

高等线性代数 Homework 10

Due: Nov. 25, 2024

姓名: 雍崔扬

学号: 21307140051

Problem 1

(Matrix Analysis 2.1 节习题 23)

给定正整数 n 和 $A = [a_1, \dots, a_n] = [a_{ij}]_{i,j=1}^n \in \mathbb{C}^{n \times n}$.

试证明 Hadamard 不等式:

$$|\det(A)|^2 \leq \prod_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \quad (\text{i.e. } |\det(A)| \leq \prod_{i=1}^n \|a_i\|_2^2)$$

Proof:

当 A 为奇异矩阵时, 命题显然成立, 当且仅当 A 的某一列 a_i 为零向量时取等.

当 A 为非奇异矩阵时, 存在唯一的 QR 分解 $A = QR$,

其中 $Q = [q_1, \dots, q_n] \in \mathbb{C}^{n \times n}$ 为酉矩阵,

而 $R = [r_1, \dots, r_n] = [r_{ij}]_{i,j=1}^n \in \mathbb{C}^{n \times n}$ 为具有正实数对角元的上三角阵.

$$A = [a_1, \dots, a_n] = [q_1, \dots, q_n] \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \end{bmatrix} = Q[r_1, \dots, r_n] = QR$$

注意到 $a_i = Qr_i$ ($i = 1, \dots, n$), 因此我们有:

$$\begin{aligned} \|a_i\|_2 &= \|Qr_i\|_2 \quad (\text{note that } l_2 \text{ norm is unitary-invariant}) \\ &= \|r_i\|_2 \\ &\geq r_{ii} \end{aligned} \quad (\forall i = 1, \dots, n)$$

当且仅当 $a_i = r_{ii}q_i$ 时取等.

于是我们有:

$$\begin{aligned} |\det(A)| &= |\det(QR)| \\ &= |\det(Q) \det(R)| \\ &= |\det(Q)| |\det(R)| \quad (\text{note that } \det(Q) = \pm 1) \\ &= 1 \cdot \prod_{i=1}^n r_{ii} \\ &\leq \prod_{i=1}^n \|a_i\|_2 \end{aligned}$$

当且仅当 $a_i = r_{ii}q_i$ ($i = 1, \dots, n$) 时取等, 即 a_1, \dots, a_n 相互正交时取等.

综上所述, Hadamard 不等式 $|\det(A)| \leq \prod_{i=1}^n \|a_i\|_2$ 成立.

当且仅当 a_1, \dots, a_n 相互正交或有某个 a_i 为零向量时取等.

Problem 2

(Matrix Analysis 2.5 节习题 32)

试证明 Euler 旋转轴定理:

若 $A \in \mathbb{R}^{3 \times 3}$ 为正交阵, 且 $\det(A) = 1$, 则存在 $v \in \mathbb{R}^3 \setminus \{0\}$ 使得 $Av = v$.

- **Lemma 1: (Matrix Analysis 定理 2.5.8)**

若 $A \in \mathbb{R}^{n \times n}$ 是正规的, 则下列命题成立:

- ① 存在实正交阵 $Q \in \mathbb{R}^{n \times n}$ 使得 $Q^T A Q$ 为拟对角阵,

每个对角块要么是 1×1 的 (对应 A 的一个实特征值),

要么是 2×2 的, 且具有特殊形式 $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ ($b > 0$) (对应 A 的一对共轭复特征值 $a \pm ib$).

这些对角块由 A 的特征值完全确定, 但可以按照任意预先指定的次序出现.

- ② 两个 n 阶实正规矩阵是实正交相似的, 当且仅当它们具有完全相同的特征值.

- **Lemma 2: (Matrix Analysis 推论 2.5.11)**

设 $A \in \mathbb{R}^{n \times n}$, 则下列命题成立:

- ① $A = A^T$ 当且仅当存在一个实正交阵 $Q \in \mathbb{R}^{n \times n}$ 使得

$$Q^T A Q = \text{diag}\{\lambda_1, \dots, \lambda_n\} \in \mathbb{R}^{n \times n},$$

其中 $\lambda_1, \dots, \lambda_n$ 为 A 的 n 个实特征值.

此外, 两个 n 阶实对称矩阵是实正交相似的, 当且仅当它们具有完全相同的特征值.

- ② $A = -A^T$ 当且仅当存在一个实正交阵 $Q \in \mathbb{R}^{n \times n}$ 和一个非负整数 $p \in \mathbb{N}$ 使得:

$$Q^T A Q = \text{diag}(0_{n-2p}) \oplus b_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus \cdots \oplus b_p \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (\text{where } b_1, \dots, b_p > 0)$$

其中 $\pm ib_1, \dots, \pm ib_p$ 为 A 的 p 对非零共轭复特征值.

此外, 两个 n 阶实反对称矩阵是实正交相似的, 当且仅当它们具有完全相同的特征值.

- ③ $A^T A = I_n$ 当且仅当存在一个实正交阵 $Q \in \mathbb{R}^{n \times n}$ 和一个非负整数 $p \in \mathbb{N}$ 使得:

$$Q^T A Q = \text{diag}\{\lambda_1, \dots, \lambda_{n-2p}\} \oplus \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) \\ -\sin(\theta_1) & \cos(\theta_1) \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \cos(\theta_p) & \sin(\theta_p) \\ -\sin(\theta_p) & \cos(\theta_p) \end{bmatrix},$$

其中 $\lambda_1, \dots, \lambda_{n-2p} \in \{1, -1\}$, 且 $\theta_1, \dots, \theta_p \in (0, \pi)$.

