5 解线性代数方程组的直接法

Direct Solutions of Linear Systems of Equations

王家兵
jbwang@scut.edu.cn

背景

- 在自然科学和工程技术中，许多问题的解决常常归结为线性代数方程组的求解.
  - 解的存在性和解的结构：线性代数
- 直接法：经过有限步计算后求得方程组精确解
- 问题描述：n阶线性代数方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

$$\boldsymbol{AX} = \boldsymbol{b}$$

$\boldsymbol{A} = (a_{ij})$是$n\times n$系数矩阵；$\boldsymbol{X} = \{x_i\}$是$n$维向量；$\boldsymbol{b} = \{b_i\}$右端项

背景

- 回顾
  - 行列式（determinant）的定义
  - 行列式的几何意义

- 若$\boldsymbol{A}$是非奇异矩阵，则方程组有唯一解.
  - 记$D = det\boldsymbol{A}$，应用Cramer法则可得

$$x_i = \frac{D_i}{D}\ (i = 1,2,\cdots,n)$$

其中，D_i是用$\boldsymbol{b}$代替$\boldsymbol{A}$中第i列而得到的相应行列式.

  - 用Cramer法则求解需要的乘除法次数
    - 计算n阶行列式需要$(n-1)\cdot n!$次乘法.
    - 共需计算$(n+1)$个行列式
    - 计算x_i需n次除法

$$N = (n^2 - 1)\cdot n! + n$$

背景

- $n = 10, N = 359251210$; $n = 20$, $N = 9.7\times10^{20}$, $n = 100, N = 9.3 \times10^{161}$.
- 2017年11月13日，全球超级计算机500强榜单公布，“神威·太湖之光”以每秒9.3亿亿次的浮点运算速度第四次夺冠：系统峰值性能每秒12.5亿亿次，持续性能每秒9.3亿亿次.
- 神威利用克莱姆法则计算n=100需要的计算时间：

$$\frac{9.3\times10^{161}}{10^{17}\times3600\times24\times365}=2.95\times10^{137}\,(\text{年})$$

□ 顺序高斯消去法

- 高斯消去法是目前计算机上一种常用的有效方法
- 消元：通过初等变换将方程组转化为一个等价的三角形方程组.

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = g_1 \\ \qquad b_{22}x_2 + \cdots + b_{2n}x_n = g_2 \\ \qquad\qquad \vdots \\ \qquad\qquad b_{nn}x_n = g_n \end{cases}$$

- 回代法求解$x_n, x_{n-1}, \cdots, x_1$

$$x_i = \frac{g_i - \sum_{j=i+1}^{n} b_{ij}x_j}{bii} \quad (i = n, n-1, \cdots, 1)$$

□ 顺序高斯消去法计算过程

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + a_{1n}^{(1)}x_n = b_1^{(1)} \\ a_{21}^{(1)}x_1 + a_{22}^{(1)}x_2 + \cdots + a_{2n}^{(1)}x_n = b_2^{(1)} \\ \qquad\qquad\qquad \vdots \\ a_{n1}^{(1)}x_1 + a_{n2}^{(1)}x_2 + \cdots + a_{nn}^{(1)}x_n = b_n^{(1)} \end{cases}$$

$$\boldsymbol{A}^{(1)}\boldsymbol{X} = \boldsymbol{b}^{(1)}$$

- 令$a_{11}^{(1)} \neq 0$,令$m_{i1} = \dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}}$，将第1行乘$-m_{i1}$加到第$i$行($i = 2,3,\cdots,n$)

$$
\begin{cases}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + a_{1n}^{(1)}x_n = b_1^{(1)} \\
\qquad\quad a_{22}^{(2)}x_2 + \cdots + a_{2n}^{(2)}x_n = b_2^{(2)} \\
\qquad\qquad\qquad \vdots \\
\qquad\quad a_{n2}^{(2)}x_2 + \cdots + a_{nn}^{(2)}x_n = b_n^{(2)}
\end{cases}
$$

其中$\begin{cases} a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1}a_{1j}^{(1)}\ (i,j = 2,3,\cdots,n) \\ b_i^{(2)} = b_i^{(1)} - m_{i1}b_1^{(1)} \end{cases}$

$$\boldsymbol{A}^{(2)}\boldsymbol{X} = \boldsymbol{b}^{(2)}$$

- 重复此过程$n-1$次，得到三角形方程组.

