

# 高等线性代数 Homework 06

Due: Oct. 21, 2024

姓名: 雍崔扬

学号: 21307140051

## Problem 1

构造可逆的线性变换

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = C \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix},$$

将下面关于  $\alpha_1, \alpha_2, \alpha_3$  的二次型化为关于  $\beta_1, \beta_2, \beta_3$  的对角二次型:

$$f(\alpha_1, \alpha_2, \alpha_3) = \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_3.$$

**Solution:**

首先我们有:

$$\begin{aligned} f(\alpha_1, \alpha_2, \alpha_3) &= \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_3 \\ &= \frac{1}{2} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}^T \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}. \end{aligned}$$

注意到对称分块矩阵的**对称行列 Gauss 消元**可以表示为:

$$\begin{bmatrix} A_{11} & A_{21}^T \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & \\ & A_{22} - A_{21}A_{11}^{-1}A_{21}^T \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{21}^T \\ & I \end{bmatrix}.$$

于是我们有:

$$\begin{aligned} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}^T \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \left( \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}^T \right) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}^T \\ &= \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}^T. \end{aligned}$$

取:

$$C^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix}^T = \frac{1}{2} \begin{bmatrix} 1 & 1 & 2 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

即:

$$C = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix},$$

则我们有:

$$\begin{aligned}
f(\alpha_1, \alpha_2, \alpha_3) &= \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_3 \\
&= \frac{1}{2} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}^T \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}^T \left( \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 2 \end{bmatrix}^T \right) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \\
&= \left( C^{-1} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \right)^T \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix} \left( C^{-1} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \right) \\
&= \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}^T \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \\
&= \beta_1^2 - \beta_2^2 - \beta_3^2.
\end{aligned}$$

上述转换通过以下线性变换完成:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = C \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \beta_1 - \beta_2 - \beta_3 \\ \beta_1 + \beta_2 - \beta_3 \\ \beta_3 \end{bmatrix}.$$

## Problem 2

已知:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 19 & 31 \\ 1 & 5 & 15 & 31 & 53 \end{bmatrix}$$

试构造一个对称矩阵  $X$  使得:

$$\begin{cases} AXA = A \\ XAX = X. \end{cases}$$

### • 一点观察:

构造  $A \in \mathbb{C}^{m \times n}$  Moore-Penrose 逆  $A^\dagger$  需要求解  $A$  的一个满秩分解  $A = XY$ , 此时它具有如下形式:

$$A^\dagger = Y^H(YY^H)^{-1}(X^HX)^{-1}X^H.$$

可以验证它满足 Penrose 方程组:

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

满秩分解通常可由 SVD 分解、QR 分解或 LU 分解得到.

这里我们选用最适合手算的 LU 分解.

### Solution:

注意到  $A$  的 LU 分解可通过行 Gauss 消元可得:

$$\begin{aligned}
A &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 19 & 31 \\ 1 & 5 & 15 & 31 & 53 \end{bmatrix} \\
&= \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & & 1 & & \\ 1 & & & 1 & \\ 1 & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 9 & 14 & 30 \\ 3 & 9 & 18 & 30 & 52 \\ 4 & 14 & 30 & 52 & \end{bmatrix} \\
&= \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & & 1 & \\ 1 & 4 & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 3 & 9 & 18 & 30 & 52 \\ 6 & 18 & 36 & 52 & \end{bmatrix} \\
&= \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

因此  $\text{rank}(A) = 3 < 5$ , 表明  $A$  是奇异矩阵,  $A^{-1}$  不存在.

所以我们需要构造  $A$  的 Moore-Penrose 广义逆.

根据上述 LU 分解我们可得到  $A$  的一个满秩分解为:

$$\begin{aligned}
Y &:= \begin{bmatrix} 1 \\ 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 6 \end{bmatrix}, \\
YY^T &= \begin{bmatrix} 1 \\ 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 2 & 3 & 4 \\ & & 1 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 19 & 31 \\ 1 & 5 & 15 & 31 & 53 \end{bmatrix} = A.
\end{aligned}$$

我们取:

$$\begin{aligned}
X &:= Y(Y^TY)^{-1}(Y^TY)^{-1}Y^T \\
&= \begin{bmatrix} 1 \\ 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 6 \end{bmatrix} \begin{bmatrix} 5 & 10 & 10 \\ 10 & 30 & 35 \\ 10 & 35 & 46 \end{bmatrix}^{-1} \begin{bmatrix} 5 & 10 & 10 \\ 10 & 30 & 35 \\ 10 & 35 & 46 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 2 & 3 & 4 \\ & & 1 & 3 & 6 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 4 & 6 \end{bmatrix} \left( \frac{1}{175} \begin{bmatrix} 155 & -110 & 50 \\ -110 & 130 & -75 \\ 50 & -75 & 50 \end{bmatrix} \right)^2 \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 2 & 3 & 4 \\ & & 1 & 3 & 6 \end{bmatrix} \\
&= \frac{1}{30625} \begin{bmatrix} 38625 & 3525 & -13075 & -11175 & 9225 \\ 3525 & 3050 & 2075 & 600 & -1375 \\ -13075 & 2075 & 8350 & 5750 & -5725 \\ -11175 & 600 & 5750 & 4275 & -3825 \\ 9225 & -1375 & -5725 & -3825 & 4325 \end{bmatrix},
\end{aligned}$$

其中 3 阶矩阵求逆可通过计算伴随矩阵除以行列式简单得到.

可以验证上述定义的  $X$  满足 **Penrose 条件** (中的两条):

$$\begin{aligned}
AXA &= (YY^T) \cdot Y(Y^TY)^{-2}Y^T \cdot (YY^T) \\
&= Y(Y^TY)(Y^TY)^{-2}(Y^TY)Y^T \\
&= YY^T \\
&= A \\
\hline
XAX &= Y(Y^TY)^{-2}Y^T \cdot (YY^T) \cdot Y(Y^TY)^{-2}Y^T \\
&= Y(Y^TY)^{-2}(Y^TY)(Y^TY)(Y^TY)^{-2}Y^T \\
&= Y(Y^TY)^{-2}Y^T \\
&= X.
\end{aligned}$$

## Problem 3

给定正整数  $n$ .

若  $A \in \mathbb{R}^{n \times n}$  是实对称矩阵, 且存在  $x, y \in \mathbb{R}^n$  使得  $(x^T Ax)(y^T Ay) < 0$ .

试证明: 存在  $\text{span}\{x, y\}$  的一组基  $\{u, v\}$  使得  $u^T Au = v^T Av = 0$ .

**Proof:**

若  $x, y$  线性相关 (即存在  $\alpha \in \mathbb{R}$  使得  $y = \alpha x$ ),

则  $(x^T Ax)(y^T Ay) = \alpha^2 (x^T Ax)^2 \geq 0$  与题干假设矛盾.

因此  $x, y$  线性无关,  $\text{span}\{x, y\}$  是  $\mathbb{R}$  上的 2 维向量空间.

假设存在  $\lambda \in \mathbb{R}$  使得  $(x + \lambda y)^T A(x + \lambda y) = 0$ , 则我们有:

$$(y^T Ay)\lambda^2 + 2(x^T Ay)\lambda + x^T Ax = 0,$$

其中  $A$  的对称性保证了  $y^T Ax = (x^T Ay)^T = x^T Ay$ .

考虑上述一元二次方程的判别式:

$$\Delta = [2(x^T Ay)]^2 - 4(y^T Ay)(x^T Ax) > 0.$$

因此它有两个不同的实数解:

$$\lambda_{1,2} = \frac{-x^T Ay \pm \sqrt{\Delta}}{y^T Ay}.$$

取:

$$\begin{cases} u := x + \lambda_1 y \\ v := x + \lambda_2 y, \end{cases}$$

则有  $u^T Au = v^T Av = 0$ .

