数据科学与工程数学基础

作业提交规范及第 2 次作业

教师：黄定江
助教：刘文辉

2022 年 11 月 18 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：“**hw2_ 学号 _ 姓名**”。其中，hw2 表示第 2 次作业。命名示例：hw2_50000000000_ 刘某某。
3. 作业提交途径：点击打开网址：**第 2 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 2 次作业

提交截至时间：**2021/11/29 12:00（中午）**

理论部分

习题 1. 判定矩阵 $\boldsymbol{C}=\begin{bmatrix}3 & 2 & -1\\ -1 & 0 & 0\\ -1 & 3 & 0\end{bmatrix}$ 和 $\boldsymbol{B}=\begin{bmatrix}0 & 2 & -1\\ -1 & 4 & -1\\ 1 & 3 & -5\end{bmatrix}$ 能否进行 $\boldsymbol{LU}$ 分解.

(1) 分析不能进行 $\boldsymbol{LU}$ 分解的原因。对于这种情况，应该做怎样的处理才能使用 LU 分解。

(2) 对于上述能分解的矩阵，试分解之。

解.

定理 0.0.1.

矩阵$A \in \mathbb{R}^{n\times n}$ 能够进行LU 分解的充分必要条件是 A 的前 n 阶顺序主子式不为 *0*。

因为矩阵 $\boldsymbol{B}$ 的一阶顺序主子式为 0，所以不能够进行 $\boldsymbol{LU}$ 分解，可以通过置换矩阵调换行的顺序进行处理。这里对矩阵 $\boldsymbol{C}$ 进行分解（具体计算过程可参考教材或课件中的例题。）:

$$\boldsymbol{C} = \begin{pmatrix} 3 & 2 & -1 \\ -1 & 0 & 0 \\ -1 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{3} & 1 & 0 \\ -\frac{1}{3} & \frac{11}{2} & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & -1 \\ 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{3}{2} \end{pmatrix} = \boldsymbol{LU}$$

习题 2. 求矩阵 $\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ 的 $\boldsymbol{LU}$ 分解。

解. （具体计算过程可参考教材或课件中的例题。）

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & 0 & -\frac{2}{3} \end{pmatrix} = \boldsymbol{LU}$$

习题 3. 求对称正定矩阵

$$\boldsymbol{A} = \begin{bmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{bmatrix}$$

的不带平方根的 *Cholesky* 分解。

解. （具体计算过程可参考教材或课件中的例题。）

$$\boldsymbol{A} = \begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{2}{5} & 1 & 0 \\ -\frac{4}{5} & -2 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{2}{5} & -\frac{4}{5} \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{LDL}^T$$

习题 4. 利用 *QR* 分解求解下述线性方程组的解（最终结果可只需写出具体矩阵与向量的乘积形式即可）:

$$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

解.

$$A=\begin{bmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}}\\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & 0\\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{6} & \sqrt{6} & \frac{7}{\sqrt{6}}\\ 0 & \sqrt{3} & \frac{1}{\sqrt{3}}\\ 0 & 0 & \sqrt{2}\end{bmatrix}=QR$$

因此，

$$\boldsymbol{x}=\begin{bmatrix}\sqrt{6} & \sqrt{6} & \frac{7}{\sqrt{6}}\\ 0 & \sqrt{3} & \frac{1}{\sqrt{3}}\\ 0 & 0 & \sqrt{2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}}\\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & 0\\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}}\end{bmatrix}^{\top}\begin{bmatrix}1\\2\\3\end{bmatrix}$$

习题 5. 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 $L_3 := \text{diag}_{3\times 3}(1,1,\sqrt{5})$ 使得 $L_3L_2L_1ML_1^{\text{T}}L_2^{\text{T}}L_3^{\text{T}} = I_3$. 我们有 $M = A^{\text{T}}A = GG^{\text{T}}$ 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 $G = U\Sigma V^{\text{T}}$，其中 $U, V \in \mathbb{R}^{n\times n}$ 正交，$\Sigma = \text{diag}_{n\times n}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{n\times n}$ 且 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. 故 $A^{\text{T}}A = GG^{\text{T}} = U\Sigma V^{\text{T}}V\Sigma U^{\text{T}} = U\Sigma^2U^{\text{T}}$，$A^{\text{T}}A$ 的奇异值为 $\sigma_1^2, \ldots, \sigma_n^2$. 因此 $\|A^{\text{T}}A\|_2 = \sigma_1^2 = \|G\|_2^2$. 同样的，我们可以得出 $\|A^{\text{T}}A\|_2 = \|A\|_2^2$.

