华东师范大学期末试卷 (A 卷)

2024－2025 学年第一学期

考试科目：数据科学与工程数学基础　　任课教师：树扬
姓　　名：________　　学　　号：________
专　　业：________　　班　　级：________

题目	一（选择题）	二	三	四	五	六	总分	阅卷人签名
得分								

题 1　(20 分) 选择题

单选题一道 3 分，多选题一道 5 分，总计 20 分。单选题不选、错选均不得分；多选题不选、错选不得分，少选得 3 分。

(1) 若 $\boldsymbol{A}=[\mathbf{a_1},\mathbf{a_2},...,\mathbf{a_n}]\in\mathbb{R}^{m\times n}$，列空间为 $Col(\boldsymbol{A})$，行空间为 $Col(\boldsymbol{A}^T)$，零空间为 $Null(\boldsymbol{A})$，左零空间为 $Null(\boldsymbol{A}^T)$。下列说法错误的是（　）

(A) $Col(\boldsymbol{A}^T)\perp Null(\boldsymbol{A})$，$Col(\boldsymbol{A})\perp Null(\boldsymbol{A}^T)$

(B) $dim(Col(\boldsymbol{A}^T))=dim(Col(\boldsymbol{A}))=rank(\boldsymbol{A})$

(C) 若 $\mathbf{x}\in\mathbb{R}^m$ 在 $Col(\boldsymbol{A})$ 上的正交投影为 $\pi(\mathbf{x})$，则对 $\forall i=1,...,n$ 有 $\mathbf{a_i^T}(\mathbf{x}-\pi(\mathbf{x}))=0$

(D) $dim(Null(\boldsymbol{A}^T))=n-rank(\boldsymbol{A})$，$dim(Null(\boldsymbol{A}))=m-rank(\boldsymbol{A})$

(2) 已知矩阵 $\boldsymbol{A}=\begin{bmatrix}1&0&-1\\-1&2&0\end{bmatrix}$，$(\boldsymbol{A}\boldsymbol{A}^T)^{-1}=\begin{bmatrix}5/9&1/9\\1/9&2/9\end{bmatrix}$，则矩阵 $\boldsymbol{A}$ 的广义逆是（　）

(A) $\begin{bmatrix} 2/3 & -1/9 \\ 1/3 & 4/9 \\ -1/3 & -1/9 \end{bmatrix}$ (B) $\begin{bmatrix} 4/9 & -1/9 \\ 2/9 & 4/9 \\ -5/9 & -1/9 \end{bmatrix}$

(C) $\begin{bmatrix} 4/9 & -1/3 \\ 2/9 & 1/3 \\ -5/9 & -1/3 \end{bmatrix}$ (D) $\begin{bmatrix} 4/9 & -1/9 \\ 2/9 & 4/9 \\ -5/9 & 1/9 \end{bmatrix}$

(3) 下面的集合不是凸集的是（ ）

(A) 一条射线，即 $\{\mathbf{x_0} + \theta\mathbf{v} \mid \theta \geq 0, \mathbf{v} \neq \mathbf{0}\}$

(B) 若 $0 < r_1 < r_2$，$\{(x,y) \mid r_1^2 \leq (x-x_0)^2 + (y-y_0)^2 \leq r_2^2\}$

(C) 设 $||\cdot||$ 是 $\mathbb{R}^n$ 中的范数，$r > 0$，$\{\mathbf{x} \mid ||\mathbf{x} - \mathbf{x_0}|| \leq r\}$

(D) 多面体 $\{\mathbf{x}|\mathbf{Ax} \leq \mathbf{b}, \mathbf{Cx} = \mathbf{d}\}$. 其中 $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{C} \in \mathbb{R}^{p\times n}, \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m, \mathbf{d} \in \mathbb{R}^p$，$\mathbf{x} \leq \mathbf{y}$ 表示向量 $\mathbf{x}$ 的每个分量均小于等于 $\mathbf{y}$ 的对应分量。

(4) 下列关于向量范数说法错误的是（ ）

(A) 设 $\boldsymbol{u}$ 为 n 维单位列向量，I_n 为 n 维单位矩阵，$\boldsymbol{A} = I_n - 2\boldsymbol{u}\boldsymbol{u}^T$。若 $\boldsymbol{Ax} = \boldsymbol{y}$，则 $||\boldsymbol{x}||_2 = ||\boldsymbol{y}||_2$

