目录

- 引言
- Guass消去法
- Guass主元素消去法
- Guass消去法的变形
- 向量和矩阵的范数
- 误差分析

Guass消去法的变形

- Gauss消去法有很多变形
  - 有的是Gauss消去法的改进、改写
  - 有的是用于某一类特殊性质矩阵的简化
- 直接三角分解法
  - 将Gauss消去法改写为紧凑形式，可以直接从矩阵$\boldsymbol{A}$的元素得到计算$\boldsymbol{L}$，$\boldsymbol{U}$元素的递推公式，而不需任何中间步骤，这就是所谓直接三角分解法
  - 一旦实现了矩阵$\boldsymbol{A}$的LU分解，那么求解式$\boldsymbol{Ax} = \boldsymbol{L}\underbrace{\boldsymbol{Ux}}_{\boldsymbol{y}} = \boldsymbol{b}$的问题就等价于以下两个三角方程组

$$
\begin{aligned}
&(1)\ \boldsymbol{Ly} = \boldsymbol{b}, \qquad \text{求}\boldsymbol{y}; \\
&(2)\ \boldsymbol{Ux} = \boldsymbol{y}, \qquad \text{求}\boldsymbol{x}.
\end{aligned}
$$

不选主元的三角分解法

- 设$\boldsymbol{A}$为非奇异矩阵，且有分解式$\boldsymbol{A}=\boldsymbol{LU}$，其中$\boldsymbol{L}$为单位下三角阵，$\boldsymbol{U}$为上三角阵，即

$$\boldsymbol{A}=\begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{nl} & \cdots & l_{n,n-1} & 1 \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix} \qquad (7.4.1)$$

- 下面说明$\boldsymbol{L},\boldsymbol{U}$的元素可以由$n$步直接计算定出，其中第$r$步定出$\boldsymbol{U}$的第$r$行和$\boldsymbol{L}$的第$r$列元素

不选主元的三角分解法（续）

$$\boldsymbol{A}=\begin{bmatrix}1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{nl} & \cdots & l_{n,n-1} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn}\end{bmatrix} \quad (7.4.1)$$

- 由式(7.4.1)，考虑$\boldsymbol{A}$的第1行元素

  $$a_{1i}=u_{1i} \quad (i=1,2,\cdots,n)$$

  于是得$\boldsymbol{U}$的第1行元素u_{1i} $(i=1,2,\cdots,n)$

- 此外，考虑$\boldsymbol{A}$的第1列元素

  $$a_{i1}=l_{i1}u_{11}, \qquad l_{i1}=a_{i1}/u_{11} \quad (i=2,\cdots,n)$$

  于是得$\boldsymbol{L}$的第1列元素l_{i1} $(i=2,\cdots,n)$

不选主元的三角分解法（续）

$$\boldsymbol{A}=\begin{bmatrix}1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{nl} & \cdots & l_{n,n-1} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn}\end{bmatrix} \quad (7.4.1)$$

- 设已定出$\boldsymbol{U}$的第1行到第$r-1$行元素与$\boldsymbol{L}$的第1列到第$r-1$列元素
- 由式(7.4.1)，考虑$\boldsymbol{A}$的第r行元素

$$a_{ri}=\sum_{k=1}^{n} l_{rk}u_{ki}=\sum_{k=1}^{r-1} l_{rk}u_{ki}+u_{ri} \qquad (当 k>r, l_{rk}=0)$$

✓ $\boldsymbol{L}$的第r行乘$\boldsymbol{U}$的第i列， $i=r,r+1,\cdots,n$

不选主元的三角分解法（续）

$$\boldsymbol{A}=\begin{bmatrix}1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{nl} & \cdots & l_{n,n-1} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn}\end{bmatrix} \qquad (7.4.1)$$

- 设已定出$\boldsymbol{U}$的第1行到第$r-1$行元素与$\boldsymbol{L}$的第1列到第$r-1$列元素
- 由式(7.4.1)，考虑$\boldsymbol{A}$的第r行元素

$$a_{ri}=\sum_{k=1}^{n} l_{rk}u_{ki}=\sum_{k=1}^{r-1} l_{rk}u_{ki}+u_{ri} \qquad (当k>r, l_{rk}=0)$$

于是得$\boldsymbol{U}$的第r行元素

$$u_{ri}=a_{ri}-\sum_{k=1}^{r-1} l_{rk}u_{ki} \qquad (i=r,r+1,\cdots,n)$$

不选主元的三角分解法（续）

$$\boldsymbol{A}=\begin{bmatrix}1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{nl} & \cdots & l_{n,n-1} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn}\end{bmatrix} \quad (7.4.1)$$

- 设已定出$\boldsymbol{U}$的第1行到第$r-1$行元素与$\boldsymbol{L}$的第1列到第$r-1$列元素
- 由式(7.4.1)，考虑$\boldsymbol{A}$的第r列元素

$$a_{ir}=\sum_{k=1}^{n} l_{ik}u_{kr}=\sum_{k=1}^{r-1} l_{ik}u_{kr}+l_{ir}u_{rr} \ (\text{当}k>r, u_{kr}=0)$$

