# 数据科学与工程数学基础

## 作业提交规范及第 4 次作业

教师：黄定江
助教：刘文辉

2023 年 1 月 7 日

### 作业提交规范

1. 作业提交形式：**pdf 格式文档，可以使用 Word、LaTeX 或手写所得到的电子文档**。建议博士生均使用 LaTeX 编写。
2. 作业命名规范：提交的电子文档必须命名为："**hw4_ 学号 _ 姓名**"。其中，hw4 表示第 4 次作业。命名示例：hw4_50000000000_ 刘某某。
3. 作业提交途径：点击打开网址：**第 4 次作业提交传送门**，无需注册和登录，直接上传作业文档即可。注意：传送门将会在截至时间点到达后自动关闭。
4. 作业更改说明：如果需要修改已经提交的作业，只要在截至日期前，再次上传更改后的作业，我将批改最新时间提交的作业。

### 第 4 次作业

提交截至时间：**2023/01/21 12:00（中午）**

理论部分

**习题 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) 骰子上点数总和为偶数。
试计算事件 $A, B, C$ 发生后所提供的信息量

**习题 2.** 证明：在多分类问题中，利用交叉熵函数作为损失函数和用 *KL* 散度作为损失函数是等价的。

**习题 3.** *(互信息)* 假设 $X_1 \to X_2 \to X_3 \to \cdots \to X_n$ 是一个马尔科夫链，即

$$p\left(x_1, x_2, \ldots, x_n\right) = p\left(x_1\right) p\left(x_2 \mid x_1\right) \cdots p\left(x_n \mid x_{n-1}\right)$$

试化简 $I\left(X_1; X_2, \ldots, X_n\right)$

**习题 4.** 假设总体 $X \sim N\left(\mu, \sigma^2\right)$ *(* $\sigma^2$ 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 $\mu$ 的取值比较集中在 $\mu_0$ 附近, 离 $\mu_0$ 越远, $\mu$ 取值的可能性越小, 于是我们假定 $\mu$ 的先验分布为正态分布

$$\pi(\mu) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp\left[-\frac{1}{2\sigma_\mu^2}\left(\mu - \mu_0\right)^2\right] \quad \left(\mu_0, \sigma_\mu \text{已知}\right)$$

求 $\mu$ 的后验概率分布。

**习题 5.** 证明*: Gauss* 概率密度函数的累积分布函数 $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du$ 是对数-凹函数. 即 $\log(\Phi(x))$ 是凹函数。

**习题 6.** 计算函数 $f(x)$ 的共轭函数，以及共轭函数的定义域。

*(1)* $f(x) = -\log x$

*(2)* $f(x) = e^x$

**习题 7.** 写出下述非线性规划的 *KKT* 条件并求解

$$\begin{aligned}
(1) \quad \mathtt{maximize} \quad & f(x) = (x-3)^2 \\
\mathtt{suject\ to} \quad & 1 \leq x \leq 5 \\
(2) \quad \mathtt{minimize} \quad & f(x) = (x-3)^2 \\
\mathtt{suject\ to} \quad & 1 \leq x \leq 5
\end{aligned}$$

**习题 8.** 考虑等式约束的最小二乘问题

$$\begin{aligned}
\mathtt{minimize} \quad & \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \\
\mathtt{suject\ to} \quad & \mathbf{G}\mathbf{x} = \mathbf{h}
\end{aligned}$$

其中 $\mathbf{A} \in \mathbb{R}^{m \times n}, \operatorname{rank}(\mathbf{A}) = n, \mathbf{G} \in \mathbb{R}^{p \times n}, \operatorname{rank}(\mathbf{G}) = p$. 给出 *KKT* 条件, 推导原问题最优解 $x^*$ 以及对偶问题最优解 $v^*$ 的表达式.

**习题 9.** 用 *Lagrange* 乘子法证明：矩阵 $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ 的 *2* 范数

$$\|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2 = 1, \boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{A}\boldsymbol{x}\|_2$$

的平方是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的最大特征值。

**习题 10.** 用 *Lagrange* 乘子法求欠定方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的最小二范数解，其中 $\boldsymbol{A} \in \mathbb{R}^{m\times n}, m \leq n, \mathrm{rank}(\boldsymbol{A}) = m$

**习题 11.** 用最速下降法和精确线搜索计算 $\min f(x) = x_1^2+x_2^2+x_3^2$ ，初始点 $x^{(0)} = (2,2,1)^T$. 当 $(f(x^{(n+1)}) - f(x^{(n)})) < 0.001$ 时迭代终止.

## 实操部分（**该题选做**）

**习题 12.** 复现 *Lec34* 讲的 *34.1.3* 中的实际应用：牛顿法求解 *Logistic* 回归。