# 期中考试 (2024 秋)

2024 年 11 月 4 日 (8 : 00 ～ 9 : 40)

## Problem 1

设 $A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}$，试证明 $\det(A\otimes B) = \det(B\otimes A)$.
其中 $\otimes$ 代表 Kronecker 乘积.

- 这个结论是显然的:
  $A\otimes B$ 和 $B\otimes A$ 是置换相似的，
  即存在置换矩阵 $P \in \mathbb{R}^{mn\times mn}$ (满足 $P^{-1} = P^{\mathrm{T}}$) 使得 $P^{\mathrm{T}}(B\otimes A)P = (A\otimes B)$
  因此我们有:

$$\begin{aligned}\det(A\otimes B) &= \det\left(P^{\mathrm{T}}(B\otimes A)P\right)\\ &= \det\left((B\otimes A)PP^{\mathrm{T}}\right) \quad (\text{note that } P^{-1} = P^{\mathrm{T}})\\ &= \det(B\otimes A)\end{aligned}$$

  实际上这源自于从 $B\otimes A$ 得到 $A\otimes B$ 的行列重排是对称的，因此行列重排次数是偶数次，行列式不变.
  但上述讨论并不严谨，也不本质.

- Kronecker 乘积的性质:

$$(A\otimes B)(C\otimes D) = (AC)\otimes(BD) \text{ if the matrix products } AB \text{ and } CD \text{ are well-defined}$$
$$\begin{cases}(A\otimes B)^{\mathrm{T}} = A^{\mathrm{T}}\otimes B^{\mathrm{T}}\\ (A\otimes B)^{-1} = A^{-1}\otimes B^{-1}\\ (A\otimes B)^{\dagger} = A^{\dagger}\otimes B^{\dagger}\end{cases}$$

---

**解法一:**
设 $A$ 的 Jordan 分解为 $A = P_1J_1P_1^{-1}$，$B$ 的 Jordan 分解为 $B = P_2J_2P_2^{-1}$.
设 $A$ 的特征值是 $\lambda_1,\dots,\lambda_m \in \mathbb{C}$ (于是 $J_1$ 的对角元为 $\lambda_1,\dots,\lambda_m$).
设 $B$ 的特征值是 $\mu_1,\dots,\mu_n \in \mathbb{C}$ (于是 $J_2$ 的对角元为 $\mu_1,\dots,\mu_n$).
则我们有:

$$\begin{aligned}\det(A\otimes B) &= \det\left((P_1J_1P_1^{-1})\otimes(P_2J_2P_2^{-1})\right)\\ &= \det\left((P_1\otimes P_2)(J_1\otimes J_2)(P_1^{-1}\otimes P_2^{-1})\right)\\ &= \det\left((P_1\otimes P_2)(J_1\otimes J_2)(P_1\otimes P_2)^{-1}\right)\\ &= \det\left((J_1\otimes J_2)(P_1\otimes P_2)^{-1}(P_1\otimes P_2)\right)\\ &= \det(J_1\otimes J_2)\\ &= \prod_{i=1}^{m}\left\{\lambda_i\prod_{j=1}^{n}\mu_j\right\}\\ &= \prod_{i,j=1}^{m,n}\lambda_i\mu_j\\ &= \prod_{j=1}^{n}\left\{\mu_j\prod_{i=1}^{m}\lambda_i\right\}\\ &= \det(J_2\otimes J_1)\\ &= \det\left((J_2\otimes J_1)(P_2\otimes P_1)^{-1}(P_2\otimes P_1)\right)\\ &= \det\left((P_2\otimes P_1)(J_2\otimes J_1)(P_2\otimes P_1)^{-1}\right)\\ &= \det\left((P_2\otimes P_1)(J_2\otimes J_1)(P_2^{-1}\otimes P_1^{-1})\right)\\ &= \det\left((P_2J_2P_2^{-1})\otimes(P_1J_1P_1^{-1})\right)\\ &= \det(B\otimes A)\end{aligned}$$

---

事实上我们有更本质的解法，
这个解法表明 $\{\mathrm{eig}(A\otimes B)\} = \{\mathrm{eig}(B\otimes A)\}$ if $A, B$ are square matrices.

**解法二:**
设 $A$ 的特征值是 $\lambda_1,\dots,\lambda_m\in\mathbb{C}$，$B$ 的特征值是 $\mu_1,\dots,\mu_n\in\mathbb{C}$.
可以证明 $A\otimes B$ 的特征值是 $\lambda_i\mu_j\ (i=1,\dots,m,\ j=1,\dots,n)$

- 设 $(x,\lambda)$ 是 $A$ 的特征对 (即满足 $Ax=x\lambda$)，而 $(y,\mu)$ 是 $B$ 的特征对 (即满足 $By=y\mu$)，
  则我们有:
  $$\begin{aligned}(A\otimes B)(x\otimes y)&=(Ax)\otimes(By)\\&=(x\lambda)\otimes(y\mu)\\&=(x\otimes y)(\lambda\mu).\end{aligned}$$
  因此 $(x\otimes y,\lambda\mu)$ 就是 $A\otimes B$ 的特征对.
  这说明 $A\otimes B$ 的特征值是 $\lambda_i\mu_j\ (i=1,\dots,m,\ j=1,\dots,n)$.