  ✓ $\boldsymbol{L}$的第i行乘$\boldsymbol{U}$的第r列， $i=r+1,r+1,\cdots,n$

不选主元的三角分解法（续）

$$\boldsymbol{A}=\begin{bmatrix}1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{nl} & \cdots & l_{n,n-1} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn}\end{bmatrix} \quad (7.4.1)$$

- 设已定出$\boldsymbol{U}$的第1行到第$r-1$行元素与$\boldsymbol{L}$的第1列到第$r-1$列元素
- 由式(7.4.1)，考虑$\boldsymbol{A}$的第r列元素

$$a_{ir}=\sum_{k=1}^{n} l_{ik}u_{kr}=\sum_{k=1}^{r-1} l_{ik}u_{kr}+l_{ir}u_{rr}\ \ (当k>r, u_{kr}=0)$$

于是得$\boldsymbol{U}$的第r列元素

$$l_{ir}=\left(a_{ir}-\sum_{k=1}^{r-1} l_{ik}u_{kr}\right)/u_{rr} \quad (i=r+1,\cdots,n)$$

用直接三角分解法解$\boldsymbol{Ax=b}$

1．计算$\boldsymbol{U}$的第1行和$\boldsymbol{L}$的第1列元素

$$u_{1i}=a_{1i}\qquad (i=1,2,\cdots,n)$$

$$l_{i1}=a_{i1}u_{11}\quad (i=2,3,\cdots,n)$$

■ 按照$r=2,3,\cdots,n$，重复下面步骤

2．计算$\boldsymbol{U}$的第r行元素

$$u_{ri}=a_{ri}-\sum_{k=1}^{r-1}l_{rk}u_{ki}\qquad (i=r,r+1,\cdots,n)\qquad (7.4.2)$$

3．计算$\boldsymbol{L}$的第r列元素

$$l_{ir}=\left(a_{ir}-\sum_{k=1}^{r-1}l_{ik}u_{kr}\right)/u_{rr}\quad (i=r+1,\cdots,n)\qquad (7.4.3)$$

用直接三角分解法解$\boldsymbol{Ax} = \boldsymbol{b}$（续）

4. 求解$\boldsymbol{Ly} = \boldsymbol{b}$

$$\begin{cases} y_1 = b_1 \\ y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k \quad (i = 2,3,\cdots,n) \end{cases} \tag{7.4.4}$$

5. 求解$\boldsymbol{Ux} = \boldsymbol{y}$

$$\begin{cases} x_n = y_n / u_{nn} \\ x_i = \left(y_i - \sum_{k=i+1}^{n} u_{ik} x_k \right) / u_{ii} \quad (i = n-1, n-2, \cdots, 1) \end{cases} \tag{7.4.5}$$

例7.5

- 用直接三角分解法解

$$\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 3 & -5 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}14\\ 18\\ 20\end{bmatrix}$$

- 根据前面步骤2和步骤3，得到

$$\boldsymbol{A}=\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 3 & -5 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\ 0 & 1 & -4\\ 0 & 0 & -24\end{bmatrix}=\boldsymbol{LU}$$

- 根据步骤4求解方程

$$\boldsymbol{Ly}=(14,18,20)^{\mathrm{T}}$$

得到

$$\boldsymbol{y}=(14,-10,-72)^{\mathrm{T}}$$

例7.5（续）

- 用直接三角分解法解

$$\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 3 & -5 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}14\\ 18\\ 20\end{bmatrix}$$

- 根据上一页步骤2和步骤3，得到

$$\boldsymbol{A}=\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 3 & -5 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\ 0 & 1 & -4\\ 0 & 0 & -24\end{bmatrix}=\boldsymbol{LU}$$

- 根据步骤5求解方程

$$\boldsymbol{Ux}=(14,-10,-72)^{\mathrm{T}}$$

得到

$$\boldsymbol{x}=(1,2,3)^{\mathrm{T}}$$

讨论

- 在用计算机计算时，由于计算好u_{ri}后a_{ri}就不用了，因此计算好$\boldsymbol{L},\boldsymbol{U}$的元素后就存放在$\boldsymbol{A}$的相应位置
  - 例如

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \to \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ l_{21} & u_{22} & u_{23} & u_{24} \\ l_{31} & l_{32} & u_{33} & u_{34} \\ l_{41} & l_{42} & l_{43} & u_{44} \end{bmatrix}$$

    最后在存放$\boldsymbol{A}$的数组中得到$\boldsymbol{L},\boldsymbol{U}$的元素

- 由直接三角分解计算公式，需要计算形如$\sum a_i b_i$的式子，可采用“双精度累加”，以提高精度

讨论（续）

- 直接分解法大约需要$n^3/3$次乘、除法运算，和Gauss消去法的计算量基本相同
- 如果已经实现了$\boldsymbol{A}=\boldsymbol{LU}$的分解计算，且$\boldsymbol{L},\boldsymbol{U}$保存在$\boldsymbol{A}$的相应位置，则用直接三角分解法解具有相同系数的方程组$\boldsymbol{Ax}=(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_m)$是相当方便的
  - 每解一个方程组$\boldsymbol{Ax}=\boldsymbol{b}_j$仅需要增加$n^2$次乘除法运算
- 矩阵$\boldsymbol{A}$的分解公式(7.4.2)、(7.4.3)又称为Doolittle分解公式

