期末考试 (2024 秋)

Problem 1

计算 A 的 Moore-Penrose 伪逆:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$$

Solution:
注意到实矩阵 A 行满秩，故其 Moore-Penrose 伪逆为:

$$\begin{aligned} A^{\dagger} &= A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1} \\ &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}^{\mathrm{T}} \left(\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}^{\mathrm{T}} \right)^{-1} \\ &= \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix} \left(\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} 30 & 70 \\ 70 & 174 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix} \cdot \frac{1}{30 \times 174 - 70 \times 70} \begin{bmatrix} 174 & -70 \\ -70 & 30 \end{bmatrix} \\ &= \frac{1}{320} \begin{bmatrix} -176 & 80 \\ -72 & 40 \\ 32 & 0 \\ 136 & -40 \end{bmatrix} \end{aligned}$$

Problem 2

设矩阵 $A(\tau)$ 的每个元素都关于 $\tau \in [0,1]$ 连续，试证明:

$$\left\| \int_0^1 A(\tau) \mathrm{d}\tau \right\|_{\mathrm{F}} \le \int_0^1 \|A(\tau)\|_{\mathrm{F}} \, \mathrm{d}\tau$$

Proof:
定义泛函 $\langle \cdot, \cdot \rangle_{\mathrm{F}}$ 为:

$$\langle A, B \rangle_{\mathrm{F}} := \operatorname{tr}\left(A^{\mathrm{H}} B\right) \quad (\forall\, A, B \in \mathbb{C}^{m \times n})$$

对于任意 $A, B, C \in \mathbb{C}^{m \times n}$ 和 $\alpha \in \mathbb{C}$

- ① 非负性: $\langle A, A \rangle_{\mathrm{F}} = \operatorname{tr}\left(A^{\mathrm{H}} A\right) = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \ge 0$
- ② 正定性: $\langle A, A \rangle_{\mathrm{F}} = 0$ 当且仅当 $A = 0_{m \times n}$
- ③ 线性性: $\langle A + B, C \rangle_{\mathrm{F}} = \operatorname{tr}\left((A+B)^{\mathrm{H}} C\right) = \operatorname{tr}\left(A^{\mathrm{H}} C\right) + \operatorname{tr}\left(B^{\mathrm{H}} C\right) = \langle A, C \rangle_{\mathrm{F}} + \langle B, C \rangle_{\mathrm{F}}$
- ④ 齐次性: $\langle \alpha A, B \rangle_{\mathrm{F}} = \alpha \langle A, B \rangle_{\mathrm{F}}$

- ⑤ 共轭对称性: $\overline{\langle A,B\rangle_{\mathrm{F}}}=\langle B,A\rangle_{\mathrm{F}}$

因此泛函 $\langle\cdot,\cdot\rangle_{\mathrm{F}}$ 是一个内积 (称为 Frobenius 内积)
于是我们有:

$$\begin{aligned}
\left\|\int_0^1 A(\tau)\mathrm{d}\tau\right\|_{\mathrm{F}}^2 &= \left\langle\int_0^1 A(x)\mathrm{d}x,\int_0^1 A(y)\mathrm{d}y\right\rangle_{\mathrm{F}} \\
&= \int_0^1\int_0^1 \langle A(x),A(y)\rangle_{\mathrm{F}}\,\mathrm{d}x\mathrm{d}y \quad (\text{use Cauchy-Schwarz inequality}) \\
&\le \int_0^1\int_0^1 \|A(x)\|_{\mathrm{F}}\|A(y)\|_{\mathrm{F}}\,\mathrm{d}x\mathrm{d}y \\
&= \left(\int_0^1 \|A(x)\|_{\mathrm{F}}\,\mathrm{d}x\right)\left(\int_0^1 \|A(y)\|_{\mathrm{F}}\,\mathrm{d}y\right) \\
&= \left(\int_0^1 \|A(\tau)\|_{\mathrm{F}}\,\mathrm{d}\tau\right)^2
\end{aligned}$$

因此 $\left\|\int_0^1 A(\tau)\mathrm{d}\tau\right\|_{\mathrm{F}}\le\int_0^1\|A(\tau)\|_{\mathrm{F}}\,\mathrm{d}\tau$

Problem 3

证明 $S_1:=\{A^2:A\in\mathbb{R}^{2\times 2},\det(A)\le 0\}$ 是 $S_2:=\{B\in\mathbb{R}^{2\times 2}:\operatorname{tr}(B)\ge 0\}$ 的真子集.

