

2025 年 11 月 11 日作业 (截止: 2025 年 11 月 18 日 08:00)

1. 若  $n$  是大于 1 的整数,  $A, B \in \mathbb{C}^{n \times n}$ . 举例说明下列情况可能发生:

(a)  $(AB)^\dagger \neq B^\dagger A^\dagger$ .

(b)  $(A^k)^\dagger \neq (A^\dagger)^k$ , 其中  $k$  是正整数.

(c)  $A^\dagger$  的非零特征值的倒数不是  $A$  的特征值.

2. 已知  $m, n, k$  是正整数,  $X \in \mathbb{C}^{m \times k}$  是列满秩矩阵,  $Y \in \mathbb{C}^{k \times n}$  是行满秩矩阵, 证明:  
 $(XY)^\dagger = Y^\dagger X^\dagger$ .

3. 若  $n$  是正整数,  $A \in \mathbb{C}^{n \times n}$  是 Hermite 矩阵, 其  $(i, j)$  元素为  $a_{i,j}$ . 证明: 每个圆盘

$$\mathcal{D}_i = \left\{ \lambda \in \mathbb{C} : |\lambda - a_{i,i}|^2 \leq \sum_{j \neq i} |a_{i,j}|^2 \right\}$$

中都至少含有  $A$  的一个特征值.

4. 若  $n$  是正整数,  $A \in \mathbb{C}^{n \times n}$  是 Hermite 矩阵,  $v \in \mathbb{C}^n$ , 证明:

$$\lambda_i(A) \leq \lambda_i(A + vv^*) \leq \lambda_{i+1}(A), \quad (1 \leq i < n).$$

5. 已知  $m, n$  是正整数且  $1 \leq m < n$ . 若  $n$  阶 Hermite 正定矩阵  $A$  可分块表示为

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

其中  $A_{11} \in \mathbb{C}^{m \times m}$ , 证明:  $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$  正定, 并且

$$\|S\|_2 \|S^{-1}\|_2 \leq \|A\|_2 \|A^{-1}\|_2.$$

6. (选做) 证明: 对于任何复矩阵  $A$  都有

$$A^\dagger = \lim_{\delta \rightarrow 0+} (A^*A + \delta I)^{-1} A^* = \lim_{\delta \rightarrow 0+} A^* (AA^* + \delta I)^{-1}.$$

7. (选做) 若  $a, b, c$  是实数, 证明:

$$-\frac{3}{2}(a^2 + b^2 + 2c^2) \leq 3ab + bc + ca \leq \frac{3 + \sqrt{13}}{4}(a^2 + b^2 + 2c^2).$$

## 参考解答

1

证明. (a) 为方便计算, 下设  $n = 3$ . 取  $A = e_1 e_1^\top$ ,  $B = e e^\top$ . 则

$$\begin{aligned} A &= I(e_1 e_1^\top)I & \implies A^\dagger &= A \\ B &= U\Lambda U^\top, \quad \Lambda = \text{diag}\{3, 0, 0\} & \implies B^\dagger &= U\Lambda^\dagger U^\top = \frac{1}{9}B \\ AB &= I(\sqrt{3}e_1 e_1^\top)U^\top & \implies (AB)^\dagger &= U(\sqrt{3}e_1 e_1^\top). \end{aligned}$$

而  $B^\dagger A^\dagger = \frac{1}{9}B e_1 e_1^\top$ . 显然  $(AB)^\dagger \neq B^\dagger A^\dagger$ .

(b) 我们需要找一个非正规矩阵. 我们希望能够找到一个幂等矩阵  $A$ , 但  $A^\dagger$  不是幂等的. 取

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

则

$$A^\dagger = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}.$$

显然  $A^2 = A$ , 但

$$(A^\dagger)^2 = \begin{bmatrix} \frac{1}{4} & 0 \\ \frac{1}{4} & 0 \end{bmatrix} \neq A^\dagger.$$

因此  $(A^2)^\dagger \neq (A^\dagger)^2$ .

(c) 我们直接复用上面的矩阵  $A$ . 其非零特征值为 1, 而  $A^\dagger$  的非零特征值为  $1/2$ , 显然 2 不是  $A$  的特征值.

可以验证, 当  $AA^\dagger = A^\dagger A$  时, 上述三种情况均不会发生. □

2

证明. 下面我们证明一个引理:

1. 列满秩矩阵  $X$  的广义逆满足  $X^\dagger X = I_k$ .
2. 行满秩矩阵  $Y$  的广义逆满足  $Y Y^\dagger = I_k$ .