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + a_{1n}^{(1)}x_n = b_1^{(1)} \\ \qquad\qquad a_{22}^{(2)}x_2 + \cdots + a_{2n}^{(2)}x_n = b_2^{(2)} \\ \qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad\qquad a_{nn}^{(n)}x_n = b_n^{(n)} \end{cases}$$

$$\boldsymbol{A}^{(n)}\boldsymbol{X} = \boldsymbol{b}^{(n)}$$

- 消元计算公式归纳：

$$\begin{cases} a_{ij}^{(k+1)} = a_{ij}^{(k)}, b_i^{(k+1)} = b_i^{(k)} \ (1 \le i \le k, i \le j \le n) \\ a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)} \ (i, j = k+1, \cdots, n) \\ b_i^{(k+1)} = b_i^{(k)} - m_{ik}b_k^{(k)} \ (i = k+1, \cdots, n) \\ a_{ij}^{(k+1)} = 0 \quad (1 \le j \le k < i \le n) \end{cases}$$

高斯消去法的计算量

- 第1次消元
  - 计算$m_{i1}(i=2,3,\cdots,n)$需要$n-1$次除法
  - 计算$a_{ij}^{(2)}=a_{ij}^{(1)}-m_{i1}a_{1j}^{(1)}(i,j=2,3,\cdots,n)$需要$(n-1)^2$次乘法和减法
  - 计算$b_i^{(2)}=b_i^{(1)}-m_{i1}b_1^{(1)}\ (i=2,3,\cdots,n)$需要$n-1$次乘法和减法

  乘除法：$(n-1)+(n-1)^2+(n-1)=n^2-1$

  加减法：$(n-1)^2+(n-1)=n(n-1)$
- 第2次消元
  - 等价于一个$n-1$阶方程组的第1次消元

- 第1次消元

  乘除法：$(n-1)+(n-1)^2+(n-1)=n^2-1$

  加减法：$(n-1)^2+(n-1)=n(n-1)$

- 整个消元过程计算量：

乘除法：$\sum_{i=0}^{n-1}[(n-i)^2-1]=\frac{2n^3+3n^2-5n}{6}$

加减法：$\sum_{i=1}^{n} i(i-1)=\frac{n^3-n}{3}$

- 回代过程计算量

乘除法：$1+\sum_{i=1}^{n-1}[(n-i)+1]=\frac{n^2+n}{2}$

加减法：$\sum_{i=1}^{n-1}[(n-i-1)+1]=\frac{n^2-n}{2}$

$$\frac{100^3}{10^{17}}=10^{-11}(\text{秒})$$

- 顺序高斯消去法的矩阵解释
- 回顾：
  - 矩阵的初等变换
  - 初等变换与行列式的关系

顺序高斯消去法的矩阵解释

- 交换矩阵的两行、两列的变换
  - 单位矩阵$\boldsymbol{I} = [\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n]$，其中$\boldsymbol{e}_i$为单位坐标向量，第$i$位为1，其它为0.
  - 交换单位矩阵$\boldsymbol{I}$的i, j两行所得矩阵记为$\boldsymbol{P}_{ij}$
    - 用$\boldsymbol{P}_{ij}$左乘矩阵$\boldsymbol{A}$，其作用是交换$\boldsymbol{A}$的第i行和第j行
    - 用$\boldsymbol{P}_{ij}$右乘矩阵$\boldsymbol{A}$，其作用是交换$\boldsymbol{A}$的第i列和第j列
- 矩阵的第$i(i = 2,3, \cdots, n)$减去m_{i1}乘第1行的变换
  - 相当于矩阵$\boldsymbol{M}_1$左乘矩阵$\boldsymbol{A}$

$$\boldsymbol{M}_1 = \begin{bmatrix} 1 & & & \\ -m_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ -m_{n1} & 0 & \cdots & 1 \end{bmatrix}$$

- 第1次消元等价于

$$\boldsymbol{A}^{(1)}\boldsymbol{X} = \boldsymbol{b}^{(1)}$$
$$\boldsymbol{M}_1\boldsymbol{A}^{(1)}\boldsymbol{X} = \boldsymbol{M}_1\boldsymbol{b}^{(1)}$$
$$\boldsymbol{A}^{(2)} = \boldsymbol{M}_1\boldsymbol{A}^{(1)}, \boldsymbol{b}^{(2)} = \boldsymbol{M}_1\boldsymbol{b}^{(1)}$$
$$\boldsymbol{A}^{(2)}\boldsymbol{X} = \boldsymbol{b}^{(2)}$$