选主元的三角分解法

- 从直接三角分解公式看出，当$\boldsymbol{u_{rr}}=0$时计算将中断或者当$\boldsymbol{u_{rr}}$绝对值很小时，按分解公式计算可能引起舍入误差的累积

- 如果$\boldsymbol{A}$非奇异，就可以通过交换$\boldsymbol{A}$的行实现矩阵$\boldsymbol{PA}$的LU分解，因此可采用与列主元素消去法类似的方法，将直接三角分解法修改为（部分）选主元的三角分解法
  - 可证明下述方法与列主元素消去法等价

选主元的三角分解法

- 设第$r-1$步分解已完成，这时有

$$
\boldsymbol{A} \rightarrow \begin{bmatrix}
u_{11} & u_{12} & \cdots & \cdots & \cdots & \cdots & u_{1n} \\
l_{21} & u_{22} & & & & & \vdots \\
l_{31} & l_{32} & \ddots & & & & \vdots \\
\vdots & & \ddots & u_{r-1,r-1} & \cdots & \cdots & u_{r-1,n} \\
\vdots & & & l_{r,r-1} & a_{rr} & \cdots & a_{rn} \\
\vdots & & & \vdots & \vdots & & \vdots \\
l_{n1} & l_{n2} & \cdots & l_{n,r-1} & a_{nr} & \cdots & a_{nn}
\end{bmatrix}.
$$

- 第r步分解需用到式(7.4.2)及式(7.4.3)

选主元的三角分解法（续）

$$u_{ri} = \left(a_{ri} - \sum_{k=1}^{r-1} l_{rk} u_{ki}\right) \quad (i = r, r+1, \cdots, n) \qquad (7.4.2)$$

$$l_{ir} = \left(a_{ir} - \sum_{k=1}^{r-1} l_{ik} u_{kr}\right) / u_{rr} \quad (i = r+1, \cdots, n) \qquad (7.4.3)$$

- 引进量

$$s_i = a_{ir} - \sum_{k=1}^{r-1} l_{ik} u_{kr} \quad (i = r, r+1, \cdots, n)$$

- 显然

$$u_{rr} = s_r$$

$$l_{ir} = s_i / s_r \quad (i = r+1, \cdots, n)$$

选主元的三角分解法（续）

- 为了避免用小的数u_{rr}作除数，选主元

$$\max_{r\le i\le n}|s_i| = |s_{i_r}|$$

- 用s_{i_r}作为u_{rr}，交换$\boldsymbol{A}$的r行与i_r行元素（将(i,j)位置的新元素仍记作l_{ij}及a_{ij}），由此再进行第r步分解计算

- 通过这样的变换，可保证

$$|l_{ir}| = \left|\frac{s_i}{s_r}\right| \le 1 \ \ (i = r+1,\cdots,n)$$

算法4 选主元的三角分解法

- 用$\boldsymbol{PA} = \boldsymbol{I}_{n-1,i_{n-1}} \cdots \boldsymbol{I}_{1i_1}\boldsymbol{A}$的三角分解冲掉$\boldsymbol{A}$，用$\mathrm{Ip}(n)$记录主行，解$\boldsymbol{x}$存放在$\boldsymbol{b}$内，$r$表示消元次数
  - 对于$r = 1,2,\cdots,n$，做到步4

1. 计算s_i

$$a_{ir} \leftarrow s_i = a_{ir} - \sum_{k=1}^{r-1} l_{ik}u_{kr} \quad (i = r, r+1, \cdots, n)$$

2. 选主元

$$|s_{i_r}| = \max_{r \le i \le n} |s_i|, \qquad \mathrm{Ip}(r) \leftarrow i_r$$

与算法3一样

3. 交换$\boldsymbol{A}$的r行与i_r行元素

$$a_{ri} \leftrightarrow a_{i_r i} \quad (i = 1,2,\cdots,n)$$

算法4 选主元的三角分解法（续）

4．计算$\boldsymbol{U}$的第r行元素，$\boldsymbol{L}$的第r列元素

$$a_{rr} = u_{rr} = s_r$$

$$a_{ir} \leftarrow l_{ir} = s_i / u_{rr} = a_{ir} / a_{rr} \quad (i = r+1, \cdots, n)$$

$$a_{ri} \leftarrow u_{ri} = a_{ri} - \sum_{k=1}^{r-1} l_{rk} u_{ki} \quad (i = r+1, \cdots, n)$$

这时有$|l_{ir}| \le 1$

- 上述计算过程完成后就实现了PA的LU分解，且$\boldsymbol{U}$保存在$\boldsymbol{A}$上三角部分，$\boldsymbol{L}$保存在$\boldsymbol{A}$的下三角部分，排列阵$\boldsymbol{P}$由Ip(n)最后记录可知