A 的特征值是 $\lambda_1, \dots, \lambda_{n-2p} \in \{1, -1\}$ 加上 $e^{\pm i\theta_1}, \dots, e^{\pm i\theta_p}$.

此外, 两个 n 阶实正交矩阵是实正交相似的, 当且仅当它们具有完全相同的特征值.

Solution: (Reference)

若 $A \in \mathbb{R}^{3 \times 3}$ 为正交阵, 且 $\det(A) = 1$,

则根据 Lemma 2 可知 A 实正交相似于 $1 \oplus \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$ (其中 $\theta \in [0, \pi]$),

即存在实正交阵 $Q = [q_1, q_2, q_3] \in \mathbb{R}^{3 \times 3}$ 使得:

$$A = Q \begin{bmatrix} \cos(\theta) & \sin(\theta) & \\ -\sin(\theta) & \cos(\theta) & \\ & & 1 \end{bmatrix} Q^T$$

取 $v = Q[0, 0, 1]^T = q_3$ (它是 Q 的第 3 列, 显然不为零向量), 则我们有:

$$Av = AQ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = Q \begin{bmatrix} \cos(\theta) & \sin(\theta) & \\ -\sin(\theta) & \cos(\theta) & \\ & & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = Q \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = v$$

命题得证.

Problem 3

给定正整数 n .

设 $A \in \mathbb{C}^{n \times n}$ 为 Hermite 正定阵, $w \in \mathbb{C}^n$ 满足 $w^H A w = 1$.

若 $H = I - 2ww^H A$, 试证明 $H^{-1} = H$, 并求解 H 的所有特征值.

- **Lemma:**

给定单位向量 $w \in \mathbb{C}^n$, 定义 Householder 矩阵 $H := I - 2ww^H$.

可以证明 H 是 Hermite 的酉矩阵:

$$\begin{aligned} H^H &= (I - 2ww^H)^H \\ &= I - 2ww^H \\ &= H \\ \hline H^H H &= H^2 \\ &= (I - 2ww^H)^2 \\ &= I - 4ww^H + 4ww^H w w^H \quad (\text{note that } w^H w = \|w\|^2 = 1) \\ &= I - 4ww^H + 4ww^H \\ &= I \end{aligned}$$

且特征值为 $n - 1$ 个 1 和 1 个 -1 (这是因为 ww^H 的特征值是 $n - 1$ 个 0 和 1 个 1).

Solution:

注意到 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 正定阵(自然是正规矩阵), 因此它可酉对角化, 且特征值均为正实数. 即存在酉矩阵 $U \in \mathbb{C}^{n \times n}$ 和对角元为正实数的对角阵 $\Lambda := \text{diag}\{\lambda_1, \dots, \lambda_n\}$ 使得 $A = U\Lambda U^H$. 我们定义:

$$\begin{aligned} \Lambda^{\frac{1}{2}} &:= \text{diag}\{\lambda_1^{\frac{1}{2}}, \dots, \lambda_n^{\frac{1}{2}}\} \\ A^{\frac{1}{2}} &:= U\Lambda^{\frac{1}{2}}U^H \end{aligned}$$

显然 $A^{\frac{1}{2}}$ 也是 Hermite 正定阵.

于是我们有 $\|A^{\frac{1}{2}}w\|^2 = w^H A^{\frac{1}{2}} A^{\frac{1}{2}} w = w^H A w = 1$ 成立.

定义 Householder 矩阵 $H_0 := I_n - 2(A^{\frac{1}{2}}w)(A^{\frac{1}{2}}w)^H = I_n - 2A^{\frac{1}{2}}ww^H A^{\frac{1}{2}}$.

根据 Lemma 可知 H_0 是 Hermite 的酉矩阵, 且特征值为 $n - 1$ 个 1 和 1 个 -1 .

注意到:

$$\begin{aligned} H &= I - 2ww^H A \\ &= A^{-\frac{1}{2}}(I - 2A^{\frac{1}{2}}ww^H A^{\frac{1}{2}})A^{\frac{1}{2}} \\ &= A^{-\frac{1}{2}}H_0A^{\frac{1}{2}} \end{aligned}$$

因此 H 与 H_0 相似(故特征值为 $n - 1$ 个 1 和 1 个 -1).

注意到:

$$\begin{aligned} H^2 &= (A^{-\frac{1}{2}}H_0A^{\frac{1}{2}})(A^{-\frac{1}{2}}H_0A^{\frac{1}{2}}) \\ &= A^{-\frac{1}{2}}H_0^2A^{\frac{1}{2}} \quad (\text{note that } H_0^2 = I_n) \\ &= A^{-\frac{1}{2}}I_nA^{\frac{1}{2}} \\ &= I_n \end{aligned}$$

因此 $H^{-1} = H$, 命题得证.

Problem 4

给定正整数 n 以及 n 阶复方阵 A .

试证明以下命题等价:

- ① A 是正规矩阵, 即 $AA^H = A^H A$
- ② A 可表示为 $A = X + iY$, 其中 X, Y 是 Hermite 阵且 $XY = YX$
- ③ $\|A\|_F^2 = \sum_{k=1}^n |\lambda_k|^2$ (其中 $\lambda_1, \dots, \lambda_n$ 是 A 的特征值)
- ④ 存在复数域上的多项式 $f(\lambda)$ 使得 $A^H = f(A)$
- ⑤ A 的任意特征向量都是 A^H 的特征向量

Proof:

$$\textcircled{1} \Leftrightarrow \textcircled{2}$$

任意复方阵 $A \in \mathbb{C}^{n \times n}$ 都可以唯一地分解为 $A = H(A) + K(A)$

其中 $H(A) = \frac{1}{2}(A + A^H)$ 称为 A 的 **Hermite 部分**, 而 $K(A) = \frac{1}{2}(A - A^H)$ 称为 A 的**反 Hermite 部分**.