- 类似地

$$\boldsymbol{M}_i = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -m_{i+1,i} & 1 & & \\ & & \vdots & & \ddots & \\ & & -m_{n,i} & & & 1 \end{bmatrix}$$

$$\boldsymbol{A}^{(i+1)} = \boldsymbol{M}_i \boldsymbol{A}^{(i)}, \qquad \boldsymbol{b}^{(i+1)} = \boldsymbol{M}_i \boldsymbol{b}^{(i)}$$

- $n-1$步以后得到的上三角矩阵$\boldsymbol{A}^{(n)}$、右端项$\boldsymbol{b}^{(n)}$和$\boldsymbol{A}^{(1)}, \boldsymbol{b}^{(1)}$满足下列关系：

$$\begin{cases} \boldsymbol{A}^{(n)} = \boldsymbol{M}_{n-1} \cdot \boldsymbol{M}_{n-2} \cdots \boldsymbol{M}_1 \boldsymbol{A}^{(1)} \\ \boldsymbol{b}^{(n)} = \boldsymbol{M}_{n-1} \cdot \boldsymbol{M}_{n-2} \cdots \boldsymbol{M}_1 \boldsymbol{b}^{(1)} \end{cases}$$

列主元高斯消去法

- 主元高斯消去法根据主元选取范围的不同
  - 列主元消去法
  - 全主元消去法
- 在顺序高斯消去法的消元过程中，若出现$a_{kk}^{(k)} = 0$，则消元无法进行
- 即使$a_{kk}^{(k)} \neq 0$，但其值很小，用它作为除数也会导致其他元素量级的巨大增长和舍入误差的扩散，引起解的失真.

例：解方程组

$$\begin{cases} 0.003000x_1 + 59.14x_2 = 59.17 \\ 5.291x_1 - 6.130x_2 = 46.78 \end{cases}$$

该方程组的精确解为$x_1^* = 10.00, x_2^* = 1.000.$

答：利用顺序高斯消去法取4位有效数字进行运算

$$m_{21} = \frac{5.291}{0.003000} = 1763.666 \approx 1764$$

将第1个方程的-1764倍加到第2个方程，得

$$\begin{cases} 0.003000x_1 + 59.14x_2 = 59.17 \\ -104300x_2 = -104400 \end{cases}$$

回代得到：$x_2 = 1.001, x_1 = -10.00$

例：解方程组

$$\begin{cases} 5.291x_1 - 6.130x_2 = 46.78 \\ 0.003000x_1 + 59.14x_2 = 59.17 \end{cases}$$

答：利用顺序高斯消去法取4位有效数字进行运算

$$m_{21} = \frac{0.003000}{5.291} = 0.0005670$$

将第1个方程的-0.0005670倍加到第2个方程得

$$\begin{cases} 5.291x_1 - 6.130x_2 = 46.78 \\ 59.14x_2 = 59.14 \end{cases}$$

通过回代得：$x_1 = 10.00, x_2 = 1.000$

- 小主元是不稳定的根源

列主元高斯消去法计算过程

1. 在系数矩阵$\boldsymbol{A}^{(1)}$的第1列元素中选出绝对值最大的元素$a_{i_1,1}^{(1)}$称之为第1列的主元，即$\left|a_{i_1,1}^{(1)}\right| = \max\limits_{1\le i\le n}\left|a_{i1}^{(1)}\right|$. 如果$i_1 \neq 1$，则交换第1行和第$i_1$行元素及相应的右端项，交换完后再进行消元

$$\boldsymbol{A}^{(2)}\boldsymbol{X} = \boldsymbol{b}^{(2)}$$

2. 在系数矩阵$\boldsymbol{A}^{(2)}$第2列$a_{22}^{(2)}$及以下$n-1$个元素中选主元$a_{i_2,2}^{(2)}$，即$\left|a_{i_2,2}^{(2)}\right| = \max\limits_{2\le i\le n}\left|a_{i2}^{(2)}\right|$.如果$i_2 \neq 2$，则交换矩阵$\boldsymbol{A}^{(2)}$中的第2行与第$i_2$行元素及相应的右端项，交换完再进行消元