根据  $\lambda_1 \neq \lambda_2$  易知  $u, v$  线性无关, 因此  $\{u, v\}$  构成  $\text{span}\{x, y\}$  的一组基.

## Problem 4

给定正整数  $n$ .

若  $A, B, A - B$  均为  $n$  阶 Hermite 正定阵.

试证明  $B^{-1} - A^{-1}$  也是 Hermite 正定阵.

- **Lemma: (矩阵乘积的谱不变性, Matrix Analysis 定理 1.3.22)**

任意给定矩阵  $A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{n \times m}$  (其中  $m \geq n$ )

则我们有  $AB \in \mathbb{C}^{m \times m}, BA \in \mathbb{C}^{n \times n}$ , 且:

$$\text{eig}(AB) = \text{eig}(BA) \cup \underbrace{\{0, \dots, 0\}}_{m-n \text{ times}},$$

即  $AB$  的  $m$  个特征值即  $BA$  的  $n$  个特征值附加上  $m - n$  个零特征值.

换句话说, 二者的特征多项式满足:  $p_{AB}(\lambda) = \lambda^{m-n} p_{BA}(\lambda)$ .

这意味着  $AB, BA$  的非零特征值是完全相同的 (包括重数), 而零特征值的个数相差  $m - n$  个.

特殊地, 当  $m = n$  时, 矩阵乘积  $AB, BA$  的特征值完全相同,

此时若  $A, B$  至少有一个是非奇异阵, 则  $AB$  和  $BA$  相似.

**Proof:**

任意给定矩阵  $A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{n \times m}$  (其中  $m \geq n$ ), 我们都有:

$$\begin{bmatrix} I_n & -B \\ & I_m \end{bmatrix} \begin{bmatrix} BA & 0 \\ A & 0_m \end{bmatrix} \begin{bmatrix} I_n & B \\ & I_m \end{bmatrix} = \begin{bmatrix} 0_n & 0 \\ A & AB \end{bmatrix}.$$

注意到  $\begin{bmatrix} I_n & -B \\ & I_m \end{bmatrix}$  的逆矩阵即为  $\begin{bmatrix} I_n & B \\ & I_m \end{bmatrix}$ .

记:

$$C_1 = \begin{bmatrix} BA & 0 \\ A & 0_m \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0_n & 0 \\ A & AB \end{bmatrix},$$

则上述等式表明  $C_1, C_2$  相似, 于是  $C_1, C_2$  的特征值完全相同.

而  $C_1$  的特征值由  $BA$  的  $n$  个特征值和  $m$  个零特征值构成.

相应地,  $C_2$  的特征值由  $AB$  的  $m$  个特征值和  $n$  个零特征值构成.

比较二者, 即可知  $AB$  的  $m$  个特征值即  $BA$  的  $n$  个特征值附加上  $m - n$  个零特征值.

特殊地, 当  $m = n$  时, 若  $A, B$  至少有一个是非奇异阵 (不妨设  $A$  非奇异),

则有  $AB = A(BA)A^{-1}$ , 表明  $AB \sim BA$ .

**Proof:**

根据  $A, B$  的 Hermite 正定性可知  $A^{-1}, B^{-1}$  存在且为 Hermite 正定阵.

因此  $B^{-1} - A^{-1}$  也是 Hermite 阵.

注意到  $A, B$  是正规矩阵 (Hermite 阵自然是正规矩阵), 因此  $A, B$  可酉对角化, 即其谱分解存在.

设  $A, B$  的谱分解为:

$$\begin{cases} A := U_1 \Lambda_1 U_1^H = U_1 \text{diag}\{\lambda_1(A), \dots, \lambda_n(A)\} U_1^H \\ B := U_2 \Lambda_2 U_2^H = U_2 \text{diag}\{\lambda_1(B), \dots, \lambda_n(B)\} U_2^H. \end{cases}$$

我们定义其 2 次根为:

$$\begin{cases} A^{1/2} := U_1 \Lambda_1^{1/2} U_1^H = U_1 \text{diag}\{\sqrt{\lambda_1(A)}, \dots, \sqrt{\lambda_n(A)}\} U_1^H \\ B^{1/2} := U_2 \Lambda_2^{1/2} U_2^H = U_2 \text{diag}\{\sqrt{\lambda_1(B)}, \dots, \sqrt{\lambda_n(B)}\} U_2^H. \end{cases}$$

根据  $A - B \succ 0$  可知  $B^{-1/2} A B^{-1/2} - I_n \succ 0$ ,

这表明  $(B^{-1/2} A^{1/2})(A^{1/2} B^{-1/2})$  的所有特征值均大于 1.

根据 **Lemma (矩阵乘积的谱不变性)** 可知  $(A^{1/2} B^{-1/2})(B^{-1/2} A^{1/2})$  的所有特征值均大于 1,

即有:

$$\begin{aligned} A^{1/2} B^{-1} A^{1/2} - I_n &= (A^{1/2} B^{-1/2})(B^{-1/2} A^{1/2}) - I_n \succ 0 \\ &\quad \Updownarrow \\ B^{-1} - A^{-1/2} I_n A^{-1/2} &= B^{-1} - A^{-1} \succ 0 \end{aligned}$$

命题得证.

给定可逆矩阵  $A, B \in \mathbb{C}^{n \times n}$ , 我们有以下恒等式:

$$B^{-1} - A^{-1} = A^{-1}(A - B)B^{-1} = B^{-1}(A - B)A^{-1}.$$

**邵老师提供的解法 1:**

$$\begin{aligned} B^{-1} - A^{-1} &= A^{-1}(A - B)B^{-1} \\ &= A^{-1}(A - B)A^{-1} + A^{-1}(A - B)(B^{-1} - A^{-1}) \\ &= A^{-1}(A - B)A^{-1} + A^{-1}(A - B)B^{-1}(A - B)A^{-1} \quad (\text{note that } A, B \text{ are Hermitian matrices}) \\ &= (A^{-1})^H(A - B)A^{-1} + ((A - B)A^{-1})^H B^{-1}((A - B)A^{-1}) \quad (\text{note that } A - B \succ 0 \text{ and } B \succ 0) \\ &\succ 0. \end{aligned}$$

**邵老师提供的解法 2:**

$$\begin{aligned}
B^{-1} - A^{-1} &= A^{-1}(A - B)B^{-1} \\
&= A^{-1}(A - B)A^{-1} + A^{-1}(A - B)(B^{-1} - A^{-1}) \\
&= A^{-1}(A - B)A^{-1} + A^{-1}(A - B)A^{-1}(A - B)B^{-1} \\
&= \cdots (\text{continuously decomposing } B^{-1} = A^{-1} + (B^{-1} - A^{-1}) = A^{-1} + A^{-1}(A - B)B^{-1}) \\
&= A^{-1} \left( \sum_{m=1}^{\infty} ((A - B)A^{-1})^m \right) \\
&= \sum_{k=1}^{\infty} A^{-1}((A - B)A^{-1})^{2k} + \sum_{k=1}^{\infty} A^{-1}((A - B)A^{-1})^{2k+1} \\
&= \sum_{k=1}^{\infty} ((A^{-1}(A - B))^k A^{-1}(((A - B)A^{-1})^k) + \sum_{k=1}^{\infty} (A^{-1}((A - B)A^{-1})^k)(A - B)((A^{-1}(A - B))^k A^{-1}) \\
&= \sum_{k=1}^{\infty} (((A - B)A^{-1})^k)^H A^{-1}(((A - B)A^{-1})^k) + \sum_{k=1}^{\infty} ((A^{-1}(A - B))^k A^{-1})^H (A - B)((A^{-1}(A - B))^k A^{-1}) \\
&\succ 0 \quad (\text{note that } A - B \succ 0 \text{ and } B \succ 0).
\end{aligned}$$

邵老师提供的解法 3:

注意到:

$$\begin{aligned}
(B^{-1} - A^{-1})B((A - B)^{-1} + B^{-1})B &= A^{-1}(A - B)B^{-1} \cdot B((A - B)^{-1} + B^{-1})B \\
&= A^{-1}I_n B + A^{-1}(A - B)I_n \\
&= A^{-1}B + I_n - A^{-1}B \\
&= I_n.
\end{aligned}$$

因此我们有:

$$B^{-1} - A^{-1} = B^{-1}((A - B)^{-1} + B^{-1})^{-1}B^{-1}.$$

根据  $[(A - B)^{-1} + B^{-1}]^{-1} \succ 0$  可知  $B^{-1} - A^{-1} = B^{-1}((A - B)^{-1} + B^{-1})^{-1}B^{-1} \succ 0$ .