通常我们会从 $K(A)$ 中提取一个 i 出来, 就得到 A 的 **Toeplitz 分解**:

$$A = H(A) + iK(A) \text{ where } \begin{cases} H(A) = \frac{1}{2}(A + A^H) \\ K(A) = \frac{1}{2i}(A - A^H) \end{cases}$$

- $\textcircled{1} \Rightarrow \textcircled{2}$

设 A 是正规矩阵, 取 $\begin{cases} X := \frac{1}{2}(A + A^H) \\ Y := \frac{1}{2i}(A - A^H) \end{cases}$ 就得到 A 的 Toeplitz 分解 $A = X + iY$

$$\begin{aligned} X + iY &= \frac{1}{2}(A + A^H) + i \cdot \frac{1}{2i}(A - A^H) \\ &= \frac{1}{2}(A + A^H) + \frac{1}{2}(A - A^H) \\ &= A \end{aligned}$$

$$\begin{aligned} X^H &= \left[\frac{1}{2}(A + A^H) \right]^H \\ &= \frac{1}{2}(A^H + A) \\ &= X \end{aligned}$$

$$\begin{aligned} Y^H &= \left[\frac{1}{2i}(A - A^H) \right]^H \\ &= -\frac{1}{2i}(A^H - A) \\ &= Y \end{aligned}$$

$$\begin{aligned} XY &= \frac{1}{2}(A + A^H) \cdot \frac{1}{2i}(A - A^H) \\ &= \frac{1}{4i}(A^2 + A^H A - A A^H - (A^H)^2) \quad (\text{note that } A^H A = A A^H) \\ &= \frac{1}{4i}(A^2 - (A^H)^2) \\ &= \frac{1}{4i}(A^2 - A^H A + A A^H - (A^H)^2) \\ &= \frac{1}{2i}(A - A^H) \cdot \frac{1}{2}(A + A^H) \\ &= YX \end{aligned}$$

因此 X, Y 是 Hermite 阵且 $XY = YX$.

- $\textcircled{1} \Leftarrow \textcircled{2}$

设 A 可表示为 $A = X + iY$, 其中 X, Y 是 Hermite 阵且 $XY = YX$.

则我们有:

$$\begin{aligned}
A^H A &= (X + iY)^H (X + iY) \\
&= (X^H - iY^H)(X + iY) \quad (\text{note that } X^H = X, Y^H = Y) \\
&= (X - iY)(X + iY) \\
&= X - iYX + iXY - i^2Y^2 \quad (\text{note that } YX = XY) \\
&= X - i^2Y^2 \\
&= X + iYX - iXY - i^2Y^2 \\
&= (X + iY)(X - iY) \\
&= (X + iY)(X + iY)^H \\
&= AA^H
\end{aligned}$$

这表明 A 是正规矩阵.

① \Leftrightarrow ③

(Schur 分解定理的推论)

正规矩阵 $A \in \mathbb{C}^{n \times n}$ (满足 $A^H A = AA^H$, 即 A 与 A^H 可交换) 可以酉对角化, 且对角元的次序可以任意指定.

反过来, 可酉对角化的矩阵 $A \in \mathbb{C}^{n \times n}$ 也一定是正规矩阵.

因此正规矩阵和可酉对角化的矩阵是等价的.

- ① \Rightarrow ③

若 A 是正规矩阵, 则 A 可酉对角化,

即存在酉矩阵 $U \in \mathbb{C}^{n \times n}$ 和对角阵 $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ 使得 $A = U\Lambda U^H$.

于是我们有:

$$\begin{aligned}
\|A\|_F^2 &= \text{tr}(A^H A) \\
&= \text{tr}((U\Lambda U^H)^H (U\Lambda U^H)) \\
&= \text{tr}(U\bar{\Lambda}U^H U\Lambda U^H) \quad (\text{where } \bar{\Lambda} = \text{diag}\{\bar{\lambda}_1, \dots, \bar{\lambda}_n\}) \\
&= \text{tr}(U\bar{\Lambda}\Lambda U^H) \\
&= \text{tr}(\bar{\Lambda}\Lambda U^H U) \\
&= \text{tr}(\bar{\Lambda}\Lambda) \\
&= \sum_{i=1}^n |\lambda_i|^2
\end{aligned}$$

- ① \Leftarrow ③

设 $\|A\|_F^2 = \sum_{i=1}^n |\lambda_i|^2$ (其中 $\lambda_1, \dots, \lambda_n$ 是 A 的特征值).

设 A 的 Schur 分解为 $A = UTU^H$,

其中 $U \in \mathbb{C}^{n \times n}$ 为酉矩阵, 上三角阵 $T = [t_{ij}] \in \mathbb{C}^{n \times n}$ 的对角元为 $\lambda_1, \dots, \lambda_n$.