$$\boldsymbol{A}^{(3)}\boldsymbol{X} = \boldsymbol{b}^{(3)}$$

3. 一般地，第k步求i_k使$\left|a_{i_k,k}^{(k)}\right| = \max\limits_{k\le i\le n}\left|a_{ik}^{(k)}\right|$，如果$i_k \ne k$，则交换第$k$行与第$i_k$行及相应的右端项，经消元得

$$\boldsymbol{A}^{(k+1)}\boldsymbol{X} = \boldsymbol{b}^{(k+1)}$$

- 列主元消去法除了每步需要按列选主元并可能进行矩阵行交换外，其消元过程与顺序高斯消去法的过程一致.
- 列主元高斯消去法的计算量
  - 乘除法和加减次数与顺序高斯消去法相同
  - 增加选主元时$n(n-1)/2$次比较以及交换方程次序所需时间.

列主元高斯消去法的矩阵解释

- 记$\left|a_{i_1,1}^{(1)}\right| = \max\limits_{1\le i\le n}\left|a_{i1}^{(1)}\right|$，则说明第1个主元在第$i_1$行，交换第1行与第$i_1$行相当于用$\boldsymbol{P}_{1,i_1}$左乘矩阵$\boldsymbol{A}^{(1)}$，然后再进行消元

$$\boldsymbol{A}^{(2)} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}, \qquad \boldsymbol{b}^{(2)} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)}$$

- 同理有

$$\boldsymbol{A}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\boldsymbol{A}^{(n-1)},$$
$$\boldsymbol{b}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\boldsymbol{b}^{(n-1)}$$

- 由此得到等价三角方程组$\boldsymbol{A}^{(n)}\boldsymbol{X} = \boldsymbol{b}^{(n)}$，其中

$$\begin{cases}\boldsymbol{A}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\cdots\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\\ \boldsymbol{b}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\cdots\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)}\end{cases}$$

例：解方程组

$$\begin{cases} 12x_1 - 3x_2 + 3x_3 = 15 \\ -18x_1 + 3x_2 - x_3 = -15 \\ x_1 + x_2 + x_3 = 6 \end{cases}$$

注：准确解为$x_1 = 1, x_2 = 2, x_3 = 3$

全主元高斯消去法

- 全主元高斯消去法与列主元高斯消去法类似，所不同的选取主元的范围不同
- 优点：
  - 全主元高斯消去法在整个系数矩阵中找绝对值最大的元素作为主元，以控制舍入误差的增长，将舍入误差控制在一个最小的范围
- 缺点：
  - 找主元和交换行列次序要花费大量的机器时间

- 全主元高斯消去法计算过程
  - 求i_1, j_1使

$$\left|a_{i_1,j_1}^{(1)}\right| = \max_{\substack{1\le i\le n\\1\le j\le n}} \left|a_{ij}^{(1)}\right|$$

    - 如果$i_1 \neq 1$，则交换第1行和第i_1行及相应的右端项（$\boldsymbol{P}_{1,i_1}$左乘$\boldsymbol{A}^{(1)}$）
    - 如果$j_1 \neq 1$，则交换第1列和第j_1列（$\boldsymbol{P}_{1,j_1}$右乘$\boldsymbol{A}^{(1)}$），并交换x_1与x_{j_1}
  - 消元得

$$\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\boldsymbol{P}_{1,j_1}\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)}$$

$$\boldsymbol{A}^{(2)}\boldsymbol{P}_{(1,j_1)}^{-1}\boldsymbol{X} = \boldsymbol{b}^{(2)}$$

- 一般地，第k步($k=1,2,\cdots,n-1$)，求i_k, j_k

$$\left|a_{i_k,j_k}^{(k)}\right| = \max_{k\le i,j\le n}\left|a_{ij}^{(k)}\right|$$