类似地，我们可以说明 $B\otimes A$ 的特征值是 $\mu_j\lambda_i\ (i=1,\dots,m,\ j=1,\dots,n)$
因此我们有:

$$\begin{aligned}\det(A\otimes B)&=\prod_{i,j=1}^{m,n}\lambda_i\mu_j\\&=\prod_{i,j=1}^{m,n}\mu_j\lambda_i\\&=\det(B\otimes A)\end{aligned}$$

命题得证.

---

事实上，当 $A\in\mathbb{C}^{m\times n},B\in\mathbb{C}^{n\times m}$ 时 $\det(A\otimes B)=\det(B\otimes A)$ 也是成立的.
**证明:**
不妨设 $m<n$，则 $A$ 的列一定线性相关，因此存在 $u\neq 0_n\in\mathbb{C}^n$ 使得 $Au=0_m$
对于任意非零向量 $v\neq 0_m\in\mathbb{C}^m$ 我们都有:

$$\begin{aligned}(A\otimes B)(u\otimes v)&=(Au)\otimes(Bv)=0_n\otimes(Bv)=0_{mn}\\(B\otimes A)(v\otimes u)&=(Bv)\otimes(Au)=(Bv)\otimes 0_n=0_{mn}\end{aligned}$$

注意到 $(u\otimes v)$ 和 $(v\otimes u)$ 均为 $mn$ 维非零向量，故 $(A\otimes B)$ 和 $(B\otimes A)$ 均为奇异矩阵.
因此我们有 $\det(A\otimes B)=\det(B\otimes A)=0$ 成立.

# Problem 2

设 $A\in\mathbb{C}^{m\times n}$，试证明 $\|A\|_2^2\leq\|A\|_1\|A\|_\infty$.

**解法一:**

- **Note:** 根据列和范数 $\|\cdot\|_1$ 与行和范数 $\|\cdot\|_\infty$ 的定义，我们有 $\|A^{\mathrm{H}}\|_1=\|A^{\mathrm{T}}\|_1=\|A\|_\infty$ 成立.
- **Lemma (谱半径定理):**
  设 $\|\cdot\|$ 是 $\mathbb{C}^{n\times n}$ 上的相容范数.
  若 $\lambda$ 是 $A\in\mathbb{C}^{n\times n}$ 的一个特征值，则我们有 $|\lambda|\leq\rho(A)\leq\|A\|$ 成立.

使用谱半径定理以及 $\|A^{\mathrm{H}}\|_1=\|A^{\mathrm{T}}\|_1=\|A\|_\infty$ 的恒等式，我们有:

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

---

**解法二 (存疑):**
首先注意到:

$$
\begin{aligned}
\|x\|_2^2 &= \sum_{i=1}^n |x_i|^2 \\
&\le \max_{1\le i\le n} |x_i| \cdot \sum_{i=1}^n |x_i| \\
&= \|x\|_\infty \|x\|_1
\end{aligned}
$$

(值得注意的是，直接使用 **Holder 不等式** $|x^{\mathrm{H}}y| \le \|x\|_p\|x\|_q$ $(\text{where } \frac{1}{p}+\frac{1}{q}=1)$ 得到上述不等式是不严谨的)

因此我们有:
(其中 $\max$ 操作符与 $\sqrt{\cdot}$ 操作符的可交换性依赖于 $f(x)=\sqrt{x}$ 是 $\mathbb{R}_+ \mapsto \mathbb{R}_+$ 的双射)

$$
\begin{aligned}
\|A\|_2 &= \max_{\|x\|_2=1} \|Ax\|_2 \\
&\le \max_{\|x\|_2=1} \sqrt{\|Ax\|_1\|Ax\|_\infty} \\
&= \sqrt{\max_{\|x\|_2=1} \|Ax\|_1\|Ax\|_\infty} \\
&\le \sqrt{\max_{\|x\|_2=1} \|Ax\|_1 \cdot \max_{\|x\|_2=1} \|Ax\|_\infty} \\
&= \sqrt{\|A\|_1\|A\|_\infty}
\end{aligned}
$$

于是我们有 $\|A\|_2^2 \le \|A\|_1\|A\|_\infty$ 成立.

# Problem 3

设 $A,B\in\mathbb{C}^{n\times n}$ 非奇异，试证明 $\text{rank}(A^{-1}-B^{-1})=\text{rank}(A-B)$

**Proof:**
注意到:

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

由于 $A^{-1},B^{-1}$ 非奇异，故我们有 $\text{rank}(A^{-1}-B^{-1})=\text{rank}(A-B)$ 成立.