证明: 我们只证明列满秩情况, 行满秩情况可由列满秩情况转置得到. 由于  $X$  列满秩,  $X^\dagger$  可表示为

$$X^\dagger = (X^* X)^{-1} X^*.$$

因此

$$X^\dagger X = (X^* X)^{-1} X^* X = I_k.$$

下面我们证明  $(XY)^\dagger = Y^\dagger X^\dagger$ . 我们直接将其代入 Penrose 方程组验证即可, 不再赘述。□

3

证明. 题中对  $\lambda$  和  $a_{i,i}$  进行了比较. 这似乎表明, 我们是考虑用  $a_{i,i}$  来近似  $\lambda$ . 右端项如果写成矩阵形式, 就是

$$\|(A - a_{i,i}I)e_i\|_2^2 = \|Ae_i - e_i a_{i,i}\|_2^2.$$

这表明, 此处我们是在用  $a_{i,i}$  来近似特征值  $\lambda$ , 而用  $e_i$  来近似特征向量  $x$ , 右端项就是残差的范数平方. 题目要我们证的结论, 就是说特征值的误差可以由残差的误差控制. 事实上, 我们有如下引理:

**引理:** 设  $A \in \mathbb{C}^{n \times n}$  是正规矩阵, 则

$$\min_{\lambda \in \text{spec}(A)} |\lambda - \mu| \leq \|Ax - x\mu\|_2, \quad \forall x \in \mathbb{C}^n, \|x\|_2 = 1, \mu \in \mathbb{C}.$$

证明: 考虑将右端项展开, 得到

$$\|Ax - x\mu\|_2^2 = x^*(A - \mu I)^*(A - \mu I)x.$$

由于  $A$  是正规矩阵,  $A - \mu I$  也是正规矩阵, 因此  $(A - \mu I)^*(A - \mu I)$  是 Hermite 矩阵. 由 Rayleigh 商定理知

$$\|Ax - x\mu\|_2^2 = R(x) \geq \min_{\lambda \in \text{spec}(A)} |\lambda - \mu|^2.$$

取平方根即得所需结论。

回到题目中, 我们直接应用引理, 取  $x = e_i$ ,  $\mu = a_{i,i}$  即可得到所需结论。□

4

证明.  $vv^*$  的特征值满足:

$$0 = \lambda_1(vv^*) = \cdots = \lambda_{n-1}(vv^*) < \lambda_n(vv^*) = \|v\|_2^2.$$

由 Weyl 不等式

$$\lambda_i(A + vv^*) \geq \lambda_i(A) + \lambda_1(vv^*) = \lambda_i(A).$$

原始形式的 Weyl 不等式告诉我们

$$\lambda_i(A + vv^*) \leq \lambda_{i+1}(A) + \lambda_{n-1}(vv^*) = \lambda_{i+1}(A).$$

如果不使用原始形式的 Weyl 不等式，我们可以考虑使用 Courant–Fischer 极小极大定理。由于我们想要的是上界，我们考虑使用

$$\lambda_i(A + vv^*) = \min_{\dim S=i} \max_{x \in S, \|x\|_2=1} x^*(A + vv^*)x.$$

取  $S_{i+1} = \text{span}(u_1, \dots, u_{i+1})$ ，其中  $u_i$  是  $A$  的第  $i$  个单位特征向量。由于我们希望把扰动项的影响降到最低，我们考虑

$$\mathcal{S}_0 = \{S_0 : S_0 \subset S_{i+1}, v^*x = 0 (\forall x \in S_0)\}.$$

显然  $\forall S_0 \in \mathcal{S}_0$ ,  $\dim S_0 \geq i$ . 因此我们可以取一个  $S \in \mathcal{S}_0$  使得  $\dim S = i$ . 由  $S \subset S_{i+1}$  可知

$$\max_{x \in S, \|x\|_2=1} x^*Ax \leq \lambda_{i+1}(A).$$

同时由  $v^*x = 0$  可知  $x^*vv^*x = 0$ . 因此

$$\lambda_i(A + vv^*) \leq \max_{x \in S, \|x\|_2=1} x^*(A + vv^*)x = \max_{x \in S, \|x\|_2=1} x^*Ax \leq \lambda_{i+1}(A).$$

综上，我们得到了所需结论。 □

5

证明. 显然  $S$  是 Schur 补，因此我们考虑下面的合同变换：

$$A = \begin{bmatrix} I & 0 \\ A_{2,1}A_{1,1}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{1,1} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & A_{1,1}^{-1}A_{1,2} \\ 0 & I \end{bmatrix}.$$

由  $A$  正定可知  $S$  也正定。为计算范数，我们考虑  $A$  的逆：

$$A^{-1} = \begin{bmatrix} I & -A_{1,1}^{-1}A_{1,2} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{1,1}^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{2,1}A_{1,1}^{-1} & I \end{bmatrix} = \begin{bmatrix} * & * \\ * & S^{-1} \end{bmatrix}$$

因此  $S^{-1}$  是  $A^{-1}$  的子块。由 Cauchy 交错定理可知  $\|S^{-1}\|_2 \leq \|A\|_2$ .