- 如果$i_k \neq k$,交换第k行与i_k行，$\boldsymbol{P}_{k,i_k}$左乘$\boldsymbol{A}^{(k)}$和右端项；如果$j_k \neq k$,交换第k列与j_k列，$\boldsymbol{P}_{k,i_k}$右乘$\boldsymbol{A}^{(k)}$，并交换x_k与x_{j_k}，交换完后进行消元：

$$\boldsymbol{M}_k\boldsymbol{P}_{k,i_k}\boldsymbol{A}^{(k)}\boldsymbol{P}_{k,j_k}\boldsymbol{P}_{k,j_k}^{-1}\cdots\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{M}_k\boldsymbol{P}_{k,i_k}\boldsymbol{b}^{(k)}$$

$$\boldsymbol{A}^{(k+1)}\boldsymbol{P}_{k,j_k}^{-1}\cdots\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{b}^{(k+1)}$$

- 全主元高斯消去法算法

① 输入$\boldsymbol{A}, \boldsymbol{b}$及$\epsilon$

② for $i = 1$ to n do $d_i = i$//记录未知量位置变化

③ 选主元消元:

for $k = 1$ to $n-1$ do

- 求i_k, j_k使得$T = \left|a_{i_k,j_k}^{(k)}\right| = \max\limits_{k \le i,j \le n} \left|a_{ij}^{(k)}\right|$
- 若$\mathrm{T} < \epsilon$，则打印"求解失败"，停止；否则
- 若$i \neq k$，则交换$\boldsymbol{A}$的第i_k行与第k行，交换b_{i_k}与b_k
- 若$j_k \neq k$，则交换$\boldsymbol{A}$的第j_k列与第k列，交换d_{j_k}与d_k

$$\text{for } i = k+1 \text{ to } n \text{ do}$$

$$T = \frac{a_{ik}}{a_{kk}};\ b_i = b_i - T \times b_k$$

$$\text{for } j = k+1 \text{ to } n \text{ do}$$

$$a_{ij} = a_{ij} - T \times a_{kj}$$

④ 回代

$$Z_n = \frac{b_n}{a_{nn}}$$

$$\text{for } i = n-1 \text{ downto } 1 \text{ do}$$

$$Z_i = \frac{b_i - \sum_{j=i+1}^{n} a_{ij} Z_j}{a_{ii}}$$

⑤ $\text{for } j = 1 \text{ to } n \text{ do}$

$$x_{d_j} = Z_j$$

打印x_i, $i = 1,2,\ldots, n$

- 全主元高斯消去法的计算量
  - 乘除法和加减次数与顺序高斯消去法相同
  - 增加选主元的比较及交换方程次序所需时间

$$\sum_{k=1}^{n-1}[(n-k+1)^2-1]=\sum_{i=2}^{n}(i^2-1)\ \text{（比较次数）}$$

□ 全主元高斯消去法的矩阵解释

- 记$\left|a_{i_1,j_1}^{(1)}\right| = \max\limits_{\substack{1\le i\le n\\1\le j\le n}}\left|a_{ij}^{(1)}\right|$，主元在第$i_1$行$j_1$列，需交换第1行和第$i_1$行及相应的右端项、交换第1列和第$j_1$列

$$\begin{cases} \boldsymbol{A}^{(2)} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\boldsymbol{P}_{1,j_1} \\ \boldsymbol{b}^{(2)} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)} \end{cases}$$

- 同理可得

$$\begin{cases} \boldsymbol{A}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\boldsymbol{A}^{(n-1)}\boldsymbol{P}_{n-1,j_{n-1}} \\ \boldsymbol{b}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\boldsymbol{b}^{(n-1)} \end{cases}$$

- 由$A^{(1)}X=b^{(1)}$两边同乘 $M_{n-1}P_{n-1,i_{n-1}}\cdots M_1P_{1,i_1}$得:

$$M_{n-1}P_{n-1,i_{n-1}}\cdots M_1P_{1,i_1}A^{(1)}X = M_{n-1}P_{n-1,i_{n-1}}\cdots M_1P_{1,i_1}b^{(1)} = b^{(n)}$$

