\angle EIF = k\sin\angle A = k'a \\ S_{\triangle DIF} &= k\sin\angle DIF = k\sin\angle B = k'b \\ S_{\triangle DIE} &= k\sin\angle DIE = k\sin\angle C = k'c, \end{aligned}$$

因此

$$I = \frac{aD + bE + cF}{a + b + c}.$$

□

7

证明. 我们选取 $\mathcal{X}$ 的一组基作为全空间的一组基。这样，$P_{\mathcal{X}}$ 可以简化为

$$P_{\mathcal{X}} = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}.$$

设

$$P_{\mathcal{Y}} = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}, \quad A = A^*, D = D^*.$$

由于 $P_{\mathcal{Y}}$ 是正投影算子，因此 $P_{\mathcal{Y}}^2 = P_{\mathcal{Y}}$，从而

$$A^2 + BB^* = A, \quad AB + BD = B, \quad B^*B + D^2 = D.$$

由经典角的定义，我们考虑 $P_{\mathcal{X}}P_{\mathcal{Y}}$ 的奇异值分解：

$$\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}^2 \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$$

因此，A 的特征值就是经典角余弦值的平方。由子空间距离的定义，我们考虑

$$\left(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \right) \left(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \right) = \begin{bmatrix} I - A & 0 \\ 0 & D \end{bmatrix}$$

因此，$P_{\mathcal{X}} - P_{\mathcal{Y}}$ 的奇异值就是 $I - A$ 和 D 的特征值的平方根，最大奇异值就是 $I - A$ 的最大特征值的平方根或 D 的最大特征值的平方根。下面我们来比较一下这两者的大小。由投影算子的性质我们知道

$$B^*B = D - D^2.$$

任取 D 的特征向量 v，对应的特征值为 λ，我们有

$$B^*Bv = Dv - D^2v = v\lambda(1 - \lambda).$$

同理，

$$BB^*(Bv) = A(Bv) - A^2(Bv) = (Bv)\lambda(1 - \lambda).$$

因此，A 和 D 的非零特征值要么相同，要么和为一。事实上，我们可以从代数和几何两个角度共同验证：A 和 D 的非零特征值和为 1. 从代数角度来看，由于 $P_{\mathcal{Y}}$ 是 k 维子空间上的投影算子，因此

$$\mathrm{tr}(P_{\mathcal{Y}}) = \mathrm{tr}(P_{\mathcal{Y}}^2) = \mathrm{tr}(P_{\mathcal{Y}}^{\star}P_{\mathcal{Y}}) = \|P_{\mathcal{Y}}\|_{\mathrm{F}}^2 = k,$$

如果 A 和 D 的非零特征值相同，那么特征值的总和为 $2\sum\lambda_i \neq k$，这与迹为 k 矛盾。因此，A 和 D 的非零特征值和为 1.

从几何角度来看，A 表示 $\mathcal{Y}$ 在 $\mathcal{X}$ 上的投影，而 D 表示 $\mathcal{Y}$ 在 $\mathcal{X}^{\perp}$ 上的投影，如果 A 的特征值是 $\cos^2\theta_i$，那么 D 的特征值就是 $\sin^2\theta_i$，两者和为 1.

于是，我们发现：$I - A$ 和 D 的特征值完全相同。因此，

$$\mathrm{dist}(\mathcal{X}, \mathcal{Y}) = \|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2 = \sqrt{\lambda_{\max}(I - A)} = \sin\theta_1.$$

□