邵老师提供的解法 4:

设  $A$  的 Cholesky 分解为  $A = LL^H$ , 其中  $L$  为对角元为正实数的下三角阵.

因此  $I - L^{-1}BL^{-H} = L^{-1}(LL^H - B)L^{-H} = L^{-1}(A - B)L^{-H} \succ 0$ ,

这表明  $L^{-1}BL^{-H}$  的所有特征值都小于 1 (谱半径  $\rho(L^{-1}BL^{-H}) < 1$ ).

结合  $L^{-1}BL^{-H}$  的正定性可知其所有特征值都在  $(0, 1)$  之间 (谱半径  $\rho(L^{-1}BL^{-H}) \in (0, 1)$ ),

因此其逆矩阵  $(L^{-1}BL^{-H})^{-1} = L^H B^{-1}L$  的所有特征值都大于 1 (谱半径  $\rho(L^H B^{-1}L) > 1$ ).

于是  $L^H B^{-1}L - I \succ 0$ .

因此我们有  $B^{-1} - A^{-1} = B^{-1} - (LL^H)^{-1} = L^{-H}(L^H B^{-1}L - I)L^{-1} \succ 0$ .

## Problem 5

(惯性指数的次可加性, Subadditivity of Inertia Index)

给定正整数  $n$ .

设  $A, B$  都是  $n$  阶 Hermite 阵.

试证明:  $I_{n^+}(A + B) \leq I_{n^+}(A) + I_{n^+}(B)$ ,

其中  $I_{n^+}(A)$  代表  $A$  的正惯性指数.

- **Lemma 1: (Courant-Fischer min-max 定理, Matrix Analysis 定理 4.2.6)**

给定 Hermite 阵  $A \in \mathbb{C}^{n \times n}$ , 特征值按非减的次序排列:  $\lambda_{\min} = \lambda_1(A) \leq \cdots \leq \lambda_n(A) = \lambda_{\max}$ .

记  $S$  为  $\mathbb{C}^n$  的子空间, 则我们有:

$$\begin{aligned}
\lambda_i &= \min_{S \subseteq \mathbb{C}^n: \dim(S)=i} \left\{ \max_{x \neq 0, x \in S} \frac{x^H A x}{x^H x} \right\} \\
&= \max_{S \subseteq \mathbb{C}^n: \dim(S)=n-i+1} \left\{ \min_{x \neq 0, x \in S} \frac{x^H A x}{x^H x} \right\}. \quad (i = 1, \dots, n)
\end{aligned}$$

**Proof:**

$A \in \mathbb{C}^{n \times n}$  是 Hermite 阵 (自然是正规矩阵), 一定可以酉对角化,

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  $A = U\Lambda U^H$ ,

其中  $U$  的列向量组  $\{u_1, u_2, \dots, u_n\}$  构成  $\mathbb{C}^n$  的一组标准正交基.

对于任意给定的  $i = 1, \dots, n$ , 定义  $U_{(i)} = \text{span}\{u_i, u_{i+1}, \dots, u_n\}$ ,

则  $U_{(i)}$  是  $\mathbb{C}^n$  的子空间, 维数  $\dim(U_{(i)}) = n - i + 1$ .

对于  $\mathbb{C}^n$  任意给定的  $i$  维子空间  $S$ , 我们都有:

$$\begin{aligned}\dim(S \cap U_{(i)}) &= \dim(S) + \dim(U_{(i)}) - \dim(S + U_{(i)}) \\ &\geq i + (n - i + 1) - n \\ &= 1.\end{aligned}$$

这表明  $S \cap U_{(i)}$  一定有非零向量, 即  $(S \cap U_{(i)}) \setminus \{0_n\} \neq \emptyset$ .

任意给定  $\mathbb{C}^n$  的  $i$  维子空间  $S$ .

不失一般性, 可取单位向量  $x \in (S \cap U_{(i)})$ , 我们都有:

$$\begin{aligned}x^H A x &= x^H (U \Lambda U^H) x \\ &= (U^H x)^H \Lambda (U^H x) \quad (\text{Denote } y := U^H x, \text{ note that } \|y\|_2 = \|U^H x\|_2 = 1) \\ &= y^H \Lambda y \quad (\text{Let } y = \sum_{k=i}^n u_k \alpha_k, \text{ where } \sum_{k=i}^n |\alpha_k|^2 = 1) \\ &= \sum_{k=i}^n |\alpha_k|^2 \lambda_k \quad (\text{note that } \lambda_i \leq \lambda_{i+1} \leq \dots \leq \lambda_n) \\ &\geq \lambda_i \sum_{k=i}^n |\alpha_k|^2 \quad (\text{note that } \sum_{k=i}^n |\alpha_k|^2 = 1) \\ &= \lambda_i.\end{aligned}$$

上式的不等号至少在  $S = \text{span}\{u_1, u_2, \dots, u_i\}$  且  $x$  与  $u_i$  线性相关时取等.

因此我们有:

$$\begin{aligned}\lambda_i &= \min_{S \subseteq \mathbb{C}^n: \dim(S)=i} \left\{ \max_{\{x|x \in S \text{ such that } \|x\|_2=1\}} x^H A x \right\} \\ &= \min_{S \subseteq \mathbb{C}^n: \dim(S)=i} \left\{ \max_{x \neq 0_n \in S} \frac{x^H A x}{x^H x} \right\}.\end{aligned} \quad (i = 1, \dots, n)$$

对  $-A$  应用上述结论即得:

(注意对  $-A$  来说, 特征值非减次序为  $-\lambda_n \leq \dots \leq -\lambda_1$ , 因此  $-\lambda_i$  是  $-A$  的第  $n - i + 1$  小的特征值)

$$\begin{aligned}-\lambda_i &= \min_{S \subseteq \mathbb{C}^n: \dim(S)=n-i+1} \left\{ \max_{\{x|x \in S \text{ such that } \|x\|_2=1\}} -x^H A x \right\} \\ &= \min_{S \subseteq \mathbb{C}^n: \dim(S)=n-i+1} \left\{ \max_{x \neq 0_n \in S} -\frac{x^H A x}{x^H x} \right\}.\end{aligned} \quad (i = 1, \dots, n)$$

于是有:

$$\begin{aligned}\lambda_i &= \min_{S \subseteq \mathbb{C}^n: \dim(S)=n-i+1} \left\{ \max_{\{x|x \in S \text{ such that } \|x\|_2=1\}} x^H A x \right\} \\ &= \min_{S \subseteq \mathbb{C}^n: \dim(S)=n-i+1} \left\{ \max_{x \neq 0_n \in S} \frac{x^H A x}{x^H x} \right\}.\end{aligned} \quad (i = 1, \dots, n)$$

命题得证.

