矩阵鉴赏

其实是矩阵论复习题

统计 91 董晟渤

2021 年 6 月 11 日

目录

考试内容

考试题型

1. 填空题: 5 题, 每题 3 分, 共 15 分;
2. 计算题与证明题: 4 题, 每题 10 分; 3 题, 每题 15 分, 共 85 分.

- **矩阵初步:** 特征值, 特征向量, 最小多项式, 初等因子, 不变因子, 行列式因子, Jordan 标准型;
- **矩阵分解:** UR 分解 (QR 分解), 满秩分解, 奇异值分解, 谱分解, LU 分解.
- **矩阵广义逆:** 线性方程组的 L-S 解, L-N 解, L-S-N 解.
- **矩阵分析:** 矩阵微分, 矩阵函数, 向量范数与矩阵范数.

Tips

数值方法和一些比较复杂的内容不考.

试卷上标注的试卷类型是 B.

填空题

- 正规矩阵的定义 ($\boldsymbol{A}\boldsymbol{A}^H = \boldsymbol{A}^H\boldsymbol{A}$), 求矩阵的谱半径 (最大特征值);
- $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\boldsymbol{A}^{2021} = \boldsymbol{O}$, 求 $|\boldsymbol{A}+\boldsymbol{I}|$;
  (此时 Jordan 标准型 $\boldsymbol{J}$ 的对角线元素全为 0, 参考题目2.)
- $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\boldsymbol{A}^2 = \boldsymbol{I}$, 求 $\sin\boldsymbol{A}$;
  (此时 Jordan 标准型 $\boldsymbol{J}$ 为单位矩阵, 从而 $\sin\boldsymbol{A} = \boldsymbol{P}\sin\boldsymbol{I}\boldsymbol{P}^{-1} = \sin 1\cdot\boldsymbol{I}$.)
- $\boldsymbol{A}, \boldsymbol{B}$ 是 n 阶酉矩阵, 求 $M = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{A} & \boldsymbol{B} \end{bmatrix}$ 的 M-P 广义逆. [1]

考前老师说可能会考算矩阵 p-范数的填空题, 猜测是在 A 卷上.

[1] 这题不太确定做对, 就不写思路了.

计算题与解答题

- 求矩阵的行列式因子，不变因子，初等因子，参考题目4;
- 满秩分解相关的证明题，参考题目9;
- 求矩阵的 QR 分解 (Gram-Smith 正交化过程);
- 设 $\boldsymbol{x}$ 是 n 维列向量，$\boldsymbol{a}=\boldsymbol{a}(\boldsymbol{x})$ 和 $\boldsymbol{b}=\boldsymbol{b}(\boldsymbol{x})$ 是 m 维函数列向量，证明

$$\frac{\mathrm{d}\boldsymbol{a}^T\boldsymbol{b}}{\mathrm{d}\boldsymbol{x}^T}=\boldsymbol{b}^T\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\boldsymbol{x}^T}+\boldsymbol{a}^T\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}\boldsymbol{x}^T};$$

  (按照微分的定义打开计算即可．)
- $\boldsymbol{A}=\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}$，求 $\cos\pi\boldsymbol{A}$，参考题目10;
- M-P 广义逆相关的证明题，参考题目12;
- 给定矩阵 $\boldsymbol{A}\in\mathbb{R}^{4\times4}$，求 M-P 广义逆; 给定向量 $\boldsymbol{x}\in\mathbb{R}^4$，用广义逆判断方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 是否相容; 在上面的基础上，求 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的最小范数解或最小二乘范数解.

目录

Example

设 n 阶方阵 $\boldsymbol{A}$ 的每个元素都是整数，证明 $\frac{1}{2}$ 不是 $\boldsymbol{A}$ 的特征值.

设 $\boldsymbol{A} = (a_{ij})_{n\times n}$，此时 $\boldsymbol{A}$ 的特征多项式

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

是首项系数为 1 的整系数多项式，若其有理根存在，则其必定为整数，从而 $\frac{1}{2}$ 不是 $\boldsymbol{A}$ 的特征值.

Example

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 证明:
(1) 方阵 $\boldsymbol{A}$ 的特征值全为 0, 当且仅当存在 $m \in \mathbb{Z}$, 使得 $\boldsymbol{A}^m = \boldsymbol{O}$.
(2) 若 $\boldsymbol{A}^m = \boldsymbol{O}$, 则 $|\boldsymbol{A}+\boldsymbol{I}| = 1$.

