

高等线性代数 Homework 12

Due: Dec. 9, 2024

姓名: 雍崔扬

学号: 21307140051

Problem 1

给定正整数 n .

设 $A = [a_{ij}] \in \mathbb{C}^{n \times n}$ 是 Hermite 阵.

试证明其关于行的每一个 Gershgorin 圆盘都至少包含 A 的一个特征值.

(关于列的 Gershgorin 圆盘有相同的结论)

- Lemma 1 (Weyl 不等式的最常用形式, Matrix Analysis 推论 4.3.15)

若 A, B 没有公共特征向量, 则下列不等式严格成立:

$$\lambda_1(B) \leq \lambda_i(A + B) - \lambda_i(A) \leq \lambda_n(B) \quad (i = 1, \dots, n)$$

注意到对于 Hermite 阵 $B \in \mathbb{C}^{n \times n}$ 来说有 $\|B\|_2 = \rho(B) = \max\{|\lambda_1(B)|, |\lambda_n(B)|\}$ 成立.

故我们还可以得到:

$$\begin{aligned} & \max_{1 \leq i \leq n} \{\lambda_i(A + B) - \lambda_i(A)\} \leq \lambda_n(B) \\ & \min_{1 \leq i \leq n} \{\lambda_i(A + B) - \lambda_i(A)\} \geq \lambda_1(B) \\ \hline & -\|B\|_2 \leq \lambda_1(B) \leq \lambda_i(A + B) - \lambda_i(A) \leq \lambda_n(B) \leq \|B\|_2 \quad (i = 1, \dots, n) \\ & \max_{1 \leq i \leq n} |\lambda_i(A + B) - \lambda_i(A)| \leq \|B\|_2 = \max\{|\lambda_1(B)|, |\lambda_n(B)|\} \end{aligned}$$

- Lemma 2 (范数不等式)

对于任意 $A \in \mathbb{C}^{m \times n}$ 都有 $\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty$ 成立.

Proof:

$$\begin{aligned} \|A\|_2^2 &= \sigma_{\max}^2(A) \\ &= \lambda_{\max}(A^H A) \\ &= \rho(A^H A) \quad (\text{use spectral theorem}) \\ &= \|A^H A\|_1 \\ &\leq \|A^H\|_1 \|A\|_1 \\ &= \|A\|_\infty \|A\|_1 \end{aligned}$$

Proof:

任意给定 $i = 1, \dots, n$.

定义 $B \in \mathbb{C}^{n \times n}$ 为 A 移除第 i 行和第 i 列所有非对角元得到的矩阵.

显然 a_{ii} 是 B 的一个特征值.

注意到 B 也是一个 Hermite 阵.

根据 Hermite 特征值的 Weyl 不等式 (Lemma 1) 可知:

存在 A 的某个特征值 $\lambda \in \mathbb{C}$ 使得 $|\lambda - a_{ii}| \leq \|A - B\|_2$.

注意到 $A - B$ 仅保留了 A 的第 i 行和第 i 列的非零元,

因此 $A - B$ 的最大绝对行和即为 A 的第 i 行的去心绝对行和:

$$\|A - B\|_\infty = \sum_{j \neq i}^n |a_{ij}|$$

而 $A - B$ 的最大绝对列和即为 A 的第 i 列的去心绝对列和,

根据 A 的 Hermite 性可知其等于 A 的第 i 行的去心绝对列和:

$$\begin{aligned}
\|A - B\|_1 &= \sum_{j \neq i}^n |a_{ji}| \quad (\text{note that } A \text{ is Hermitian}) \\
&= \sum_{j \neq i}^n |\bar{a}_{ij}| \\
&= \sum_{j \neq i}^n |a_{ij}|
\end{aligned}$$

因此我们有:

$$\begin{aligned}
|\lambda - a_{ii}| &\leq \|A - B\|_2 \quad (\text{utilize Lemma 2}) \\
&\leq \sqrt{\|A - B\|_1 \|A - B\|_\infty} \\
&= \sqrt{\sum_{j \neq i}^n |a_{ij}| \sum_{j \neq i}^n |a_{ij}|} \\
&= \sum_{j \neq i}^n |a_{ij}|
\end{aligned}$$

这表明 A 关于行的每一个 Gershgorin 圆盘都至少包含 A 的一个特征值,
其中 A 关于行的第 $i = 1, \dots, n$ 个 Gershgorin 圆盘的定义为:

$$\begin{aligned}
G_i(A) &:= \{z \in \mathbb{C} : |z - a_{ii}| \leq \text{Row}'_i(A)\} \\
\text{where } \text{Row}'_i(A) &:= \sum_{j \neq i}^n |a_{ij}|
\end{aligned}$$

Problem 2

给定正整数 n .