Proof:
$A\in\mathbb{R}^{2\times 2}$ 的特征方程为 $\lambda^2-\operatorname{tr}(A)\lambda+\det(A)=0$
若 $\det(A)\le 0$，则判别式 $\Delta=\operatorname{tr}^2(A)-4\det(A)\ge 0$，说明 A 具有两个实特征值 $\lambda_1,\lambda_2\in\mathbb{R}$
因此 $\operatorname{tr}\left(A^2\right)=\lambda_1^2+\lambda_2^2\ge 0$
这说明 $S_1\subseteq S_2$

考虑以下示例:

$$A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$$
$$B=A^2=\begin{bmatrix}1&2\\0&1\end{bmatrix}$$

显然 $\det(A)=1>0$，故 $B=A^2\notin S_1$
但 $\operatorname{tr}(B)=2\ge 0$，表明 $B\in S_2$
这说明 S_1 是 S_2 的真子集.

Problem 4

设 $M\in\mathbb{C}^{n\times n}$ 为 Hermite 正定阵.
若 $A\in\mathbb{C}^{n\times n}$ 满足 $AMA^{\mathrm{H}}=A^{\mathrm{H}}MA$，试证明 A 可合同对角化.

Proof:
由于 $M\in\mathbb{C}^{n\times n}$ 为 Hermite 正定阵，故 $M^{\frac{1}{2}}\in\mathbb{C}^{n\times n}$ 是定义良好的，且是 Hermite 正定阵.
注意到:

$$\begin{aligned}
(M^{\frac{1}{2}}AM^{\frac{1}{2}})(M^{\frac{1}{2}}AM^{\frac{1}{2}})^{\mathrm{H}} &= M^{\frac{1}{2}}(AMA^{\mathrm{H}})M^{\frac{1}{2}} \\
&= M^{\frac{1}{2}}(A^{\mathrm{H}}MA)M^{\frac{1}{2}} \\
&= (M^{\frac{1}{2}}AM^{\frac{1}{2}})^{\mathrm{H}}(M^{\frac{1}{2}}AM^{\frac{1}{2}})
\end{aligned}$$

这表明 $M^{\frac{1}{2}}AM^{\frac{1}{2}}$ 是正规矩阵，等价于存在谱分解 $M^{\frac{1}{2}}AM^{\frac{1}{2}} = U\Lambda U^{\mathrm{H}}$
其中 $U \in \mathbb{C}^{n\times n}$ 为酉矩阵，Λ 为对角阵.
因此我们有:

$$\begin{aligned}(M^{\frac{1}{2}}U)^{\mathrm{H}}A(M^{\frac{1}{2}}U) &= U^{\mathrm{H}}(M^{\frac{1}{2}}AM^{\frac{1}{2}})U \\ &= U^{\mathrm{H}}(U\Lambda U^{\mathrm{H}})U \\ &= \Lambda\end{aligned}$$

这表明 A 可合同对角化.

Problem 5

试求 $x_0 \in \mathbb{R}^n$ 在超平面 $\{x \in \mathbb{R}^n : w^{\mathrm{T}}x + b = 0\}$ 上的正交投影.

Solution:
根据超平面的定义可知 $w \neq 0_n$
考虑以下优化问题:

$$\min_{w^{\mathrm{T}}x+b=0} \frac{1}{2}\|x_0 - x\|_2^2$$

定义 Lagrange 函数:

$$L(x,\lambda) := \frac{1}{2}\|x_0 - x\|_2^2 + \lambda(w^{\mathrm{T}}x + b)$$

其关于 x 的梯度为:

$$\nabla_x L(x,\lambda) := (x_0 - x) + \lambda w$$

根据 KKT 必要条件可列出以下方程:

$$\begin{cases}\nabla_x L(x,\lambda) = (x_0 - x) + \lambda w = 0_n \\ w^{\mathrm{T}}x + b = 0\end{cases}$$

