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、徐艺玮

2023 年 10 月 24 日

作业提交规范

1. 作业提交形式：建议使用 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。若使用手写拍照的方式，请务必规整好各题解答的图片，并整合在一个 PDF 文档中，只发图片格式的作业概不批改！
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：52200000000_刘某某。
3. 作业提交途径：点击打开每次作业的传送门地址：**第 3 次作业提交传送门**，无需注册和登录，按要求输入个人学号和姓名，然后上传作业文档即可。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；若发现存在抄袭的作业时，相似的两份作业均会被记为 0 分。

第 3 次作业

提交截至时间：**2023/10/24 下周二 12:00（中午）**

理论部分 (范数与二次型)

习题 1. 完成以下阅读材料即可。

- **为什么需要 *LU* 分解？**

  为什么不直接求解方程组 $Ax = b$？如果了解 LU 分解所需要的计算量与高斯消去法解方程组所需的计算量，便可知计算复杂度均为 $O(n^3)$。似乎 LU 分解并没有任何优势。然而，在实际的工程问题中，我们经常会遇到的是这样的问题，需要求解这样一系列的方程组（例如时序的）:

  $$Ax = b_1,\ Ax = b_2,\ \cdots,\ Ax = b_r.$$

  其中的系数矩阵是不变的，只是右边的常数项在改变。这时，如果直接使用高斯消去法求解各个方程组的话，则对应的计算复杂度为 $O(rn^3)$。若使用 LU 分解，则只需要的运算量级为 $O(n^3 + rn^2)$，这是因为 LU 分解在解方程组时只需要 $O(n^2)$。

  或许，你会提出直接计算出 A^{-1}，然后再利用它去求解各个方程组。这时所需要的计算量的效果似乎与 LU 分解是相似的。事实上，在实际计算中往往不会直接计算矩阵的逆。这里给出其中一个原因：通常实际工程中矩阵 A 的维度很大，但有一个优点是它是稀疏的，例如一个“带状的”矩阵。使用 LU 分解，可以保持矩阵的稀疏性 (因此，还可以提高运算速度)。然而，直接求解矩阵的逆，则会丢失矩阵的稀疏性。因此，在数据的存储上直接求逆存在明显的劣势。

- ***Gauss* 消去与 *LU* 分解**

  还记得初中学过的求解二元一次方程组的消元法吗？这就是 LU 分解与 *Cholesky* 分解的全部。假设我们有如下二元一次方程组

  $$\begin{cases} 2x_1 + 4x_2 = 5 \\ 4x_1 + 5x_2 = 3 \end{cases}$$

  现在我们考虑方程组的求解。在中学时，我们会通过将第一个等式乘以 -2 加到第二个等式，则得到

  $$\begin{cases} 2x_1 + 4x_2 = 5 \\ 0x_1 - 3x_2 = -7 \end{cases}$$

  这个行变换对应的矩阵为 $\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$，它的逆矩阵就是 $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$。另外，变换后的系数矩阵为 $\begin{pmatrix} 2 & 4 \\ 0 & -3 \end{pmatrix}$。

如果我们使用 LU 分解，则可得到

$$\begin{pmatrix} 2 & 4 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 0 & -3 \end{pmatrix}$$

从上可以看出，高斯消去的变换矩阵的逆恰好对应 LU 分解的 L 矩阵，变换后的系数矩阵恰好对应 U 矩阵。

习题 2. 对矩阵 $\begin{pmatrix} 5 & 3 & -1 & 3 \\ 0 & 1 & 1 & -2 \\ -5 & -3 & 4 & -4 \\ 0 & 1 & 1 & 0 \end{pmatrix}$ 进行 LU 分解。

解. （具体计算过程可参考教材或课件中的例题。）

$$\boldsymbol{A} = \begin{pmatrix} 5 & 3 & -1 & 3 \\ 0 & 1 & 1 & -2 \\ -5 & -3 & 4 & -4 \\ 0 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 & 3 & -1 & 3 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \boldsymbol{LU}$$

2

习题 3. 设 A 对称且 $a_{11} \neq 0$, 并假经过一步 *Gauss* 消去之后, A 具有如下形式

$$\begin{bmatrix} a_{11} & a_1^{\mathrm{T}} \\ \mathbf{0} & A_2 \end{bmatrix}$$

证明 A_2 仍是对称阵.

解. 记矩阵 A 和高斯变换矩阵分别为：

$$A = \begin{bmatrix} a_{11} & a_1^{\mathrm{T}} \\ a_1 & A_1 \end{bmatrix}, G = \begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ l_1 & I \end{bmatrix}$$

0

则

$$GA = \begin{bmatrix} a_{11} & a_1^{\mathrm{T}} \\ -a_{11}l_1 + a_1 & -l_1 a_1^{\mathrm{T}} + A_1 \end{bmatrix}$$

易知 $-a_{11}l_1 + a_1 = \mathbf{0}$ 以及 $A_2 = -l_1 a_1^{\mathrm{T}} + A_1$。由第一个等式可得 $l_1 = 1/a_{11}a_1$，代入第二个等式知

$$A_2 = -\frac{1}{a_{11}} a_1 a_1^{\mathrm{T}} + A_1.$$

故可知 A_2 仍是对称矩阵。

习题 4. 证明上三角矩阵与上三角矩阵的乘积仍是上三角矩阵。

解. 假设 A, B 为上三角矩阵，且其乘积为 C。下证 C 为上三角矩阵，只需证 $C_{ij}, (i > j)$ 时为 0。易知，

$$C_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{j} A_{ik}B_{kj} + \sum_{k=j+1}^{n} A_{ik}B_{kj}$$

因为 A, B 为上三角矩阵，所以右边第一项为中 $A_{ik} = 0$，右边第二项中 $B_{kj} = 0$。故得证。

2

习题 5. 用 *Householder* 方法求矩阵 $A=\begin{bmatrix}1&1\\2&0\\2&1\end{bmatrix}$ 的 *QR* 分解。

解. 令 $\alpha_1=(1,2,2)^T$，$a_1=\|\alpha_1\|_2=3$，则

$$w_1=\frac{\alpha_1-a_1\boldsymbol{e}_1}{\|\alpha_1-a_1\boldsymbol{e}_1\|_2}=\frac{1}{2\sqrt{3}}(-2,2,2)^T$$

故有

$$H_1=I-2w_1w_1^T=\begin{bmatrix}1/3&2/3&2/3\\2/3&1/3&-2/3\\2/3&-2/3&1/3\end{bmatrix}$$

3

此时，

$$H_1A=\begin{bmatrix}3&1\\0&0\\0&1\end{bmatrix}$$

再令 $\beta_1=(0,1)^T$，$b_1=\|\beta_1\|_2=1$，则

$$w_2=\frac{\beta_1-b_1\boldsymbol{e}_1}{\|\beta_1-b_1\boldsymbol{e}_1\|_2}=\frac{1}{\sqrt{2}}(-1,1)^T$$

故有

$$\hat{H}_2=I-2w_2w_2^T=\begin{bmatrix}0&1\\1&0\end{bmatrix}$$

记

$$H_2=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$$

此时

$$H_2H_1A=\begin{bmatrix}3&1\\0&1\\0&0\end{bmatrix}\triangleq R$$

因此，$A=QR$，其中 $Q=H_1^TH_2^T$。

习题 6. 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 $\|A^{\mathrm{T}}A\|_2=\|A\|_2^2=\|G\|_2^2$

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 $\|A^{\mathrm{T}}A\|_2=\|A\|_2^2$.