(B) 若 $\boldsymbol{x} \in \mathbb{R}^n$，则 $||\boldsymbol{x}||_2 \leq ||\boldsymbol{x}||_1 \leq n||\boldsymbol{x}||_\infty$

(C) 若 $\boldsymbol{x} \in \mathbb{R}^n$，$p > 0$，则 $(\sum_{i=1}^n x_i^p)^{\frac{1}{p}}$ 是向量的 l_p 范数

(D) $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, ..., \boldsymbol{p}_n) \in \mathbb{R}^{n\times n}$ 为非奇异矩阵，则对于 $\forall \boldsymbol{x} \in \mathbb{R}^n$，$||\boldsymbol{Px}||_1 \leq \underset{1\leq j\leq n}{max}||\boldsymbol{p}_j||_1 \cdot ||\boldsymbol{x}||_1$

(5) 考虑一个线性映射 $\Phi : \mathbb{R}^2 \to \mathbb{R}^3$，其在标准基（基矩阵为单位阵）下的变换矩阵为：$\begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 4 & 2 \end{bmatrix}$ 我们寻找一组新的基下的 Φ 的变换矩阵。令新的基分别为：

$$\tilde{\boldsymbol{B}}=\left(\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right),\tilde{\boldsymbol{C}}=\left(\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right)$$

通过计算可得：

$$\tilde{\boldsymbol{B}}^{-1}=\begin{pmatrix}1&0\\-1&1\end{pmatrix},\quad \tilde{\boldsymbol{C}}^{-1}=\frac{1}{2}\begin{pmatrix}1&1&-1\\1&-1&1\\-1&1&1\end{pmatrix}$$

请问下面哪项为新的基下的变换矩阵（　）

$(A)\ \frac{1}{2}\begin{pmatrix}-1&3\\7&1\\5&3\end{pmatrix}$ $(B)\ \begin{pmatrix}-1&3\\7&1\\5&3\end{pmatrix}$ $(C)\ \begin{pmatrix}-5&5\\1&4\\-2&5\end{pmatrix}$ $(D)\frac{1}{2}\begin{pmatrix}-7&3\\5&1\\-1&3\end{pmatrix}$

(6)【多选】设矩阵 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$，它的完全奇异值分解为 $\boldsymbol{A}=\boldsymbol{U}\sum\boldsymbol{V}^T$，紧奇异值分解为 $\boldsymbol{A}=\boldsymbol{U}_r\sum_r\boldsymbol{V}_r^T$，$r=rank(\boldsymbol{A})$，下列关于 SVD(奇异值分解) 的说法错误的是（　）

(A) 对矩阵 $\boldsymbol{A}$ 的奇异值分解中，$\boldsymbol{U},\boldsymbol{V}$ 矩阵是唯一的

(B) $rank(\boldsymbol{A})=rank(\boldsymbol{A}^T\boldsymbol{A})=rank(\boldsymbol{A}\boldsymbol{A}^T)$

(C) $\boldsymbol{A}$ 的奇异值分解可以表示为：$\boldsymbol{A}=\sum_{i=1}^r\sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^\top$。在 $r\geq 2$ 时，令 $\boldsymbol{w}\in\mathbb{R}^n$ 为一个向量且满足：$\boldsymbol{w}=\alpha\boldsymbol{v}_1+\beta\boldsymbol{v}_2$，则 $\boldsymbol{A}\boldsymbol{w}=\alpha\sigma_1\boldsymbol{u}_1+\beta\sigma_2\boldsymbol{u}_2$

(D) $\forall i=r+1,...m,j=1,...n$，$\boldsymbol{u}_i^T\boldsymbol{A}\boldsymbol{v}_j\neq 0$

(E) 利用截断 SVD 方法，寻找秩为 k（$k<r$）的矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times n}$ 使得 $||\boldsymbol{A}-\boldsymbol{X}||_F$ 最小。寻找得到的最优矩阵就是在紧奇异值分解中对 $\sum_r$ 任意地选择 k 个奇异值 σ_i 和其对应的 $\boldsymbol{U}_r$、$\boldsymbol{V}_r$ 中的向量 $\boldsymbol{u}_i,\boldsymbol{v}_i$，再将选择到的 $\sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^\top$ 累加求和即可。

