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：陈诺、刘文辉

2022 年 10 月 22 日

作业提交规范

1. 作业提交形式：**使用 Word 或 LaTeX 编写所得到的电子文档**。若使用 Word 编写，将其另存为 PDF 形式，然后提交 PDF 文档。若使用 LaTeX 编写，将其编译成 PDF 形式，然后提交 Tex 和 PDF 两个文档。
2. 作业命名规范：提交的电子文档必须命名为：“**学号 _ 姓名**”。命名示例：10175501112_ 陈诺。
3. 作业提交途径：点击打开每次作业的传送门网址：**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（切记保持同名），即可覆盖已有作业。
5. 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 4 次作业

提交截至时间：**2022/10/28 周五 12:00（中午）**

理论部分

习题 1. 定义

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

(i) 给出矩阵 $A^{\mathrm{T}}A$ 的 *Cholesky* 分解 $A^{\mathrm{T}}A = GG^{\mathrm{T}}$

(ii) 试说明 $\|A^{\mathrm{T}}A\|_2 = \|A\|_2^2 = \|G\|_2^2$

解. *(i)* 记

$$M = A^{\mathrm{T}}A = \begin{pmatrix} 14 & -2 & 0 \\ -2 & 6 & -4 \\ 0 & -4 & 3 \end{pmatrix}$$

消除 M 的第一列中的对角线条目

$$L_1M = \begin{pmatrix} \sqrt{14} & -\frac{2}{\sqrt{14}} & 0 \\ 0 & \frac{40}{7} & -4 \\ 0 & -4 & 3 \end{pmatrix} \qquad L_1 = \begin{pmatrix} \frac{1}{\sqrt{14}} & 0 & 0 \\ \frac{1}{7} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

L_1M 右乘 L_1^{T}

$$L_1ML_1^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{40}{7} & -4 \\ 0 & -4 & 3 \end{pmatrix}$$

消除 $L_1ML_1^{\mathrm{T}}$ 的第二列中的对角线条目

$$L_2L_1ML_1^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2\sqrt{70}}{7} & -\frac{\sqrt{70}}{5} \\ 0 & 0 & \frac{1}{5} \end{pmatrix} \qquad L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{70}}{20} & 0 \\ 0 & \frac{7}{10} & 1 \end{pmatrix}$$

L_1M 右乘 $L_2L_1ML_1^{\mathrm{T}}L_2^{\mathrm{T}}$

$$L_2L_1ML_1^{\mathrm{T}}L_2^{\mathrm{T}} = \mathrm{diag}_{3\times 3}\left(1, 1, \frac{1}{5}\right).$$

令 $L_3 := \mathrm{diag}_{3\times 3}(1, 1, \sqrt{5})$ 使得 $L_3L_2L_1ML_1^{\mathrm{T}}L_2^{\mathrm{T}}L_3^{\mathrm{T}} = I_3$. 我们有 $M = A^{\mathrm{T}}A = GG^{\mathrm{T}}$ 其中

$$G = L_1^{-1}L_2^{-1}L_3^{-1} = \begin{pmatrix} \sqrt{14} & 0 & 0 \\ -\frac{\sqrt{14}}{7} & \frac{2\sqrt{70}}{7} & 0 \\ 0 & -\frac{\sqrt{70}}{5} & \frac{\sqrt{5}}{5} \end{pmatrix}$$

(ii) 令 $G = U\Sigma V^{\mathrm{T}}$ ，其中 $U, V \in \mathbb{R}^{n\times n}$ 正交，$\Sigma = \mathrm{diag}_{n\times n}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{n\times n}$ 且 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. 故 $A^{\mathrm{T}}A = GG^{\mathrm{T}} = U\Sigma V^{\mathrm{T}}V\Sigma U^{\mathrm{T}} = U\Sigma^2U^{\mathrm{T}}$ ，$A^{\mathrm{T}}A$ 的奇异值为 $\sigma_1^2, \ldots, \sigma_n^2$. 因此 $\|A^{\mathrm{T}}A\|_2 = \sigma_1^2 = \|G\|_2^2$. 同样的，我们可以得出 $\|A^{\mathrm{T}}A\|_2 = \|A\|_2^2$.