将 $x = x_0 + \lambda w$ 代入 $w^{\mathrm{T}}x + b = 0$ 就得到:

$$w^{\mathrm{T}}x + b = w^{\mathrm{T}}(x_0 + \lambda w) + b = w^{\mathrm{T}}x_0 + \lambda\|w\|_2^2 + b = 0$$

解得 $\lambda = -\frac{w^{\mathrm{T}}x_0 + b}{\|w\|_2^2}$
于是我们有:

$$\begin{aligned}x &= x_0 + \lambda w \\ &= x_0 - \frac{w^{\mathrm{T}}x_0 + b}{\|w\|_2^2}w\end{aligned}$$

这是优化问题的最优解，也即 x_0 在超平面 $\{x \in \mathbb{R}^n : w^{\mathrm{T}}x + b = 0\}$ 上的正交投影.

Problem 6

试证明 $\det(\exp(A+B)) = \det(\exp(A))\det(\exp(B))$

- **Lemma 1:** [(reference)]
  若 $AB = BA$，则 $\exp(A)\exp(B) = \exp(A+B)$
  取 $B = -A$ 可得 $\exp(A)\exp(-A) = \exp(A-A) = 0_{n\times n}$
  这说明 $(\exp(A))^{-1} = \exp(-A)$

- **Lemma 2:** [(行列式的导数/Jacobi 公式)]
可以证明对于足够小的 t 我们有 $\det(A+tB)=\det(A)+t\operatorname{tr}(\operatorname{adj}(A)B)+O(t^2)$ 成立.
换言之，对于 $A(t)$ 我们有:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\det(A(t)) &= \operatorname{tr}\left(\operatorname{adj}(A)\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right) \quad (\text{note that } \operatorname{adj}(X)=\det(X)X^{-1}) \\
&= \operatorname{tr}\left(\det(X)X^{-1}\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right) \\
&= \det(X)\operatorname{tr}\left(X^{-1}\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right)
\end{aligned}
$$

- **Lemma 3:** $\det(\exp(A))=\exp(\operatorname{tr}(A))$ [(reference)]
定义关于参数 t 的函数 $f(t):=\det(\exp(tA))$，并对 t 求导:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}f(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\{\det(\exp(tA))\} \quad (\text{utilize Lemma 1}) \\
&= \det(\exp(tA))\cdot\operatorname{tr}\left((\exp(tA))^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA)\right) \quad (\text{note that } \frac{\mathrm{d}}{\mathrm{d}t}\exp(tA)=\exp(tA)\odot A) \\
&= \det(\exp(tA))\cdot\operatorname{tr}\left((\exp(tA))^{-1}\cdot\exp(tA)\odot A\right) \\
&= \det(\exp(tA))\cdot\operatorname{tr}(A) \\
&= f(t)\operatorname{tr}(A)
\end{aligned}
$$

于是我们得到了常微分方程 $\frac{\mathrm{d}}{\mathrm{d}t}f(t)=f(t)\operatorname{tr}(A)$
其解为 $f(t)=C\exp(\operatorname{tr}(A)t)$
根据 $f(0)=1$ 可知常数 $C=1$
因此我们有 $\det(\exp(tA))=f(t)=\exp(\operatorname{tr}(A)t)$
取 $t=1$ 即有 $\det(\exp(A))=\exp(\operatorname{tr}(A))$

Proof:

$$
\begin{aligned}
\det(\exp(A+B)) &= \exp(\operatorname{tr}(A+B)) \\
&= \exp(\operatorname{tr}(A)\operatorname{tr}(B)) \\
&= \exp(\operatorname{tr}(A))\exp(\operatorname{tr}(B)) \\
&= \det(\exp(A))\det(\exp(B))
\end{aligned}
$$

Problem 7

设 $A\in\mathbb{R}^{n\times n}$ 为非负矩阵.
试证明 $\rho(A)\leq\rho(\frac{1}{2}(A+A^{\mathrm{T}}))$