• **Lemma 2: (Weyl 不等式, Matrix Analysis 定理 4.3.1)**

给定 Hermite 阵  $A, B \in \mathbb{C}^{n \times n}$ .

设  $\{\lambda_i(A)\}_{i=1}^n, \{\lambda_i(B)\}_{i=1}^n, \{\lambda_i(A+B)\}_{i=1}^n$  为  $A, B, A+B$  的非减次序的特征值.

任意给定  $i = 1, 2, \dots, n$ .

- ① 对于任意  $j = 1, \dots, i$  都有  $\lambda_j(A) + \lambda_{1+i-j}(B) \leq \lambda_i(A+B)$  成立.

上式对某一对  $(i, j)$  取等, 当且仅当存在非零向量  $x$  使得:

$$\begin{cases} Ax = x\lambda_j(A) \\ Bx = x\lambda_{1+i-j}(B) \\ (A+B)x = x\lambda_i(A+B). \end{cases}$$

若  $A, B, A+B$  不存在公共特征向量, 则上述不等式都是严格不等式.

- ② 对于任意  $j = i, \dots, n$  都有  $\lambda_i(A+B) \leq \lambda_j(A) + \lambda_{n+i-j}(B)$  成立.

上式对某一对  $(i, j)$  取等, 当且仅当存在非零向量  $x$  使得

$$\begin{cases} Ax = x\lambda_j(A) \\ Bx = x\lambda_{n+i-j}(B) \\ (A+B)x = x\lambda_i(A+B). \end{cases}$$

若  $A, B, A+B$  不存在公共特征向量, 则上述不等式都是严格不等式.

**Proof:**

任意给定  $i = 1, 2, \dots, n$ .

**首先证明**  $\forall j = 1, \dots, i, \lambda_j(A) + \lambda_{1+i-j}(B) \leq \lambda_i(A+B)$ :

对于任意给定的  $j = 1, 2, \dots, i$ ,

设  $S_1, S_2, S_3$  分别为  $\mathbb{C}^n$  的  $(n-j+1), (n-(1+i-j)+1), i$  维子空间, 于是有:

$$\begin{aligned} \dim(S_1 \cap S_2 \cap S_3) &\geq \dim(S_1) + \dim(S_2) + \dim(S_3) - (3-1)\dim(\mathbb{C}^n) \\ &= (n-j+1) + (n-(1+i-j)+1) + i - 2n \\ &= 1. \end{aligned}$$

故  $(S_1 \cap S_2 \cap S_3) \setminus \{0_n\} \neq \emptyset$ .

因此可取  $x_0 \neq 0_n \in (S_1 \cap S_2 \cap S_3)$ ,

则根据 **Courant-Fischer min-max 定理**可知:

$$\begin{aligned} \begin{cases} \lambda_i(A) = \min_{S \subseteq \mathbb{C}^n: \dim(S)=i} \left\{ \max_{x \neq 0_n \in S} \frac{x^H A x}{x^H x} \right\} \\ \lambda_i(A) = \max_{S \subseteq \mathbb{C}^n: \dim(S)=n-i+1} \left\{ \min_{x \neq 0_n \in S} \frac{x^H A x}{x^H x} \right\} \end{cases} \\ \Downarrow \\ \lambda_j(A) + \lambda_{1+i-j}(B) &\leq \frac{x_0^H A x_0}{x_0^H x_0} + \frac{x_0^H B x_0}{x_0^H x_0} \quad (\text{note that } x_0 \in S_1 \text{ and } x_0 \in S_2) \\ &= \frac{x_0^H (A+B) x_0}{x_0^H x_0} \quad (\text{note that } x_0 \in S_3) \\ &\leq \lambda_i(A+B). \end{aligned}$$

**其次证明**  $\forall j = i, \dots, n, \lambda_i(A+B) \leq \lambda_j(A) + \lambda_{n+i-j}(B)$ :

对于任意给定的  $j = i, \dots, n$ ,

设  $S_1, S_2, S_3$  分别为  $\mathbb{C}^n$  的  $j, (n+i-j), (n-i+1)$  维子空间, 于是有:

$$\begin{aligned} \dim(S_1 \cap S_2 \cap S_3) &\geq \dim(S_1) + \dim(S_2) + \dim(S_3) - (3-1)\dim(\mathbb{C}^n) \\ &= j + (n+i-j) + (n-i+1) \\ &= 1. \end{aligned}$$

故  $(S_1 \cap S_2 \cap S_3) \setminus \{0_n\} \neq \emptyset$ .

因此可取  $x_0 \neq 0_n \in (S_1 \cap S_2 \cap S_3)$ ,

则根据 **Courant-Fischer min-max 定理**可知:

$$\begin{aligned} \begin{cases} \lambda_i(A+B) = \min_{S \subseteq \mathbb{C}^n: \dim(S)=i} \left\{ \max_{x \neq 0_n \in S} \frac{x^H (A+B) x}{x^H x} \right\} \\ \lambda_i(A+B) = \max_{S \subseteq \mathbb{C}^n: \dim(S)=n-i+1} \left\{ \min_{x \neq 0_n \in S} \frac{x^H (A+B) x}{x^H x} \right\} \end{cases} \\ \Downarrow \\ \lambda_i(A+B) &\leq \frac{x_0^H (A+B) x_0}{x_0^H x_0} \quad (\text{note that } x_0 \in S_3) \\ &= \frac{x_0^H A x_0}{x_0^H x_0} + \frac{x_0^H B x_0}{x_0^H x_0} \quad (\text{note that } x_0 \in S_1 \text{ and } x_0 \in S_2) \\ &\leq \lambda_j(A) + \lambda_{n+i-j}(B). \end{aligned}$$

命题得证.

**Proof:**

要证明  $I_{n^+}(A+B) \leq I_{n^+}(A) + I_{n^+}(B)$ ,

等价于证明  $A+B$  的第  $I_{n^+}(A) + I_{n^+}(B)$  大的特征值  $\lambda_{n-I_{n^+}(A)-I_{n^+}(B)}(A+B) \leq 0$ .

这可有 Weyl 不等式直接得到:



$$\begin{aligned} \text{Weyl inequality: } \forall j = i, \dots, n, \quad \lambda_i(A+B) &\leq \lambda_j(A) + \lambda_{n+i-j}(B) \\ \text{Substituting } i = n - \mathbf{I}_{n^+}(A) - \mathbf{I}_{n^+}(B) \text{ and } j = n - \mathbf{I}_{n^+}(A) \text{ we obtain:} \\ \lambda_{n-\mathbf{I}_{n^+}(A)-\mathbf{I}_{n^+}(B)}(A+B) &\leq \lambda_{n-\mathbf{I}_{n^+}(A)}(A) + \lambda_{n-\mathbf{I}_{n^+}(B)}(B) \\ &< 0 + 0 \\ &= 0. \end{aligned}$$

命题得证.

**邵老师提供的简单证明:**

我们可以使用合同变换

$$\begin{bmatrix} I & I \\ & I \end{bmatrix} \begin{bmatrix} A & \\ & B \end{bmatrix} \begin{bmatrix} I & \\ & I \end{bmatrix} = \begin{bmatrix} A & B \\ & B \end{bmatrix} \begin{bmatrix} I & \\ & I \end{bmatrix} = \begin{bmatrix} A+B & B \\ & B \end{bmatrix}$$

得到以下等价关系:

$$\begin{bmatrix} A & \\ & B \end{bmatrix} \sim \begin{bmatrix} A+B & B \\ & B \end{bmatrix}.$$

根据 Cauchy 交错原理可知子矩阵的特征值交错排列在原矩阵特征值之间, 因此其惯性指数也小于等于原矩阵的惯性指数.