(1) 若 $\boldsymbol{A}$ 的特征值全为 0, 则存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J}$, 其中 $\boldsymbol{J}$ 的对角线上的元素全为 0, 从而 $\boldsymbol{J}^n = \boldsymbol{O}$, 进而

$$\boldsymbol{A}^n = \boldsymbol{P}\boldsymbol{J}^n\boldsymbol{P}^{-1} = \boldsymbol{O}.$$

若存在 $m \in \mathbb{Z}$, 使得 $\boldsymbol{A}^m = \boldsymbol{O}$, 对 $\boldsymbol{A}$ 的任一特征值 λ, 设其对应的非零特征向量为 $\boldsymbol{x}$, 则有 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, 从而

$$\boldsymbol{A}^m\boldsymbol{x} = \lambda^m\boldsymbol{x} = \boldsymbol{0},$$

解得 $\lambda = 0$, 从而 $\boldsymbol{A}$ 的特征值全为 0.
(2) 计算得

$$|\boldsymbol{A}+\boldsymbol{I}| = |\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1} + \boldsymbol{P}\boldsymbol{I}\boldsymbol{P}^{-1}| = |\boldsymbol{P}| \cdot |\boldsymbol{J}+\boldsymbol{I}| \cdot |\boldsymbol{P}^{-1}| = 1.$$

Example

记 $[\boldsymbol{X},\boldsymbol{Y}]=\boldsymbol{XY}-\boldsymbol{YX}$, 证明对任意的二阶方阵 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, 都有

$$[[\boldsymbol{A},\boldsymbol{B}]^2,\boldsymbol{C}]=\boldsymbol{O}.$$

注意到

$$\mathrm{tr}[\boldsymbol{A},\boldsymbol{B}]=\mathrm{tr}\boldsymbol{AB}-\mathrm{tr}\boldsymbol{BA}=0.$$

设 $[\boldsymbol{A},\boldsymbol{B}]$ 的特征值为 λ_1,λ_2, 则 $\lambda_1+\lambda_2=0$.
(1) $\lambda_1=\lambda_2=0$, 设 $[\boldsymbol{A},\boldsymbol{B}]=\boldsymbol{PJP}^{-1}$, 根据 $\boldsymbol{J}^2=\boldsymbol{O}$ 可知 $[\boldsymbol{A},\boldsymbol{B}]^2=\boldsymbol{O}$.
(2) $\lambda_1=-\lambda_2\neq 0$, 设 $[\boldsymbol{A},\boldsymbol{B}]=\boldsymbol{PJP}^{-1}$, 则

$$[\boldsymbol{A},\boldsymbol{B}]^2=\boldsymbol{P}\begin{bmatrix}\lambda_1^2 & 0\\ 0 & \lambda_1^2\end{bmatrix}\boldsymbol{P}^{-1}=\boldsymbol{P}\lambda_1^2\boldsymbol{I}\boldsymbol{P}^{-1}=\lambda_1^2\boldsymbol{I},$$

从而 $[\boldsymbol{A},\boldsymbol{B}]^2$ 与任意的二阶矩阵可交换.

Example

设

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & -2\\ -2 & -2 & 3\\ -1 & -1 & 1\end{bmatrix},$$

求 $\boldsymbol{A}$ 的: (1) 特征值; (2) 不变因子; (3) 行列式因子; (4) 初等因子; (5) 最小多项式; (6) Jordan 标准型.

计算得

$$\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix}\lambda-1 & -1 & 2\\ 2 & \lambda+2 & -3\\ 1 & 1 & \lambda-1\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \lambda^3\end{bmatrix},$$

$\boldsymbol{A}$ 的特征值为 $\lambda=0$, 不变因子为 $1,1,\lambda^3$, 行列式因子为 $1,1,\lambda^3$, 初等因子为 λ^3, 最小多项式为 $f(\lambda)=\lambda^3$, Jordan 标准型为

$$\boldsymbol{J} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Tips

数字矩阵 $\boldsymbol{A}$ 的**特征矩阵** $\lambda \boldsymbol{I} - \boldsymbol{A}$ 经初等变换后可以化为 Smith 标准形

$$\lambda \boldsymbol{I} - \boldsymbol{A} \to \begin{bmatrix} d_1(\lambda) & 0 & \cdots & 0 \\ 0 & d_2(\lambda) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n(\lambda) \end{bmatrix}, \tag{1}$$

其中 $d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)$ 称为矩阵 $\boldsymbol{A}$ 的**不变因子**. 并且对任意的 $1 \leq i \leq n-1$, 都有 $d_i(\lambda) \mid d_{i+1}(\lambda)$. 在这里我们也指出, $d_n(\lambda)$ 即为 $\boldsymbol{A}$ 的**最小多项式**.