设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 阵, λ 是 A 的 m 重特征值.

试证明 λ 是 A 的所有 $n - 1$ 阶主子阵的特征值, 且重数至少为 $m - 1$.

- (Cauchy 交错定理, Matrix Analysis 定理 4.3.17)

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

考虑 A 的 $n - 1$ 主子阵 $B = A_{(1:n-1, 1:n-1)} \in \mathbb{C}^{(n-1) \times (n-1)}$, 并记:

$$A = \begin{bmatrix} B & y \\ y^H & a \end{bmatrix}.$$

特征值按非减的次序排列: $\lambda_1(B) \leq \dots \leq \lambda_{n-1}(B)$.

则我们有如下的交错性质:

$$\begin{aligned}
\lambda_1(A) \leq \lambda_1(B) \leq \lambda_2(A) \leq \dots \leq \lambda_{n-1}(A) \leq \lambda_{n-1}(B) \leq \lambda_n(A) \\
\Leftrightarrow \\
\lambda_i(A) \leq \lambda_i(B) \leq \lambda_{i+1}(A) \quad (\forall i = 1, \dots, n-1)
\end{aligned}$$

其中 $\lambda_i(A) = \lambda_i(B)$ 成立当且仅当存在非零向量 $z \in \mathbb{C}^{n-1}$ 使得:

$$\begin{cases} Bz = z\lambda_i(B) \\ Bz = z\lambda_i(A) \\ y^H z = 0. \end{cases}$$

而 $\lambda_i(B) = \lambda_{i+1}(A)$ 成立当且仅当存在非零向量 $z \in \mathbb{C}^{n-1}$ 使得:

$$\begin{cases} Bz = z\lambda_i(B) \\ Bz = z\lambda_{i+1}(A) \\ y^H z = 0. \end{cases}$$

若 B 没有与 y 正交的特征向量, 则上述不等式均为严格不等式.

Proof:

设 A 的特征值按非减次序排列: $\lambda_1 \leq \dots \leq \lambda_n$.

设 B 是 A 的任意一个 $n - 1$ 阶主子阵, 其特征值也按非减次序排列: $\mu_1 \leq \dots \leq \mu_{n-1}$.

设 λ 是 A 的 m 重特征值, 记 $\lambda_{i-m+1} = \dots = \lambda_i = \lambda$.

根据 Cauchy 交错定理可知:

$$\begin{aligned} \lambda = \lambda_{i-m+1} &\leq \mu_{i-m+1} \leq \lambda_{i-m+2} \leq \cdots \leq \lambda_{i-1} = \mu_{i-1} = \lambda_i = \lambda \\ &\Downarrow \\ \lambda = \lambda_{i-m+1} &= \mu_{i-m+1} = \lambda_{i-m+2} = \cdots = \lambda_{i-1} = \mu_{i-1} = \lambda_i = \lambda \end{aligned}$$

因此 λ 至少是 B 的 $m - 1$ 重特征值.

(存疑: 至多为 B 的 m 重特征值? 还是 $m + 1$ 重?)

Problem 3

(秩 1 Hermite 摆动的交错定理, Matrix Analysis 推论 4.3.9)

给定正整数 n .

设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 阵, $z \in \mathbb{C}^n$.

试证明对于任意 $1 \leq i \leq n - 1$ 都有:

$$\lambda_i(A) \leq \lambda_i(A + zz^H) \leq \lambda_{i+1}(A).$$

- Lemma: (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:

注意到 $n = 1$ 的情况是平凡的, $z = 0_n$ 的情况也是平凡的.

