第四章 矩阵分解

第 10 讲 LU 分解

黄定江

DaSE @ ECNU

djhuang@dase.ecnu.edu.cn

10.1.1 引入

- 线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的求解是矩阵计算的核心问题。
- 由于三角形方程组简单易于求解，一个基本的思路就是把一般的线性方程组的求解转化为三角方程组的求解。
- 可以通过矩阵的三角分解，也即 LU 分解来实现！

定义 1

设矩阵 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$，对矩阵 $\boldsymbol{A}$ 进行 LU 分解是指求上三角矩阵 $\boldsymbol{U}$ 和单位下三角矩阵 $\boldsymbol{L}$ 使得

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$$

写成分量的形式即

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ & u_{22} & u_{23} & \cdots & u_{2n} \\ & & u_{33} & \cdots & u_{3n} \\ & & & \ddots & \vdots \\ & & & & u_{nn} \end{pmatrix}$$

- 假设矩阵 $\boldsymbol{A}$ 可以写成若干个秩 1 矩阵和的形式：

$$\boldsymbol{A}=\boldsymbol{l}_1\boldsymbol{u}_1^{\mathrm{T}}+\boldsymbol{l}_2\boldsymbol{u}_2^{\mathrm{T}}+\cdots+\boldsymbol{l}_r\boldsymbol{u}_r^{\mathrm{T}}=\sum_{i=1}^{r}\boldsymbol{l}_i\boldsymbol{u}_i^{\mathrm{T}}$$

  其中 r 为矩阵 $\boldsymbol{A}$ 的秩。

- 若 $\boldsymbol{A}$ 是满秩，也即 $r=n$，则令

$$L=(\boldsymbol{l}_1,\boldsymbol{l}_2,\ldots,\boldsymbol{l}_n),U=(\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_n)^T$$

  那么就有

$$\boldsymbol{A}=\sum_{i=1}^{r}\boldsymbol{l}_i\boldsymbol{u}_i^{\mathrm{T}}=(\boldsymbol{l}_1,\boldsymbol{l}_2,\ldots,\boldsymbol{l}_n)\begin{pmatrix}\boldsymbol{u}_1^{\mathrm{T}}\\ \boldsymbol{u}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{u}_n^{\mathrm{T}}\end{pmatrix}=LU$$

- 如果我们进一步假设 $\boldsymbol{l}_i$ 和 $\boldsymbol{u}_i$ 的前 $i-1$ 个元素均为 0，并且 $\boldsymbol{l}_i$ 的第 i 个元素为 1，那么我们实际上就得到了 $\boldsymbol{A}$ 的 LU 分解。

我们现在来考虑

$$\boldsymbol{l}_1 \boldsymbol{u}_1^T = \boldsymbol{A} - (0, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_n) \begin{pmatrix} 0 \\ \boldsymbol{u}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{u}_n^{\mathrm{T}} \end{pmatrix}.$$

- 我们知道 $\boldsymbol{l}_1$ 的第一个元素为 1，所以 $\boldsymbol{u}_1$ 就是 $\boldsymbol{l}_1 \boldsymbol{u}_1^T$ 的第一行的行向量。
- 另外一方面

$$(0, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_n) \begin{pmatrix} 0 \\ \boldsymbol{u}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{u}_n^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & * \end{pmatrix} \begin{pmatrix} 0 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & * \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & * \end{pmatrix}$$

的第一行和第一列均为 0。

$$
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
a_{21} & * & * & \cdots & * \\
a_{31} & * & * & \cdots & * \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_{n1} & * & \cdots & \cdots & *
\end{pmatrix}
=
\begin{pmatrix}
1 \\ l_{21} \\ l_{31} \\ \vdots \\ l_{n1}
\end{pmatrix}
\begin{pmatrix}
u_{11} & u_{12} & u_{13} & \cdots & u_{1n}
\end{pmatrix}
+
\begin{pmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & * & * & \cdots & * \\
0 & * & * & \cdots & * \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & * & \cdots & \cdots & *
\end{pmatrix}
$$

$$
=
\begin{pmatrix}
u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\
l_{21} * u_{11} & * & * & \cdots & * \\
l_{31} * u_{11} & * & * & \cdots & * \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
l_{n1} * u_{11} & * & \cdots & \cdots & *
\end{pmatrix}
$$

所以 $\boldsymbol{u}_1$ 就是 $\boldsymbol{A}$ 的第一行。而 $\boldsymbol{l}_1$ 则是 $\boldsymbol{A}$ 的第一列除以 u_{11} 也就是 a_{11} 得到的。