由于  $A+B$  是  $\begin{bmatrix} A+B & B \\ B & B \end{bmatrix}$  的子矩阵, 故我们有:

$$\begin{aligned} \mathbf{I}_{n^+}(A+B) &\leq \mathbf{I}_{n^+} \left( \begin{bmatrix} A+B & B \\ B & B \end{bmatrix} \right) \\ &= \mathbf{I}_{n^+} \left( \begin{bmatrix} A & \\ & B \end{bmatrix} \right) \\ &= \mathbf{I}_{n^+}(A) + \mathbf{I}_{n^+}(B). \end{aligned}$$

## Problem 6 (optional)

在平面直角坐标系中, 由方程  $y = x + x^{-1}$  定义的曲线是双曲线. 试构造一个正交变换  $Q \in \mathbb{R}^{2 \times 2}$  使得在由

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = Q \begin{bmatrix} x \\ y \end{bmatrix}$$

定义的新坐标下  $y = x + x^{-1}$  转化为双曲线的标准方程.

**Solution:**

定义  $Q$  为复数  $e^{i\theta} = \cos(\theta) + i \sin(\theta)$  的实矩阵表示 (设  $\theta \in (0, \pi/2)$ ):

$$\begin{aligned} Q &:= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \\ Q^{-1} &= \frac{1}{\det(Q)} \begin{bmatrix} \cos(\theta) & -(-\sin(\theta)) \\ -(\sin(\theta)) & \cos(\theta) \end{bmatrix}^T = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}. \end{aligned}$$

于是我们有:

$$\begin{bmatrix} x \\ y \end{bmatrix} = Q^{-1} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} \cos(\theta)\tilde{x} - \sin(\theta)\tilde{y} \\ \sin(\theta)\tilde{x} + \cos(\theta)\tilde{y} \end{bmatrix}.$$

因此有:

$$\begin{aligned} y &= x + x^{-1} \\ &\Updownarrow \\ \sin(\theta)\tilde{x} + \cos(\theta)\tilde{y} &= \cos(\theta)\tilde{x} - \sin(\theta)\tilde{y} + \frac{1}{\cos(\theta)\tilde{x} - \sin(\theta)\tilde{y}} \\ &\Updownarrow \\ (\sin(\theta) - \cos(\theta))\tilde{x} + (\cos(\theta) + \sin(\theta))\tilde{y} &= \frac{1}{\cos(\theta)\tilde{x} - \sin(\theta)\tilde{y}} \end{aligned}$$

我们令:

$$\begin{aligned}
& \frac{\cos(\theta) + \sin(\theta)}{\sin(\theta) - \cos(\theta)} = \frac{\sin(\theta)}{\cos(\theta)} \\
& \quad \Downarrow \\
& (\cos(\theta) + \sin(\theta))^2 = 2\sin^2(\theta) \quad (\text{note that } \theta \in (0, \frac{\pi}{2})) \\
& \quad \Downarrow \\
& \cos(\theta) + \sin(\theta) = \sqrt{2}\sin(\theta) \\
& \quad \Downarrow \\
& \begin{cases} \sin(\theta) = \frac{\sqrt{2+\sqrt{2}}}{2} \\ \cos(\theta) = \frac{\sqrt{2-\sqrt{2}}}{2} \end{cases}
\end{aligned}$$

于是我们有:

$$\begin{aligned}
& (\sin(\theta) - \cos(\theta))\tilde{x} + (\cos(\theta) + \sin(\theta))\tilde{y} = \frac{1}{\cos(\theta)\tilde{x} - \sin(\theta)\tilde{y}} \\
& \quad \Downarrow \\
& (\sin(\theta) - \cos(\theta))\cos(\theta)\tilde{x}^2 - (\cos(\theta) + \sin(\theta))\sin(\theta)\tilde{y}^2 = 1 \\
& \quad \Downarrow \\
& \frac{\sqrt{2}-1}{2}\tilde{x}^2 - \frac{\sqrt{2}+1}{2}\tilde{y}^2 = 1
\end{aligned}$$

此时我们有:

$$Q = \begin{bmatrix} \frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2} \\ -\frac{\sqrt{2+\sqrt{2}}}{2} & \frac{\sqrt{2-\sqrt{2}}}{2} \end{bmatrix},$$

并满足:

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = Q \begin{bmatrix} x \\ y \end{bmatrix}.$$

## Problem 7 (optional)

给定正整数  $n$  和向量  $x, y \in \mathbb{R}^n$  (其中  $x \neq 0_n$ ).

试证明:  $y^T x > 0$  的充要条件是存在对称正定阵  $A \in \mathbb{R}^{n \times n}$  使得  $y = Ax$ .

• **Lemma:**

若非零向量  $x, y \in \mathbb{R}^n$  满足  $\|x\|_2 = \|y\|_2$ , 定义

$$\begin{aligned}
w &:= \frac{x - y}{\|x - y\|_2} \\
H &:= I_n - 2ww^T
\end{aligned}$$

则我们有  $y = Hx$  成立.

**Proof:**

$$\begin{aligned}
Hx &= (I_n - 2ww^T)x \\
&= \left\{ I_n - 2 \frac{(x - y)(x - y)^T}{(x - y)^T(x - y)} \right\} x \\
&= x - 2 \frac{(x - y)^T x}{x^T y - y^T x - x^T y - y^T y} (x - y) \quad (\text{note that } \|y\|_2 = \|x\|_2) \\
&= x - 2 \frac{x^T x - x^T y}{2(x^T x - x^T y)} (x - y) \\
&= x - (x - y) \\
&= y
\end{aligned}$$

**必要性证明:**

若存在对称正定阵  $A \in \mathbb{R}^{n \times n}$  使得  $y = Ax$ , 则我们有  $y^T x = x^T Ax > 0$  成立.

**充分性证明:**

现假设  $y^T x > 0$ , 我们希望构造一个对称正定阵  $A \in \mathbb{R}^{n \times n}$  使得  $y = Ax$ .

• **构造法 1:**

$$A := I_n - \frac{xx^T}{x^T x} + \frac{yy^T}{y^T x}$$


---


$$Ax = \left( I_n - \frac{xx^T}{x^T x} + \frac{yy^T}{y^T x} \right) x = x - x + y = y$$

显然  $A$  是对称阵, 只需说明  $A$  正定即可:

首先  $I_n - \frac{xx^T}{x^T x}$  半正定, 仅有一个特征值为 0, 对应特征向量为  $x$ , 其余特征值为 1.

因此  $v^T (I_n - \frac{xx^T}{x^T x}) v = 0$  当且仅当  $v \in \text{span}\{x\}$  成立.

而  $\frac{yy^T}{y^T x}$  同样半正定, 但由于

$$x^T \left( \frac{yy^T}{y^T x} \right) x = x^T y > 0,$$

故对于任意  $v \in \text{span}\{x\} \setminus \{0_n\}$  我们都有

$$v^T \left( \frac{yy^T}{y^T x} \right) v > 0$$

因此对于任意  $v \neq 0_n \in \mathbb{R}^n$  我们都有

$$v^T A v = v^T \left( I_n - \frac{xx^T}{x^T x} + \frac{yy^T}{y^T x} \right) v > 0,$$

这表明  $A$  是正定阵.

• **构造法 2:**

$$\alpha := \frac{2x^T x}{y^T x} > 0$$

$$z := y\alpha - x$$

$$A := \frac{1}{\alpha} \left( I + \frac{zz^T}{x^T x} \right)$$

显然  $A$  是对称正定阵.