习题 2. 对 $k \in \mathbb{N}_0$, 定义

$$A = \begin{pmatrix} -8 & 5 & 1 \\ -4 & 7 & 5 \\ -8 & 5 & 1 \\ -4 & 7 & 5 \end{pmatrix}, \quad \gamma_k = \inf_{\substack{M \in \mathbb{R}^{3\times 4} \\ \mathrm{rk}(M) \le k}} \left\| A^{\mathrm{T}} - M \right\|_2$$

(i) 计算矩阵 A 的 *SVD* 分解 $A = U\Sigma V^{\mathrm{T}}$，并使 $2U$ 为 *Hadamard* 矩阵

(ii) 使用 *(i)* 中的结论, 求 $\operatorname{rank}(A), \mathcal{R}(A), \mathcal{N}(A), \|A\|_2, \|A\|_F$

(iii) 对每个 $k \in \mathbb{N}_0$, 计算 γ_k 并找出矩阵 $A_k \in \mathbb{R}^{3\times 4}$ 使得 $\operatorname{rank}(A_k) \le k$ 且 $\left\| A^{\mathrm{T}} - A_k \right\|_2 = \gamma_k$

解. *(i)*

$$A^{\mathrm{T}} A = 4 \begin{pmatrix} 40 & -34 & -14 \\ -34 & 37 & 20 \\ -14 & 20 & 13 \end{pmatrix}$$

特征多项式 $p_{A^{\mathrm{T}}A}(z) = \det\left(zI_3 - A^{\mathrm{T}}A\right) = z(z-36)(z-324)$. $A^{\mathrm{T}}A$ 特征值为

$$\lambda_1 = 324, \quad \lambda_2 = 36, \quad \lambda_3 = 0$$

$\lambda_i, i \in \{1,2,3\}$ 对应 $A^{\mathrm{T}}A$ 的特征空间 $E_{\lambda_i} = \mathcal{N}\left(\lambda_i I_3 - A^{\mathrm{T}}A\right)$ 分别为 $E_{\lambda_1} = \operatorname{span}\left((-2,2,1)^{\mathrm{T}}\right), E_{\lambda_2} = \operatorname{span}\left((2,1,2)^{\mathrm{T}}\right), E_{\lambda_3} = \operatorname{span}\left((1,2,-2)^{\mathrm{T}}\right)$. 取归一化特征向量

$$v_1 = \frac{1}{3}(-2,2,1)^{\mathrm{T}}, \quad v_2 = \frac{1}{3}(2,1,2)^{\mathrm{T}}, \quad v_3 = \frac{1}{3}(1,2,-2)^{\mathrm{T}}$$

令 $V = (v_1 \,|v_2|\, v_3)$

$$\sigma_1 = \sqrt{\lambda_1} = 18, \quad \sigma_2 = \sqrt{\lambda_2} = 6, \quad \sigma_3 = \sqrt{\lambda_3} = 0$$

并令 $\Sigma = \operatorname{diag}_{4\times 3}(\sigma_1, \sigma_2, \sigma_3)$. 然后找到正交矩阵 $U = (u_1 \,|u_2|\, u_3 \mid u_4) \in \mathbb{R}^{4\times 4}$ 使得 $Av_i = \sigma_i u_i$ *for all* $i \in \{1,2,3\}$. 其中

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{2}(1,1,1,1)^{\mathrm{T}}, \quad u_2 = \frac{Av_2}{\sigma_2} = \frac{1}{2}(-1,1,-1,1)^{\mathrm{T}}.$$

计算 $\mathcal{N}\left((u_1 \mid u_2)^{\mathrm{T}}\right) = \operatorname{span}\left((1,0,-1,0)^{\mathrm{T}}, (0,1,0,-1)^{\mathrm{T}}\right)$ 并得到

$$u_3 = \frac{1}{2}(1,1,-1,-1)^{\mathrm{T}}$$

(我们希望 $2U$ 是一个 Hadamard 矩阵, 需要在 $\mathcal{N}\left((u_1 \mid u_2)^{\mathrm{T}}\right)$ 找个单位向量，其元素只含 $\pm\frac{1}{2}$.)