- 我们记 $\tilde{\boldsymbol{A}}^{(0)} = \boldsymbol{A}$，当 $i \geq 1$ 时，记
$$\tilde{\boldsymbol{A}}^{(i)} = \boldsymbol{A} - \sum_{j=1}^{i} \boldsymbol{l}_j \boldsymbol{u}_j^{\mathrm{T}}$$
- 根据上面对于 $\boldsymbol{u}_1$ 和 $\boldsymbol{l}_1$ 的推导，我们很容易将其应用到 $\boldsymbol{u}_i$ 和 $\boldsymbol{l}_i$ 上。
- 也就是说，$\boldsymbol{u}_i$ 就是 $\tilde{\boldsymbol{A}}^{(i-1)}$ 的第 i 行。
- 而 $\boldsymbol{l}_i$ 则是 $\tilde{\boldsymbol{A}}^{(i-1)}$ 的第 i 列除以 $\tilde{a}_{ii}^{(i-1)}$ 得到的。

例 1

求矩阵 $\boldsymbol{A}$

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 10\end{pmatrix}$$

的 LU 分解。

解

记 $\tilde{\boldsymbol{A}}^{(0)} = \boldsymbol{A}$，令 $\boldsymbol{u}_1$ 是 $\boldsymbol{A}$ 的第 *1* 行，$\boldsymbol{l}_1$ 是 $\boldsymbol{A}$ 的第 *1* 列除以 u_{11}。则

$$\boldsymbol{l}_1 \boldsymbol{u}_1^{\mathrm{T}} = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 8 & 2 \\ 7 & 14 & 21 \end{pmatrix}$$

$$\tilde{\boldsymbol{A}}^{(1)} = \boldsymbol{A} - \boldsymbol{l}_1 \boldsymbol{u}_1^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 4 & 8 & 2 \\ 7 & 14 & 21 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix}$$

容易看出 $\tilde{\boldsymbol{A}}^{(1)}$ 的第 2 行是 $\boldsymbol{A}$ 的第 2 行减去其第一行的 4 倍，$\tilde{\boldsymbol{A}}^{(1)}$ 的第 3 行是 $\boldsymbol{A}$ 的第 3 行减去其第一行的 7 倍。

$$\tilde{\boldsymbol{A}}^{(1)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix}$$

令 $\boldsymbol{u}_2$ 是 $\tilde{\boldsymbol{A}}^{(1)}$ 的第 2 行，$\boldsymbol{l}_2$ 是 $\tilde{\boldsymbol{A}}^{(1)}$ 的第 2 列除以 u_{22}。则

$$\boldsymbol{l}_2\boldsymbol{u}_2^T = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} 0 & -3 & -6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix}$$

$$\tilde{\boldsymbol{A}}^{(2)} = \boldsymbol{A} - \boldsymbol{l}_1\boldsymbol{u}_1^T - \boldsymbol{l}_2\boldsymbol{u}_2^T = \tilde{\boldsymbol{A}}^{(1)} - \boldsymbol{l}_2\boldsymbol{u}_2^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

容易看出 $\tilde{\boldsymbol{A}}^{(2)}$ 的第 3 行就是 $\tilde{\boldsymbol{A}}^{(1)}$ 的第 3 行减去其第二行的 2 倍。

$$\tilde{\boldsymbol{A}}^{(2)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

令 $\boldsymbol{u}_3$ 是 $\tilde{\boldsymbol{A}}^{(2)}$ 的第 3 行，$\boldsymbol{l}_3$ 是 $\tilde{\boldsymbol{A}}^{(2)}$ 的第 3 列除以 u_{33}。即

$$\boldsymbol{l}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$

所以

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}$$

- 可以看出从 $\boldsymbol{A}$ 得到 $\boldsymbol{U}$ 的过程等价于对 $\boldsymbol{A}$ 进行初等行变换。具体地说，$\boldsymbol{u}_k$ 是通过将矩阵 $\boldsymbol{A}$ 的第 k 行分别减去 $\boldsymbol{A}$ 的前 $k-1$ 行的若干倍得到的。因此，我们可以利用初等行变换将矩阵进行 LU 分解：
- 步骤 1 利用初等行变换（某一行加其它行的倍数）化矩阵 $\boldsymbol{A}$ 为阶梯型矩阵 $\boldsymbol{U}$，即

$$\boldsymbol{A} = \boldsymbol{A}^{(0)} \xrightarrow{\boldsymbol{L}_1} [\ \ ] \xrightarrow{\boldsymbol{L}_2} \cdots \xrightarrow{\boldsymbol{L}_{k-1}} [\ \ ] \cdots \xrightarrow{\boldsymbol{L}_{n-1}} [\ \ ] = \boldsymbol{U}$$