算法4 选主元的三角分解法（续）

计算$\boldsymbol{Pb}$，对$\boldsymbol{b}$顺序执行前面的换行操作

5\. $i = 1,2,\cdots,n-1$

（1）$t \leftarrow \mathrm{Ip}(i)$

（2）如果$i = t$，则转（3）；否则$b_i \leftrightarrow b_t$

（3）继续循环

6\. 求解$\boldsymbol{Ly} = \boldsymbol{Pb}$

$$b_i \leftarrow b_i - \sum_{k=1}^{i-1} l_{ik} b_k \quad (i = 2,3,\cdots,n)$$

7\. 求解$\boldsymbol{Ux} = \boldsymbol{y}$

$$b_n \leftarrow b_n / u_{nn}$$

$$b_i \leftarrow \left(b_i - \sum_{k=i+1}^{n} u_{ik} b_k\right) / u_{ii} \quad (i = n-1,\cdots,1)$$

矩阵求逆

- 利用算法4的结果，可计算$\boldsymbol{A}$的逆矩阵

$$\boldsymbol{PA} = \boldsymbol{LU}$$

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

- 利用$\boldsymbol{PA}$的三角分解计算$\boldsymbol{A}^{-1}$步骤
  1. 计算上三角阵的逆矩阵$\boldsymbol{U}^{-1}$
  2. 计算$\boldsymbol{U}^{-1}\boldsymbol{L}^{-1}$
  3. 交换$\boldsymbol{U}^{-1}\boldsymbol{L}^{-1}$列（利用 $\mathrm{Ip}(n)$ 最后记录）
  - 上述方法求$\boldsymbol{A}^{-1}$大约需要n^3次乘法运算

平方根法

- 考虑系数矩阵是对称正定的
- 平方根法：利用对称正定矩阵的三角分解而得到的求解对称正定方程组的一种有效方法
  - 在计算机上广泛应用平方根法解此类方程组
- 设$\boldsymbol{A}$为对称矩阵，且$\boldsymbol{A}$的所有顺序主子式均不为零，由定理7.3知，$\boldsymbol{A}$可唯一分解为

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}=\begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{nl} & \cdots & l_{n,n-1} & 1 \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix} \quad (7.4.1)$$

对称矩阵的三角分解

- 为了利用$\boldsymbol{A}$的对称性，将$\boldsymbol{U}$再分解

$$\boldsymbol{U}=\begin{bmatrix} u_{11} & & & \\ & u_{22} & & \\ & & \ddots & \\ & & & u_{nn} \end{bmatrix}\begin{bmatrix} 1 & \frac{u_{12}}{u_{11}} & \cdots & \frac{u_{1n}}{u_{11}} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \frac{u_{n-1,n}}{u_{n-1,n-1}} \\ & & & 1 \end{bmatrix}=\boldsymbol{D}\boldsymbol{U}_0$$

  - $\boldsymbol{D}$为对角阵，$\boldsymbol{U}_0$为单位上三角阵
  - 因此 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{U}_0$
  - 此外 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{U}_0^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}})$
  - 根据定理7.3中分解的唯一性，可知

$$\boldsymbol{U}_0^{\mathrm{T}}=\boldsymbol{L}\ \Rightarrow \boldsymbol{A}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}$$

对称矩阵的三角分解（续）

- **定理7.7**（对称阵的三角分解定理）设$\boldsymbol{A}$为n阶对称阵，且$\boldsymbol{A}$的所有顺序主子式均不为零，则$\boldsymbol{A}$可唯一分解为

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}$$

其中$\boldsymbol{L}$为单位下三角阵，$\boldsymbol{D}$为对角阵

- 接下来考虑$\boldsymbol{A}$为对称正定矩阵的情况，重温7.2节的引理（需要用到证明的中间结论）
- **引理** 约化的主元素$a_{ii}^{(i)}\neq 0$（$i=1,2,\cdots,k$）的充要条件是矩阵$\boldsymbol{A}$的顺序主子式$D_i\neq 0$（$i=1,2,\cdots,k$）

对称矩阵的三角分解（续）

- 现设$\boldsymbol{A}$为对称正定矩阵，首先说明$\boldsymbol{A}$的分解式$\boldsymbol{A}=\boldsymbol{LDL}^{\mathrm{T}}$中$\boldsymbol{D}$的对角元素$d_i$均为正数
  - $\boldsymbol{A}$为对称正定矩阵，那么$\boldsymbol{A}$的顺序主子式$D_k > 0\ (k=1,2,\dots,n-1)$
  - 由7.2节的引理证明过程可知，

$$D_k = \begin{vmatrix} a_{11}^{(1)} & \cdots & a_{1k}^{(1)} \\ & \ddots & \vdots \\ & & a_{kk}^{(k)} \end{vmatrix} = a_{11}^{(1)} \cdots a_{kk}^{(k)} = d_1 \cdots d_k$$