---

不放心的话，我们可以这样说明:

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

因此我们有:

$$
\begin{aligned}
\text{rank}(A^{-1}-B^{-1}) &= \text{rank}(A^{-1}(B-A)B^{-1}) \\
&\le \min\{\text{rank}(A^{-1}),\text{rank}(B-A),\text{rank}(B^{-1})\} \\
&= \min\{n,\text{rank}(B-A),n\} \\
&= \text{rank}(B-A) \\
&= \text{rank}(A-B) \\
\hline
\text{rank}(A-B) &= \text{rank}(B-A) \\
&= \text{rank}(A(A^{-1}-B^{-1})B) \\
&\le \min\{\text{rank}(A),\text{rank}(A^{-1}-B^{-1}),\text{rank}(B)\} \\
&= \min\{n,\text{rank}(A^{-1}-B^{-1}),n\} \\
&= \text{rank}(A^{-1}-B^{-1})
\end{aligned}
$$

于是我们有 $\text{rank}(A^{-1}-B^{-1})=\text{rank}(A-B)$ 成立.

---

**另解: (待纠正)**
SMW 公式 $(A+UV^{\mathrm{H}})^{-1}=A^{-1}-A^{-1}U(I+V^{\mathrm{H}}A^{-1}U)^{-1}V^{\mathrm{H}}A^{-1}$
注意到 $A+UV^{\mathrm{H}}=A^{-1}(I+A^{-1}UV^{\mathrm{H}})$ 可逆时，$I+V^{\mathrm{H}}A^{-1}U$ 也可逆 (矩阵乘积交换的谱不变性)
设 $B-A=UV^{\mathrm{H}}$
于是有:

$$
\begin{aligned}
B^{-1} - A^{-1} &= -A^{-1}U(I + V^{\mathrm{H}}A^{-1}U)^{-1}V^{\mathrm{H}}A^{-1} \\
\mathrm{rank}(B^{-1} - A^{-1}) &= \mathrm{rank}(UV^{\mathrm{H}}) = \mathrm{rank}(B - A)
\end{aligned}
$$

## Problem 4

设 $A, B \in \mathbb{C}^{n\times n}$ 满足 $\mathrm{rank}(A) + \mathrm{rank}(B) < n$，试证明 $A, B$ 存在公共特征对.

**解法一:**
直观上我们有:

$$
\mathrm{rank}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) \le \mathrm{rank}\left(\begin{bmatrix} A & \\ & B \end{bmatrix}\right) = \mathrm{rank}(A) + \mathrm{rank}(B) < n
$$

因此线性方程组 $\begin{bmatrix} A \\ B \end{bmatrix} x = \begin{bmatrix} 0_n \\ 0_n \end{bmatrix}$ 有非零解 $x$，于是就得到公共特征对 $(x, 0)$，满足 $\begin{cases} Ax = 0_n = x\cdot 0 \\ Bx = 0_n = x\cdot 0 \end{cases}$

**解法二:**
这个解法与解法一没有本质的不同，仅仅是陈述方式有区别而已.
我们有:

$$
\begin{aligned}
&\dim(\mathrm{Ker}(A)\cap \mathrm{Ker}(B)) \\
&= \dim(\mathrm{Ker}(A)) + \dim(\mathrm{Ker}(B)) - \dim(\mathrm{Ker}(A) + \mathrm{Ker}(B)) \\
&= (n - \mathrm{rank}(A)) + (n - \mathrm{rank}(B)) - \dim(\mathrm{Ker}(A) + \mathrm{Ker}(B)) \\
&= 2n - (\mathrm{rank}(A) + \mathrm{rank}(B)) - \dim(\mathrm{Ker}(A) + \mathrm{Ker}(B)) \quad \left(\text{note that } \begin{cases} \mathrm{rank}(A) + \mathrm{rank}(B) < n \\ \dim(\mathrm{Ker}(A) + \mathrm{Ker}(B)) \le n \end{cases}\right) \\
&> 2n - n - n \\
&= 0
\end{aligned}
$$

因此 $\mathrm{Ker}(A)\cap\mathrm{Ker}(B) \ne \{0_n\}$
即存在非零向量 $x \in \mathrm{Ker}(A)\cap\mathrm{Ker}(B)$
于是 $x$ 满足 $\begin{cases} Ax = 0_n = x\cdot 0 \\ Bx = 0_n = x\cdot 0 \end{cases}$，表明 $(x, 0)$ 是 $A, B$ 的公共特征对.

## Problem 5

求解以下矩阵的 Jordan 标准型:

$$
\exp(A) := \exp\left(\begin{bmatrix} 1 & 0 & 0 & 1 & & \\ & 1 & 0 & 0 & 2 & \\ & & 2 & 0 & 0 & 3 \\ & & & 2 & 0 & 0 \\ & & & & 3 & 0 \\ 4 & & & & & 3 \end{bmatrix}\right)
$$

- 直观上我们很容易看出 $A$ 的特征值至少有 $1, 2, 3$
  因为对角元 $1, 2, 3$ 均有独占 $A$ 的某一行或一列，因此 $A - I, A - 2I, A - 3I$ 的行列式必然为零.