- 由 $A^{(n)}=M_{n-1}P_{n-1,i_{n-1}}\cdots M_1P_{1,i_1}A^{(1)}P_{1,j_1}\cdots P_{n-1,j_{n-1}}$
- 得 $A^{(n)}P_{n-1,j_{n-1}}^{-1}\cdots P_{1,j_1}^{-1}=M_{n-1}P_{n-1,i_{n-1}}\cdots M_1P_{1,i_1}A^{(1)}$
- 上式同乘X，有:

$$A^{(n)}P_{n-1,j_{n-1}}^{-1}\cdots P_{1,j_1}^{-1}X = M_{n-1}P_{n-1,i_{n-1}}\cdots M_1P_{1,i_1}A^{(1)}X = b^{(n)}$$

$$Z=P_{n-1,j_{n-1}}^{-1}\cdots P_{1,j_1}^{-1}X$$

$$A^{(n)}Z=b^{(n)}$$

$$\Rightarrow X=P_{1,j_1}P_{2,j_2}\cdots P_{n-1,j_{n-1}}Z$$

- 顺序高斯消去法可把方程组$A^{(1)}X=b^{(1)}$等价转化为一个上三角型方程组$A^{(n)}X=b^{(n)}$.
- $A^{(n)}=M_{n-1}M_{n-2}\cdots M_1A^{(1)}$
- $A^{(1)}=M_1^{-1}M_2^{-1}\cdots M_{n-1}^{-1}A^{(n)}$
  - $M_i^{-1}(i=1,2,\cdots,n-1)$均为下三角矩阵且对角线元素为1，则$M_1^{-1}M_2^{-1}\cdots M_{n-1}^{-1}$为对角线元素为1的下三角矩阵，记
  - $L=M_1^{-1}M_2^{-1}\cdots M_{n-1}^{-1}$
  - $U=A^{(n)}$
  - 只要方程$AX=b$的所有顺序主子式不为零，则A一定可以分解成$A=LU$

定理： 如果n阶矩阵A的所有顺序主子式不为零，则A有唯一的LU分解.

- 如果$A = LU$，则可求解方程组$AX = b$如下：
- $AX = b,\ \ A = LU$
- $LU\ X = b$
- 可记$U\ X = Y$，则

$$LY = b,\quad y_1 = b_1, y_i = b_i - \sum_{j=1}^{i-1} l_{ij} y_j\ (i = 2, \cdots, n)$$

$$UX = Y, x_n = \frac{y_n}{u_{nn}}, x_i = \frac{\left(y_i - \sum_{j=i+1}^{n} u_{ij x_j}\right)}{u_{ii}}\ (i = n-1, \cdots, 1)$$

直接LU分解法

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

- $u_{1j} = a_{1j}\ (j = 1,2,\cdots,n)$：第1行对应相等

  $l_{i1} = \frac{a_{i1}}{u_{11}}, (i = 2,3,\dots,n)$ ：由第**1**列对应相等

- 一般地，设$\boldsymbol{U}$矩阵的前$k-1$行和$\boldsymbol{L}$矩阵的前$k-1$列已经求出，则

$$a_{kj}=\sum_{m=1}^{k-1} l_{km}u_{mj}+u_{kj} \Longrightarrow$$

$$u_{kj}=a_{kj}-\sum_{m=1}^{k-1} l_{km}u_{mj} \; (j=k,k+1,\cdots,n)$$

$$a_{ik}=\sum_{m=1}^{k-1} l_{im}u_{mk}+l_{ik}u_{kk} \Longrightarrow$$

$$l_{ik}=\frac{a_{ik}-\sum_{m=1}^{k-1} l_{im}u_{mk}}{u_{kk}} \; (i=k+1,\cdots,n)$$

- 直接LU分解法的计算量（乘除法）
  - LU分解

计算$\boldsymbol{U}$ 的计算量：$\sum_{k=2}^{n}(n-k+1)\cdot(k-1)=\frac{n^3-n}{6}$

计算$\boldsymbol{L}$的计算量：$\sum_{k=1}^{n-1}(n-k)\cdot k=\frac{n^3-n}{6}$

  - 方程组求解

求解$\boldsymbol{LY}=\boldsymbol{b}$的计算量：$\sum_{i=2}^{n}(i-1)=\frac{n(n-1)}{2}$

求解$\boldsymbol{U}\,\boldsymbol{X}=\boldsymbol{Y}$的计算量：$\sum_{i=1}^{n}(n-i+1)=\frac{n(n+1)}{2}$

总的计算量：$\frac{n^3+3n^2-n}{3}$

例：用LU分解法求解方程组

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

- 列主元LU分解法
  - 直接进行LU分解时，u_{kk}可能为零，或者绝对值接近零，从而导致出现溢出，或者因以一个小的数作为除数，引起舍入误差的累积，最后引起解的失真.
  - 如果$\boldsymbol{A}$非奇异，可以采用列主元LU分解.