  - 因此

$$d_1 = D_1 > 0 \qquad d_i = \frac{D_i}{D_{i-1}} > 0\ (i=2,3,\dots,n)$$

对称矩阵的三角分解（续）

- 根据$\boldsymbol{D}$的非负性

$$\boldsymbol{D}=\begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}=\begin{bmatrix} \sqrt{d_1} & & \\ & \ddots & \\ & & \sqrt{d_n} \end{bmatrix}\begin{bmatrix} \sqrt{d_1} & & \\ & \ddots & \\ & & \sqrt{d_n} \end{bmatrix}=\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{D}^{\frac{1}{2}}$$

- 结合定理7.7得到

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}=\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{L}^{\mathrm{T}}=(\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}})(\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}})^{\mathrm{T}}=\boldsymbol{L}_1\boldsymbol{L}_1^{\mathrm{T}}$$

其中$\boldsymbol{L}_1=\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}}$为下三角阵

- **定理7.8**（对称正定矩阵的三角分解） 如果$\boldsymbol{A}$为n阶对称正定矩阵，则存在一个实的非奇异下三角阵$\boldsymbol{L}$使$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}$，当限定$\boldsymbol{L}$的对角元素为正时，这种分解是唯一的

计算对称正定矩阵的三角分解

- 下面用直接分解方法来确定计算$\boldsymbol{L}$元素的递推公式
  - 由于

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$=\begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}\begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ & l_{22} & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & l_{nn} \end{bmatrix}$$

其中$l_{ii}>0\ (i=1,2,\cdots,n)$

计算对称正定矩阵的三角分解（续）

- 考虑下三角元素，即$a_{ij}\ (i \geq j)$，根据矩阵乘法

$$a_{ij} = \sum_{k=1}^{n} l_{ik} l_{jk} = \sum_{k=1}^{j-1} l_{ik} l_{jk} + l_{jj} l_{ij}$$

其中第二个等号由于$l_{jk} = 0\ (j < k)$

- 依据上式，可以逐列计算矩阵$\boldsymbol{L}$的元素

$$\boldsymbol{L} = \begin{bmatrix} l_{11} & & & \\ l_{21} & \boxed{l_{22}} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}$$

$$l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2\right)^{\frac{1}{2}} \quad (7.4.7)$$

$$l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}\, l_{jk}\right) / l_{jj}$$

平方根法计算公式

- 解对称正定方程组$\boldsymbol{Ax} = \boldsymbol{b}$的平方根法计算公式
  - 对于$j = 1,2,\cdots,n$

1. 计算l_{jj}

$$l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2\right)^{\frac{1}{2}}$$

2. 按列计算l_{ij}

$$l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}\, l_{jk}\right)/l_{jj}, \quad i = j+1,\cdots,n$$

平方根法计算公式（续）

- 对于方程组$\boldsymbol{Ax}=\boldsymbol{b}$，如果对系数矩阵进行了分解$\boldsymbol{A}=\boldsymbol{LL}^T$，则变为求解两个三角方程组

3．求解$\boldsymbol{Ly}=\boldsymbol{b}$

$$y_i=\left(b_i-\sum_{k=1}^{i-1} l_{ik}y_k\right)\Big/ l_{ii} \qquad i=1,2,\cdots,n$$

4．求解$\boldsymbol{L}^T\boldsymbol{x}=\boldsymbol{y}$

$$x_i=\left(b_i-\sum_{k=i+1}^{n} l_{ki}x_k\right)\Big/ l_{ii} \quad i=n,n-1,\cdots,1$$

数值稳定性

- 根据前的推导

$$l_{jj} = (a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2)^{\frac{1}{2}} \Rightarrow a_{jj} = \sum_{k=1}^{j} l_{jk}^2, \quad j = 1,2,\cdots,n$$

- 因此

$$l_{jk}^2 \le a_{jj} \le \max_{1\le j\le n} a_{jj} \Rightarrow \max_{j,k} l_{jk}^2 \le \max_{1\le j\le n} a_{jj}$$

- 上面分析说明，分解过程中元素l_{jk}的数量级不会增长且对角元素l_{jj}恒为正数，于是不选主元素的平方根法是一个数值稳定的方法

时空复杂度

- 当求出$\boldsymbol{L}$的第j列元素时，$\boldsymbol{L}^{\mathrm{T}}$的第$j$行元素亦就算出，所以平方根法约需$\frac{n^3}{6}$次乘除法运算，大约为一般直接LU分解法计算量的一半

- 由于$\boldsymbol{A}$为对称阵，因此在用计算机计算时只需存储$\boldsymbol{A}$的下三角部分，共需要存储$\frac{n(n+1)}{2}$个元素，可用一维数组存放

- $\boldsymbol{L}$的元素存放在$\boldsymbol{A}$的相应位置

避免开方的分解

- 根据$l_{jj}=(a_{jj}-\sum_{k=1}^{j-1}l_{jk}^2)^{\frac{1}{2}}$ 可知平方根法需要用到开方运算
- 为了避免开方运算，下面用定理7.7的分解式