Tips

矩阵 $\lambda \boldsymbol{I}-\boldsymbol{A}$ 的所有 k 阶子式的最大公因数称为 k 阶**行列式因子**，记作 $D_k(\lambda)$. 根据公式1可以看出，不变因子与行列式因子之间的关系为

$$\begin{cases} D_1(\lambda)=d_1(\lambda), \\ D_2(\lambda)=d_1(\lambda)d_2(\lambda), \\ \quad \cdots, \\ D_n(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda). \end{cases}$$

另一方面，也有

$$\begin{cases} d_1(\lambda)=D_1(\lambda), \\ d_2(\lambda)=\dfrac{D_2(\lambda)}{D_1(\lambda)}, \\ \quad \cdots, \\ d_n(\lambda)=\dfrac{D_n(\lambda)}{D_{n-1}(\lambda)}. \end{cases}$$

Tips

把矩阵 $\boldsymbol{A}$ 的每个次数大于零的不变因子分解成互不相同的首项为 1 的一次因式方幂的乘积, 所有这些一次因子方幂 (相同的必须按出现的次数计算) 称为矩阵 $\boldsymbol{A}$ 的**初等因子**. 矩阵 $\boldsymbol{A}$ 的初等因子 $(\lambda-\lambda_i)^{\alpha_i}$ 所对应的 Jordan 块为

$$\boldsymbol{J}(\lambda_i,\alpha_i)=\begin{bmatrix}\lambda_i & 1 & \cdots & 0\\ 0 & \lambda_i & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_i\end{bmatrix}_{\alpha_i\times\alpha_i},$$

从而根据矩阵的所有初等因子, 可以写出矩阵的 **Jordan 标准型**.

Example

设

$$A=\begin{bmatrix}1&0&0&1\\1&1&0&2\\0&0&1&3\\0&0&0&2\end{bmatrix},$$

求 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}$ 为 Jordan 标准型.

计算得

$$\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix}\lambda-1&0&0&-1\\-1&\lambda-1&0&-2\\0&0&\lambda-1&-3\\0&0&0&\lambda-2\end{bmatrix},$$

其行列式因子分别为 $1,1,\lambda-1,(\lambda-1)^3(\lambda-2)$, 从而不变因子为 $1,1,\lambda-1,(\lambda-1)^2(\lambda-2)$, 初等因子为 $\lambda-1,(\lambda-1)^2,\lambda-2$,

据此得到矩阵的 Jordan 标准型为

$$\boldsymbol{J}=\begin{bmatrix}1&0&0&0\\0&1&1&0\\0&0&1&0\\0&0&0&2\end{bmatrix}.$$

设 $\boldsymbol{P}=[\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{P}_3,\boldsymbol{P}_4]$, 且 $\boldsymbol{AP}=\boldsymbol{PJ}$, 从而

$$[\boldsymbol{AP}_1,\boldsymbol{AP}_2,\boldsymbol{AP}_3,\boldsymbol{AP}_4]=[\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{P}_2+\boldsymbol{P}_3,2\boldsymbol{P}_4],$$

取

$$\boldsymbol{P}=\begin{bmatrix}0&0&1&1\\1&1&0&3\\1&0&0&3\\0&0&0&1\end{bmatrix},$$

则 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{J}$.

目录

Example

设

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix},$$

求 A 的谱分解式.

计算得

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 \\ 1 & \lambda - 4 \end{vmatrix} = (\lambda - 2)(\lambda - 3),$$

则特征值 $\lambda_1 = 3, \lambda_2 = 2$, 令 $\varphi_1(\lambda) = \lambda - 2, \varphi_2(\lambda) = \lambda - 3$, 则有

$$P_1 = \frac{\varphi_1(A)}{\varphi_1(3)} = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}, \quad P_2 = \frac{\varphi_2(A)}{\varphi_2(2)} = \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix},$$

从而 A 的谱分解式为

$$A = 3 \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} + 2 \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}.$$

Example

设

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

求 A 的奇异值分解式.

计算得

$$A^H A = A^2 = \begin{bmatrix} 5 & 4 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 9 \end{bmatrix},$$

以及 $|\lambda I - A^H A| = (\lambda - 1)(\lambda - 9)^2$, 从而 $\delta_1 = \delta_2 = 3$, $\delta_3 = 1$, 记

$$S = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

此时 $\boldsymbol{A}^H\boldsymbol{A}$ 的特征值 $9, 1$ 所对的单位正交向量组为

$$\boldsymbol{v}_1 = (0,0,1)^T, \quad \boldsymbol{v}_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad \boldsymbol{v}_3 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right),$$

其恰好构成 $\mathbb{R}^3$ 的一组基, 记

$$\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)^T = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix},$$

又计算得

$$\boldsymbol{U} = \boldsymbol{A}\boldsymbol{V}\boldsymbol{S}^{-1} = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix}.$$

从而 $\boldsymbol{A}$ 的奇异值分解式为

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^H = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

Example

设 $\boldsymbol{A}$ 的奇异值分解为

$$\boldsymbol{A}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Sigma} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{V}^H,$$

证明 $\boldsymbol{U}$ 的列向量是 $\boldsymbol{A}\boldsymbol{A}^H$ 的特征向量, 而 $\boldsymbol{V}$ 的列向量是 $\boldsymbol{A}^H\boldsymbol{A}$ 的特征向量.