- 假若第$k-1$步的分解已完成，在进行第k步分解时，为避免出现小的u_{kk}作除数

$$S_i = a_{ik} - \sum_{m=1}^{k-1} l_{im} u_{mk}$$

- 选取行号i_k，使$|S_{i_k}| = \max_{k \le i \le n} |S_i|$，若$i_k \neq k$，则对调$A$的第$k$行与第$i_k$行，再求$u_{kj}(j = k, \cdots, n)$和$l_{ik}(i = k+1, \cdots, n)$

定义：正定矩阵——$n \times n$的实对称矩阵$\boldsymbol{A}$是正定的当且仅当对于所有的非零实系数向量$\boldsymbol{X}$，都有$\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X} > \boldsymbol{0}$

定理 以下陈述等价，均为正定矩阵的充分必要条件

1、对于所有的非零实系数向量$\boldsymbol{X}$，都有$\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X} > \boldsymbol{0}$；

2、A为正定矩阵当且仅当其所有特征值$\lambda_i > 0$；

3、A为正定矩阵当且仅当A的所有顺序主子式A_k大于0；

4、A为正定矩阵当且仅当所有pivots（无行交换）$d_k > 0$。

定理：设A是n阶对称矩阵，且A的所有顺序主子式均不为零，则A有如下分解：

$$A = LDL^T$$

其中L是单位下三角阵；D为对角阵，且分解唯一.

定理：n阶对称正定矩阵A一定有Cholesky分解：

$$A = L_1 L_1^T$$

当L_1（实的非奇异下三角矩阵）的对角线为正时，矩阵的Cholesky分解唯一.

- Cholesky分解计算公式
  - 设$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^T$，即

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

其中$a_{ij}=a_{ji}(i,j=1,2,\cdots,n), l_{ii}>0(i=1,2,\cdots,n)$

1. 由矩阵乘法有$a_{11} = l_{11}^2$，且$a_{i1} = l_{i1}l_{11}$，故求得

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

2. 一般地，设$\boldsymbol{L}$矩阵的前$k-1$列元素已求出，则由矩阵乘法得

$$\sum_{m=1}^{k-1} l_{km}^2 + l_{kk}^2 = a_{kk} \quad \sum_{m=1}^{k-1} l_{im}l_{km} + l_{ik}l_{kk} = a_{ik} \Longrightarrow$$

$$\begin{cases} l_{kk} = \sqrt{a_{kk} - \sum\limits_{m=1}^{k-1} l_{km}^2} \\ l_{ik} = \left(a_{ik} - \sum\limits_{m=1}^{k-1} l_{im}l_{km}\right)/l_{kk} \end{cases} \quad (i = k+1,\cdots,n; k = 2,\cdots,n)$$

- 分解公式中的每一步都有开方运算，故又称Cholesky方法为平方根法.
- $a_{kk}=\sum_{m=1}^{k-1} l_{km}^2+l_{kk}^2 \Rightarrow l_{km}^2 \le a_{kk} \le \max\limits_{1\le k\le n} a_{kk}$
  - 分解过程中元素l_{km}的平方不会超过$\boldsymbol{A}$的最大对角元，数量级不会增长，舍入误差的放大受到限制
  - 平方根法求解对称正定方程组时可以不考虑选主元的问题，是一个数值稳定算法
- 用平方根法求解$\boldsymbol{AX}=\boldsymbol{b}$所需的乘除法次数为$\frac{1}{6}(n^3+9n^2-4n)$，另外$n$次开平方运算，乘除法次数大约只是高斯消去法的一半左右.

- 为了避免平方根法的开方运算，可对$\boldsymbol{A}$作LDL^T分解，分解方法可借助与LU分解法.
  - 当$\boldsymbol{A}$的所有顺序主子式不为零且对称时，$\boldsymbol{A}$的LU分解存在且唯一，并且

$$\boldsymbol{U} = \boldsymbol{D}\boldsymbol{L}^T$$

其中$\boldsymbol{D} = diag(u_{11}, u_{22}, \cdots, u_{nn}), l_{ij} = \frac{u_{ji}}{u_{ii}}$.