最后，计算 $\mathcal{N}\left((u_1 \,|u_2|\, u_3)^{\mathrm{T}}\right) = \operatorname{span}\left((-1,1,1,-1)^{\mathrm{T}}\right)$ 并得到

$$u_4 = \frac{1}{2}(-1,1,1,-1)^{\mathrm{T}}.$$

(我们希望 $2U$ 是一个 Hadamard 矩阵, 需要在 $\mathcal{N}\left((u_1 \,|u_2|\, u_3)^{\mathrm{T}}\right)$ 找个单位向量, 其元素只含 $\pm\frac{1}{2}$.)

因此 A 的 *SVD* 分解为:

$$A = \begin{pmatrix} -8 & 5 & 1 \\ -4 & 7 & 5 \\ -8 & 5 & 1 \\ -4 & 7 & 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 18 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix}^{\mathrm{T}} = U\Sigma V^{\mathrm{T}}.$$

(ii) A 有两个非零奇异值故 $\operatorname{rank}(A)=2$.

$$\mathcal{R}(A)=\operatorname{span}(u_1,\dots,u_r)=\operatorname{span}(u_1,u_2)=\operatorname{span}\left(\frac{1}{2}(1,1,1,1)^{\mathrm{T}},\frac{1}{2}(-1,1,-1,1)^{\mathrm{T}}\right),$$

$$\mathcal{N}(A)=\operatorname{span}(v_{r+1},\dots,v_3)=\operatorname{span}(v_3)=\operatorname{span}\left(\frac{1}{3}(1,2,-2)^{\mathrm{T}}\right)$$

$\|A\|_2=\sigma_1=18$, $\|A\|_F=\sqrt{\sum_{i=1}^r\sigma_i^2}=\sqrt{\sigma_1^2+\sigma_2^2}=6\sqrt{10}$.

(iii) 例用 *(i)* 中 A 的 *SVD* 分解，记 $A^{\mathrm{T}}=\tilde U\tilde\Sigma\tilde V^{\mathrm{T}}$ ，其中 $\tilde U=V,\tilde\Sigma=\Sigma^{\mathrm{T}},\tilde V=U$:

$$A^{\mathrm{T}}=\begin{pmatrix}-\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\\\frac{2}{3}&\frac{1}{3}&\frac{2}{3}\\\frac{1}{3}&\frac{2}{3}&-\frac{2}{3}\end{pmatrix}\begin{pmatrix}18&0&0&0\\0&6&0&0\\0&0&0&0\end{pmatrix}\begin{pmatrix}\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}\end{pmatrix}^{\mathrm{T}}=\tilde U\tilde\Sigma\tilde V^{\mathrm{T}}.$$

记 $\tilde U=(\tilde u_1|\tilde u_2|\tilde u_3)$, $\tilde V=(\tilde v_1|\tilde v_2|\tilde v_3\mid\tilde v_4)$, $\tilde\Sigma=\operatorname{diag}_{3\times4}(\tilde\sigma_1,\tilde\sigma_2,\tilde\sigma_3)$, 其中 $\tilde u_1,\tilde u_2,\tilde u_3\in\mathbb{R}^3$,$\tilde v_1,\tilde v_2,\tilde v_3,\tilde v_4\in\mathbb{R}^4$,$\tilde\sigma_1\ge\tilde\sigma_2\ge\tilde\sigma_3\ge0$.