题 2 (12 分) 完成以下问题:

(1) (4 分) $A = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 0 & -3 \end{bmatrix}$ 分别求 A 的 l_1 范数，1 范数，l_∞ 范数，∞ 范数

(2) (5 分) 求向量 $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ 在矩阵 $M = \begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 4 & 2 \end{bmatrix}$ 的列空间上的正交投影。

【已知：$(M^T M)^{-1} = \frac{1}{72}\begin{bmatrix} 7 & -5 \\ -5 & 7 \end{bmatrix}$ 】

(3) (3 分) 设 $\mathbf{B} \in \mathbb{R}^{n\times n}$，$I$ 是 n 阶单位矩阵，$||\mathbf{B}||$ 是关于 $\mathbf{B}$ 的矩阵 l_2 范数。已知 $||\mathbf{B}|| < 1$，$I - \mathbf{B}$ 可逆，证明：$(1 - ||\mathbf{B}||)\cdot||(I - \mathbf{B})^{-1}|| \le n$。

【提示：$I = (I - \mathbf{B})^{-1}\cdot(I - \mathbf{B}) = (I - \mathbf{B})^{-1} - (I - \mathbf{B})^{-1}\mathbf{B}$，

则 $(I - \mathbf{B})^{-1} = I + (I - \mathbf{B})^{-1}\mathbf{B}$ 】

题 3 (17 分)

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 3 & 1 \\ 0 & 4 & -2 \\ 2 & 1 & -1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 17 \\ 6 \\ 1 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$$

(1) (2 分) 矩阵 $\mathbf{A}$ 能否进行 QR 分解，为什么？直接写出结论及原因即可。

(2) (6 分) 求矩阵 $\mathbf{B}$ 的 QR 分解。

(3) (5 分) 利用 (2) 中的分解结果来求解方程组 $\mathbf{Bx} = \mathbf{c}$

(4) (4 分) 利用正规化方程组，求解 $\mathbf{A}$ 和 $\mathbf{d}$ 所对应的最小二乘问题 $\min\limits_{\mathbf{x}}||\mathbf{Ax}-\mathbf{d}||_2$ 的全部解。【对正规化方程组的求解方法不限】

题 4 (15 分)

(1) (6 分) 给定矩阵 $\mathbf{A}=\begin{bmatrix}1 & 3 & 2\\ -1 & 2 & 3\\ 4 & 2 & -2\end{bmatrix}$，分别求其完全 SVD 和紧 SVD。

【已知：$\mathbf{A}^{\mathbf{T}}\mathbf{A}=\begin{bmatrix}18 & 9 & -9\\ 9 & 17 & 8\\ -9 & 8 & 17\end{bmatrix}$，$\mathbf{A}^{\mathbf{T}}\mathbf{A}$ 的特征值 $\lambda_1=27,\lambda_2=25,\lambda_3=0$，特征向量为 $\boldsymbol{q}_1=\begin{bmatrix}-2 & -1 & 1\end{bmatrix}^T,\boldsymbol{q}_2=\begin{bmatrix}0 & 1 & 1\end{bmatrix}^T,\boldsymbol{q}_3=\begin{bmatrix}-1 & 1 & -1\end{bmatrix}^T$】

(2) (4 分) 假设 $\boldsymbol{M}$ 是任意一个非奇异 $n\times n$ 的矩阵，已知其奇异值分解（SVD）为 $\boldsymbol{M}=\boldsymbol{U}\Sigma\boldsymbol{V}^\top$，其中 $\boldsymbol{U}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n]$，$\Sigma=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_n)$，$\boldsymbol{V}=[\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n]$。请写出 $\boldsymbol{M}$ 的逆矩阵的 SVD 分解。