计算得

$$\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{U}=\begin{bmatrix}\boldsymbol{\Sigma}^2 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{U}$$

其中 $\boldsymbol{\Sigma}^2$ 的对角线元素是 $\boldsymbol{A}\boldsymbol{A}^H$ 的非零特征值, 从而 $\boldsymbol{U}$ 的列向量是 $\boldsymbol{A}\boldsymbol{A}^H$ 的特征向量, 同样地, 计算得

$$\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}=\begin{bmatrix}\boldsymbol{\Sigma}^2 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{V}$$

其中 $\boldsymbol{\Sigma}^2$ 的对角线元素也是 $\boldsymbol{A}^H\boldsymbol{A}$ 的非零特征值, 从而 $\boldsymbol{V}$ 的列向量是 $\boldsymbol{A}^H\boldsymbol{A}$ 的特征向量.

Example

设 $\boldsymbol{A} \in \mathbb{C}_r^{m\times n}$, $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{D} = \boldsymbol{C}_1\boldsymbol{D}_1$ 均为其满秩分解, 其中, $\boldsymbol{C}$ 与 $\boldsymbol{C}_1$ 为 $m\times r$ 的列满秩矩阵, $\boldsymbol{D}$ 与 $\boldsymbol{D}_1$ 为 $r\times n$ 的行满秩矩阵, 证明:
(1) 存在 r 阶可逆矩阵 $\boldsymbol{K}$, 使 $\boldsymbol{C} = \boldsymbol{C}_1\boldsymbol{K}$, $\boldsymbol{D} = \boldsymbol{K}^{-1}\boldsymbol{D}_1$;
(2) $\boldsymbol{D}^H(\boldsymbol{D}\boldsymbol{D}^H)^{-1}(\boldsymbol{C}^H\boldsymbol{C})^{-1}\boldsymbol{C}^H = \boldsymbol{D}_1^H(\boldsymbol{D}_1\boldsymbol{D}_1^H)^{-1}(\boldsymbol{C}_1^H\boldsymbol{C}_1)^{-1}\boldsymbol{C}_1^H$.

(1) 根据 $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{D} = \boldsymbol{C}_1\boldsymbol{D}_1$ 知, $\boldsymbol{A}$ 的列向量都可以由 $\boldsymbol{C}$ 和 $\boldsymbol{C}_1$ 的列向量组线性表出. 又由 $\boldsymbol{C}$ 和 $\boldsymbol{C}_1$ 知它们的列向量组线性无关. 从而 $\boldsymbol{C}$ 和 $\boldsymbol{C}_1$ 的列向量组是 $\boldsymbol{A}$ 的列向量组的极大线性无关组, 因此这两个矩阵的列向量组是等价的, 从而存在 r 阶可逆矩阵 $\boldsymbol{K}$, 使得 $\boldsymbol{C} = \boldsymbol{C}_1\boldsymbol{K}$. 代入得

$$\boldsymbol{C}\boldsymbol{D} = \boldsymbol{C}_1\boldsymbol{K}\boldsymbol{D} = \boldsymbol{C}_1\boldsymbol{D}_1,$$

从而 $\boldsymbol{C}_1^H\boldsymbol{C}_1\boldsymbol{K}\boldsymbol{D} = \boldsymbol{C}_1^H\boldsymbol{C}_1\boldsymbol{D}_1$, 其中 $\boldsymbol{C}_1^H\boldsymbol{C}_1$ 是 $r\times r$ 可逆矩阵, 故 $\boldsymbol{K}\boldsymbol{D} = \boldsymbol{D}_1$.

(2) 根据 $\boldsymbol{C}=\boldsymbol{C}_1\boldsymbol{K}$, $\boldsymbol{D}=\boldsymbol{K}^{-1}\boldsymbol{D}_1$ 得

$$\begin{aligned}(\boldsymbol{C}^H\boldsymbol{C})^{-1}\boldsymbol{C}^H &= (\boldsymbol{K}^H\boldsymbol{C}_1^H\boldsymbol{C}_1\boldsymbol{K})^{-1}\boldsymbol{K}^H\boldsymbol{C}_1^H \\ &= \boldsymbol{K}^{-1}(\boldsymbol{C}_1^H\boldsymbol{C}_1)^{-1}\boldsymbol{C}_1^H\end{aligned}$$