下面我们证明  $y = Ax$ :

$$\begin{aligned} z^T x &= (y\alpha - x)^T x \\ &= y^T x \cdot \alpha - x^T x \\ &= y^T x \cdot \frac{2x^T x}{y^T x} - x^T x \\ &= x^T x \end{aligned}$$


---


$$\begin{aligned} y &= Ax \\ &= \frac{1}{\alpha} \left( I + \frac{zz^T}{x^T x} \right) x \\ &= \frac{1}{\alpha} x + \frac{1}{\alpha} \frac{z^T x}{x^T x} z \quad (\text{note that } z^T x = x^T x \text{ and } z = y\alpha - x) \\ &= \frac{1}{\alpha} x + \frac{1}{\alpha} \frac{x^T x}{x^T x} (y\alpha - x) \\ &= \frac{1}{\alpha} x + \frac{1}{\alpha} (y\alpha - x) \\ &= y. \end{aligned}$$

• **构造法 3:**

由于  $x \neq 0_n$ , 故存在非奇异阵  $P \in \mathbb{R}^{n \times n}$  使得  $Px = e_1$ .

具体来说, **Lemma** 指导我们这样构造  $P$ :

$$w := \frac{x - \|x\|_2 e_1}{\|x - \|x\|_2 e_1\|_2}$$

$$H := I_n - 2ww^T$$

$$P := \frac{1}{\|x\|_2} H$$

记  $\tilde{y} := P^{-T}y = [\tilde{y}_1, \dots, \tilde{y}_n]^T$ , 考虑其第一个元素:

$$\tilde{y}_1 = e_1^T \tilde{y} = (Px)^T (P^{-T}y) = x^T P^T P^{-T}y = x^T y > 0.$$

我们可取一个对称阵  $B \in \mathbb{R}^{n \times n}$ , 其第一列为  $\tilde{y}$  (第一行对应为  $\tilde{y}^T$ ),

显然它满足  $Be_1 = \tilde{y}$ .

由于其右下  $n - 1$  阶主子阵可以任取 (不过要保证对称性),

故我们可令其为对角阵, 且对角元取得足够大使得  $B$  为正定阵 (所有顺序主子式都大于 0).

取  $A := P^T B P$ , 显然它是对称正定阵.

同时我们有:

$$\begin{aligned} Ax &= P^T B P x \quad (\text{note that } Px = e_1) \\ &= P^T B e_1 \quad (\text{note that } B e_1 = \tilde{y}) \\ &= P^T \tilde{y} \quad (\text{note that } \tilde{y} = P^{-T}y) \\ &= P^T (P^{-T}y) \\ &= y. \end{aligned}$$

综上所述, 命题得证.

#### 错误的做法 1:

现假设  $y^T x > 0$

当  $x, y$  线性相关 (即存在  $\alpha > 0$  使得  $y = \alpha x$ ) 时, 我们可取  $A = \alpha I_n$ , 即满足  $y = Ax$

当  $x, y$  线性无关时, 我们定义:

$$\alpha := \frac{\|y\|_2}{\|x\|_2}$$

$$\tilde{y} := \frac{1}{\alpha} y \quad (\text{so that } \|\tilde{y}\|_2 = \|x\|_2)$$

并构造以下 Householder 变换:

$$w := \frac{x - \tilde{y}}{\|x - \tilde{y}\|_2}$$

$$H := I_n - 2ww^T$$

根据 **Lemma** 可知  $Hx = \tilde{y}$

定义  $A := \alpha H$ , 则我们有  $Ax = \alpha Hx = \alpha \tilde{y} = y$  成立.

但  $A$  并不是一个对称正定阵.

#### 错误的做法 2:

当  $x, y$  线性无关时, 我们定义:

$$\alpha := \frac{\|y\|_2}{\|x\|_2}$$

$$\tilde{y} := \frac{1}{\alpha} y \quad (\text{so that } \|\tilde{y}\|_2 = \|x\|_2)$$

$$z := \frac{1}{2}(x + y)$$

$$\beta := \frac{\|z\|_2}{\|x\|_2}$$

$$\tilde{z} := \frac{1}{\beta} z \quad (\text{so that } \|\tilde{z}\|_2 = \|x\|_2)$$

并构造以下 Householder 变换:

$$w_1 := \frac{x - \tilde{z}}{\|x - \tilde{z}\|_2}$$

$$H_1 := I_n - 2w_1w_1^T$$

$$w_2 := \frac{\tilde{y} - \tilde{z}}{\|\tilde{y} - \tilde{z}\|_2}$$

$$H_2 := I_n - 2w_2w_2^T$$

根据 **Lemma** 可知  $\begin{cases} H_1x = \tilde{z} \\ H_2\tilde{z} = \tilde{y} \end{cases}$

定义  $A := \alpha H_2 H_1$ , 则我们有  $Ax = \alpha H_2 H_1 x = \alpha H_2 \tilde{z} = \alpha \tilde{y} = y$  成立.

但是这样定义的  $A$  并不是一个对称正定阵 (甚至不一定是对称阵)

### 错误的做法 3:

当  $x, y$  线性无关时, 基于 [Rodrigues' rotation formula](#) 我们可以构造如下旋转矩阵:

$$\alpha := \frac{\|y\|_2}{\|x\|_2}$$

$$\tilde{y} := \frac{1}{\alpha}y \text{ (so that } \|\tilde{y}\|_2 = \|x\|_2)$$

$$H := 2 \frac{(x + \tilde{y})(x + \tilde{y})^T}{(x + \tilde{y})^T(x + \tilde{y})} - I_n$$

$$A := \alpha H$$

可以证明  $Ax = \alpha Hx = \alpha \tilde{y} = y$

但  $A$  并不是对称正定阵.

### 错误的做法 4:

当  $x, y$  线性无关时, 我们可通过如下步骤构造对称正定阵  $A$  使得  $y = Ax$ :

- ① 首先构造一个实对称正交阵 (Householder 变换)  $H \in \mathbb{R}^{n \times n}$  使得  $Hx = \|x\|_2 e_1$   
其中  $e_1$  是  $\mathbb{R}^n$  的第 1 个标准正交基向量.  
具体来说, **Lemma** 指导我们这样构造  $H$ :

$$w := \frac{x - \|x\|_2 e_1}{\|x - \|x\|_2 e_1\|_2}$$

$$H := I_n - 2ww^T$$

根据  $\frac{x}{\|x\|_2} = \frac{1}{\|x\|_2} H^T \|x\|_2 e_1 = H^T e_1 = H e_1$  可知  $H$  的第 1 列为  $\frac{x}{\|x\|_2}$ , 第 1 行为  $\frac{x^T}{\|x\|_2}$

记  $\tilde{y} := Hy$ , 则其第一个元素  $\tilde{y}_1 = e_1^T \tilde{y} = e_1^T Hy = \frac{x^T}{\|x\|_2} y > 0$

- ② 其次构造  $n - 1$  个 Givens 变换  $G_2, \dots, G_n$  使得:

$$G_i(\|x\|_2 e_1) = \frac{1}{n-1} \tilde{y}_1 e_1 + \tilde{y}_i e_i \quad (i = 2, \dots, n)$$

$$\text{where } G_i = I_n + (\cos(\theta_i) - 1)(e_1 e_1^T + e_i e_i^T) + \sin(\theta_i)(e_1 e_i^T - e_i e_1^T)$$

$$\text{so that } (G_2 + \dots + G_n)\|x\|_2 e_1 = (n-1) \frac{1}{n-1} \tilde{y}_1 e_1 + \tilde{y}_2 e_2 + \dots + \tilde{y}_n e_n$$

$$= \tilde{y}_1 e_1 + \tilde{y}_2 e_2 + \dots + \tilde{y}_n e_n$$

$$= \tilde{y}$$

单独提取第 1,  $i$  ( $i = 2, \dots, n$ ) 行列可以让我们看得更加清楚:

$$\begin{bmatrix} \cos(\theta_i) & \sin(\theta_i) \\ -\sin(\theta_i) & \cos(\theta_i) \end{bmatrix} \begin{bmatrix} \|x\|_2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{n-1} \tilde{y}_1 \\ \tilde{y}_i \end{bmatrix} \Rightarrow \begin{cases} \cos(\theta_i) = \frac{1}{(n-1)\|x\|_2} \tilde{y}_1 > 0 \\ \sin(\theta_i) = -\frac{1}{\|x\|_2} \tilde{y}_i \end{cases}$$

注意到  $e_1 e_i^T$  和  $e_i e_1^T$  的特征值均为 0, 而  $e_1 e_1^T$  和  $e_i e_i^T$  分别有一个特征值为 1, 其余特征值为 0  
因此  $\cos(\theta_i) > 0$  就保证了  $G_i$  ( $i = 2, \dots, n$ ) 是正定阵.