**Solution:**
首先求解 $A$ 的特征值:

$$
\begin{aligned}
&\det(\lambda I - A)\\
&= \begin{vmatrix}
\lambda-1 & 0 & 0 & -1 & & \\
 & \lambda-1 & 0 & 0 & -2 & \\
 & & \lambda-2 & 0 & 0 & -3 \\
 & & & \lambda-2 & 0 & 0 \\
 & & & & \lambda-3 & 0 \\
-4 & & & & & \lambda-3
\end{vmatrix}\\
&= (\lambda-1)\begin{vmatrix}
\lambda-1 & 0 & 0 & -2 & \\
 & \lambda-2 & 0 & 0 & -3 \\
 & & \lambda-2 & 0 & 0 \\
 & & & \lambda-3 & 0 \\
 & & & & \lambda-3
\end{vmatrix} + (-1)^{6+1}(-4)\begin{vmatrix}
0 & 0 & -1 & & \\
\lambda-1 & 0 & 0 & -2 & \\
 & \lambda-2 & 0 & 0 & -3 \\
 & & \lambda-2 & 0 & 0 \\
 & & & \lambda-3 & 0
\end{vmatrix}\\
&= (\lambda-1)(\lambda-1)(\lambda-2)(\lambda-2)(\lambda-3)(\lambda-3) + 4\cdot(-1)^{2+1}(\lambda-1)\begin{vmatrix}
0 & -1 & & \\
\lambda-2 & 0 & 0 & -3 \\
 & \lambda-2 & 0 & 0 \\
 & & \lambda-3 & 0
\end{vmatrix}\\
&= (\lambda-1)^2(\lambda-2)^2(\lambda-3)^2 - 4(\lambda-1)(-1)^{2+1}(\lambda-2)\begin{vmatrix}
-1 & & \\
\lambda-2 & 0 & 0 \\
 & \lambda-3 & 0
\end{vmatrix}\\
&= (\lambda-1)^2(\lambda-2)^2(\lambda-3)^2 + 4(\lambda-1)(\lambda-2)\cdot 0\\
&= (\lambda-1)^2(\lambda-2)^2(\lambda-3)^2
\end{aligned}
$$

令 $\det(\lambda I - A) = (\lambda-1)^2(\lambda-2)^2(\lambda-3)^2 = 0$
便解得 $A$ 的特征值为 $1, 2, 3$，代数重数均为 $2$

下面求解特征向量，得到几何重数:

- 考虑特征值 $1$:

$$
A - I = \begin{bmatrix}
0 & 0 & 0 & 1 & & \\
 & 0 & 0 & 0 & 2 & \\
 & & 1 & 0 & 0 & 3 \\
 & & & 1 & 0 & 0 \\
 & & & & 2 & 0 \\
4 & & & & & 2
\end{bmatrix} \Rightarrow \mathrm{Ker}(A-I) = \mathrm{span}\left\{\begin{bmatrix}1\\0\\6\\0\\0\\-2\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\\0\end{bmatrix}\right\}
$$

  因此特征值 $1$ 的几何重数是 $2$
  也可根据 $\mathrm{rank}(A-I) = 4$ 反推特征值 $1$ 的几何重数是 $2$

- 考虑特征值 $2$:

$$
A - 2I = \begin{bmatrix}
-1 & 0 & 0 & 1 & & \\
 & -1 & 0 & 0 & 2 & \\
 & & 0 & 0 & 0 & 3 \\
 & & & 0 & 0 & 0 \\
 & & & & 1 & 0 \\
4 & & & & & 1
\end{bmatrix} \Rightarrow \mathrm{Ker}(A-I) = \mathrm{span}\left\{\begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}\right\}
$$

  因此特征值 $2$ 的几何重数是 $1$
  也可根据 $\mathrm{rank}(A-I) = 5$ 反推特征值 $2$ 的几何重数是 $1$

- 考虑特征值 $3$:

$$
A - I = \begin{bmatrix}
-2 & 0 & 0 & 1 & & \\
 & -2 & 0 & 0 & 2 & \\
 & & -1 & 0 & 0 & 3 \\
 & & & -1 & 0 & 0 \\
 & & & & 0 & 0 \\
4 & & & & & 0
\end{bmatrix} \Rightarrow \mathrm{Ker}(A-I) = \mathrm{span}\left\{\begin{bmatrix}0\\0\\3\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\1\\0\end{bmatrix}\right\}
$$

  因此特征值 $3$ 的几何重数是 $2$
  也可根据 $\mathrm{rank}(A-3I) = 4$ 反推特征值 $3$ 的几何重数是 $2$

综上所述，我们得到 $A$ 的 Jordan 标准型 $J$ 为:

$$J=\begin{bmatrix}1&&&&&\\&1&&&&\\&&2&1&&\\&&&2&&\\&&&&3&\\&&&&&3\end{bmatrix}$$