于是我们有:

$$\begin{aligned}
\|A\|_F^2 &= \text{tr}(A^H A) \\
&= \text{tr}((UTU^H)^H(UTU^H)) \\
&= \text{tr}(UT^H U^H U T U^H) \\
&= \text{tr}(UT^H T U^H) \\
&= \text{tr}(T^H T U^H) \\
&= \text{tr}(T^H T) \\
&= \|T\|_F^2 \\
&= \sum_{i=1}^n |\lambda_i|^2 + \sum_{i < j} |t_{ij}|^2 \\
&= \sum_{i=1}^n |\lambda_i|^2
\end{aligned}$$

因此我们有 $\sum_{i < j} |t_{ij}|^2 = 0$, 说明 T 为对角阵.

故 A 的 Schur 分解 $A = UTU^H$ 实际上就是 A 的谱分解

$$A = U \text{diag}\{\lambda_1, \dots, \lambda_n\} U^H = U \Lambda U^H.$$

于是我们有:

$$\begin{aligned}
A^H A &= (U \Lambda U^H)^H (U \Lambda U^H) \\
&= U \bar{\Lambda} U^H U \Lambda U^H \\
&= U \bar{\Lambda} \Lambda U^H \\
&= U \Lambda \bar{\Lambda} U^H \\
&= (U \Lambda U^H)(U \bar{\Lambda} U^H) \\
&= (U \Lambda U^H)(U \Lambda U^H)^H \\
&= AA^H
\end{aligned}$$

这意味着 A 是正规矩阵.

① \Leftrightarrow ④

Lagrange 插值多项式: ([reference](#))

给定 $(x_1, y_1), \dots, (x_n, y_n)$ (设 x_1, \dots, x_n 互不相同)

一个不超过 $n - 1$ 次多项式 $p(x)$ 构造如下:

$$p(x) := \sum_{i=1}^n y_i p_i(x) \text{ where } p_i(x) := \prod_{1 \leq j \leq n, j \neq i} \frac{x - x_j}{x_i - x_j} \quad (\forall i = 1, \dots, n)$$

可以证明:

$$\begin{aligned}
p_i(x_j) &= \begin{cases} 0 & \text{if } j \neq i \\ 1 & \text{if } j = i \end{cases} \quad (\forall i = 1, \dots, n) \\
p(x_i) &= y_i \quad (\forall i = 1, \dots, n)
\end{aligned}$$

- ① \Rightarrow ④

若 A 是正规矩阵, 则存在谱分解 $A = U \Lambda U^H$.

设 A 的互不相同的特征值为 $\lambda_1, \dots, \lambda_d$, 代数重数分别为 n_1, \dots, n_d (满足 $n_1 + \dots + n_d = n$).

定义 Lagrange 插值多项式 (不超过 $d - 1$ 次):

$$p(\lambda) := \sum_{i=1}^d \bar{\lambda}_i p_i(\lambda) \text{ where } p_i(\lambda) := \prod_{1 \leq j \leq d, j \neq i} \frac{\lambda - \lambda_j}{\bar{\lambda}_i - \bar{\lambda}_j} \quad (\forall i = 1, \dots, d)$$

则我们有:

$$p_i(\lambda_j) = \begin{cases} 0 & \text{if } j \neq i \\ 1 & \text{if } j = i \end{cases} (\forall i = 1, \dots, d)$$

$$p(\lambda_i) = \bar{\lambda}_i (\forall i = 1, \dots, d)$$

因此我们有:

$$\begin{aligned} p(A) &= p(U\Lambda U^H) \\ &= Up(\Lambda)U^H \\ &= U\bar{\Lambda}U^H \\ &= (U\Lambda U^H)^H \\ &= A^H \end{aligned}$$

- ① \Leftarrow ④

设 A 的 Schur 分解为 $A = UTU^H$.

其中 $U \in \mathbb{C}^{n \times n}$ 为酉矩阵, 上三角阵 $T = [t_{ij}] \in \mathbb{C}^{n \times n}$ 的对角元为 $\lambda_1, \dots, \lambda_n$.

若存在复数域上的多项式 $p(\lambda)$ 使得 $A^H = p(A)$, 则我们有:

$$\begin{aligned} p(A) &= p(UTU^H) \\ &= Up(T)U^H \quad (\text{note that } p(A) = A^H) \\ &= A^H \\ &= (UTU^H)^H \\ &= UT^H U^H \end{aligned}$$

于是我们有 $Up(T)U^H = UT^H U^H$ 成立, 即有 $p(T) = T^H$ 成立.

注意到 $p(T)$ 是上三角阵, 而 T^H 为下三角阵, 故 T 为对角阵.

故 A 的 Schur 分解 $A = UTU^H$ 实际上就是 A 的谱分解

$$A = U\text{diag}\{\lambda_1, \dots, \lambda_n\}U^H = U\Lambda U^H.$$

这意味着 A 是正规矩阵.

① \Leftrightarrow ⑤

- ① \Rightarrow ⑤

若 A 是正规矩阵, 则存在谱分解 $A = U\Lambda U^H$.

注意到 A^H 的谱分解即为 $A^H = U\bar{\Lambda}U^H$.

上述两式可以等价表示为:

$$AU = U\Lambda$$

$$A^H U = U\bar{\Lambda}$$

这表明 U 的 n 个列向量都是 A, A^H 的公共特征向量.

因此 A 的任意特征向量都是 A^H 的特征向量.

- ① \Leftarrow ⑤

(Matrix Analysis 2.5 节习题 34) ([reference](#))

设 A 的 Schur 分解为 $A = UTU^H$.

其中 $U = [u_1, \dots, u_n] \in \mathbb{C}^{n \times n}$ 为酉矩阵, 上三角阵 $T = [t_{ij}] \in \mathbb{C}^{n \times n}$ 的对角元为 $\lambda_1, \dots, \lambda_n$

注意到 u_1 是 A 关于 λ_1 的特征向量.