而由  $S$  的定义我们知道

$$S = A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2} = A_{2,2} - A_{1,2}^*A_{1,1}^{-1}A_{1,2}.$$

由  $A_{1,1}$  正定，有

$$A_{2,2} \succeq S \implies \|S\|_2 \leq \|A_{2,2}\|_2 \leq \|A\|_2.$$

综上，

$$\|S\|_2 \|S^{-1}\|_2 \leq \|A\|_2 \|A^{-1}\|_2.$$

□

这个证法充分利用了 Schur 补的性质。这个证法是助教在课上提到的。我第一次做这道题时，使用了另外一种方法：

证明. 要证  $\|S\|_2\|S^{-1}\|_2 \leq \|A\|_2\|A^{-1}\|_2$ ，我们只需证

$$\frac{\lambda_m(S)}{\lambda_1(S)} \leq \frac{\lambda_n(A)}{\lambda_1(A)}.$$

我们只需证

$$\lambda_1(A) \leq \lambda_1(S), \quad \lambda_m(S) \leq \lambda_n(A).$$

由 Rayleigh 商定理可知

$$\begin{aligned} \lambda_1(A) &= \min_{\|z\|_2=1} \begin{bmatrix} x \\ y \end{bmatrix}^* A \begin{bmatrix} x \\ y \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{C}^n \\ &= \min_{\|z\|_2=1} \begin{bmatrix} x \\ y \end{bmatrix}^* \begin{bmatrix} I & 0 \\ A_{2,1}A_{1,1}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{1,1} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & A_{1,1}^{-1}A_{1,2} \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \min_{\|z\|_2=1} \begin{bmatrix} x + A_{1,1}^{-1}A_{1,2}y \\ y \end{bmatrix}^* \begin{bmatrix} A_{1,1} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} x + A_{1,1}^{-1}A_{1,2}y \\ y \end{bmatrix} \\ &= \min_{\|z\|_2=1} (x + A_{1,1}^{-1}A_{1,2}y)^* A_{1,1} (x + A_{1,1}^{-1}A_{1,2}y) + y^* S y \\ &\leq \min_{\|y\|_2 \leq 1} y^* S y \quad \text{取 } x = -A_{1,1}^{-1}A_{1,2}y \\ &\leq \frac{y^* S y}{y^* y} = \lambda_1(S). \end{aligned}$$

同理可证

$$\lambda_n(A) \geq \lambda_m(S).$$

因此

$$\|S\|_2\|S^{-1}\|_2 \leq \|A\|_2\|A^{-1}\|_2.$$

□

6

证明. 考虑  $A$  的奇异值分解  $A = U\Sigma V^*$ ，则有

$$A^*A = V\Sigma^*\Sigma V^*, \quad AA^* = U\Sigma\Sigma^*U^*.$$

因此

$$(A^*A + \delta I)^{-1}A^* = V(\Sigma^*\Sigma + \delta I)^{-1}\Sigma^*U^* \rightarrow V\Sigma^\dagger U^* = A^\dagger, \quad \text{as } \delta \rightarrow 0.$$

同理

$$A^*(AA^* + \delta I)^{-1} \rightarrow A^\dagger, \quad \text{as } \delta \rightarrow 0.$$

因此

$$\lim_{\delta \rightarrow 0^+} (A^*A + \delta I)^{-1}A^* = \lim_{\delta \rightarrow 0^+} A^*(AA^* + \delta I)^{-1} = A^\dagger.$$

□

7

证明. 考虑矩阵

$$A = \begin{bmatrix} 0 & \frac{3}{2} & \frac{3}{2} \\ \frac{3}{2} & 0 & \frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix}$$

则原问题可化为

$$x^*(-1.5B)x \leq x^*Ax \leq x^*\left(\frac{3+\sqrt{13}}{4}B\right)x, \quad \forall x \in \mathbb{R}^3.$$

我们只需证

$$-\frac{3}{2} \leq \frac{x^*Ax}{x^*Bx} \leq \frac{3+\sqrt{13}}{4}, \quad \forall x \in \mathbb{R}^3 \setminus \{0\}.$$

这实际上是广义特征值的数值域问题。我们只需求解特征方程

$$Ax = Bx\lambda.$$

我们直接考虑行列式

$$\det(A - \lambda B) = 0$$

化简得

$$(2\lambda + 3)(4\lambda^2 - 6\lambda - 1) = 0.$$

因此广义特征值为

$$\lambda_1 = -\frac{3}{2}, \quad \lambda_2 = \frac{3-\sqrt{13}}{4}, \quad \lambda_3 = \frac{3+\sqrt{13}}{4}.$$

因此

$$-\frac{3}{2}(a^2 + b^2 + 2c^2) \leq 3ab + bc + ca \leq \frac{3+\sqrt{13}}{4}(a^2 + b^2 + 2c^2).$$

□