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

$$=\begin{bmatrix}1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1\end{bmatrix}\begin{bmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{bmatrix}\begin{bmatrix}1 & l_{21} & \cdots & l_{n1} \\ & 1 & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1\end{bmatrix}$$

避免开方的分解（续）

- 考虑下三角元素，即$a_{ij}\ (i \geq j)$，根据矩阵乘法

$$a_{ij} = \sum_{k=1}^{n} (\boldsymbol{LD})_{ik} \left(\boldsymbol{L}^{\mathrm{T}}\right)_{kj} = \sum_{k=1}^{n} l_{ik} d_k l_{jk} = \sum_{k=1}^{j-1} l_{ik} d_k l_{jk} + l_{ij} d_j$$

其中第二个等号由于$l_{jk} = 0\ (j < k)$、$l_{jj} = 1$

- 依据上式，可以逐行计算矩阵$\boldsymbol{L}$的元素

$$\boldsymbol{L} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \qquad l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}\right) \Big/ d_j$$

$$(j = 1,2,\cdots,i-1)$$

避免开方的分解（续）

- 考虑下三角元素，即$a_{ij}\ (i \geq j)$，根据矩阵乘法

$$a_{ij} = \sum_{k=1}^{n} (\boldsymbol{LD})_{ik} \left(\boldsymbol{L}^{\mathrm{T}}\right)_{kj} = \sum_{k=1}^{n} l_{ik} d_k l_{jk} = \sum_{k=1}^{j-1} l_{ik} d_k l_{jk} + l_{ij} d_j$$

其中第二个等号由于$l_{jk} = 0\ (j < k)$、$l_{jj} = 1$

- 依据上式，可以得到$\boldsymbol{D}$的对角元素公式

$$\boldsymbol{D} = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \boxed{\ddots} & \\ & & & d_n \end{bmatrix} \qquad d_i = a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 d_k$$

改进的平方根法计算公式

- 为避免重复计算，令$t_{ij}=l_{ij}d_j$，根据上页得到以下按行计算$\boldsymbol{L},\boldsymbol{T}$元素的公式
  - 对于$i=1,2,\cdots,n$

1. 按行计算t_{ij}

$$t_{ij}=a_{ij}-\sum_{k=1}^{j-1}t_{ik}l_{jk}\quad j=1,2,\cdots,i-1$$

2. 按行计算l_{ij}

$$l_{ij}=t_{ij}/d_j\quad j=1,2,\cdots,i-1$$

3. 计算d_i

$$d_i=a_{ii}-\sum_{k=1}^{i-1}t_{ik}l_{ik}$$

改进的平方根法计算公式（续）

- 为了节省空间
  - 在计算出$\boldsymbol{T}=\boldsymbol{LD}$的第$i$行元素$t_{ij}$ ($j=1,2,\cdots,i-1$)后，存放在$\boldsymbol{A}$的第i行相应位置
  - 然后再计算$\boldsymbol{L}$的第i行元素，存放在$\boldsymbol{A}$的第i行
  - $\boldsymbol{D}$的对角元素存放在$\boldsymbol{A}$的相应位置

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{21} & a_{22} & a_{32} & a_{42} \\ a_{31} & a_{32} & a_{33} & a_{43} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \to \begin{bmatrix} d_1 & & & \\ l_{21} & d_2 & & \\ l_{31} & l_{32} & d_3 & \\ t_{41} & t_{42} & t_{43} & t_{44} \end{bmatrix} \to \begin{bmatrix} d_1 & & & \\ l_{21} & d_2 & & \\ l_{31} & l_{32} & d_3 & \\ l_{41} & l_{42} & l_{43} & d_4 \end{bmatrix}$$

- 对称正定矩阵$\boldsymbol{A}$按$\boldsymbol{LDL}^{\mathrm{T}}$分解和按$\boldsymbol{LL}^{\mathrm{T}}$分解计算量差不多，但$\boldsymbol{LDL}^{\mathrm{T}}$分解不需要开方计算

改进的平方根法计算公式（续）

- 对于方程组$\boldsymbol{Ax}=\boldsymbol{b}$，如果对系数矩阵进行了分解$\boldsymbol{A}=\boldsymbol{LDL}^{\mathrm{T}}$，则变为求解两个三角方程组

4．求解$\boldsymbol{Ly}=\boldsymbol{b}$

$$\begin{cases} y_1 = b_1 \\ y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k \quad i = 2, \cdots, n \end{cases}$$

5．求解$\boldsymbol{DL}^T\boldsymbol{x}=\boldsymbol{y}$

$$\begin{cases} x_n = y_n / d_n \\ x_i = y_i / d_i - \sum_{k=i+1}^{n} l_{ki} x_k \quad i = n-1, \cdots, 2, 1 \end{cases}$$