若 A 的任意特征向量都是 A^H 的特征向量, 则我们有:

$$\begin{aligned}
T &= U^H A U \\
&= [u_i^H A u_j]_{i,j=1}^n \\
&= \left[\begin{array}{c|ccc} u_1^H A u_1 & u_1^H A u_2 & \cdots & u_1^H A u_n \\ \hline u_2^H A u_1 & u_2^H A u_2 & \cdots & u_2^H A u_n \\ \vdots & \vdots & & \vdots \\ u_n^H A u_1 & u_n^H A u_2 & \cdots & u_n^H A u_n \end{array} \right] \quad (\text{note that } A u_1 = u_1 \lambda_1 \text{ and } u_1^H A = \lambda_1 u_1^H) \\
&= \left[\begin{array}{c|ccc} \lambda_1 u_1^H u_1 & \lambda_1 u_1^H u_2 & \cdots & \lambda_1 u_1^H u_n \\ \hline \lambda u_2^H u_1 & u_2^H A u_2 & \cdots & u_2^H A u_n \\ \vdots & \vdots & & \vdots \\ \lambda u_n^H u_1 & u_n^H A u_2 & \cdots & u_n^H A u_n \end{array} \right] \\
&= \left[\begin{array}{c|cccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & u_2^H A u_2 & \cdots & u_2^H A u_n \\ \vdots & \vdots & & \vdots \\ 0 & u_n^H A u_2 & \cdots & u_n^H A u_n \end{array} \right]
\end{aligned}$$

对比 Schur 型 T 的结构, 我们知道 T 的第 1 行的非对角元均为零, 且对于第 2 列有:

$$\begin{cases} u_2^H A u_2 = \lambda_2 \\ u_i^H A u_2 = 0 \ (\forall i \in \{1, \dots, n\} \setminus \{2\}) \end{cases}$$

无论 λ_2 是否为零, 我们都有 $A u_2 \perp \text{span}\{u_1, u_3, \dots, u_n\} = \text{span}\{u_2\}^\perp$.

因此 $A u_2 \in \text{span}\{u_2\}$, 表明 u_2 是 A 的特征向量.

根据 $u_2^H A u_2 = \lambda_2$ 可知 u_2 是 A 关于 λ_2 的特征向量, 故有:

$$\begin{aligned}
u_i^H A u_2 &= \lambda_2 u_i^H u_2 = \begin{cases} \lambda_2 & \text{if } i = 2 \\ 0 & \text{if } i = 3, \dots, n \end{cases} \\
u_2^H A u_j &= \lambda_2 u_2^H u_j = \begin{cases} \lambda_2 & \text{if } j = 2 \\ 0 & \text{if } j = 3, \dots, n \end{cases}
\end{aligned}$$

于是 Schur 型 T 的第 2 行的非对角元均为零.

以此类推, 可知 Schur 型 $T = \text{diag}\{\lambda_1, \dots, \lambda_n\}$.

故 A 的 Schur 分解 $A = U T U^H$ 实际上就是 A 的谱分解

$A = U \text{diag}\{\lambda_1, \dots, \lambda_n\} U^H = U \Lambda U^H$.

这意味着 A 是正规矩阵.

(失败的尝试)

设 A 的任意特征向量都是 A^H 的特征向量.

于是对于 A 的任意特征对 (λ, x) 都有:

$$\begin{cases} Ax = x\lambda \\ A^H x = x\bar{\lambda} \Leftrightarrow x^H A = \lambda x^H \end{cases}$$

这表明 A 的任意特征向量都是 A (对应特征值) 的左特征向量, 即 A 的任意特征对 (λ, x) 都是正规的.

现考虑 A 的不同特征值的特征对 $(\lambda, x), (\mu, y)$ (其中 $\lambda \neq \mu$)

利用共轭算子的性质我们有:

$$\begin{aligned}
(\lambda - \mu)\langle x, y \rangle &= \langle \lambda x, y \rangle - \langle x, \bar{\mu}y \rangle \\
&= \langle Ax, y \rangle - \langle x, A^H y \rangle \\
&= 0
\end{aligned}$$

根据 $\lambda \neq \mu$ 可知 $\langle x, y \rangle = 0$, 这表明 A 的不同特征值对应的特征向量相互正交.

当 A 的特征值互不相同时, 它一定具有 n 个互不相同的特征向量, 因而存在谱分解.
此时可得 A 是正规矩阵.

(存疑) 当 A 具有重特征值时, 它可由方阵序列 $\{A_n\}$ 逼近.

其中 A_n ($\forall n \in \mathbb{Z}_+$) 只有单特征值, 且 A_n^H 和 A_n 具有相同的特征向量 (因而是正规矩阵,
存在谱分解 $A_n = U_n \Lambda_n U_n^H$)

此时可得 A 是正规矩阵.

(疑点: 有重特征值时能保证 A 有 n 个线性无关的特征向量吗?

此外, 能保证相同特征值的线性无关的特征向量相互正交吗?

因为我们的目标是证明 A 有 n 个相互正交的特征向量)

Problem 5

(单边 Jacobi 迭代)

计算如下矩阵的奇异值分解:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 0 & -2 & 1 \end{bmatrix}.$$

• Lemma:

我们可以这样将两个线性相关的向量 $x, \alpha x \in \mathbb{R}^n$ 正交化:

$$[x \quad \alpha x] \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = [\sqrt{\alpha^2 + 1}x \quad 0_n] \text{ where } \begin{cases} c = \frac{1}{\sqrt{\alpha^2 + 1}} \\ s = \frac{-\alpha}{\sqrt{\alpha^2 + 1}} \end{cases}$$

Solution:

注意到 A 的列向量组中, 第 1, 2, 3 列相互正交, 而第 4 列与第 1 列线性相关.