$k=0$: 注意到 $\left\{M\in\mathbb{R}^{3\times4}:\operatorname{rk}(M)\le0\right\}=\{0_{3\times4}\}$. 记 $A_0:=0_{3\times4}$ 并有 $\gamma_0=\left\|A^{\mathrm{T}}\right\|_2=\tilde\sigma_1=18$

$k=1$: 利用 *Eckhart-Young-Mirsky* 定理,

$$A_1=\tilde\sigma_1\tilde u_1\tilde v_1^{\mathrm{T}}=18\begin{pmatrix}-\frac{2}{3}\\\frac{2}{3}\\\frac{1}{3}\end{pmatrix}\begin{pmatrix}\frac{1}{2}&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\end{pmatrix}=\begin{pmatrix}-6&-6&-6&-6\\6&6&6&6\\3&3&3&3\end{pmatrix}$$

并有 $\gamma_1=\left\|A^{\mathrm{T}}-A_1\right\|_2=\tilde\sigma_2=6$.

$k\ge2$: 因为 $\operatorname{rank}\left(A^{\mathrm{T}}\right)=2$, 记 $A_k:=A^{\mathrm{T}}$, 对任意 $k\in\mathbb{N}_{\ge2}$ 都有 $\gamma_k=0$.

习题 3.

(i) 假设 A 可逆，根据 A 的 *SVD* 结果给出 A^{-1} 的 *SVD* 分解 (提示:$Av_i=\sigma_iu_i\ \forall i\in\{1,\dots,n\}$)

(ii) 假设 Q 是正交阵. 给出 Q 的 *SVD* 分解及其奇异值

(iii) 假设 $A=QBQ^{\mathrm{T}}$ ，其中 $Q\in\mathbb{R}^{n\times n}$, 说明 A 和 B 有相同奇异值

解. *(i)* 记 A 的 *SVD* 分解为

$$A=U\Sigma V^{\mathrm{T}}=(u_1|\cdots|u_n)\left[\operatorname{diag}_{n\times n}(\sigma_1,\dots,\sigma_n)\right](v_1|\cdots|v_n)^{\mathrm{T}}$$

其中 $U,V\in\mathbb{R}^{n\times n}$ 正交, $\sigma_1\ge\sigma_2\ge\cdots\ge\sigma_n\ge0$. 因为 A 可逆，有 $\operatorname{rank}(A)=n$, $\sigma_i>0$, $\forall i\in\{1,\dots,n\}$. 又由 $A=U\Sigma V^{\mathrm{T}}$ 有 $Av_i=\sigma_iu_i\ \forall i\in\{1,\dots,n\}$ ，因此 $A^{-1}u_i=A^{-1}\left(\frac{1}{\sigma_i}Av_i\right)=\frac{1}{\sigma_i}v_i$(其中 $\frac{1}{\sigma_n}\ge\cdots\ge\frac{1}{\sigma_2}\ge\frac{1}{\sigma_1}>0$) 故

$$A^{-1}=(v_n|\cdots|v_2\mid v_1)\left[\operatorname{diag}_{n\times n}\left(\frac{1}{\sigma_n},\dots,\frac{1}{\sigma_2},\frac{1}{\sigma_1}\right)\right](u_n|\cdots|u_2\mid u_1)^{\mathrm{T}}=(VP)\left(P\Sigma^{-1}P\right)(UP)^{\mathrm{T}},$$

记 $P := (e_n|\cdots|e_2 \mid e_1) \in \mathbb{R}^{n\times n}$. *(由于* $PP^{\mathrm{T}} = P^2 = I_n$，$P$ 是正交阵，且 VP, UP 是正交阵.*)*

(ii) 由于 $Q \in \mathbb{R}^{n\times n}$ 正交，$Q = QI_nI_n^T$ 即 Q 的 *SVD* 分解. 显然，所有奇异值为 *1*.

(iii) 记 B 的 *SVD* 分解为 $B = U\Sigma V^{\mathrm{T}}$，则 $A = QBQ^{\mathrm{T}} = QU\Sigma V^{\mathrm{T}}Q^{\mathrm{T}} = (QU)\Sigma(QV)^{\mathrm{T}}$. 显然 A 和 B 有相同奇异值.