因此我们只需考虑 $n \geq 2$ 且 $z \neq 0_n$ 的情况即可.

此时 $B = zz^H$ 恰好有一个正特征值 $z^H z$.

因此我们一定有 $\lambda_1(B) = 0 = \lambda_{n-1}(B)$ 成立.

根据完整形式的 Weyl 不等式 (Lemma) 可知:

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

上式对某个 i 取等, 当且仅当 B 奇异且存在非零向量 x 使得:

$$\begin{cases} Ax = x\lambda_{i+1}(A) \\ Bx = 0_n \\ (A + B)x = x\lambda_i(A + B). \end{cases}$$

- ② 对于任意 $i = 1, \dots, n$ 都有 $\lambda_i(A) = \lambda_i(A) + \lambda_1(B) \leq \lambda_i(A + B)$ 成立.

上式对某个 i 取等, 当且仅当 B 奇异且存在非零向量 x 使得:

$$\begin{cases} Ax = x\lambda_i(A) \\ Bx = 0_n \\ (A + B)x = x\lambda_i(A + B). \end{cases}$$

上述结论可以写成更紧凑的形式:

$$\begin{aligned} \lambda_i(A) &\leq \lambda_i(A + B) \leq \lambda_{i+1}(A) \quad (i = 1, \dots, n - 1) \\ \lambda_n(A) &\leq \lambda_n(A + B) \end{aligned}$$

命题得证.

Problem 4

(Hermite 正定阵的 Schur 补更加良态)

给定正整数 m, n (满足 $1 \leq m < n$).

设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 正定阵, 且分块为:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (\text{where } A_{11} \in \mathbb{C}^{m \times m})$$

试证明 Schur 补 $S := A_{22} - A_{21}A_{11}^{-1}A_{12}$ 正定, 且有: ([reference](#))

$$\kappa_2(S) = \|S\|_2\|S^{-1}\|_2 \leq \|A\|_2\|A^{-1}\|_2 = \kappa_2(A)$$

- (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: $\lambda_{\min} = \lambda_1 \leq \dots \leq \lambda_n = \lambda_{\max}$

记 S 为 \mathbb{C}^n 的子空间, 则我们有:

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

Proof:

Hermite 分块矩阵的共轭对称行列 Gauss 消元可以表示为:

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

注意到这是一个合同变换.

根据 A 的 Hermite 正定性可知 $\tilde{A} = A_{11} \oplus (A_{22} - A_{21}A_{11}^{-1}A_{21}^H)$ 是 Hermite 正定阵.

因此 Schur 补 $S := A_{22} - A_{21}A_{11}^{-1}A_{21}^H$ 作为 \tilde{A} 的主子阵也是 Hermite 正定阵.

注意到:

$$\begin{aligned} \begin{bmatrix} A_{11} & A_{21}^H \\ A_{21} & A_{22} \end{bmatrix} &= \begin{bmatrix} 0_{m \times m} & \\ & A_{22} - A_{21}A_{11}^{-1}A_{21}^H \end{bmatrix} + \begin{bmatrix} A_{11} & A_{21}^H \\ A_{21} & A_{21}A_{11}^{-1}A_{21}^H \end{bmatrix} \\ &= \begin{bmatrix} 0_{m \times m} & \\ & A_{22} - A_{21}A_{11}^{-1}A_{21}^H \end{bmatrix} + \begin{bmatrix} A_{11}^{\frac{1}{2}} \\ A_{21}A_{11}^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} A_{11}^{\frac{1}{2}} & A_{11}^{-\frac{1}{2}}A_{21}^H \\ A_{21}A_{11}^{-\frac{1}{2}} & \end{bmatrix} \\ &= \begin{bmatrix} 0_{m \times m} & \\ & A_{22} - A_{21}A_{11}^{-1}A_{21}^H \end{bmatrix} + \begin{bmatrix} A_{11}^{\frac{1}{2}} \\ A_{21}A_{11}^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} A_{11}^{\frac{1}{2}} \\ A_{21}A_{11}^{-\frac{1}{2}} \end{bmatrix}^H \\ &= 0_{m \times m} \oplus S + XX^H \\ &= B + C \end{aligned}$$

其中 $X = \begin{bmatrix} A_{11}^{\frac{1}{2}} \\ A_{21}A_{11}^{-\frac{1}{2}} \end{bmatrix} \in \mathbb{C}^{n \times m}$ 的秩为 m , 因此 $C = XX^H$ 的秩为 $m < n$.