现在我们可以定义  $A := H^T(G_2 + \dots + G_n)H$

则我们有:

$$Hy = \tilde{y} = (G_2 + \dots + G_n)\|x\|_2 e_1 = (G_2 + \dots + G_n)Hx$$

$$\Leftrightarrow$$

$$y = H^T(G_2 + \dots + G_n)Hx = Ax$$

不过  $A$  虽然正定, 但不是对称阵.

## Problem 8 (optional)

### (Lowner-Heinz 不等式)

给定正整数  $n$ .

设  $A, B, A - B$  均为  $n$  阶 Hermite 正定阵.

试证明:  $A^{1/2} - B^{1/2}$  也是 Hermite 正定阵.

举例说明  $A^2 - B^2$  不一定是 Hermite 正定阵.

- 当  $AB = BA$  时, 我们有  $A^2 - B^2 = (A - B)(A + B) \succ 0$ .

### 一个不严谨的做法:

注意到  $A, B$  是正规矩阵, 因此  $A, B$  可酉对角化, 即其谱分解存在.

设  $A, B$  的谱分解为

$$\begin{cases} A := U_1 \Lambda_1 U_1^H = U_1 \text{diag}\{\lambda_1(A), \dots, \lambda_n(A)\} U_1^H \\ B := U_2 \Lambda_2 U_2^H = U_2 \text{diag}\{\lambda_1(B), \dots, \lambda_n(B)\} U_2^H. \end{cases}$$

我们定义其 2 次根为

$$\begin{cases} A^{1/2} := U_1 \Lambda_1^{1/2} U_1^H = U_1 \text{diag}\{\sqrt{\lambda_1(A)}, \dots, \sqrt{\lambda_n(A)}\} U_1^H \\ B^{1/2} := U_2 \Lambda_2^{1/2} U_2^H = U_2 \text{diag}\{\sqrt{\lambda_1(B)}, \dots, \sqrt{\lambda_n(B)}\} U_2^H. \end{cases}$$

根据  $A - B \succ 0$  可知  $I_n - A^{-1}B \succ 0$  (由于  $A^{-1}B$  不一定 Hermite, 故这个写法不严谨), 即  $A^{-1}B$  的谱半径  $\rho(A^{-1}B) < 1$ .

因此我们有:

$$\begin{aligned} (\rho(A^{-1/2}B^{1/2}))^2 &\leq \|A^{-1/2}B^{1/2}\|_2^2 \quad (\text{spectral theorem}) \\ &= \rho((A^{-1/2}B^{1/2})^H(A^{-1/2}B^{1/2})) \\ &= \rho(B^{1/2}A^{-1}B^{1/2}) \\ &= \rho(B^{-1/2} \cdot B^{1/2}A^{-1}B^{1/2} \cdot B^{1/2}) \quad (\text{similarity transformation does not change eigenvalues}) \\ &= \rho(A^{-1}B) \\ &< 1. \end{aligned}$$

因此  $\rho(A^{-1/2}B^{1/2}) < 1$ , 即有  $I_n - A^{-1/2}B^{1/2} \succ 0$  (由于  $A^{-1/2}B^{1/2}$  不一定 Hermite, 故这个写法不严谨), 这表明  $A^{1/2} - B^{1/2} \succ 0$ .

根据上面的推导我们知道  $\rho(A^{-1}B) \leq \rho(A^{-2}B^2)$ .

显然  $\rho(A^{-1}B)$  不能保证  $\rho(A^{-2}B^2)$  成立.

因此这道题是在说:

$$A^2 \succ B^2 \succ 0 \Rightarrow A \succ B \succ 0.$$

且其逆命题不成立:

$$A = \begin{bmatrix} 2+\varepsilon & 1 \\ 1 & 3 \end{bmatrix} \text{ where } \varepsilon > 0 \text{ is a small positive number}$$

$$B = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$$

$$A - B = \begin{bmatrix} 1+\varepsilon & 1 \\ 1 & 2 \end{bmatrix} \succ 0 \quad (\forall \varepsilon > 0)$$

$$A^2 - B^2 = \begin{bmatrix} (2+\varepsilon)^2 + 1 & 5+\varepsilon \\ 5+\varepsilon & 10 \end{bmatrix} - \begin{bmatrix} 1 & \\ & 4 \end{bmatrix} = \begin{bmatrix} (2+\varepsilon)^2 & 5+\varepsilon \\ 5+\varepsilon & 6 \end{bmatrix}$$

我们发现  $\det(A^2 - B^2) = 6(2+\varepsilon)^2 - (5+\varepsilon)^2 \rightarrow -1 \quad (\varepsilon \rightarrow 0)$ .

因此存在某个  $\varepsilon = \varepsilon_0$  (例如  $\varepsilon = 0.01$ , 此时  $\det(A^2 - B^2) = -0.8575$ ) 使得  $A^2 - B^2$  不是正定阵.

实际上我们还可推得:

$$A^2 \succeq B^2 \succeq 0$$

$$\Downarrow$$

$$A \succeq B \succeq 0.$$

这是因为对于任意  $t_1 > t_2 > 0$  都有:

$$A^2 \succeq B^2 \succeq 0$$

$$\Downarrow$$

$$(A + t_1 I)^2 \succ (B + t_2 I)^2 \succ 0$$

$$\Downarrow$$

$$(A + t_1 I) \succ (B + t_2 I) \succ 0.$$

对  $(A + t_1 I) \succ (B + t_2 I) \succ 0$  取极限  $t_1, t_2 \rightarrow 0$  即得  $A \succeq B \succeq 0$ .

其逆命题同样不成立 (反例待补充).

邵老师提供的解法 1:

- 我们首先说明任意 Hermite 正定阵  $A$  和 Hermite 阵  $B$  可以通过合同变换同时对角化:  
根据惯性定理可知存在非奇异阵  $P_0 \in \mathbb{C}^{n \times n}$  使得  $P_0^H A P_0 = I_n$ .  
注意到  $P_0^H B P_0$  仍是 Hermite 阵, 故存在酉矩阵  $U \in \mathbb{C}^{n \times n}$  使得  $D_B := U^H (P_0^H B P_0) U$  为对角阵.  
取  $P = P_0 U$ , 则我们有:

$$P^H A P = U^H (P_0^H A P_0) U = U^H I_n U = I_n$$

$$P^H B P = U^H (P_0^H B P_0) U = D_B.$$

- 因此对于 Hermite 正定阵  $A^{1/2}$  和  $B^{1/2}$ , 存在非奇异阵  $P \in \mathbb{C}^{n \times n}$  使得

$$\begin{cases} D_A := P^H A^{1/2} P \\ D_B := P^H B^{1/2} P \end{cases}$$

都是对角阵.