单边 Jacobi 迭代的思想指导我们使用一个正交变换 $Q_{1,4}$ 将第 4 列化为零:

$$Q_{1,4} := \begin{bmatrix} \frac{\sqrt{2}}{2} & & & -\frac{\sqrt{2}}{2} \\ & 1 & & \\ & & 1 & \\ \frac{\sqrt{2}}{2} & & & \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$AQ_{1,4} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & & & -\frac{\sqrt{2}}{2} \\ & 1 & & \\ & & 1 & \\ \frac{\sqrt{2}}{2} & & & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 1 & 1 & 0 \\ \sqrt{2} & -1 & 1 & 0 \\ \sqrt{2} & 0 & -2 & 0 \end{bmatrix}$$

这样 $AQ_{1,4}$ 的列向量就相互正交了.

此时可取:

$$U := \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & -\frac{\sqrt{6}}{3} \end{bmatrix}$$

$$\Sigma := \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix}$$

$$V := Q_{1,4} = \begin{bmatrix} \frac{\sqrt{2}}{2} & & -\frac{\sqrt{2}}{2} \\ & 1 & \\ & & 1 \\ \frac{\sqrt{2}}{2} & & \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$A = U\Sigma V^T$$

为保证奇异值从大到小排列，我们可以交换 Σ 的 $(2, 2), (3, 3)$ 位置，并对应交换 U 的 2, 3 列和 V 的 2, 3 行：

$$U := \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{3} & 0 \end{bmatrix}$$

$$\Sigma := \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{6} \\ \sqrt{2} & 0 \end{bmatrix}$$

$$V := \begin{bmatrix} \frac{\sqrt{2}}{2} & & -\frac{\sqrt{2}}{2} \\ 0 & 1 & \\ 1 & 0 & \\ \frac{\sqrt{2}}{2} & & \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$A = U\Sigma V^T$$

Matlab 代码验证：

```
A = [1, 1, 1, 1;
      1, -1, 1, 1;
      1, 0, -2, 1];

U = [sqrt(3)/3, sqrt(6)/6, sqrt(2)/2;
      sqrt(3)/3, sqrt(6)/6, -sqrt(2)/2;
      sqrt(3)/3, -sqrt(6)/3, 0];

Sigma = [sqrt(6), 0, 0, 0;
          0, sqrt(6), 0, 0;
          0, 0, sqrt(2), 0];

V = [sqrt(2)/2, 0, 0, -sqrt(2)/2;
      0, 0, 1, 0;
      0, 1, 0, 0;
      sqrt(2)/2, 0, 0, sqrt(2)/2];

disp("U * Sigma * V' :");
disp(U * Sigma * V');
```

运行结果:

```
U * Sigma * V':  
1.0000 1.0000 1.0000 1.0000  
1.0000 -1.0000 1.0000 1.0000  
1.0000 0 -2.0000 1.0000
```

Problem 6 (optional)

(Euler's Fixed Point Theorem.pdf) & (Rodrigues' Rotation Formula) & (Matrix Computation 习题 6.4.7)

若 $a = [a_1, a_2, a_3]^T$ 是 \mathbb{R}^3 中的单位向量.

试证明绕 a 逆时针旋转 θ 的旋转变换为:

$$Q = \begin{bmatrix} a_1^2(1 - \cos(\theta)) + \cos(\theta) & a_1a_2(1 - \cos(\theta)) - a_3 \sin(\theta) & a_1a_3(1 - \cos(\theta)) + a_2 \sin(\theta) \\ a_1a_2(1 - \cos(\theta)) + a_3 \sin(\theta) & a_2^2(1 - \cos(\theta)) + \cos(\theta) & a_2a_3(1 - \cos(\theta)) - a_1 \sin(\theta) \\ a_1a_3(1 - \cos(\theta)) - a_2 \sin(\theta) & a_2a_3(1 - \cos(\theta)) + a_1 \sin(\theta) & a_3^2(1 - \cos(\theta)) + \cos(\theta) \end{bmatrix}$$

- **Background (叉积)**

定义 $a \in \mathbb{R}^3$ 的叉积算子:

$$[a]_{\times} := \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

可以证明对于任意与 a 线性无关的 $b \in \mathbb{R}^3$ 都有 $[a]_{\times} b$ 垂直于 $\text{span}\{a, b\}$.

具体来说 $[a]_{\times} b = \|a\| \|b\| \sin(\angle(a, b)) \cdot n$,

其中 $\angle(a, b) \in [0, \pi]$, 单位向量 n 的方向由右手法则确定.

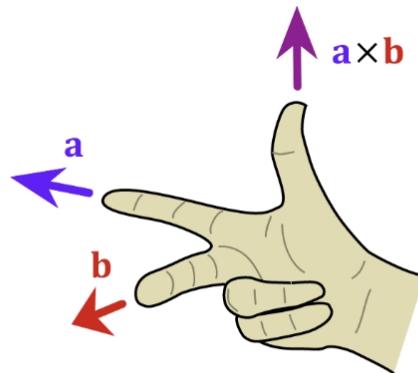


图3 通过右手法则确定叉积方向

Solution:

注意到 a 是单位向量, 因此 $\text{span}\{a\}$ 的正投影算子为 aa^T , 而 $\text{span}\{a\}^\perp$ 的正投影算子为 $I - aa^T$.
考虑向量 $x \in \mathbb{R}^3$

- ① 平行于旋转轴 a 的部分: $x_{\parallel} = aa^T x$
- ② 垂直于旋转轴 a 的部分: $x_{\perp} = (I - aa^T)x$

当向量 x 绕旋转轴 a 逆时针旋转 θ 时, x_{\parallel} 保持不变, 而 x_{\perp} 发生改变.