因此 $\exp{(J)}$ 为:

$$\exp{(J)}=\begin{bmatrix}e&&&&&\\&e&&&&\\&&e^2&e^2&&\\&&&e^2&&\\&&&&e^3&\\&&&&&e^3\end{bmatrix}$$

要求 $\exp(A)$ 的 Jordan 矩阵，即要求 $\exp{(J)}$ 的 Jordan 矩阵.
我们取对角相似变换 $D=\operatorname{diag}\{1,1,1,\frac{1}{e^2},1,1\}$ 即得到 $\exp(J)$ 的 Jordan 标准型:

$$\begin{aligned}J_{\exp(\mathrm{A})}&:=D^{-1}\exp{(J)}D\\&=\begin{bmatrix}1&&&&&\\&1&&&&\\&&1&&&\\&&&e^2&&\\&&&&1&\\&&&&&1\end{bmatrix}\begin{bmatrix}e&&&&&\\&e&&&&\\&&e^2&e^2&&\\&&&e^2&&\\&&&&e^3&\\&&&&&e^3\end{bmatrix}\begin{bmatrix}1&&&&&\\&1&&&&\\&&1&&&\\&&&\frac{1}{e^2}&&\\&&&&1&\\&&&&&1\end{bmatrix}\\&=\begin{bmatrix}e&&&&&\\&e&&&&\\&&e^2&1&&\\&&&e^2&&\\&&&&e^3&\\&&&&&e^3\end{bmatrix}\end{aligned}$$

# Problem 6

设 $A\in\mathbb{C}^{n\times n}$ 为 Hermite 半正定阵，试证明存在下三角阵 $L\in\mathbb{C}^{n\times n}$ 使得 $A=LL^{\mathrm{H}}$

**解法一:**
注意到 $A$ 是一个 Hermite 阵 (自然是正规矩阵)，故 $A$ 可酉对角化.
即存在酉矩阵 $U\in\mathbb{C}^{n\times n}$ 使得 $\Lambda=U^{\mathrm{H}}AU$ 为对角阵，且对角元均为实数.
又注意到 $A$ 是半正定的，故 $\Lambda=\operatorname{diag}\{\lambda_1,\ldots,\lambda_n\}$ 的对角元均为非负实数.

定义: $\begin{cases}\Lambda^{\frac{1}{2}}=\operatorname{diag}\{\sqrt{\lambda_1},\ldots,\sqrt{\lambda_n}\}\\A^{\frac{1}{2}}=U\Lambda^{\frac{1}{2}}U^{\mathrm{H}}\end{cases}$
显然 $A^{\frac{1}{2}}$ 也是 Hermite 半正定阵
设其 $\mathrm{QR}$ 分解为 $A^{\frac{1}{2}}=QR$
其中 $Q\in\mathbb{C}^{n\times n}$ 为酉矩阵，而 $R\in\mathbb{C}^{n\times n}$ 为具有非负实对角元的上三角阵 (不唯一)
则我们有:

$$
\begin{aligned}
A &= U\Lambda U^{\mathrm{H}} \\
&= U\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}U^{\mathrm{H}} \\
&= U\Lambda^{\frac{1}{2}}U^{\mathrm{H}}U\Lambda^{\frac{1}{2}}U^{\mathrm{H}} \\
&= A^{\frac{1}{2}}A^{\frac{1}{2}} \quad (\text{note that } A^{\frac{1}{2}} \text{ is also Hermitian}) \\
&= (A^{\frac{1}{2}})^{\mathrm{H}}A^{\frac{1}{2}} \quad (\text{note that } A^{\frac{1}{2}} = QR) \\
&= (QR)^{\mathrm{H}}(QR) \\
&= R^{\mathrm{H}}Q^{\mathrm{H}}QR \quad (\text{note that } Q^{\mathrm{H}}Q = I_n) \\
&= R^{\mathrm{H}}R
\end{aligned}
$$

记 $L := R^{\mathrm{H}}$，则我们有 $A = R^{\mathrm{H}}R = LL^{\mathrm{H}}$ (不唯一)
命题得证.

**解法二:**
我们对 $A$ 施加小扰动 $\varepsilon I$ $(0<\varepsilon<1)$ 得到 Hermite 正定阵 $A+\varepsilon I$
根据课上的结论，存在唯一的下三角阵 $L(\varepsilon) \in \mathbb{C}^{n\times n}$ 使得 $A+\varepsilon I = L(\varepsilon)L(\varepsilon)^{\mathrm{H}}$
注意到 $\|L(\varepsilon)\|_2^2 = \|L(\varepsilon)L(\varepsilon)^{\mathrm{H}}\|_2 = \|A+\varepsilon I\|_2 \leq \|A+I\|_2$
因此 $L(\varepsilon)$ $(0<\varepsilon<1)$ 关于 $\varepsilon$ 是有界的.
于是存在收敛子列 $\{L(\varepsilon_n)\}_{n=1}^{\infty}$ (其中 $\lim_{n\to\infty}\varepsilon_n = 0$)，其极限 $L$ 自然也是下三角阵，满足 $A = LL^{\mathrm{H}}$.