那么 $\boldsymbol{L}_{n-1}\cdots\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{A} = \boldsymbol{U}$。这里 $\boldsymbol{A}$ 经过 $\boldsymbol{L}_1, \boldsymbol{L}_2, \cdots, \boldsymbol{L}_{k-1}$ 得到 $\boldsymbol{A}^{(k-1)}$，$\boldsymbol{L}_k$ 将 $\boldsymbol{A}^{(k-1)}$ 的第 k 行的 $-l_{ik}$ 倍 $(i = k+1, \cdots, n)$，分别加到第 i 行，使得第 i 行的第 k 列元素都为 0。为了计算这样的 l_{ik}, 需要计算 $\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$。我们把其中 $\boldsymbol{A}^{(k-1)}$ 的第 k 行，第 k 列的元素即 $a_{kk}^{(k-1)}$ 称为主元。
- 步骤 2 对单位阵执行与步骤 1 相应的初等行变换的逆变换，得单位下三角矩阵 $\boldsymbol{L}$,

$$\boldsymbol{I} \xrightarrow{\boldsymbol{L}_{n-1}^{-1}} [\ \ ] \xrightarrow{\boldsymbol{L}_{n-2}^{-1}} \cdots \xrightarrow{\boldsymbol{L}_{k-1}^{-1}} [\ \ ] \cdots \xrightarrow{\boldsymbol{L}_1^{-1}} [\ \ ] = \boldsymbol{L}$$

$\boldsymbol{L} = \boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\cdots\boldsymbol{L}_{n-1}^{-1}$。
- 输出 LU 分解由 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ 给出。

在上述步骤 1 中，$\boldsymbol{L}_k$ 的一般形式可以表示为：

$$\boldsymbol{L}_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{n,k} & & & 1 \end{bmatrix} = \boldsymbol{I} - \boldsymbol{l}_k \boldsymbol{e}_k^T,$$

其中

$$\boldsymbol{l}_k = (0, \cdots, 0, l_{k+1,k}, \cdots, l_{n,k})^T$$

我们把这种类型的初等下三角阵称作 **Gauss 变换**，而称向量 $\boldsymbol{l}_k$ 为 **Gauss 向量**。基于 Gauss 变换的 LU 分解计算方法也称为 **Gauss 消去法**。

Gauss 变换 $\boldsymbol{L}_k$ 具有许多良好的性质：

- 对于一个给定的向量 $x=(x_1,\ldots,x_n)^T\in\mathbb{R}^n$，我们有

$$\boldsymbol{L}_k x=(x_1,\ldots,x_k,x_{k+1}-x_k l_{k+1,k},\ldots,x_n-x_k l_{n,k})^{\mathrm{T}}$$

由此立即可知，只要取

$$l_{ik}=\frac{x_i}{x_k},\, i=k+1,\ldots,n$$

便有 $\boldsymbol{L}_k x=(x_1,\ldots,x_k,0,\ldots,0)^{\mathrm{T}}$，这里我们要求 $x_k\neq 0$。

- Gauss 变换 $\boldsymbol{L}_k$ 的逆易求解。因为 $\boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{l}_k=0$，所以

$$(\boldsymbol{I}-\boldsymbol{l}_k\boldsymbol{e}_k^{\mathrm{T}})(\boldsymbol{I}+\boldsymbol{l}_k\boldsymbol{e}_k^{\mathrm{T}})=\boldsymbol{I}-\boldsymbol{l}_k\boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{l}_k\boldsymbol{e}_k^{\mathrm{T}}=\boldsymbol{I}$$

即

$$\boldsymbol{L}_k^{-1}=\boldsymbol{I}+\boldsymbol{l}_k\boldsymbol{e}_k^T$$

- Gauss 变换作用于矩阵 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$ 就相当于对该矩阵进行秩 1 修正，也即

$$\boldsymbol{L}_k\boldsymbol{A}=(\boldsymbol{I}-\boldsymbol{l}_k\boldsymbol{e}_k^{\mathrm{T}})\boldsymbol{A}=\boldsymbol{A}-\boldsymbol{l}_k(\boldsymbol{e}_k^T\boldsymbol{A})$$