具体来说 x_{\perp} 变为 $\cos(\theta)x_{\perp} + \sin(\theta)[a]_{\times}x$.

因此我们有:

$$\begin{aligned}x_{\text{rotate}} &= x_{\parallel} + \cos(\theta)x_{\perp} + \sin(\theta)[a]_{\times}x \\&= (aa^T)x + \cos(\theta)(I - aa^T)x + \sin(\theta)[a]_{\times}x\end{aligned}$$

于是对应的线性算子为:

$$\begin{aligned}Q &= aa^T + \cos(\theta)(I - aa^T) + \sin(\theta)[a]_{\times} \\&= (1 - \cos(\theta))aa^T + \cos(\theta)I + \sin(\theta)[a]_{\times} \\&= (1 - \cos(\theta)) \begin{bmatrix} a_1^2 & a_1a_2 & a_1a_3 \\ a_2a_1 & a_2^2 & a_2a_3 \\ a_3a_1 & a_3a_2 & a_3^2 \end{bmatrix} + \cos(\theta) \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} + \sin(\theta) \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \\&= \begin{bmatrix} a_1^2(1 - \cos(\theta)) + \cos(\theta) & a_1a_2(1 - \cos(\theta)) - a_3 \sin(\theta) & a_1a_3(1 - \cos(\theta)) + a_2 \sin(\theta) \\ a_1a_2(1 - \cos(\theta)) + a_3 \sin(\theta) & a_2^2(1 - \cos(\theta)) + \cos(\theta) & a_2a_3(1 - \cos(\theta)) - a_1 \sin(\theta) \\ a_1a_3(1 - \cos(\theta)) - a_2 \sin(\theta) & a_2a_3(1 - \cos(\theta)) + a_1 \sin(\theta) & a_3^2(1 - \cos(\theta)) + \cos(\theta) \end{bmatrix}\end{aligned}$$

Problem 7 (optional)

(极分解的变形)

试证明任意复方阵都可以分解为一个正投影矩阵和一个非奇异矩阵的乘积.

- 直观来说, 有限维 Euclid 空间上的一个线性变换 A 可以分解为:
先作用一个同构 (由非奇异矩阵表征), 再作用一个 $\text{Range}(A)$ 的 (正) 投影算子.

Proof:

设 A 的奇异值分解为:

$$A = U\Sigma V^H = [U_1, U_2] \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{bmatrix} [V_1, V_2]^H,$$

其中 $r = \text{rank}(A)$, 而 $U_1, V_1 \in \mathbb{C}^{n \times r}$ 和 $U_2, V_2 \in \mathbb{C}^{n \times (n-r)}$ 列标准正交.

定义 A 的 Moore-Penrose 伪逆 $A^\dagger := V_1^H \Sigma_r^{-1} U_1$, 它满足 Penrose 方程组:

$$\begin{cases} AA^\dagger A = A \\ A^\dagger AA^\dagger = A^\dagger \\ (AA^\dagger)^H = AA^\dagger \\ (A^\dagger A)^H = A^\dagger A \end{cases}$$

根据 Penrose 方程组可知 AA^\dagger 自伴且幂等, 因而是正交投影算子, 且投影空间为 $\text{Range}(A)$.

注意到 $AA^\dagger = (U_1 \Sigma_r V_1^H)(V_1 \Sigma_r^{-1} U_1^H) = U_1 U_1^H$.

$$\begin{aligned}A &= U\Sigma V^H \\&= [U_1, U_2] \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{bmatrix} [V_1, V_2]^H \\&= [U_1, U_2] \begin{bmatrix} I_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{bmatrix} \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & I_{n-r} \end{bmatrix} [V_1, V_2]^H \\&= [U_1, U_2] \begin{bmatrix} I_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (n-r)} \end{bmatrix} [U_1, U_2]^H [U_1, U_2] \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & I_{n-r} \end{bmatrix} [V_1, V_2]^H \\&= (U_1 U_1^H) \begin{bmatrix} U_1 \Sigma_r V_1^H & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & U_2 V_2^H \end{bmatrix}\end{aligned}$$

这样就表示为一个正交投影算子 $U_1 U_1^H$ 和一个非奇异矩阵 $U_1 \Sigma_r V_1^H \oplus U_2 V_2^H$ 的乘积.

(失败的尝试)

定义 A 的 Moore-Penrose 伪逆 $A^\dagger := V_1^H \Sigma_r^{-1} U_1$, 它满足 Penrose 方程组:

$$\begin{cases} AA^\dagger A = A \\ A^\dagger AA^\dagger = A^\dagger \\ (AA^\dagger)^H = AA^\dagger \\ (A^\dagger A)^H = A^\dagger A \end{cases}$$

根据 Penrose 方程组可知 AA^\dagger 自伴且幂等, 因而是正交投影算子, 且投影空间为 $\text{Range}(A)$.

注意到 $AA^\dagger = (U_1 \Sigma_r V_1^H)(V_1 \Sigma_r^{-1} U_1^H) = U_1 U_1^H$.