**解法三:**
我们对 Hermite 半正定阵 $A$ 进行 Gauss 消元，便可得到 $A = LU$
其中 $L$ 为单位下三角阵，而 $U$ 为上三角阵.
值得说明的是，若出现对角元为 $0$ 的情况，
则根据 $A$ 的半正定性可知这个 $0$ 所在行列的所有元素均为 $0$.
因此我们可以将全 $0$ 行列重排到左上角，将矩阵降阶，继续对降阶后的矩阵进行 Gauss 消元.

将 $U$ 的对角元提取出来得到 $D$，根据 $A$ 的 Hermite 性我们断言 $U = DL^{\mathrm{H}}$
即有 $A = LDL^{\mathrm{H}} = (LD^{\frac{1}{2}})(LD^{\frac{1}{2}})^{\mathrm{H}} = \widetilde{L}\widetilde{L}^{\mathrm{H}}$
其中 $\widetilde{L} := LD^{\frac{1}{2}}$ 即为所求的下三角阵.

# Problem 7

设 $A \in \mathbb{C}^{n\times n}$，试证明级数 $\sum_{k=0}^{\infty}(A^{\mathrm{H}})^kA^k$ 收敛当且仅当 $\rho(A)<1$

- 尽管谱半径函数 $\rho(\cdot)$ 并不是 $\mathbb{C}^{n\times n}$ 上的相容范数
  但对于任意 $A \in \mathbb{C}^{n\times n}$，$\rho(A)$ 都是所有相容范数取值的下确界.
  **(Matrix Analysis 引理 $5.6.10$)**
  任意给定 $A \in \mathbb{C}^{n\times n}$ 和 $\varepsilon>0$，都存在一个 $\mathbb{C}^{n\times n}$ 上的相容范数 $\|\cdot\|_{\varepsilon}$ 使得 $\rho(A) \leq \|A\|_{\varepsilon} \leq \rho(A)+\varepsilon$
- 尽管说当 $k\to\infty$ 时 $A^k$ 的单个元素的性状与 $\rho(A)^k$ 的性状相仿是不够准确的，
  但对于任意相容范数 $\|\cdot\|$，序列 $\{\|A^k\|\}$ 的确都有这个渐近性质.
  **(Gelfand 公式, Matrix Analysis 推论 $5.6.14$)**
  若 $\|\cdot\|$ 是 $\mathbb{C}^{n\times n}$ 上的一个相容范数，则对于任意 $A \in \mathbb{C}^{n\times n}$ 我们都有:

  $$\rho(A) = \lim_{k\to\infty}\|A^k\|^{\frac{1}{k}}$$

  这表明极限情况下特征值决定一切.

**Proof:**

- **必要性:**
  设级数 $\sum_{k=0}^{\infty}(A^{\mathrm{H}})^kA^k$ 收敛，及其极限为 $S := \sum_{k=0}^{\infty}(A^{\mathrm{H}})^kA^k$
  对于 $A$ 的任意特征对 $(x,\lambda)$，我们都有:

  $$
  \begin{aligned}
  x^{\mathrm{H}}Sx &= x^{\mathrm{H}}\left\{\sum_{k=0}^{\infty}(A^{\mathrm{H}})^kA^k\right\}x \\
  &= \sum_{k=0}^{\infty}\bar{\lambda}^k\lambda^k \\
  &= \sum_{k=0}^{\infty}|\lambda|^{2k} \\
  &< \infty
  \end{aligned}
  $$

于是有 $|\lambda| < 1$
注意到 $\lambda$ 可以是 $A$ 的任意特征值，因此 $\rho = \max_{1\leq i\leq n} |\lambda_i| < 1$

- **充分性:**
  设 $\rho(A) < 1$
  根据 **Matrix Analysis 引理** $5.6.10$ 可知存在相容范数 $\|\cdot\|$ 使得 $\|A\| < 1$

  注意到 $\rho(A^{\mathrm{H}}) = \rho(A) < 1$
  因此根据 **Gelfand 公式**可知 $\lim_{k\to\infty} \|(A^{\mathrm{H}})^k\| = \lim_{k\to\infty}(\rho(A^{\mathrm{H}}))^k = 0$
  这表明 $\{\|(A^{\mathrm{H}})^k\|\}$ 是有界序列，即存在 $M > 0$ 使得 $\|(A^{\mathrm{H}})^k\| \leq M\ (\forall\, k \in \mathbb{Z}_+)$