习题 6. 对 $k \in \mathbb{N}_0$, 定义

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

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

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

(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 $\text{rank}(A_k) \leq k$ 且 $\left\|A^{\text{T}} - A_k\right\|_2 = \gamma_k$

解. *(i)*

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

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

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

$\lambda_i, i \in \{1,2,3\}$ 对应 $A^{\text{T}}A$ 的特征空间 $E_{\lambda_i} = \mathcal{N}\left(\lambda_iI_3 - A^{\text{T}}A\right)$ 分别为 $E_{\lambda_1} = \text{span}\left((-2,2,1)^{\text{T}}\right), E_{\lambda_2} = \text{span}\left((2,1,2)^{\text{T}}\right), E_{\lambda_3} = \text{span}\left((1,2,-2)^{\text{T}}\right)$. 取归一化特征向量

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

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 $\Sigma = \text{diag}_{4\times 3}(\sigma_1, \sigma_2, \sigma_3)$. 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 $Av_i = \sigma_iu_i$ *for all* $i \in \{1,2,3\}$. 其中

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

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

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

(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 $\operatorname{rank}(A) = 2$.

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

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

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

(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 $\tilde{U} = (\tilde{u}_1 \,|\tilde{u}_2|\, \tilde{u}_3)$, $\tilde{V} = (\tilde{v}_1 \,|\tilde{v}_2|\, \tilde{v}_3 \mid \tilde{v}_4)$, $\tilde{\Sigma} = \operatorname{diag}_{3\times 4}(\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3)$, 其中 $\tilde{u}_1, \tilde{u}_2, \tilde{u}_3 \in \mathbb{R}^3$,$\tilde{v}_1, \tilde{v}_2, \tilde{v}_3, \tilde{v}_4 \in \mathbb{R}^4$,$\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \tilde{\sigma}_3 \geq 0$.
$k = 0$: 注意到 $\left\{M \in \mathbb{R}^{3\times 4} : \operatorname{rk}(M) \leq 0\right\} = \{0_{3\times 4}\}$. 记 $A_0 := 0_{3\times 4}$ 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 $\gamma_1 = \left\|A^{\mathrm{T}} - A_1\right\|_2 = \tilde{\sigma}_2 = 6$.
$k \geq 2$: 因为 $\operatorname{rank}\left(A^{\mathrm{T}}\right) = 2$, 记 $A_k := A^{\mathrm{T}}$, 对任意 $k \in \mathbb{N}_{\geq 2}$ 都有 $\gamma_k = 0$.

习题 7.

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

(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 是正交阵, 说明 A 和 B 有相同奇异值

解. *(i)* 记 A 的 *SVD* 分解为

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

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

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

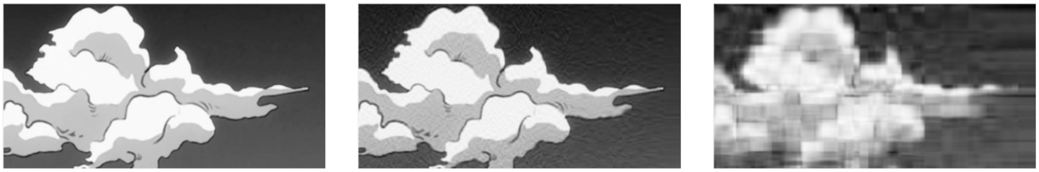
记 $P := (e_n|\cdots|e_2 \mid e_1) \in \mathbb{R}^{n\times n}$. *(*由于 $PP^{\mathrm{T}} = P^2 = I_n$，$P$ 是正交阵，且 VP, UP 是正交阵.*)*

(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.

实操部分

习题 8. 将一张图片（图片自选）利用奇异值分解完成图像的压缩。示例图如下：



(a) 原图　　(b) 提取 50 个奇异值的图像　　(c) 提取 10 个奇异值的图像

图 1: 使用 SVD 进行图像压缩