记 $I_n = [e_1, \dots, e_n]$ 为 \mathbb{C}^n 默认的标准正交基构成的基矩阵.

注意到 U 的列向量构成 \mathbb{C}^n 的一组标准正交基,

因此存在非奇异矩阵 $B \in \mathbb{C}^{n \times n}$ 使得 $U = EB$.

于是任意 $x \in \mathbb{C}^n$ 在标准正交基 U 下的坐标为 Bx .

Problem 8 (optional)

考虑闭区间 $[0, 1]$ 上的连续实值函数构成的线性空间 $C([0, 1])$.

在 $C([0, 1])$ 上定义内积:

(注意 $C([0, 1])$ 在此内积(诱导的范数)下不是完备的, 因此不是 Hilbert 空间)

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt \quad (\forall f, g \in C([0, 1]))$$

试证明 \mathcal{V} 中存在序列 $\{e_n\}$ 使得 $\langle e_i, e_j \rangle = \delta_{i,j}$ (其中 $\delta_{i,j} = \mathbb{1}(i=j)$ 为 Kronecker 记号).

- 实数域上的标准正交 Schauder 基为 $\{1, \sqrt{2} \cos(2\pi t), \sqrt{2} \sin(2\pi t), \sqrt{2} \cos(4\pi t), \sqrt{2} \sin(4\pi t), \dots\}$
- 复数域上的标准正交 Schauder 基为 $\{1, e^{-i2\pi t}, e^{i2\pi t}, e^{-i4\pi t}, e^{i4\pi t}, \dots\}$

Proof:

考虑序列 $\{e_n(t) = \sqrt{2} \cos(2\pi nt) : n \in \mathbb{Z}_+\}$

$$\begin{aligned} \langle e_n, e_m \rangle &= \int_0^1 e_n(t)e_m(t)dt \\ &= \int_0^1 \sqrt{2} \cos(2\pi nt) \sqrt{2} \cos(2\pi mt)dt \\ &= \int_0^1 \cos(2\pi(n+m)t)dt + \int_0^1 \cos(2\pi(n-m)t)dt \end{aligned}$$

注意到对于任意 $k \in \mathbb{Z}$ 我们都有:

$$\int_0^1 \cos(2\pi kt)dt = \begin{cases} \int_0^1 1 dt = 1 & \text{if } k = 0 \\ \frac{1}{2\pi k} \sin(2\pi kt)|_{t=0}^{t=1} = 0 & \text{if } k \neq 0 \end{cases}$$

于是我们有:

$$\begin{aligned} \langle e_n, e_m \rangle &= \int_0^1 \cos(2\pi(n+m)t)dt + \int_0^1 \cos(2\pi(n-m)t)dt \\ &= 0 + \delta_{n,m} \\ &= \delta_{n,m} \end{aligned}$$

考虑序列 $\{f_n(t) = \sqrt{2} \sin(2\pi nt) : n \in \mathbb{Z}_+\}$

$$\begin{aligned}
\langle f_n, f_m \rangle &= \int_0^1 f_n(t) f_m(t) dt \\
&= \int_0^1 \sqrt{2} \sin(2\pi nt) \sqrt{2} \sin(2\pi mt) dt \\
&= \int_0^1 \cos(2\pi(n-m)t) dt - \int_0^1 \cos(2\pi(n+m)t) dt
\end{aligned}$$

注意到对于任意 $k \in \mathbb{Z}$ 我们都有:

$$\int_0^1 \cos(2\pi kt) dt = \begin{cases} \int_0^1 1 d\theta = 1 & \text{if } k = 0 \\ \frac{1}{2\pi k} \sin(2\pi kt) \Big|_{t=0}^{t=1} = 0 & \text{if } k \neq 0 \end{cases}$$

于是我们有:

$$\begin{aligned}
\langle f_n, f_m \rangle &= \int_0^1 \cos(2\pi(n-m)t) dt - \int_0^1 \cos(2\pi(n+m)t) dt \\
&= \delta_{n,m} - 0 \\
&= \delta_{n,m}
\end{aligned}$$

考虑序列 $\{e_n(t) = \sqrt{2} \cos(2\pi nt) : n \in \mathbb{Z}_+\}$ 和 $\{f_n(t) = \sqrt{2} \sin(2\pi nt) : n \in \mathbb{Z}_+\}$

$$\begin{aligned}
\langle e_n, f_m \rangle &= \int_0^1 e_n(t) f_m(t) dt \\
&= \int_0^1 \sqrt{2} \cos(2\pi nt) \sqrt{2} \sin(2\pi mt) dt \\
&= \int_0^1 \sin(2\pi(n+m)t) dt + \int_0^1 \sin(2\pi(n-m)t) dt
\end{aligned}$$

注意到对于任意 $k \in \mathbb{Z}$ 我们都有:

$$\int_0^1 \sin(2\pi kt) dt = \begin{cases} \int_0^1 1 d\theta = 1 & \text{if } k = 0 \\ -\frac{1}{2\pi k} \cos(2\pi kt) \Big|_{t=0}^{t=1} = 0 & \text{if } k \neq 0 \end{cases}$$

于是我们有:

$$\begin{aligned}
\langle e_n, f_m \rangle &= \int_0^1 \sin(2\pi(n+m)t) dt + \int_0^1 \cos(2\pi(n-m)t) dt \\
&= 0 + \delta_{n,m} \\
&= \delta_{n,m}
\end{aligned}$$

The End