- **(一般非负矩阵的 Perron 定理, Matrix Analysis 定理 $8.3.1$)**
若 $A\in\mathbb{R}^{n\times n}$ 是非负矩阵，则 $\rho(A)$ 是 A 的一个特征值 (称为 **Perron 根**)，
且存在一个非负的非零向量 $x\in\mathbb{R}^n$ 使得 $Ax=\rho(A)x$
- **(Rayleigh 商定理, Matrix Analysis 定理 $4.2.2$)**
设 $A\in\mathbb{C}^{n\times n}$ 是 Hermite 阵，特征值按非减的次序排列: $\lambda_{\min}=\lambda_1(A)\leq\cdots\leq\lambda_n(A)=\lambda_{\max}$
给定整数 $1\leq i_1<\cdots<i_m\leq n$，设 $x_{i_1},\ldots,x_{i_m}$ 是标准正交的，且
$Ax_{i_k}=x_{i_k}\lambda_{i_k}\ (k=1,\ldots,m)$
记 $S:=\operatorname{span}\{x_{i_1},\ldots,x_{i_m}\}$，则我们有:

$$
\begin{aligned}
\lambda_{i_1} &= \min_{0_n\neq x\in S}\frac{x^{\mathrm{H}}Ax}{x^{\mathrm{H}}x} \\
&= \min_{\{x|x\in S\text{ such that }\|x\|_2=1\}} x^{\mathrm{H}}Ax \\
&\leq \max_{\{x|x\in S\text{ such that }\|x\|_2=1\}} x^{\mathrm{H}}Ax \\
&= \max_{0_n\neq x\in S}\frac{x^{\mathrm{H}}Ax}{x^{\mathrm{H}}x} \\
&= \lambda_{i_m}
\end{aligned}
$$

对于任意满足 $\|x\|_2=1$ 的 $x\in S$ 都有 $\lambda_{i_1}\leq x^{\mathrm{H}}Ax\leq\lambda_{i_m}$ 成立.
左侧和右侧的不等号的取等条件分别为 $Ax=x\lambda_{i_1}$ 和 $Ax=x\lambda_{i_m}$
特殊地，对于任意满足 $\|x\|_2=1$ 的 $x\in\mathbb{C}^n$ 都有 $\lambda_{\min}\leq x^{\mathrm{H}}Ax\leq\lambda_{\max}$ 成立.
左侧和右侧的不等号的取等条件分别为 $Ax=x\lambda_{\min}$ 和 $Ax=x\lambda_{\max}$
(它可以根据 A 的谱分解直接证明，原理与一般情况的证明是一致的)

几何解释:
连续实值函数 $f(x)=x^{\mathrm{H}}Ax$ 在单位球面 $\{x\in\mathbb{C}^n:\|x\|_2=1\}$ (这是个紧集) 上的最大值为 $\lambda_{\max}$，最小值为 $\lambda_{\min}$

Proof:
根据非负矩阵的 Perron 定理可知 $\rho(A)$ 是 A 的一个特征值 (即 Perron 根),
且存在一个非负的非零向量 $x_0\in\mathbb{R}^n$ 使得 $Ax_0=\rho(A)x_0$.
根据 Rayleigh 商定理可知:

$$
\begin{aligned}
\rho\left(\frac{1}{2}(A+A^{\mathrm{T}})\right) &\geq \frac{x_0^{\mathrm{T}}\left(\frac{1}{2}(A+A^{\mathrm{T}})\right)x_0}{x_0^{\mathrm{T}}x_0} \\
&= \frac{1}{2}\cdot\frac{x_0^{\mathrm{T}}Ax_0+x_0^{\mathrm{T}}A^{\mathrm{T}}x_0}{x_0^{\mathrm{T}}x_0} \quad (\text{note that } Ax_0=\rho(A)x_0) \\
&= \frac{1}{2}\cdot\frac{\rho(A)x_0^{\mathrm{T}}x_0+\rho(A)x_0^{\mathrm{T}}x_0}{x_0^{\mathrm{T}}x_0} \\
&= \rho(A)
\end{aligned}
$$

Problem 8

设 $A\in\mathbb{C}^{n\times n}$ 是 Hermite 正定阵.
试求:

$$
\min_{V\in\mathbb{C}^{n\times k}:V^{\mathrm{H}}V=I_k}\det\left(V^{\mathrm{H}}AV\right)
$$

- **(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: $\lambda_1(A)\leq\cdots\leq\lambda_n(A)$
  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 $A=\begin{bmatrix}B & y\\ y^{\mathrm{H}} & a\end{bmatrix}$
  特征值按非减的次序排列: $\lambda_1(B)\leq\cdots\leq\lambda_{n-1}(B)$
  则我们有如下的交错性质:

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

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

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

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.