  对于任意 $\varepsilon > 0$，存在正整数 $N = \lceil \log_{\|A\|}(\frac{\varepsilon}{M}(1-\|A\|)) \rceil$ 使得:

$$\begin{aligned}
\left\|\sum_{k=N}^{\infty}(A^{\mathrm{H}})^k A^k\right\| &\leq \sum_{k=N}^{\infty}\|(A^{\mathrm{H}})^k\|\|A^k\| \quad (\text{note that } \|(A^{\mathrm{H}})^k\| \leq M\ (\forall\, k \in \mathbb{Z}_+)) \\
&= M\sum_{k=N}^{\infty}\|A^k\| \\
&\leq M\sum_{k=N}^{\infty}\|A\|^k \\
&= M\cdot\|A\|^N\cdot\frac{1}{1-\|A\|} \quad (\text{note that } 0 < \|A\| < 1 \text{ and } N = \lceil \log_{\|A\|}(\frac{\varepsilon}{M}(1-\|A\|)) \rceil) \\
&\leq M\cdot\frac{\varepsilon}{M}(1-\|A\|)\cdot\frac{1}{1-\|A\|} \\
&= \varepsilon
\end{aligned}$$

  因此级数 $\sum_{k=0}^{\infty}(A^{\mathrm{H}})^k A^k$ 收敛.

命题得证.

# Problem 8

给定正整数 $m, n$，设 $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$ 为 Hermite 正定阵.
试证明:

$$\det\left(\frac{A_1 + \cdots + A_m}{m}\right) \geq \det(A_1)^{\frac{1}{m}} \cdots \det(A_m)^{\frac{1}{m}}$$

**解法一:**
对数-行列式函数 (Log-determinant) $\begin{cases} f(X) = \log(\det(X)) \\ \mathrm{dom}(f) = \mathbb{S}_{++}^n \end{cases}$ **是凹函数**.
我们证明其凹性的方法是将其转化为任意 "直线"上的单变量函数:
对于任意给定的 $\begin{cases} X \in \mathbb{S}_{++}^n \\ V \in \mathbb{S}^n \end{cases}$ 考虑关于 $t$ 的函数 $\begin{cases} g(t) = f(X+tV) \\ \mathrm{dom}(g) = \{t \in \mathbb{R} : X + tV \in \mathrm{dom}(f) = \mathbb{S}_{++}^n\} \end{cases}$
记 $X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$ 的特征值为 $\lambda_1, \ldots, \lambda_n$，则对于任意 $t \in \mathrm{dom(g)}$ 都有:

$$\begin{aligned}
g(t) &= f(X+tV) \\
&= \log(\det(X+tV)) \\
&= \log\{\det(X^{\frac{1}{2}}(I_n + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})X^{\frac{1}{2}})\} \\
&= \log\{\det((I_n + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})X)\} \\
&= \sum_{i=1}^{n}\log(1+t\lambda_i) + \log(\det(X))
\end{aligned}$$

故有 $\begin{cases} g'(t) = \sum_{i=1}^n \frac{\lambda_i}{1+t\lambda_i} \\ g''(t) = -\sum_{i=1}^n \frac{\lambda_i^2}{(1+t\lambda_i)^2} \leq 0 \end{cases}$
说明任意 "直线"上的单变量函数 $g(t)$ 都是凹函数，因此 $f(X)$ 为凹函数.

因此我们有:

$$
\begin{aligned}
\log\left(\det\left(\frac{A_1+\cdots+A_m}{m}\right)\right) &= f\left(\frac{A_1+\cdots+A_m}{m}\right) \quad (\text{use concavity of } f(X)=\log(\det(X))) \\
&\geq \frac{1}{m}f(A_1)+\cdots+\frac{1}{m}f(A_m) \\
&= \frac{1}{m}\log(\det(A_1))+\cdots+\frac{1}{m}\log(\det(A_m)) \\
&= \log\left(\det(A_1)^{\frac{1}{m}}\cdots\det(A_m)^{\frac{1}{m}}\right)
\end{aligned}
$$

因此我们有 $\det\left(\frac{A_1+\cdots+A_m}{m}\right) \geq \det(A_1)^{\frac{1}{m}}\cdots\det(A_m)^{\frac{1}{m}}$ 成立.

---

**解法二:**

- 首先我们证明两个 Hermite 正定阵 $A, B$ 可以同时合同对角化:
  设 $A$ 的 Cholesky 分解为 $A = LL^{\mathrm{H}}$，其中 $L$ 为 (唯一的) 具有正实数对角元的下三角阵.
  注意到 $L^{-1}BL^{-H}$ 仍为 Hermite 阵，因而有谱分解 $L^{-1}BL^{-H} = Q\Lambda Q^{\mathrm{H}}$
  记 $C = LQ$，则我们有:

$$
\begin{aligned}
A &= LL^{\mathrm{H}} = LQQ^{\mathrm{H}}L^{\mathrm{H}} = (LQ)(LQ)^{\mathrm{H}} = CC^{\mathrm{H}} \\
B &= LQ\Lambda Q^{\mathrm{H}}L^{\mathrm{H}} = (LQ)\Lambda(LQ)^{\mathrm{H}} = C\Lambda C^{\mathrm{H}}
\end{aligned}
$$

  这表明两个 Hermite 正定阵 $A, B$ 可以同时合同对角化.
  (实际上如果 $B$ 退化为 Hermite 阵，结论也成立)