于是我们有 $\lambda_1(C) = 0$ 和 $\dim(\text{Ker}(Q)) = n - m > 0$.

设 Hermite 阵的特征值按非减次序排列.

一方面, 根据 Weyl 不等式我们有:

$$\begin{aligned} \lambda_n(A) - \lambda_n(B) &\geq \lambda_1(C) = 0 \\ &\Updownarrow \\ \lambda_n(B) &\leq \lambda_n(A) \end{aligned}$$

另一方面, 注意到:

$$\begin{aligned}
\lambda_1(A) &= \lambda_1(B + C) \quad (\text{use Rayleigh theorem}) \\
&= \min_{\|x\|_2=1} x^H(B + C)x \\
&\leq \min_{\substack{x \in \text{Ker}(C) \\ \|x\|_2=1}} x^H(B + C)x \\
&= \min_{\substack{x \in \text{Ker}(C) \\ \|x\|_2=1}} x^H B x \\
&\leq \max_{V \subset \mathbb{C}^n: \dim(V)=n-m} \min_{\substack{x \in V \\ \|x\|_2=1}} x^H B x \quad (\text{use Courant-Fischer theorem}) \\
&= \lambda_{m+1}(B)
\end{aligned}$$

综上所述，我们有：

$$\begin{aligned}
\kappa_2(S) &= \|S\|_2 \|S^{-1}\|_2 \\
&= \frac{\sigma_{\max}(S)}{\sigma_{\min}(S)} \quad (\text{note that } S \text{ is Hermitian positive definite matrix}) \\
&= \frac{\lambda_{\max}(S)}{\lambda_{\min}(S)} \quad (\text{note that } B = 0_{m \times m} \oplus S \text{ and } S \text{ is positive definite}) \\
&= \frac{\lambda_n(B)}{\lambda_{m+1}(B)} \quad (\text{note that } \begin{cases} \lambda_n(B) \leq \lambda_n(A) \\ \lambda_{m+1}(B) \geq \lambda_1(A) \end{cases}) \\
&\leq \frac{\lambda_n(A)}{\lambda_1(A)} \quad (\text{note that } A \text{ is Hermitian positive definite matrix}) \\
&= \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)} \\
&= \|A\|_2 \|A\|_2^{-1} \\
&= \kappa_2(A)
\end{aligned}$$

命题得证。

Problem 5

给定正整数 n .

设对称阵 $A \in \mathbb{R}^{n \times n}$ 的最小特征值 $\lambda_1(A) < 0$.

试证明：

$$\min_{x \in \mathbb{R}^n} \|A + xx^T\|_F^2 = \sum_{k=2}^n (\lambda_k(A))^2$$

上述结论可以推广到复数域。

这通过证法一比较方便，而通过证法二需要定义复数域的求导法则（基本上将转置变为共轭转置就行了）。

- **(Hoffman-Wielandt 不等式的最常用形式)**

设 $A, E \in \mathbb{C}^{n \times n}$ 均为 Hermite 阵。

设 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ 是 A 的以非减次序排列的特征值，

而 $\hat{\lambda}_1, \dots, \hat{\lambda}_n \in \mathbb{R}$ 是 $A + E$ 的以非减次序排列的特征值。

记 $\Delta\Lambda := \text{diag}\{\hat{\lambda}_1 - \lambda_1, \dots, \hat{\lambda}_n - \lambda_n\}$ ，

则我们有：

$$\sum_{i=1}^n |\hat{\lambda}_i - \lambda_i|^2 = \|\Delta\Lambda\|_F^2 \leq \|E\|_F^2 = \text{tr}(E^H E)$$