对角占优的三对角方程组

- 在一些实际问题中，例如船体数学放样中建立三次样条函数，都会要求解系数矩阵为对角占优的三对角方程组

$$\begin{bmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{bmatrix}$$

简记作

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{f} \tag{7.4.12}$$

对角占优的三对角方程组（续）

$$A = \begin{bmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{bmatrix}$$

- A满足对角占优条件
  1. $|b_1| > |c_1| > 0$
  2. $|b_i| \geq |a_i| + |c_i|, \quad a_i c_i \neq 0 \quad (i = 2, \cdots, n-1)$
  3. $|b_n| > |a_n| > 0$
- 下面利用矩阵的直接三角分解法来推导解三对角方程组(7.4.12)的计算公式

对角占优的矩阵三角分解

- 由系数阵$\boldsymbol{A}$的特点，可以将$\boldsymbol{A}$分解为两个三角阵的乘积，即

$$\boldsymbol{A}=\begin{bmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{bmatrix}=\overset{\boldsymbol{L}}{\begin{bmatrix} \alpha_1 & & & & \\ \gamma_2 & \alpha_2 & & & \\ & \gamma_3 & \alpha_3 & & \\ & & \ddots & \ddots & \\ & & & \gamma_n & \alpha_n \end{bmatrix}}\overset{\boldsymbol{U}}{\begin{bmatrix} 1 & \beta_1 & & & \\ & 1 & \beta_2 & & \\ & & 1 & \ddots & \\ & & & \ddots & \beta_{n-1} \\ & & & & 1 \end{bmatrix}}$$

  - $\boldsymbol{L}$为下三角阵，$\boldsymbol{U}$为单位上三角阵
  - $\alpha_i,\beta_i,\gamma_i$为待定系数，比较公式两端可得

$$\begin{cases} b_1=\alpha_1, & c_1=\alpha_1\beta_1 \\ a_i=\gamma_i, & b_i=\gamma_i\beta_{i-1}+a_i \qquad i=2,3,\cdots,n \\ c_i=\alpha_i\beta_i, & i=2,3,\cdots,n-1 \end{cases} \tag{7.4.14}$$

对角占优的矩阵三角分解（续）

$$\begin{cases} b_1 = \alpha_1, & c_1 = \alpha_1\beta_1 \\ a_i = \gamma_i, & b_i = \gamma_i\beta_{i-1} + a_i \qquad i = 2,3,\cdots,n \\ c_i = \alpha_i\beta_i, & i = 2,3,\cdots,n-1 \end{cases} \tag{7.4.14}$$

- 上式可完全确定$\{\alpha_i\},\{\beta_i\},\{\gamma_i\}$，实现了$\boldsymbol{A}$的LU分解

$$\alpha_1 = b_1, \qquad \beta_1 = \frac{c_1}{\alpha_1} = \frac{c_1}{b_1}$$

$$\gamma_i = a_i, \qquad i = 2,3,\dots,n$$

$$\alpha_i = b_i - \gamma_i\beta_{i-1} = b_i - a_i\beta_{i-1}, \qquad i = 2,3,\dots,n$$

$$\beta_i = \frac{c_i}{\alpha_i} = \frac{c_i}{b_i - a_i\beta_{i-1}}, \qquad i = 2,3,\dots,n$$

分解后系数的性质

$$c_1 = \alpha_1\beta_1$$
$$c_i = \alpha_i\beta_i, i = 2,3,\cdots,n-1$$

- 为了分析算法的稳定性，需要刻画分解后系数的大小

1. γ_i由a_i的性质决定

$$\gamma_i = a_i, \qquad i = 2,3,\ldots,n$$

2. 从上页可知$\beta_i = c_i/\alpha_i$，下面用归纳法证明

$$|\alpha_i| > |c_i| \neq 0 \Leftrightarrow 0 < |\beta_i| = \left|\frac{c_i}{\alpha_i}\right| < 1, \qquad (i = 1,2,\cdots,n-1)$$

- 当$i = 1$时

$$|\beta_1| = \left|\frac{c_1}{\alpha_1}\right| = \left|\frac{c_1}{b_1}\right| \xRightarrow{|b_1|>|c_1|>0} 0 < |\beta_1| < 1$$

$$b_i = \gamma_i \beta_{i-1} + a_i$$
$$a_i = \gamma_i$$

分解后系数的性质（续）

- 假设$i-1$时， $0 < |\beta_{i-1}| = |c_{i-1}/\alpha_{i-1}| < 1$，

$$\left.\begin{matrix} |\alpha_i| = |b_i - \gamma_i \beta_{i-1}| \\ a_i = \gamma_i \end{matrix}\right\} \Rightarrow \left.\begin{matrix} |\alpha_i| = |b_i - a_i \beta_{i-1}| \geq |b_i| - |a_i \beta_{i-1}| \\ |\beta_{i-1}| = |c_{i-1}/\alpha_{i-1}| < 1 \end{matrix}\right\}$$

$$\left.\begin{matrix} \Rightarrow |\alpha_i| > |b_i| - |a_i| \\ |b_i| \geq |a_i| + |c_i| \end{matrix}\right\} \Rightarrow \left.\begin{matrix} |\alpha_i| > |c_i| \\ c_i \neq 0 \end{matrix}\right\} \Rightarrow 0 < |\beta_i| = \left|\frac{c_i}{\alpha_i}\right| < 1$$

