数据科学与工程数学基础

作业提交规范及第 13 次作业

教师：黄定江
助教：陈诺、刘文辉

2022 年 4 月 12 日

作业提交规范

1. 作业提交形式：**练习本或笔记本**（建议统一使用一般的**练习本**即可，不接收以纸张的方式书写的作业）。
2. 作业书写说明：
   (a) 可以讨论，禁止抄袭！
   (b) 练习本封面至少包含两方面信息：**姓名**和**学号**
   (c) 每一次的作业**请另起一页**，并在**第一行标明第几次作业**。例如“第 13 次作业”；
   (d) 每一题请**标注题号**，无需抄题，直接解答；
   (e) 题与题之间**请空一行**；
   (f) 不要求字好，但要求书写整体清晰易读。
3. 作业提交途径：纸质作业交给**学习委员**，由学习委员**按学号顺序**收齐后统一在截止日期前交到**助教实验室**。**单数周**布置的作业交到助教刘文辉处**数学馆西 109**；**双数周**布置的作业交到助教陈诺处**地理馆 353**。
4. 作业评分说明：正常提交作业的按照实际评分记录；逾期补交作业的根据逾期情况在实际评分基础上酌情扣分；未交作业的当次作业记为 0 分。

第 13 次作业

提交截至时间：**暂定 2022/04/23 下周五 20:00（晚上）**

理论部分 (最小二乘问题)

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

解. 该 LS 问题的解就是下列正规化方程组的解:

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$$

即

$$\begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} 9 \\ 12 \end{bmatrix}$$

解得：$\boldsymbol{x} = (-1, 1)^T$ 对于非满秩的 A，也可以先行变换后消去多余行再对 LS 问题求解。

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。

解. 该 LS 问题的解就是下列正规化方程组的解:

$$\begin{bmatrix} 6 & 3 & 1 & 1 \\ 3 & 9 & 3 & 3 \\ 1 & 3 & 1 & 1 \\ 1 & 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \boldsymbol{x}_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 1 \\ 1 \end{bmatrix}$$

初等行变换得到同解方程组

$$\begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 15 & 5 & 5 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \boldsymbol{x}_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

从而

$$\begin{bmatrix} 1 & 3 \\ 0 & 15 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} 1 - \boldsymbol{x}_3 - \boldsymbol{x}_4 \\ 2 - 5\boldsymbol{x}_3 - 5\boldsymbol{x}_5 \end{bmatrix}$$

即

$$\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{2}{15} - \frac{1}{3}\boldsymbol{x}_3 - \frac{1}{3}\boldsymbol{x}_4 \end{bmatrix}$$

其中 $\boldsymbol{x}_3, \boldsymbol{x}_4 \in R$

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 $(\boldsymbol{A}\boldsymbol{X})^T = \boldsymbol{A}\boldsymbol{X}$.

证明. 由 $\boldsymbol{b}$ 的任意性, 取 $\boldsymbol{b}$ 分别为 $\boldsymbol{A}$ 的每一列 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$, 则显然, 若 $\boldsymbol{x}$ 极小化 $\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{a}_i\|_2$, $\boldsymbol{x}$ 可以取第 i 个元素为 1, 其余元素为 0 的向量, 因此 $\boldsymbol{X}$ 使得 $\boldsymbol{x} = \boldsymbol{X}\boldsymbol{a}_i$ 最小化的 $\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{a}_i - \boldsymbol{a}_i\|_2 = 0$, 这样 $\boldsymbol{A}\boldsymbol{X}\boldsymbol{a}_i = \boldsymbol{a}_i$, 即

$$\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} = \boldsymbol{A}$$

因为对每一个 $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{x} = \boldsymbol{X}\boldsymbol{b}$ 均极小化 $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2$。有 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\boldsymbol{b} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$。

由于 $\boldsymbol{b}$ 的任意性, 有 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{A}^{\mathrm{T}}$, 等式两边同时乘以 $\boldsymbol{X}^{\mathrm{T}}$, 有

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$$

即

$$(\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{X}) = (\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}$$

所以

$$\boldsymbol{A}\boldsymbol{X} = (\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{X}) = (\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}$$

证毕。 □

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 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$

解. 设 $f(\alpha) = \|\boldsymbol{A}(\boldsymbol{x} + \alpha\boldsymbol{w}) - \boldsymbol{b}\|_2^2$, 由于 $\boldsymbol{x} \in \boldsymbol{X}_{LS}$, 说明当 $\alpha = 0$ 时, 函数取极小点。由于 f 是关于 α 的二次函数, 故在 $\alpha = -\frac{2\boldsymbol{w}^T\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})}{2\alpha^2\|\boldsymbol{A}\boldsymbol{w}\|_2^2}$ 取得极值点。代入 $\alpha = 0$, 有

$$\boldsymbol{w}^T\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = 0$$

又由于 $\boldsymbol{w}$ 的任意性, 有

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$$