例 2

求矩阵 $\boldsymbol{A}$ 的 LU 分解。

$$\boldsymbol{A}=\begin{pmatrix}1&2&3\\4&5&6\\7&8&10\end{pmatrix}$$

解

$\boldsymbol{A}\to\boldsymbol{U}$

$$\underset{\boldsymbol{A}^{(0)}}{\begin{pmatrix}1&2&3\\4&5&6\\7&8&10\end{pmatrix}}\xrightarrow[R_3-(\frac{7}{1})R_1]{R_2-(\frac{4}{1})R_1}\underset{\boldsymbol{A}^{(1)}}{\begin{pmatrix}1&2&3\\0&-3&-6\\0&-6&-11\end{pmatrix}}\xrightarrow{R_3-(\frac{-6}{-3})R_2}\underset{\boldsymbol{A}^{(2)}}{\begin{pmatrix}1&2&3\\0&-3&-6\\0&0&1\end{pmatrix}}$$

初等行变换 $\xrightarrow[R_3-(\frac{7}{1})R_1]{R_2-(\frac{4}{1})R_1}$ 即对矩阵 $\boldsymbol{A}^{(0)}$ 左乘一个初等矩阵

$$\boldsymbol{L}_1=\begin{pmatrix}1&0&0\\-4&1&0\\-7&0&1\end{pmatrix}$$

初等行变换 $\xrightarrow{R_3-(\frac{-6}{-3})R_2}$ 即对 $\boldsymbol{A}^{(1)}$ 左乘一个初等矩阵

$$\boldsymbol{L}_2=\begin{pmatrix}1&0&0\\0&1&0\\0&-2&1\end{pmatrix}$$

即 $\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{A}=\boldsymbol{U}$。所以 $\boldsymbol{A}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\boldsymbol{U}$，显然

$$\boldsymbol{L}_1^{-1}=\begin{pmatrix}1&0&0\\4&1&0\\7&0&1\end{pmatrix},\boldsymbol{L}_2^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\0&2&1\end{pmatrix}$$

可得

$$\boldsymbol{L}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}=\begin{pmatrix}1&0&0\\4&1&0\\7&2&1\end{pmatrix}$$

Algorithm 1 LU分解

```
L = I, U = O
for k = 1 to n − 1 do
    for i = k + 1 to n do
        l_ik = a_ik / a_kk    %更新L的第k列
    end for
    for j = k to n do
        u_kj = a_kj    %更新U的第k行
    end for
    for i = k + 1 to n do
        for j = k + 1 to n do
            a_ij = a_ij − l_ik u_ik    %更新矩阵A(k + 1 : n, k + 1 : n)
        end for
    end for
end for
```

图 1: LU 分解

通过以上的算法，我们会得到唯一形式的 LU 分解。我们可以证明如下定理。

定理 1

[LU 分解的唯一性] 如果 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 非奇异，并且其 LU 分解存在，则 $\boldsymbol{A}$ 的 LU 分解是唯一的，且 $\det(\boldsymbol{A}) = u_{11}u_{22}\cdots u_{nn}$。

证明.
令 $\boldsymbol{A} = \boldsymbol{L}_1\boldsymbol{U}_1$ 和 $\boldsymbol{A} = \boldsymbol{L}_2\boldsymbol{U}_2$ 是非奇异矩阵 $\boldsymbol{A}$ 的两个 LU 分解，则 $\boldsymbol{L}_1\boldsymbol{U}_1 = \boldsymbol{L}_2\boldsymbol{U}_2$。
由于 $\boldsymbol{L}_2^{-1}\boldsymbol{L}_1$ 是下三角矩阵，并且 $\boldsymbol{U}_2\boldsymbol{U}_1^{-1}$ 是上三角矩阵，所以这两个矩阵必定都等于单位矩阵，否则它们们不可能相等。就是说，$\boldsymbol{L}_1 = \boldsymbol{L}_2, \boldsymbol{U}_1 = \boldsymbol{U}_2$，即 LU 分解是唯一的。
若 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$，则 $\det(\boldsymbol{A}) = \det(\boldsymbol{L}\boldsymbol{U}) = \det(\boldsymbol{L})\det(\boldsymbol{U}) = \det(\boldsymbol{U}) = u_{11}u_{22}\cdots u_{nn}$。 □

