数据科学与工程数学基础

作业提交规范及第 3 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：“**hw3_ 学号 _ 姓名**”。其中，hw3 表示第 3 次作业。命名示例：hw3_50000000000_ 刘某某。
3. 作业提交途径：点击打开网址：**第 3 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 3 次作业

提交截至时间：**2023/01/21 12:00（中午）**

理论部分

习题 1. 设 $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$，$\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 用正规化方法求对应的 LS 问题的解。

习题 2. 设 $\boldsymbol{A}=\begin{bmatrix}1&3&1&1\\2&0&0&0\\1&0&0&0\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 用任意一种方法求对应的 *LS* 问题的全部解。

习题 3. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 且存在 $\boldsymbol{X}\in\mathbb{R}^{n\times m}$ 使得对每一个 $\boldsymbol{b}\in\mathbb{R}^{m}$,$\boldsymbol{x}=\boldsymbol{X}\boldsymbol{b}$ 均极小化 $\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2$. 证明 $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}=\boldsymbol{A}$ 和 $(\boldsymbol{A}\boldsymbol{X})^T=\boldsymbol{A}\boldsymbol{X}$.

习题 4. 利用等式

$$\|\boldsymbol{A}(\boldsymbol{x}+\alpha\boldsymbol{w})-\boldsymbol{b}\|_2^2=\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2+2\alpha\boldsymbol{w}^T\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})+\alpha^2\|\boldsymbol{A}\boldsymbol{w}\|_2^2$$

证明：如果 $\boldsymbol{x}\in\boldsymbol{X}_{LS}$, 那么 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^T\boldsymbol{b}$

习题 5.

$$A:=\begin{pmatrix}5&-1&1\\-1&2&0\\1&0&3\end{pmatrix}$$

记 $\Lambda(A)=\{\lambda_1,\lambda_2,\lambda_3\}\subseteq\mathbb{C}$ *with* $|\lambda_1|\geq|\lambda_2|\geq|\lambda_3|$.
(i) 使用 *Gerschgorin* 圆盘定理, 证明 $\frac{|\lambda_1|}{|\lambda_3|}\leq 7$.（注：由于 A 为对称矩阵，$\frac{|\lambda_1|}{|\lambda_3|}$ 为 A 的条件数）
(ii)（编程题，提交代码）使用幂法与反幂法计算 $\frac{|\lambda_1|}{|\lambda_3|}$

习题 6. 构建模型使得预测值与真实值的误差最小常用向量 2-范数度量，求解模型过程中需要计算梯度，求梯度：

- $f(A)=\frac{1}{2}\|Ax+b-y\|_2^2$，求 $\frac{\partial f}{\partial A}$
- $f(x)=\frac{1}{2}\|Ax+b-y\|_2^2$，求 $\frac{\partial f}{\partial x}$

其中 $A\in R^{m\times n}$，$x\in R^n$，$b,y\in R^m$

习题 7. 二次型是数据分析中常用函数，求 $\frac{\partial x^TAx}{\partial x}$，$\frac{\partial x^TAx}{\partial A}$，其中 $A\in R^{m\times m}$，$x\in R^m$

习题 8. 利用迹微分法求解 $\frac{\partial Tr(W^{-1})}{\partial W}$，其中 $W\in R^{m\times m}$

习题 9. $(\exp(\boldsymbol{z}))_i=\exp(\boldsymbol{z}_i)$，$(\log(\boldsymbol{z}))_i=\log(\boldsymbol{z}_i)$ $f(\boldsymbol{z})=\frac{\exp(\boldsymbol{z})}{\mathbf{1}^T\exp(\boldsymbol{z})}$ 称为 *softmax* 函数，，如果 $\boldsymbol{q}=f(\boldsymbol{z}),J=-\boldsymbol{p}^Tlog(\boldsymbol{q})$，其中 $\boldsymbol{p},\boldsymbol{q},\boldsymbol{z}\in\mathbb{R}^n$，并且 $\mathbf{1}^{\mathrm{T}}\boldsymbol{p}=1$，

- 证：$\frac{\partial J}{\partial \boldsymbol{z}}=\boldsymbol{q}-\boldsymbol{p}$
- 若 $\boldsymbol{z}=\boldsymbol{W}\boldsymbol{x}$，其中 $\boldsymbol{W}\in\mathbb{R}^{n\times m}$,$\boldsymbol{x}\in\mathbb{R}^m$，$\frac{\partial J}{\partial \boldsymbol{W}}=(\boldsymbol{q}-\boldsymbol{p})\boldsymbol{x}^T$ 是否成立。

习题 10. 以下内容是利用极大似然估计求解多元正态分布模型的关键步骤:$L=-\frac{Nd}{2}ln(2\pi)-\frac{N}{2}ln|\boldsymbol{\Sigma}|-\frac{1}{2}\sum_t(\boldsymbol{x}_t-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_t-\boldsymbol{\mu})$，$L$ 是对数似然，N 为样本数，d 为样本维数，$\Sigma\in\mathbb{R}^{d\times d}$ 为协方差矩阵，$\mu\in\mathbb{R}^d$ 为期望向量。

1) 求 $\frac{\partial L}{\partial \boldsymbol{\mu}}$

2) 当 $\mu=\frac{1}{N}\sum_t\boldsymbol{x}_t$ 时，求 $\frac{\partial L}{\partial \boldsymbol{\Sigma}}$，并求使 $\frac{\partial L}{\partial \boldsymbol{\Sigma}}=0$ 成立的 $\boldsymbol{\Sigma}$。

实操部分

习题 11. 随机生成一个对称半正定的矩阵，利用幂法迭代求得其最大特征值。