于是我们有:

$$\begin{aligned} P^H (A - B) P &= P^H (A^{1/2} A^{1/2} - B^{1/2} B^{1/2}) P \\ &= P^H (P^{-H} D_A P^{-1} P^{-H} D_A P^{-1} - P^{-H} D_B P^{-1} P^{-H} D_B P^{-1}) P \\ &= D_A (P^{-1} P^{-H}) D_A - D_B (P^{-1} P^{-H}) D_B \quad (\text{denote } X := P^{-1} P^{-H}) \\ &= D_A X D_A - D_B X D_B \quad (\text{note that } A - B \succ 0) \\ &\succ 0. \end{aligned}$$

记:

$$\begin{cases} D_A := \text{diag}\{d_1^{(1)}, \dots, d_n^{(1)}\} \\ D_B := \text{diag}\{d_1^{(2)}, \dots, d_n^{(2)}\}. \end{cases}$$

考虑  $D_A X D_A - D_B X D_B$  的对角元, 我们有:

$$(d_{ii}^{(1)})^2 x_{ii} - (d_{ii}^{(2)})^2 x_{ii} = ((d_{ii}^{(1)})^2 - (d_{ii}^{(2)})^2) x_{ii} > 0 \quad (i = 1, \dots, n).$$

注意到非奇异阵  $X := P^{-1} P^{-H}$  是 Hermite 正定阵,

因此  $X = [x_{ij}]$  的对角元  $x_{ii}$  ( $i = 1, \dots, n$ ) 均为正实数.

于是我们有  $d_{ii}^{(1)} > d_{ii}^{(2)}$  ( $i = 1, \dots, n$ ) 成立, 进而有  $D_A - D_B \succ 0$  成立.

最终我们得到:

$$\begin{aligned} A^{1/2} - B^{1/2} &= P^{-H} D_A P^{-1} - P^{-H} D_B P^{-1} \\ &= P^{-H} (D_A - D_B) P^{-1} \\ &\succ 0. \end{aligned}$$

得证  $A^{1/2} - B^{1/2}$  正定.

邵老师提供的解法 2:

注意到  $A^{1/2} - B^{1/2}$  是 Hermite 阵, 故其特征值均为实数.

设  $\lambda$  为  $A^{1/2} - B^{1/2}$  的最小特征值,  $x \neq 0_n$  为对应的特征向量, 满足  $(A^{1/2} - B^{1/2})x = x\lambda$ ,

则我们有:

$$\begin{aligned}
x^H(A-B)x &= x^H Ax - x^H Bx \\
&= x^H Ax - (B^{1/2}x)^H (B^{1/2}x) \quad (\text{note that } (A^{1/2} - B^{1/2})x = x\lambda \Rightarrow B^{1/2}x = (A^{1/2} - \lambda I_n)x) \\
&= x^H Ax - [(A^{1/2} - \lambda I_n)x]^H (A^{1/2} - \lambda I_n)x \\
&= x^H [A - (A^{1/2} - \lambda I_n)^2]x \\
&= 2\lambda x^H A^{1/2}x - \lambda^2 x^H Ax \quad (\text{note that } A - B \succ 0) \\
&> 0.
\end{aligned}$$

根据  $A \succ 0$  可知  $A^{1/2} \succ 0$ , 进而结合  $x \neq 0_n$  可知:

$$\begin{cases} x^H Ax > 0 \\ x^H A^{1/2}x > 0. \end{cases}$$

因此我们有:

$$\lambda > \frac{\lambda^2 x^H Ax}{2x^H A^{1/2}x} > 0.$$

因此  $A^{1/2} - B^{1/2}$  的最小特征值  $\lambda > 0$ , 表明  $A^{1/2} - B^{1/2}$  正定.

### 邵老师提供的解法 3:

可以证明对于任意正定阵  $A$  都有:

$$A^{1/2} = \frac{2}{\pi} \int_0^\infty (I + t^2 A^{-1})^{-1} dt.$$

设  $A$  的谱分解为  $A = U \Lambda U^H$ .

要验证上式, 根据

$$U^H A^{1/2} U = \frac{2}{\pi} \int_0^\infty (I + t^2 \Lambda^{-1}) dt$$

可知仅需对标量  $a > 0$  验证

$$a^{1/2} = \frac{2}{\pi} \int_0^\infty (1 + t^2 a^{-1})^{-1} dt$$

成立即可.

$$\begin{aligned}
\frac{2}{\pi} \int_0^\infty (1 + t^2 a^{-1})^{-1} dt &= \frac{2}{\pi} \int_0^\infty \frac{\sqrt{a}}{1 + u^2} du \quad (\text{denote } u := \frac{t}{\sqrt{a}}) \\
&= \sqrt{a} \cdot \frac{2}{\pi} \arctan(u) \Big|_0^\infty \\
&= \sqrt{a} \cdot \frac{2}{\pi} \cdot \left( \frac{\pi}{2} - 0 \right) \\
&= \sqrt{a}.
\end{aligned}$$

因此由  $A - B \succ 0$  得  $(I + t^2 B^{-1}) - (I + t^2 A^{-1})$  正定,

进而有  $(I + t^2 A^{-1})^{-1} - (I + t^2 B^{-1})^{-1}$  正定,

于是我们有:

$$\begin{aligned}
A^{1/2} - B^{1/2} &= \frac{2}{\pi} \int_0^\infty (I + t^2 A^{-1})^{-1} dt - \frac{2}{\pi} \int_0^\infty (I + t^2 B^{-1})^{-1} dt \\
&= \frac{2}{\pi} \int_0^\infty ((I + t^2 A^{-1})^{-1} - (I + t^2 B^{-1})^{-1}) dt \\
&\succ 0.
\end{aligned}$$

因此  $A^{1/2} - B^{1/2}$  也是 Hermite 正定阵.

利用邵老师提供的解法 3 的技术, 我们可以得到更一般的结论.

当  $0 < \alpha < 1$  时, 我们有:

$$\begin{aligned}
A^\alpha &= \frac{\int_0^\infty (I + t^{1/\alpha} A^{-1})^{-1} dt}{\int_0^\infty (1 + t^{1/\alpha})^{-1} dt} \\
&= \frac{\sin(\alpha\pi)}{\alpha\pi} \int_0^\infty (I + t^{1/\alpha} A^{-1})^{-1} dt.
\end{aligned}$$



因此由  $A - B \succ 0$  得  $(I + t^{1/\alpha} B^{-1}) - (I + t^{1/\alpha} A^{-1})$  正定,  
 进而有  $(I + t^{1/\alpha} A^{-1})^{-1} - (I + t^{1/\alpha} B^{-1})^{-1}$  正定,  
 于是我们有:

$$\begin{aligned} A^\alpha - B^\alpha &= \frac{2}{\pi} \int_0^\infty (I + t^{1/\alpha} A^{-1})^{-1} dt - \frac{2}{\pi} \int_0^\infty (I + t^{1/\alpha} B^{-1})^{-1} dt \\ &= \frac{2}{\pi} \int_0^\infty [(I + t^{1/\alpha} A^{-1})^{-1} - (I + t^{1/\alpha} B^{-1})^{-1}] dt \\ &\succ 0. \end{aligned}$$

因此  $A^\alpha - B^\alpha$  ( $0 < \alpha < 1$ ) 也是 Hermite 正定阵.

总之, 若  $A, B, A - B \succ 0$ , 则对于任意  $\alpha \in (0, 1)$  我们都有  $A^\alpha - B^\alpha \succ 0$  成立.

这个结论称为 **Lowner-Heinz 不等式**.

**The End**