(3) (5 分) 已知矩阵 $\boldsymbol{M}\in\mathbb{R}^{m\times n}$ 的元素非负，$r=rank(\boldsymbol{M})$，其奇异值分解为 $\boldsymbol{M}=\boldsymbol{U}\Sigma\boldsymbol{V}^\top$，其中 $\boldsymbol{U}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_r]$，$\Sigma=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_r)$，$\boldsymbol{V}=[\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_r]$。求拼接矩阵 $\boldsymbol{A}=\begin{bmatrix}O & \boldsymbol{M}\\ \boldsymbol{M}^T & O\end{bmatrix}$ 的非零特征值和其对应的特征向量。【结果用 $\boldsymbol{u}_i,\boldsymbol{v}_i,\sigma_i$ 相关形式表示】

题 5 (19 分)

(1) (2 分) 已知 $\mathbf{X}\in\mathbb{R}^{n\times n}$ 非奇异，求证：$d(\mathbf{X}^{-1})=-\mathbf{X}^{-1}d\mathbf{X}\mathbf{X}^{-1}$。

(2) (5 分) 利用迹微分法求函数 $f(\mathbf{X})=Tr(\mathbf{X}^\top\mathbf{X}^{-1}\mathbf{A})$ 关于变量 $\mathbf{X}$ 的梯度矩阵，其中 $\mathbf{X}\in\mathbb{R}^{n\times n}$ 非奇异，$\mathbf{A}\in\mathbb{R}^{n\times n}$ 是常数矩阵。

(3) (7 分) 考虑一个两层的全连接神经网络：

$$\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{x})=\mathrm{ReLU}(\boldsymbol{A}_2(\mathrm{ReLU}(\boldsymbol{A}_1\boldsymbol{x}+\boldsymbol{b}_1))+\boldsymbol{b}_2)$$

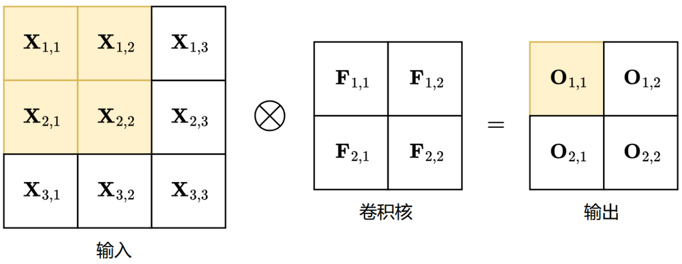
ReLU 的含义：若 $\boldsymbol{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\in\mathbb{R}^n$，则 $\mathrm{ReLU}(\boldsymbol{x})=\begin{bmatrix}\mathrm{ReLU}(x_1)\\ \vdots\\ \mathrm{ReLU}(x_n)\end{bmatrix}$，其中

$$\mathrm{ReLU}(x_i)=\begin{cases}0, & \text{if } x_i<0\\ x_i, & \text{if } x_i\geq 0\end{cases}$$

已知：

$$\boldsymbol{A}_1=\begin{bmatrix}2 & 3\\ -2 & 1\\ 3 & -1\end{bmatrix},\boldsymbol{A}_2=\begin{bmatrix}3 & -1 & 0\\ 1 & -2 & 2\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}0\\ 0\\ -3\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-7\\ 3\end{bmatrix}$$

假设输入为 $\boldsymbol{x}=(-1,2)^T$，并且对应的真实输出为 $\hat{\boldsymbol{y}}=(0,1)^T$，采用平方损失 $L=\frac{1}{2}\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|_2^2$。试计算函数 L 关于 $\boldsymbol{b}_1$ 的梯度。



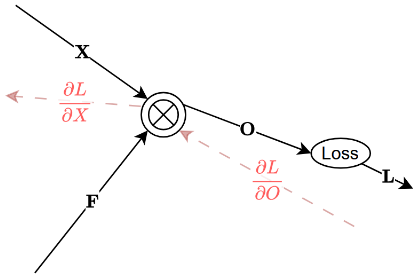
$$\mathbf{O}_{1,1}=\mathbf{X}_{1,1}\times\mathbf{F}_{1,1}+\mathbf{X}_{1,2}\times\mathbf{F}_{1,2}+\mathbf{X}_{2,1}\times\mathbf{F}_{2,1}+\mathbf{X}_{2,2}\times\mathbf{F}_{2,2}$$