上述定理表明:

Hermite 阵无论是加边扩充还是删边约简，其新旧特征值必定是交错的.
当然，加边扩充和删边约简不一定要在最后一行和最后一列进行，它可以在一行和对应的列进行.

归纳法指出:

考虑 Hermite 阵 A 的 $n-m$ 主子阵 $B=A_{(1:n-m,1:n-m)}\in\mathbb{C}^{(n-m)\times(n-m)}$
特征值按非减的次序排列: $\lambda_1(B)\leq\cdots\leq\lambda_{n-m}(B)$
并记 $A=\begin{bmatrix} B & C \\ C^{\mathrm{H}} & D \end{bmatrix}$ (其中 $C\in\mathbb{C}^{(n-m)\times m}, D\in\mathbb{C}^{m\times m}$)
则我们有如下的交错性质:

$$\lambda_i(A)\leq\lambda_i(B)\leq\lambda_{i+m}(A)\quad(\forall\, i=1,\ldots,n-m)$$

其中 $\lambda_i(A)=\lambda_i(B)$ 成立当且仅当存在非零向量 $z\in\mathbb{C}^{n-m}$ 使得 $\begin{cases} Bz=z\lambda_i(B) \\ Bz=z\lambda_i(A) \\ C^{\mathrm{H}}z=0_m \end{cases}$

而 $\lambda_i(B)=\lambda_{i+m}(A)$ 成立当且仅当存在非零向量 $z\in\mathbb{C}^{n-m}$ 使得 $\begin{cases} Bz=z\lambda_i(B) \\ Bz=z\lambda_{i+m}(A) \\ C^{\mathrm{H}}z=0 \end{cases}$

若 B 没有与 C 的列向量组正交的特征向量，则上述不等式均为严格不等式.

Solution:

设 $A\in\mathbb{C}^{n\times n}$ 是 Hermite 正定阵，$1\leq k\leq n$
设 $u_1,\ldots,u_n\in\mathbb{C}^n$ 标准正交
记 $U=[u_1,\ldots,u_n]\in\mathbb{C}^{n\times n}$ 和 $V=[u_1,\ldots,u_k]\in\mathbb{C}^{n\times k}$
记 $B:=V^{\mathrm{H}}AV=[u_i^{\mathrm{H}}Au_j]_{i,j=1}^k\in\mathbb{C}^{k\times k}$ (它就是 $U^{\mathrm{H}}AU$ 的 k 阶顺序主子阵)
注意到 $U^{\mathrm{H}}AU$ 的特征值与 A 的特征值完全相同.
设 A,B 的特征值以非减次序排列，则根据 Cauchy 交错定理我们有:

$$0<\lambda_i(A)\leq\lambda_i(B)=\lambda_i(V^{\mathrm{H}}AV)\leq\lambda_{i+(n-k)}(A)\quad(\forall\, i=1,\ldots,k)$$

这也说明对于任意列标准正交的 $V\in\mathbb{C}^{n\times k}$ 我们都有:

$$0<\prod_{i=1}^k\lambda_i(A)\leq\prod_{i=1}^k\lambda_i(V^{\mathrm{H}}AV)=\det\left(V^{\mathrm{H}}AV\right)\leq\prod_{i=1}^k\lambda_{n-k+i}(A)$$

注意到左右不等号均可取等，于是我们顺理成章地得到:

$$\min_{V\in\mathbb{C}^{n\times k}:V^{\mathrm{H}}V=I_k}\det\left(V^{\mathrm{H}}AV\right)=\prod_{i=1}^k\lambda_i(A)$$
$$\max_{V\in\mathbb{C}^{n\times k}:V^{\mathrm{H}}V=I_k}\det\left(V^{\mathrm{H}}AV\right)=\prod_{i=1}^k\lambda_{n-k+i}(A)$$

The End