然而，利用上述算法进 LU 分解并不一定总是能够进行下去。

例 3

求矩阵 $\boldsymbol{A}$ 的 LU 分解。

$$\boldsymbol{A}=\begin{pmatrix}1&2&0\\1&2&1\\0&2&0\end{pmatrix}$$

解

$\boldsymbol{A}\rightarrow\boldsymbol{U}$

$$\begin{pmatrix}1&2&0\\1&2&1\\0&2&0\end{pmatrix}\xrightarrow{R_2-(\frac{1}{1})R_1}\begin{pmatrix}1&2&0\\0&0&1\\0&2&0\end{pmatrix}$$

由于第一次初等变换后得到矩阵 $\boldsymbol{L}_1\boldsymbol{A}=\begin{pmatrix}1&2&0\\0&0&1\\0&2&0\end{pmatrix}$ 的主元 $a_{22}^{(1)}=0$，无法进行下一步初等行变换，也就无法继续进行 LU 分解。

定理 2

矩阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 能够进行 *LU* 分解的充分必要条件是 $\boldsymbol{A}$ 的前 $n-1$ 个主元均不为 *0*。

例 4

设矩阵

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 0 \\ 4 & 7 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

将矩阵 $\boldsymbol{A}$ 进行 *LU* 分解可以得到

$$\begin{pmatrix} 2 & 3 & 0 \\ 4 & 7 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ 2 & 1 & \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & 0 \\ & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

观察 $\boldsymbol{U}$ 可以知道矩阵 $\boldsymbol{A}$ 的第 *3* 个主元为 *0*。

我们自然地会提出疑问：那么什么时候主元会为 0？并且当主元为 0 时，又该如何处理？

定理 3

设矩阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$，假设通过上述 *LU* 分解算法能得到 $\boldsymbol{A}^{(k-1)}$，则主元 $a_{kk}^{(k-1)}$ 不为零的充分必要条件是 $\boldsymbol{A}$ 的 k 阶顺序主子式 $|\boldsymbol{A}_k|$ 不为零。

证明.

在上述 LU 分解算法中，我们只将矩阵的第 i 行的若干倍加到第 k 行 (其中 $k > i$)，这个变换并不改变矩阵的各顺序主子式的值。

也就是说

$$|\boldsymbol{A}_k| = \prod_{i=1}^{k} a_{ii}^{(i-1)}$$

我们得到 $\boldsymbol{A}^{(k-1)}$，说明 $a_{ii}^{(i-1)} \neq 0, (i = 1, \cdots, k-1)$，因此 $a_{kk}^{(k-1)}$ 不为零，等价于 $\boldsymbol{A}$ 的 k 阶顺序主子式 $|\boldsymbol{A}_k|$ 不为零。

$\square$

例 5

在例3中，$\boldsymbol{A}$ 的 2 阶顺序主子式为 0，因此 $a_{22}^{(1)}=0$，那么当主元为 0 时如何继续分解矩阵？

$$\boldsymbol{A}=\begin{pmatrix}1&2&0\\1&2&1\\0&2&0\end{pmatrix}$$

解

对于出现主元为 0 的矩阵使用初等行变换中的行交换。

$\boldsymbol{A}\to\boldsymbol{U}$

$$\underbrace{\begin{pmatrix}1&2&0\\1&2&1\\0&2&0\end{pmatrix}}_{\boldsymbol{A}^{(0)}}\xrightarrow{R_2-(\frac{1}{1})R_1}\underbrace{\begin{pmatrix}1&2&0\\0&0&1\\0&2&0\end{pmatrix}}_{\boldsymbol{A}^{(1)}}\xrightarrow{R_2\longleftrightarrow R_3}\underbrace{\begin{pmatrix}1&2&0\\0&2&0\\0&0&1\end{pmatrix}}_{\boldsymbol{A}^{(2)}}$$

第一次初等变换 $\xrightarrow{R_2-(\frac{1}{1})R_1}$ 即对矩阵 $\boldsymbol{A}^{(0)}$ 左乘

$$\boldsymbol{L}_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

第二次初等变换 $\xrightarrow{R_2 \longleftrightarrow R_3}$ 即对矩阵 $\boldsymbol{A}^{(1)}$ 左乘

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

即 $\boldsymbol{P}\boldsymbol{L}_1\boldsymbol{A} = \boldsymbol{U}$。

所以 $\boldsymbol{A}=(\boldsymbol{P}\boldsymbol{L}_1)^{-1}\boldsymbol{U}=\boldsymbol{L}_1^{-1}\boldsymbol{P}^{-1}\boldsymbol{U}$。

