2020 级华东师范大学本科生期末考试

《数据科学与工程数学基础》笔试试卷 2022.9

学院：数据科学与工程学院　　考试形式：闭卷　　所需时间：120 分钟

考生姓名：_______　　学号：_______　　专业：_______　　任课教师：黄定江

题目	一	二	三	四	五	六	七	八	总分
满分	14	12	13	13	13	12	12	11	100
得分									
阅卷人									

题 1 (14') 已知矩阵

$$\boldsymbol{M}=\begin{pmatrix}1 & 0\\ -2 & 3\\ 0 & 1\end{pmatrix}$$

(1) 计算矩阵 $\boldsymbol{M}$ 的 1 范数；

(2) 计算矩阵 $\boldsymbol{M}$ 的 2 范数；

(3) 证明：对任意 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$，由

$$\|\boldsymbol{A}\|_{m_\infty}:=\max_{1\le i\le m,1\le j\le n}|a_{ij}|$$

定义的 $\|\cdot\|_{m_\infty}$ 是 $\mathbb{R}^{m\times n}$ 上的（广义）矩阵范数。

题 2 (12’) 已知矩阵

$$M=\begin{pmatrix}1 & -1 & 0\\ 2 & 4 & 1\\ 4 & 2 & 1\end{pmatrix}$$

(1) 求矩阵 $\boldsymbol{M}$ 的零空间；

(2) 求向量 $\boldsymbol{x}=(1,1,1)^\top$ 在矩阵 $\boldsymbol{M}$ 的列空间上的正交投影；

(3) 利用 Householder 变换，将非零向量 $\boldsymbol{x}=(1,1,1)^\top$ 变为标准向量 $\boldsymbol{e}_1=(0,1,0)^\top$ 的 $\sqrt{3}$ 倍，请写出这个 Householder 变换矩阵。

题 3 (13’) 矩阵

$$A=\begin{pmatrix}3 & 2 & 2\\ 2 & 3 & -2\end{pmatrix}$$

(1) 计算矩阵 $\boldsymbol{A}$ 的 SVD 分解；

(2) 假设 $\boldsymbol{B}$ 是一个 $n\times d$ 的矩阵，矩阵 $\boldsymbol{M}$ 是 $(n+d)\times(n+d)$ 定义为

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{O} & \boldsymbol{B}^\top\\ \boldsymbol{B} & \boldsymbol{O}\end{pmatrix}$$

显然 $\boldsymbol{M}$ 是对称矩阵。请证明矩阵 $\boldsymbol{M}$ 的对角化会产生 $\boldsymbol{B}$ 的奇异值分解所需要的所有信息。

题 4 (13’) 完成下列函数求导或梯度：

(1) 求激活函数 $\sigma(x)=\frac{1}{1+e^{-x}}$ 的导数；

(2) 求函数 $f(\boldsymbol{X})=Tr(\boldsymbol{X}^{-1})$ 对 $\boldsymbol{X}$ 的梯度，其中 $\boldsymbol{X}\in\mathbb{R}^{m\times n}$。

(3) 若非奇异矩阵 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$，$\boldsymbol{x}\in\mathbb{R}^{n\times 1}$，$\boldsymbol{y}\in\mathbb{R}^{n\times 1}$，求函数 $f(\boldsymbol{A})=\boldsymbol{y}^\top\boldsymbol{A}^{-1}\boldsymbol{x}$ 对 $\boldsymbol{A}$ 的梯度。

题 5 (13') 计算和证明以下问题：

(1) 同时抛 2 颗骰子，事件 A, B, C 分别表示为

(A) 仅有一个骰子是 3

(B) 至少一个骰子是 4

(C) 骰子上点数总和为 7

试计算事件 A, B, C 发生后所提供的信息量；

(2) 证明联合熵和条件熵有如下关系：

$$H(X,Y) = H(X) + H(Y \mid X)$$

题 6 (12') 求解以下问题：

(1) 假设总体 $X \sim N\left(\mu, \sigma_0^2\right)$ (σ_0^2 已知), $X_1, X_2, \ldots, X_n$ 为来自总体 X 的样本, 由过去的经验和知识, 我们可以确定 μ 的取值比较集中在 $\hat{\mu}$ 附近, 离 $\hat{\mu}$ 越远, μ 取值的可能性越小, 于是我们假定 μ 的先验分布为正态分布

$$\pi(\mu) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left[-\frac{1}{2\hat{\sigma}^2}(\mu - \hat{\mu})^2\right] \quad (\hat{\mu}, \hat{\sigma}\text{已知})$$

求 μ 的后验概率分布。

(2) 假设总体 $X \sim P(\lambda), X_1, X_2, \ldots, X_n$ 为来自总体 X 的样本, 假定 λ 的先验分布为伽玛分布 $\Gamma(\alpha, \beta)$, 求 λ 的后验期望估计（泊松分布的分布律：$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$，泊松分布的期望为 λ；伽马分布的概率密度函数：$f(x;\alpha,\beta) = \frac{x^{\alpha-1}e^{-\beta x}\beta^\alpha}{\Gamma(\alpha)}$，伽马分布的期望为 α/β）。

题 7 (12') 求证下列与凸集或凸函数相关的问题：

(1) 证明：设 $f : \mathbb{R}^n \to \mathbb{R}^m$ 是仿射变换, 即 $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m$, 则凸集在 f 下的像是凸集:

$$S \subseteq \mathbb{R}^n \text{ 为凸集} \Rightarrow f(S) \overset{\text{def}}{=} \{f(\boldsymbol{x}) \mid \boldsymbol{x} \in S\} \text{ 为凸集;}$$

(2) 判定函数 $f(x)=\max(\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}\|_2,\|\boldsymbol{x}^\top\boldsymbol{x}\|_1),\boldsymbol{A}\in\mathbb{R}^{m\times n},\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{b}\in\mathbb{R}^m$ 是否为凸函数，并说明理由。

题 8 (11’) 考虑优化问题

$$\min_{\boldsymbol{x}} \quad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{x}$$

其中 $\boldsymbol{W}\in S_+^n$ 是对称半正定矩阵。

(1) 求出在任意点 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^\top$ 处沿负梯度方向迭代的最佳步长 α。

(2) 若对上述优化问题增加约束条件

$$x_i^2=1,\quad i=1,\cdots,n$$

请写出此时约束优化问题的拉格朗日函数 $L(\boldsymbol{x},\boldsymbol{\lambda})$，并计算出约束优化问题的拉格朗日对偶函数。