证法一：

设 A 的谱分解为 $A = U\Lambda U^T$ 。

根据 Frobenius 范数的酉不变性我们有：

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \|A + xx^T\|_F^2 &= \min_{x \in \mathbb{R}^n} \|U^T(A + xx^T)U\|_F^2 \\
&= \min_{x \in \mathbb{R}^n} \|\Lambda + (U^T x)(U^T x)^T\|_F^2 \quad (\text{denote } y := U^T x) \\
&= \min_{y \in \mathbb{R}^n} \|\Lambda + yy^T\|_F^2 \quad (\text{use Hoffman-Wielandt inequality}) \\
&\geq \min_{y \in \mathbb{R}^n} \sum_{i=1}^n |\lambda_i(\Lambda) - \lambda_i(-yy^T)|^2 \quad (\text{note that } \begin{cases} \lambda_i(\Lambda) = \lambda_i(A) & (i = 1, \dots, n) \\ \lambda_1(-yy^T) = -y^T y = -\|y\|_2^2 \\ \lambda_2(-yy^T) = \dots = \lambda_n(-yy^T) = 0 \end{cases}) \\
&= \min_{y \in \mathbb{R}^n} \left\{ [\lambda_1(A) - (-\|y\|_2^2)]^2 + \sum_{i=2}^n (\lambda_i(A))^2 \right\} \quad (\text{The minimum is reached when } \|y\|_2 = -\sqrt{\lambda_1(A)}) \\
&= \sum_{i=2}^n (\lambda_i(A))^2
\end{aligned}$$

上述不等号当且仅当 $y = \pm\sqrt{-\lambda_1(A)}e_1$ 时取等.

因此我们有:

$$\min_{x \in \mathbb{R}^n} \|A + xx^T\|_F^2 = \sum_{i=2}^n (\lambda_i(A))^2$$

证法二:

设 A 的谱分解为 $A = U\Lambda U^T$.

根据 Frobenius 范数的酉不变性我们有:

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \|A + xx^T\|_F^2 &= \min_{x \in \mathbb{R}^n} \|U^T(A + xx^T)U\|_F^2 \\
&= \min_{x \in \mathbb{R}^n} \|\Lambda + (U^T x)(U^T x)^T\|_F^2 \quad (\text{denote } y := U^T x) \\
&= \min_{y \in \mathbb{R}^n} \|\Lambda + yy^T\|_F^2 \\
&= \min_{y \in \mathbb{R}^n} \text{tr}\{(\Lambda + yy^T)(\Lambda + yy^T)\} \\
&= \min_{y \in \mathbb{R}^n} \text{tr}\{\Lambda^2 + 2yy^T\Lambda + yy^Tyy^T\} \\
&= \min_{y \in \mathbb{R}^n} \{2y^T\Lambda y + \|y\|_2^4\} + \text{tr}(\Lambda^2) \\
&= \min_{y \in \mathbb{R}^n} \{2y^T\Lambda y + \|y\|_2^4\} + \sum_{i=1}^n (\lambda_i(A))^2
\end{aligned}$$

要证明 $\min_{x \in \mathbb{R}^n} \|A + xx^T\|_F^2 = \sum_{k=2}^n (\lambda_k(A))^2$,
等价于证明 $\min_{y \in \mathbb{R}^n} \{2y^T\Lambda y + \|y\|_2^4\} = -(\lambda_1(A))^2$.

计算 $2y^T\Lambda y + \|y\|_2^4$ 关于 y 的梯度为:

$$\begin{aligned}
\nabla_y \{2y^T\Lambda y + \|y\|_2^4\} &= 4\Lambda y + 4\|y\|_2^2 y \\
&= 4(\Lambda + \|y\|_2^2)y
\end{aligned}$$

令 $\nabla_y \{2y^T\Lambda y + \|y\|_2^4\} = 4(\Lambda + \|y\|_2^2)y = 0_n$, 即有:

$$\begin{bmatrix} (\lambda_1 + \|y\|_2^2)y_1 \\ \vdots \\ (\lambda_n + \|y\|_2^2)y_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

设 $\lambda_1, \dots, \lambda_k$ 为负特征值(且互不相同), $\lambda_{k+1}, \dots, \lambda_n$ 为非负特征值.

则方程 $\lambda_i + \|y\|_2^2 = 0$ ($i = 1, \dots, k$) 只能有一个成立.