$$\mathbf{O}_{1,2}=\mathbf{X}_{1,2}\times\mathbf{F}_{1,1}+\mathbf{X}_{1,3}\times\mathbf{F}_{1,2}+\mathbf{X}_{2,2}\times\mathbf{F}_{2,1}+\mathbf{X}_{2,3}\times\mathbf{F}_{2,2}$$

$$\mathbf{O}_{2,1}=\mathbf{X}_{2,1}\times\mathbf{F}_{1,1}+\mathbf{X}_{2,2}\times\mathbf{F}_{1,2}+\mathbf{X}_{3,1}\times\mathbf{F}_{2,1}+\mathbf{X}_{3,2}\times\mathbf{F}_{2,2}$$

$$\mathbf{O}_{2,2}=\mathbf{X}_{2,2}\times\mathbf{F}_{1,1}+\mathbf{X}_{2,3}\times\mathbf{F}_{1,2}+\mathbf{X}_{3,2}\times\mathbf{F}_{2,1}+\mathbf{X}_{3,3}\times\mathbf{F}_{2,2}$$

(a)卷积过程演示



(b)卷积的反向传播演示

(4) (5 分) 卷积是常用的数学运算，运算过程中，卷积核矩阵 $\mathbf{F}$ 在输入矩阵 $\mathbf{X}$ 上滑动，卷积核每滑动到与输入矩阵的某一子矩阵重叠时，卷积核与该子

矩阵对应位置元素相乘再累加，得到输出结果在该位置的值。以步长（每次滑动的距离）等于 1 为例，其得到输出矩阵 $\mathbf{O}$ 的过程和公式如图所示。

已知，输入 $\mathbf{X} = \begin{bmatrix} 3 & 1 & 2 \\ 7 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = (x_{ij})_{3\times 3}$，卷积核 $\mathbf{F} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} = (f_{ij})_{2\times 2}$

根据卷积过程易得输出 $\mathbf{O} = \begin{bmatrix} 0 & 4 \\ 14 & 9 \end{bmatrix}$，$L = Loss(\mathbf{O})$ 是关于 $\mathbf{O}$ 的某种损失函数。现在假设 $\frac{\partial L}{\partial \mathbf{O}} = \begin{bmatrix} 3 & 7 \\ 6 & 8 \end{bmatrix}$，请据此求解 $\frac{\partial L}{\partial x_{11}}$ 和 $\frac{\partial L}{\partial \mathbf{X}}$

题 6　(17 分)

(1) (5 分) 判断函数 $f(\mathbf{x}) = \max(\|\mathbf{A}\mathbf{x}+\mathbf{b}\|_2, \sqrt{\mathbf{x}^T\mathbf{x}})+\frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x}$（其中 $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$，$\mathbf{P}$ 为 n 阶半正定矩阵）是否为凸函数，并说明理由。

(2) (4 分) 考虑优化问题 $min f(\mathbf{x}) = x_1^2 + 4x_1x_2$，从初始点 $\mathbf{x}^{(0)} = (1,0)^T$ 出发，写出用梯度下降法迭代一步的过程，迭代时采用精确线搜索方法。

(3) (4 分) 利用二阶最优性条件找到问题 $min f(\mathbf{x}) = x_1^2 + 4x_2^2 + 2x_1x_2$ 的全局最优解。

(4) (4 分) 证明 $f(\mathbf{x}) = (\prod_{k=1}^{n} x_k)^{\frac{1}{n}}$，($\mathbf{x} \in \mathbb{R}^n$ 且 $x_i > 0$) 是凹函数。

【提示，已知:

$$\nabla f(\mathbf{x}) = \left(\frac{f(\mathbf{x})}{nx_1}, \frac{f(\mathbf{x})}{nx_2}, \ldots, \frac{f(\mathbf{x})}{nx_n}\right)^\top, \quad \nabla^2 f(\mathbf{x}) = -\frac{f(\mathbf{x})}{n^2}\begin{bmatrix} \frac{n-1}{x_1^2} & -\frac{1}{x_1x_2} & \cdots & -\frac{1}{x_1x_n} \\ -\frac{1}{x_2x_1} & \frac{n-1}{x_2^2} & \cdots & -\frac{1}{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{x_nx_1} & -\frac{1}{x_nx_2} & \cdots & \frac{n-1}{x_n^2} \end{bmatrix}.$$】