- 现在证明命题对 $m = 2$ 的情况成立:
  设 $A, B$ 为两个 Hermite 正定阵.
  根据上面的结论可知，存在非奇异阵 $C \in \mathbb{C}^{n\times n}$ 使得 $\begin{cases} A = CC^{\mathrm{H}} \\ B = C\Lambda C^{\mathrm{H}} \end{cases}$ (其中 $\Lambda = \mathrm{diag}\{\lambda_1,\ldots,\lambda_n\}$)
  于是我们有:

$$
\begin{aligned}
\det\left(\frac{A+B}{2}\right) &= \det\left(\frac{CC^{\mathrm{H}}+C\Lambda C^{\mathrm{H}}}{2}\right) \\
&= \det(C)\det(C^{\mathrm{H}})\det\left(\frac{I+\Lambda}{2}\right) \\
&= \det(C)\det(C^{\mathrm{H}})\prod_{i=1}^{n}\frac{1+\lambda_i}{2} \\
&\geq \det(C)\det(C^{\mathrm{H}})\prod_{i=1}^{n}\sqrt{\lambda_i} \\
&= \sqrt{\det(C)\det(C^{\mathrm{H}})\cdot\det(C)\det(\Lambda)\det(C^{\mathrm{H}})} \\
&= \sqrt{\det(CC^{\mathrm{H}})\cdot\det(C\Lambda C^{\mathrm{H}})} \\
&= \sqrt{\det(A)\det(B)} \\
&= \det(A)^{\frac{1}{2}}\det(B)^{\frac{1}{2}}
\end{aligned}
$$

  因此命题对 $m = 2$ 的情况成立.

- 对于 $m = 2^k\ (k \in \mathbb{Z}_+)$ 的情况，我们也很容易处理.
  例如 $m = 2^2 = 4$ 时，我们有:

$$
\begin{aligned}
\det\left(\frac{A_1+A_2+A_3+A_4}{4}\right) &= \det\left(\frac{\frac{A_1+A_2}{2}+\frac{A_3+A_4}{2}}{2}\right) \quad (\text{use the conclusion of case } m=2) \\
&\geq \det\left(\frac{A_1+A_2}{2}\right)^{\frac{1}{2}}\det\left(\frac{A_1+A_2}{2}\right)^{\frac{1}{2}} \quad (\text{use the conclusion of case } m=2 \text{ again}) \\
&\geq \det(A_1)^{\frac{1}{4}}\det(A_2)^{\frac{1}{4}}\det(A_3)^{\frac{1}{4}}\det(A_4)^{\frac{1}{4}}
\end{aligned}
$$

  以此类推，可知命题对 $m = 2^k\ (k \in \mathbb{Z}_+)$ 的情况也成立.

- 现固定 $k \in \mathbb{Z}_+$，考虑 $m = 2^k + p\ (p = 1,\ldots,2^k-1)$ 的情况:
  记 $\widetilde{m} = 2^{k+1}$，并定义:

$$B := \frac{A_1 + \cdots + A_m}{m}$$
$$A_{m+1} = \cdots = A_{\widetilde{m}} := B$$

容易验证:

$$\begin{aligned}
\frac{A_1 + \cdots + A_{\widetilde{m}}}{\widetilde{m}} &= \frac{A_1 + \cdots + A_m + A_{m+1} + \cdots + A_{\widetilde{m}}}{\widetilde{m}} \\
&= \frac{mB + (\widetilde{m} - m)B}{\widetilde{m}} \\
&= \frac{\widetilde{m}B}{\widetilde{m}} \\
&= B
\end{aligned}$$

根据前面的结论，我们有:

$$\begin{aligned}
\det(B)^{\widetilde{m}} &= \det\left(\frac{A_1 + \cdots + A_{\widetilde{m}}}{\widetilde{m}}\right)^{\widetilde{m}} \quad (\text{note that } \widetilde{m} = 2^{k+1}) \\
&\geq [\det(A_1)^{\frac{1}{\widetilde{m}}} \cdots \det(A_{\widetilde{m}})^{\frac{1}{\widetilde{m}}}]^{\widetilde{m}} \\
&= \det(A_1) \cdots \det(A_{\widetilde{m}}) \\
&= \det(A_1) \cdots \det(A_m) \det(A_{m+1}) \cdots \det(A_{\widetilde{m}}) \quad (\text{note that } A_{m+1} = \cdots = A_{\widetilde{m}} = B) \\
&= \det(A_1) \cdots \det(A_m) \cdot \det(B)^{\widetilde{m} - m}
\end{aligned}$$

这样我们就得到了 $\det(B)^m \geq \det(A_1) \cdots \det(A_m)$
于是有:

$$\begin{aligned}
\det\left(\frac{A_1 + \cdots + A_m}{m}\right) &= \det(B) \\
&\geq [\det(A_1) \cdots \det(A_m)]^{\frac{1}{m}} \\
&= \det(A_1)^{\frac{1}{m}} \cdots \det(A_m)^{\frac{1}{m}}
\end{aligned}$$

因此命题对一般的 $m$ 也成立.

**The End**