3. 从上面的证明可知，对于$i = 2,3,\cdots,n-1$

$$|\alpha_i| > |b_i| - |a_i| \geq |c_i| > 0, \qquad |\alpha_n| > |b_n| - |a_n| > 0$$

同时，对于$i = 2,3,\cdots,n$

$$|\alpha_i| = |b_i - \gamma_i \beta_{i-1}| = |b_i - a_i \beta_{i-1}| \leq |b_i| + |a_i \beta_{i-1}| < |b_i| + |a_i|$$

分解后系数的性质（续）

- 总结上述讨论，有以下定理
- **定理7.9** 设有三对角方程组$\boldsymbol{Ax}=\boldsymbol{f}$，其中$\boldsymbol{A}$满足对角占优条件，则$\boldsymbol{A}$为非奇异矩阵且分解后的系数$\{\gamma_i\}$、$\{\beta_i\}$和$\{\alpha_i\}$满足

  1. $\gamma_i = a_i \qquad (i = 2,3,\dots,n)$
  2. $0 < |\beta_i| < 1 \qquad (i = 1,2,\dots,n-1)$
  3. $0 < |c_i| \le |b_i| - |a_i| < |\alpha_i| < |b_i| + |a_i| \qquad (i = 2,3,\cdots,n-1)$

     $0 < |b_n| - |a_n| < |\alpha_n| < |b_n| + |a_n|$

     $\alpha_1 = b_1$

追赶法

- 得到$\boldsymbol{A}=\boldsymbol{LU}$的分解之后，求解$\boldsymbol{Ax}=\boldsymbol{f}$等价于解两个三角方程组

1. 计算$\{\beta_i\}$的递推公式

$$\beta_1=\frac{c_1}{b_1},\qquad \beta_i=\frac{c_i}{b_i-a_i\beta_{i-1}},\quad i=2,3,\ldots,n-1$$

2. 解$\boldsymbol{Ly}=\boldsymbol{f}$

$$\boldsymbol{L}=\begin{bmatrix}\alpha_1 & & & & \\ \gamma_2 & \alpha_2 & & & \\ & \gamma_3 & \alpha_3 & & \\ & & \ddots & \ddots & \\ & & & \gamma_n & \alpha_n\end{bmatrix}$$

$$y_1=\frac{f_1}{\alpha_1}=\frac{f_1}{b_1}$$

$$y_i=\frac{f_i-\gamma_i y_{i-1}}{\alpha_i}=\frac{f_i-a_i y_{i-1}}{b_i-a_i\beta_{i-1}},$$

$$i=2,3,\ldots,n$$

追赶法（续）

3. 解$\boldsymbol{U}\boldsymbol{x}=\boldsymbol{y}$

$$\boldsymbol{U}=\begin{bmatrix} 1 & \beta_1 & & & \\ & 1 & \beta_2 & & \\ & & 1 & \ddots & \\ & & & \ddots & \beta_{n-1} \\ & & & & 1 \end{bmatrix}$$

$$x_n = y_n$$

$$x_i = y_i - \beta_i x_{i+1},$$

$$i = n-1, n-2, \dots, 1$$

- 将计算系数$\beta_1 \to \beta_2 \to \cdots \to \beta_{n-1}$及$y_1 \to y_2 \to \cdots \to y_n$的过程称为追的过程
- 将计算方程组的解$x_n \to x_{n-1} \to \cdots \to x_1$的过程称为赶的过程

追赶法（续）

3. 解$\boldsymbol{Ux}=\boldsymbol{y}$

$$\boldsymbol{U}=\begin{bmatrix} 1 & \beta_1 & & & \\ & 1 & \beta_2 & & \\ & & 1 & \ddots & \\ & & & \ddots & \beta_{n-1} \\ & & & & 1 \end{bmatrix}$$

$$x_n = y_n$$

$$x_i = y_i - \beta_i x_{i+1},$$
$$i = n-1, n-2, \ldots, 1$$

- 定理7.9表明，追赶法计算公式中不会出现中间结果数量级的巨大增长和舍入误差的严重累积

- 因此，追赶法是一个数值稳定的方法

时空复杂度

- 追赶法公式实际上就是把Gauss消去法用到求解三对角方程组上的结果
- 由于$\boldsymbol{A}$特别简单，因此使得求解的计算公式也非常简单，计算量仅为$5n-4$次乘除法运算
- 另外增加解一个方程组$\boldsymbol{Ax}=\boldsymbol{f}_2$仅需增加$3n-2$次乘除运算，易见追赶法的计算量比较小
- 在用计算机计算时，只需用三个一维数组分别存储$\boldsymbol{A}$的三条对角线元素$\{a_i\}$，$\{b_i\}$和$\{c_i\}$，此外还需要用两组存储单元保存$\{\beta_i\}$，$\{y_i\}$或$\{x_i\}$