同理 $\boldsymbol{D}^H(\boldsymbol{D}\boldsymbol{D}^H)^{-1}=\boldsymbol{D}_1^H(\boldsymbol{D}_1\boldsymbol{D}_1^H)^{-1}\boldsymbol{K}$, 从而

$$\begin{aligned}\boldsymbol{D}^H(\boldsymbol{D}\boldsymbol{D}^H)^{-1}(\boldsymbol{C}^H\boldsymbol{C})^{-1}\boldsymbol{C}^H &= \boldsymbol{D}_1^H(\boldsymbol{D}_1\boldsymbol{D}_1^H)^{-1}\boldsymbol{K}\boldsymbol{K}^{-1}(\boldsymbol{C}_1^H\boldsymbol{C}_1)^{-1}\boldsymbol{C}_1^H \\ &= \boldsymbol{D}_1^H(\boldsymbol{D}_1\boldsymbol{D}_1^H)^{-1}(\boldsymbol{C}_1^H\boldsymbol{C}_1)^{-1}\boldsymbol{C}_1^H.\end{aligned}$$

Tips

不能直接由 $\boldsymbol{C}_1\boldsymbol{K}\boldsymbol{D}=\boldsymbol{C}_1\boldsymbol{D}_1$ 直接得到 $\boldsymbol{K}\boldsymbol{D}=\boldsymbol{D}_1$. 这是因为不能由 $\boldsymbol{X}\boldsymbol{Y}=\boldsymbol{O}$ 推出 $\boldsymbol{X}=\boldsymbol{O}$ 或 $\boldsymbol{Y}=\boldsymbol{O}$.
这个等式可以说明矩阵的 Moore-Penrose 广义逆矩阵是唯一的.

目录

Example

设
$$A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 2 & 4 \end{bmatrix},$$
求 $A\{1\}$.

计算得
$$\begin{bmatrix} A & I \\ I & O \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 2 & 4 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 1 \\ 1 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

从而
$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}, \quad PAQ = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

据此写出 $\boldsymbol{A}\{1\}$ 中的矩阵 $\boldsymbol{A}^{(1)}$ 的形式为

$$\boldsymbol{A}^{(1)}=\boldsymbol{Q}\begin{bmatrix}1 & a\\ b & c\\ d & e\end{bmatrix}\boldsymbol{P}=\begin{bmatrix}1 & -2\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & a & b\\ c & d & e\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -2 & 0 & 1\end{bmatrix},$$

其中 a,b,c,d,e 是任意数, 从而

$$\boldsymbol{A}\{1\}=\left\{\begin{bmatrix}1 & -2\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & a & b\\ c & d & e\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -2 & 0 & 1\end{bmatrix}: a,b,c,d,e\in\mathbb{C}\right\}.$$

特别地, 取 $a=b=c=d=e=0$, 可以得到一个特殊的 $\{1\}$-逆

$$\boldsymbol{A}^{(1)}=\begin{bmatrix}1 & -2\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -2 & 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}.$$

Tips

设 $\boldsymbol{A}\in\mathbb{C}_r^{m\times n}$, 且 $\boldsymbol{P},\boldsymbol{Q}$ 分别为 m 阶和 n 阶非奇异方阵, 满足

$$\boldsymbol{PAQ}=\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix},$$

则可以构造

$$\boldsymbol{A}^{(1)}=\boldsymbol{Q}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{G}_{12}\\ \boldsymbol{G}_{21} & \boldsymbol{G}_{22}\end{bmatrix}\boldsymbol{P},$$

其中 $\boldsymbol{G}_{ij}$ 为任意矩阵. 从而我们可以对矩阵

$$\begin{bmatrix}\boldsymbol{A} & \boldsymbol{I}_2\\ \boldsymbol{I}_3 & \boldsymbol{O}\end{bmatrix}$$

进行初等变换, 将 $\boldsymbol{A}$ 化为标准形后, $\boldsymbol{I}_2$ 和 $\boldsymbol{I}_3$ 分别是 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$.

Tips

由上可以构造新的广义逆矩阵. 其中根据 $\mathrm{rank}(\boldsymbol{A}^{(1,2)}) = \mathrm{rank}(\boldsymbol{A})$, 可以构造

$$\boldsymbol{A}^{(1,2)} = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{G}_{12} \\ \boldsymbol{G}_{21} & \boldsymbol{G}_{21}\boldsymbol{G}_{12} \end{bmatrix}\boldsymbol{P},$$

其中 $\boldsymbol{G}_{ij}$ 是任意矩阵.
接下来, Urguhart 定理给出, 对任意的 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, 都有

$$\boldsymbol{X} = (\boldsymbol{A}^H\boldsymbol{A})^{(1)}\boldsymbol{A}^H \in \boldsymbol{A}\{1,2,3\}, \quad \boldsymbol{Y} = \boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{A}^H)^{(1)} \in \boldsymbol{A}\{1,2,4\}.$$