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\xrightarrow{R_2\longleftrightarrow R_3}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\xrightarrow{R_2+(\frac{1}{1})R_1}\begin{pmatrix}1&0&0\\1&0&1\\0&1&0\end{pmatrix}$$

$$\boldsymbol{A}=(\boldsymbol{P}\boldsymbol{L}_1)^{-1}\boldsymbol{U}=\begin{pmatrix}1&0&0\\1&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&2&0\\0&2&0\\0&0&1\end{pmatrix}$$

虽然 $(\boldsymbol{P}\boldsymbol{L}_1)^{-1}$ 不是一个下三角矩阵，但是 $\boldsymbol{P}(\boldsymbol{P}\boldsymbol{L}_1)^{-1}$ 是下三角矩阵。并且

$$\boldsymbol{P}\boldsymbol{A}=(\boldsymbol{P}(\boldsymbol{P}\boldsymbol{L}_1)^{-1})\boldsymbol{U}$$

这说明我们只需要对 $\boldsymbol{A}$ 的行重新排列，就可以对重新排列后的矩阵进行 LU 分解。

- 为了避免在 LU 分解过程中主元为零，在每次对 $\boldsymbol{A}^{(i-1)}$ 做初等变换 $\boldsymbol{L}_i$ 前判断主元是否为零。
- 若为零，交换 $\boldsymbol{A}^{(i-1)}$ 第 i 行与 $a_{ji}^{(i-1)} \neq 0$ 的第 $j(j \geq i)$ 行，使 $a_{ji}^{(i-1)}$ 成为主元。
- 记这个行交换的初等变换矩阵为 $\boldsymbol{P}_i$(若不需要交换行则 $\boldsymbol{P}_i = I$)。然后再做初等变换 $\boldsymbol{L}_i$ 得到 $\boldsymbol{A}^i$，重复上面的过程，最终得到上三角矩阵 $\boldsymbol{U}$。即

$$\boldsymbol{L}_n\boldsymbol{P}_n\boldsymbol{L}_{n-1}\boldsymbol{P}_{n-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1\boldsymbol{A} = \boldsymbol{U}.$$

- 这样我们得到了 $\boldsymbol{A}$ 的分解：

$$\boldsymbol{A} = (\boldsymbol{L}_n\boldsymbol{P}_n\boldsymbol{L}_{n-1}\boldsymbol{P}_{n-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1)^{-1}\boldsymbol{U}.$$

- 虽然 $(\boldsymbol{L}_n\boldsymbol{P}_n\boldsymbol{L}_{n-1}\boldsymbol{P}_{n-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1)^{-1}$ 不是一个下三角阵，但是如果我们先对 $\boldsymbol{A}$ 的按如下方式先重新排列各行，有

$$\boldsymbol{P}_n\boldsymbol{P}_{n-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A} = \boldsymbol{P}_n\boldsymbol{P}_{n-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1(\boldsymbol{L}_n\boldsymbol{P}_n\boldsymbol{L}_{n-1}\boldsymbol{P}_{n-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1)^{-1}\boldsymbol{U}$$

- 可以证明

$$\boldsymbol{P}_n\boldsymbol{P}_{n-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1(\boldsymbol{L}_n\boldsymbol{P}_n\boldsymbol{L}_{n-1}\boldsymbol{P}_{n-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1)^{-1}$$

  是一个下三角矩阵。

Algorithm 2 列主元LU分解

```
L = I, U = O
p = [1 : n]   %记录行变换矩阵P
for k = 1 to n − 1 do
    if a_kk = 0 then
        for i = k + 1 to n do
            if a_ik ≠ 0 then
                for j = 1 to n do
                    tmp = a_kj, a_kj = a_ij, a_ij = tmp   %交换第k行与第i行
                end for
                p_k = i   %更新行变换矩阵P
            end if
        end for
    end if
    for i = k + 1 to n do
        l_ik = a_ik / a_kk   %更新L的第k列
    end for
    for j = k to n do
        u_kj = a_kj   %更新U的第k行
    end for
    for i = k + 1 to n do
        for j = k + 1 to n do
            a_ij = a_ij − a_ik a_kj   %更新矩阵A(k + 1 : n, k + 1 : n)
        end for
    end for
end for
```

本讲小结

LU 分解

- Gauss 变换法
- 选主元的三角分解

将用于线性方程组的直接求解！