因此驻点只可能是 $y = \pm\sqrt{-\lambda_i}e_i$ ($i = 1, \dots, k$).

对应的目标函数值 $2y^T\Lambda y + \|y\|_2^4 = 2(-\lambda_i)\lambda_i + (-\lambda_i)^2 = -\lambda_i^2$ ($i = 1, \dots, k$).

注意到 $\lim_{|y_i| \rightarrow \infty} \{2y^T\Lambda y + \|y\|_2^4\} = \infty$ ($\forall i = 1, \dots, n$),

因此我们有:

$$\min_{y \in \mathbb{R}^n} \{2y^T\Lambda y + \|y\|_2^4\} = \min_{i=1, \dots, k} \{-(\lambda_i(A))^2\} = -(\lambda_1(A))^2$$

对于负特征值 $\lambda_1, \dots, \lambda_k$ 中有重特征值的情况,

我们只需使用一系列负特征值没有重特征值的矩阵 $\{\Lambda_k\}$ 来逼近 Λ 即可.

综上所述, 命题得证.

Problem 6 (optional)

(Matrix Computation 2.5.3 & 6.4.3 节: distance between equidimensional subspaces)

给定正整数 m, n (满足 $1 \leq m \leq n$).

设 \mathcal{X} 和 \mathcal{Y} 均为 \mathbb{C}^n 的 m 维子空间, 对应的正交投影算子为 $P_{\mathcal{X}}$ 和 $P_{\mathcal{Y}}$.

定义 \mathcal{X} 和 \mathcal{Y} 之间的距离为:

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

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

Solution:

设 \mathcal{X} 的一组标准正交基为 $\{x_1, \dots, x_m\}$ (记 $X = [x_1, \dots, x_m] \in \mathbb{C}^{n \times m}$).

设 \mathcal{Y} 的一组标准正交基为 $\{y_1, \dots, y_m\}$ (记 $Y = [y_1, \dots, y_m] \in \mathbb{C}^{n \times m}$).

定义矩阵 $A = [\langle x_i, y_j \rangle] = X^H Y \in \mathbb{C}^{m \times m}$.

设其奇异值分解为 $A = U \Sigma V^H$,

其中奇异值 $0 \leq \sigma_1 \leq \dots \leq \sigma_m \leq 1$ 按非减次序排列 (注意: 与奇异值的约定习惯相反).

我们定义 $\angle_k(\mathcal{X}, \mathcal{Y}) = \arccos(\sigma_k)$ 为 \mathcal{X} 和 \mathcal{Y} 之间的第 k 个经典角 (principal/canonical angle).

不失一般性, 设 \mathcal{X} 和 \mathcal{Y} 对应的正交投影算子 $P_{\mathcal{X}}$ 和 $P_{\mathcal{Y}}$ 为:

$$\begin{aligned} P_{\mathcal{X}} &:= XX^H \\ P_{\mathcal{Y}} &:= YY^H \end{aligned}$$

将 X, Y 补全为:

$$\begin{aligned} \tilde{X} &:= [X, X_{\perp}] \\ \tilde{Y} &:= [Y, Y_{\perp}] \end{aligned}$$

可以证明: (Matrix Computation 定理 2.5.1)

(存疑: 中间的步骤有些不严谨)

$$\begin{aligned} \text{dist}(\mathcal{X}, \mathcal{Y}) &:= \|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2 \\ &= \|X^H Y_{\perp}\|_2 \\ &= \|X^H (\tilde{Y} - [Y, 0_{n \times m}])\|_2 \\ &= \|X^H (I_n - YY^H) \tilde{Y}\|_2 \\ &= \|X^H (I_n - YY^H)\|_2 \\ &= \|XX^H (I_n - YY^H)\|_2 \\ &= \|P_{\mathcal{X}} (I_n - P_{\mathcal{Y}})\|_2 \\ &= \sqrt{1 - \cos^2(\angle_1(\mathcal{X}, \mathcal{Y}))} \\ &= \sin(\angle_1(\mathcal{X}, \mathcal{Y})) \end{aligned}$$

即第一经典角的正弦值.

The End