当 $\boldsymbol{A}$ 列满秩时, $\boldsymbol{A}^H\boldsymbol{A}$ 满秩; 而当 $\boldsymbol{A}$ 行满秩时, $\boldsymbol{A}\boldsymbol{A}^H$ 满秩.
最后, 设 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{R}$ 为 $\boldsymbol{A}$ 的满秩分解, 则有

$$\boldsymbol{A}^+ = \boldsymbol{R}^+\boldsymbol{L}^+ = \boldsymbol{R}^H(\boldsymbol{R}\boldsymbol{R}^H)^{-1}(\boldsymbol{L}^H\boldsymbol{L})^{-1}\boldsymbol{L}^H.$$

Example

设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{A}^+$ 为广义逆, 证明 $(\boldsymbol{A}^H)^+ = (\boldsymbol{A}^+)^H$.

也即验证 $(\boldsymbol{A}^+)^H$ 是 $\boldsymbol{A}^H$ 的广义逆, 首先

$$\boldsymbol{A}^H(\boldsymbol{A}^+)^H\boldsymbol{A}^H = (\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A})^H = \boldsymbol{A}^H;$$

其次

$$(\boldsymbol{A}^+)^H\boldsymbol{A}^H(\boldsymbol{A}^+)^H = (\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+)^H = (\boldsymbol{A}^+)^H;$$

接下来

$$(\boldsymbol{A}^H(\boldsymbol{A}^+)^H)^H = \boldsymbol{A}^+\boldsymbol{A} = (\boldsymbol{A}^+\boldsymbol{A})^H = \boldsymbol{A}^H(\boldsymbol{A}^+)^H;$$

最后

$$((\boldsymbol{A}^+)^H\boldsymbol{A}^H)^H = \boldsymbol{A}\boldsymbol{A}^+ = (\boldsymbol{A}\boldsymbol{A}^+)^H = (\boldsymbol{A}^+)^H\boldsymbol{A}^H.$$

从而 $(\boldsymbol{A}^+)^H$ 是 $\boldsymbol{A}^H$ 的广义逆.

Example

设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, 且存在 $\boldsymbol{G} \in \mathbb{C}^{n\times m}$, $\boldsymbol{S} \in \mathbb{C}^{m\times m}$, $\boldsymbol{T} \in \mathbb{C}^{n\times n}$, 使得

$$\boldsymbol{AGA} = \boldsymbol{A}, \quad \boldsymbol{G} = \boldsymbol{A}^H\boldsymbol{S}, \quad \boldsymbol{G} = \boldsymbol{TA}^H,$$

证明 $\boldsymbol{G} = \boldsymbol{A}^+$.

首先证明 $\boldsymbol{A}^+$ 满足上述条件, 根据 $(\boldsymbol{AA}^+)^H = (\boldsymbol{A}^+)^H\boldsymbol{A}^H = \boldsymbol{AA}^+$, 可得

$$\boldsymbol{A}^H\boldsymbol{AA}^+ = (\boldsymbol{AA}^+\boldsymbol{A})^H = \boldsymbol{A}^H,$$

从而 $\boldsymbol{A}^+ = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H = \boldsymbol{TA}^H$, 同理有 $\boldsymbol{G} = \boldsymbol{A}^H(\boldsymbol{AA}^H)^{-1} = \boldsymbol{A}^H\boldsymbol{S}$.
其次证明满足上面条件的 $\boldsymbol{G}$ 是唯一的. 假设还存在 $\boldsymbol{G}_1 \in \mathbb{C}^{n\times m}$, $\boldsymbol{S}_1 \in \mathbb{C}^{m\times m}$, $\boldsymbol{T}_1 \in \mathbb{C}^{n\times n}$ 满足等式, 令 $\boldsymbol{X} = \boldsymbol{G} - \boldsymbol{G}_1$, $\boldsymbol{U} = \boldsymbol{S} - \boldsymbol{S}_1$, $\boldsymbol{V} = \boldsymbol{T} - \boldsymbol{T}_1$, 接下来只要说明 $\boldsymbol{X} = \boldsymbol{O}$ 即可.

根据条件有

$$\boldsymbol{AXA}=\boldsymbol{O},\quad \boldsymbol{X}=\boldsymbol{A}^H\boldsymbol{U},\quad \boldsymbol{X}=\boldsymbol{V}\boldsymbol{A}^H.$$

首先计算得

$$\begin{aligned}(\boldsymbol{XA})^H\boldsymbol{XA}&=\boldsymbol{A}^H\boldsymbol{X}^H\boldsymbol{XA}\\&=\boldsymbol{A}^H\boldsymbol{X}^H\boldsymbol{A}^H\boldsymbol{UA}\\&=(\boldsymbol{AXA})^H\boldsymbol{UA}\\&=\boldsymbol{O},\end{aligned}$$

从而 $\boldsymbol{XA}=\boldsymbol{O}$, 接下来计算得

$$\boldsymbol{XX}^H=\boldsymbol{XAV}^H=\boldsymbol{O},$$

从而 $\boldsymbol{X}=\boldsymbol{O}$, 也即 $\boldsymbol{G}=\boldsymbol{G}_1$.

目录

Example

设

$$\boldsymbol{A}=\begin{bmatrix}1 & 1\\ 0 & 2\end{bmatrix},$$

求 $\cos \boldsymbol{A}$.

计算得

$$\lambda \boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix}\lambda-1 & -1\\ 0 & \lambda-2\end{bmatrix}\to\begin{bmatrix}1 & 0\\ 0 & (\lambda-1)(\lambda-2)\end{bmatrix},$$

则特征值 $\lambda_1=1$, $\lambda_2=2$, 则有

$$\boldsymbol{J}=\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix},\quad \boldsymbol{P}=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix},\quad \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}.$$

从而

$$\cos \boldsymbol{A}=\boldsymbol{P}\cos \boldsymbol{J}\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}\cos 1 & 0\\ 0 & \cos 2\end{bmatrix}\begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}=\begin{bmatrix}\cos 1 & \cos 2-\cos 1\\ 0 & \cos 2\end{bmatrix}.$$

Tips

设函数 $f(x)$ 解析, 矩阵 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$, 其中 $\boldsymbol{J}=\operatorname{diag}\{\boldsymbol{J}_1,\boldsymbol{J}_2,\cdots,\boldsymbol{J}_m\}$, 则有

$$f(\boldsymbol{A})=\boldsymbol{P}f(\boldsymbol{J})\boldsymbol{P}^{-1}=\boldsymbol{P}\operatorname{diag}\{f(\boldsymbol{J}_1),f(\boldsymbol{J}_2),\cdots,f(\boldsymbol{J}_m)\}\boldsymbol{P}^{-1}.$$

其中对于 Jordan 块

$$\boldsymbol{J}_k=\begin{bmatrix}\lambda_k & 1 & \cdots & 0\\ 0 & \lambda_k & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_k\end{bmatrix}_{n_k\times n_k},$$

有

$$f(\boldsymbol{J}_k)=\begin{bmatrix}f(\lambda_k) & f'(\lambda_k) & \cdots & \dfrac{f^{(n_k-1)}(\lambda_k)}{(n_k-1)!}\\ 0 & f(\lambda_k) & \cdots & \dfrac{f^{(n_k-2)}(\lambda_k)}{(n_k-2)!}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & f(\lambda_k)\end{bmatrix}.$$

Example

设

$$\boldsymbol{A}(x) = \begin{bmatrix} x^2 & 0 \\ 1 & x \end{bmatrix},$$

求 $\dfrac{\mathrm{d}\boldsymbol{A}^{-1}(x)}{\mathrm{d}x}$.

计算得

$$\boldsymbol{A}^{-1}(x) = \begin{bmatrix} \dfrac{1}{x^2} & 0 \\ -\dfrac{1}{x^3} & \dfrac{1}{x} \end{bmatrix}, \quad \frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x} = \begin{bmatrix} 2x & 0 \\ 0 & 1 \end{bmatrix},$$

从而

$$\frac{\mathrm{d}\boldsymbol{A}^{-1}(x)}{\mathrm{d}x} = -\boldsymbol{A}^{-1}(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}\boldsymbol{A}^{-1}(x) = \begin{bmatrix} -\dfrac{2}{x^3} & 0 \\ \dfrac{3}{x^4} & -\dfrac{1}{x^2} \end{bmatrix}.$$

Example

设 $\boldsymbol{X}=[x_{ij}]_{n\times n}$，求: (1) $\dfrac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}\mathrm{tr}\boldsymbol{X}$; (2) $\dfrac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}\det\boldsymbol{X}$.

(1) 根据 $\mathrm{tr}\boldsymbol{X}=\sum\limits_{i=1}^{n}x_{ii}$ 得

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}\mathrm{tr}\boldsymbol{X}=\begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{bmatrix}=\boldsymbol{I}.$$

(2) 根据 $\boldsymbol{X}\boldsymbol{X}^*=\det\boldsymbol{X}\boldsymbol{I}$, 两边对 $\boldsymbol{X}$ 求导得

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}\det\boldsymbol{X}=\boldsymbol{X}^*.$$

Example

设 $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{F} = (f_{ij})_{s\times m}$, 其中 f_{ij} 是关于 $\boldsymbol{x}$ 的 n 元函数, 证明

$$\mathrm{d}\boldsymbol{F} = \sum_{k=1}^{n} \frac{\partial \boldsymbol{F}}{\partial x_k} \mathrm{d}x_k.$$

计算得

$$\begin{aligned} \mathrm{d}\boldsymbol{F} &= (\mathrm{d}f_{ij})_{s\times m} \\ &= \left(\sum_{k=1}^{n} \frac{\partial f_{ij}}{\partial x_k} \mathrm{d}x_k \right)_{s\times m} \\ &= \sum_{k=1}^{n} \left(\frac{\partial f_{ij}}{\partial x_k} \right)_{s\times m} \mathrm{d}x_k \\ &= \sum_{k=1}^{n} \frac{\partial \boldsymbol{F}}{\partial x_k} \mathrm{d}x_k. \end{aligned}$$

Example

设 $\boldsymbol{A}(x)=[a_{ij}(x)]_{m\times n}$ 在 $[a,b]$ 可积, 证明

$$\left\|\int_a^b \boldsymbol{A}(x)\mathrm{d}x\right\|_1 \leq \int_a^b \|\boldsymbol{A}(x)\|_1 \mathrm{d}x.$$

此时

$$\begin{aligned}
\left\|\int_a^b \boldsymbol{A}(x)\mathrm{d}x\right\|_1 &= \max_{1\leq j\leq n}\sum_{i=1}^m \left|\int_a^b a_{ij}(x)\mathrm{d}x\right| \\
&\leq \max_{1\leq j\leq n}\sum_{i=1}^m \int_a^b |a_{ij}(x)|\,\mathrm{d}x \\
&= \int_a^b \max_{1\leq j\leq n}\sum_{i=1}^m |a_{ij}(x)|\,\mathrm{d}x \\
&= \int_a^b \|\boldsymbol{A}(x)\|_1 \mathrm{d}x.
\end{aligned}$$

Tips

设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^T$ 是向量, 对 $p\geq 1$, 定义 p-范数

$$\|\boldsymbol{x}\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}},$$

特别地, $\|\boldsymbol{x}\|_\infty=\max\limits_{1\leq i\leq n}|x_i|$; 设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$, 对 $p\geq 1$, 定义 p-范数

$$\|\boldsymbol{A}\|_p=\max_{\|\boldsymbol{x}\|_p=1}\|\boldsymbol{A}\boldsymbol{x}\|_p.$$

特别地, $\|\boldsymbol{A}\|_1=\max\limits_{1\leq j\leq n}\sum\limits_{i=1}^{m}|a_{ij}|$ 为列和的最大值, $\|\boldsymbol{A}\|_\infty=\max\limits_{1\leq i\leq m}\sum\limits_{j=1}^{n}|a_{ij}|$ 为行和的最大值, $\|\boldsymbol{A}\|_2$ 为 $\boldsymbol{A}$ 的最大奇异值.

Example

设 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 对称正定, 对任意的 $\boldsymbol{\alpha} \in \mathbb{R}^n$, 定义

$$\|\boldsymbol{\alpha}\| = \sqrt{\boldsymbol{\alpha}^T \boldsymbol{A}\boldsymbol{\alpha}},$$

验证 $\|\cdot\|$ 是 R^n 中的向量范数.

根据 $\boldsymbol{A}$ 对称正定知, 存在可逆矩阵 $\boldsymbol{B}$, 使得 $\boldsymbol{A} = \boldsymbol{B}^T\boldsymbol{B}$. 从而

$$\|\boldsymbol{\alpha}\| = \sqrt{\boldsymbol{\alpha}^T \boldsymbol{B}^T \boldsymbol{B}\boldsymbol{\alpha}} = \|\boldsymbol{B}\boldsymbol{\alpha}\|_2,$$

其中 $\|\cdot\|_2$ 表示 2-范数, 从而 $\|\cdot\|$ 满足范数的定义.

Example

证明

$$\sum_{k=1}^{\infty} \frac{k^2}{2^k} \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}^k$$

收敛，并求其和.

考虑级数 $\sum_{k=1}^{\infty} k^2 x^k$，可以求出其收敛半径 $R = \dfrac{1}{\lim\limits_{k\to\infty} \sqrt[k]{k^2}} = 1$. 又矩阵 $\dfrac{1}{2}\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$ 的谱半径

$$\rho\left(\frac{1}{2}\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}\right) = \frac{1}{2} < 1,$$

从而原级数收敛. 并且根据 $\sum_{k=1}^{\infty} k^2 x^k = \dfrac{1-3x}{(1-x)^3}$，计算得

$$\sum_{k=1}^{\infty} \frac{k^2}{2^k} \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}^k = \begin{bmatrix} 1 & 24 \\ 0 & 1 \end{bmatrix}.$$

谢谢大家！