  - 利用LU分解先计算出$\boldsymbol{U}$的第k行，由$\boldsymbol{U}$的第k行得到$\boldsymbol{D}$的第k个对角元素及$\boldsymbol{L}$的第k列.

$$d_i = u_{ii} \qquad l_{ik} = \frac{u_{ki}}{u_{kk}}$$

  - LDL^T分解的乘除法计算次数 与平方根法大致相同

例：用LDL^T分解法求方程组

$$\begin{bmatrix} 5 & -4 & 1 \\ -4 & 6 & -4 \\ 1 & -4 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

【略】

□ 向量范数

- 向量范数是n维欧几里德空间中长度概念的推广.

定义：设向量$\boldsymbol{X} \in \boldsymbol{R}^n$，若与$\boldsymbol{X}$对应的非负实数$||\boldsymbol{X}||$满足下面三个条件：

1. 非负性：$\forall \boldsymbol{X} \in \boldsymbol{R}^n$，有$\|\boldsymbol{X}\| \geq 0$，且$\|\boldsymbol{X}\| = 0$当且仅当$\boldsymbol{X} = 0$
2. 齐次性：$\forall \boldsymbol{X} \in \boldsymbol{R}^n$，$\alpha \in \boldsymbol{R}$有$\|\alpha\boldsymbol{X}\| = |\alpha| \cdot \|\boldsymbol{X}\|$
3. 三角不等式：$\forall \boldsymbol{X}, \boldsymbol{Y} \in \boldsymbol{R}^n$有$\|\boldsymbol{X} + \boldsymbol{Y}\| \leq \|\boldsymbol{X}\| + \|\boldsymbol{Y}\|$

则称$||\boldsymbol{X}||$为向量$\boldsymbol{X}$的范数.

- 常用范数有:
  - 向量的1范数： $\|\boldsymbol{X}\|_1 = \sum_{i=1}^{n}|x_i|$
  - 向量的2范数： $\|\boldsymbol{X}\|_2 = (\sum_{i=1}^{n}|x_i|^2)^{\frac{1}{2}}$
  - 向量的∞范数： $\|\boldsymbol{X}\|_\infty = \max_{1\le i\le n}|x_i|$
  - 向量的p范数： $\|\boldsymbol{X}\|_p = (\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}$

定义： 设$\|\cdot\|_\alpha$和$\|\cdot\|_\beta$是$\boldsymbol{R}^n$上两种范数，如果存在与$\boldsymbol{X}$无关的两个正常数C_1和C_2，使不等式

$$C_1\|\boldsymbol{X}\|_\alpha \le \|\boldsymbol{X}\|_\beta \le C_2\|\boldsymbol{X}\|_\alpha \quad (\forall \boldsymbol{X} \in \boldsymbol{R}^n)$$

成立，则称范数$\|\cdot\|_\alpha$和$\|\cdot\|_\beta$等价.

定理： 有限维空间上任何两种范数均等价.

定义：设$\{\boldsymbol{X}^{(k)}\}$是$\boldsymbol{R}^n$中的向量序列，若有向量$\boldsymbol{X}^* \in \boldsymbol{R}^n$，使

$$\lim_{k\to\infty}\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^*\|=0$$

则称$\{\boldsymbol{X}^{(k)}\}$收敛于$\boldsymbol{X}^*$，记为$\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}=\boldsymbol{X}^*$

定理：设$\{\boldsymbol{X}^{(k)}\}$是$\boldsymbol{R}^n$中的向量序列，则$\lim\limits_{k\to\infty}\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^*\|=0$当且仅当

$$\lim_{k\to\infty}x_j^{(k)}=x_j^*$$

其中$x_j^{(k)}$和x_j^*分别表示$\boldsymbol{X}^{(k)}$和$\boldsymbol{X}^*$的第j个分量.

□ 矩阵范数

定义：设矩阵$\boldsymbol{A} \in \boldsymbol{R}^{n\times n}$，若$\boldsymbol{A}$的非负实数$\|\boldsymbol{A}\|$满足下列四个条件：

1. 非负性：$\forall \boldsymbol{A} \in \boldsymbol{R}^{n\times n}$，有$\|\boldsymbol{A}\| \geq 0$，且$\|\boldsymbol{A}\| = 0$当且仅当$\mathbf{A} = 0$
2. 齐次性：$\forall \boldsymbol{A} \in \boldsymbol{R}^{n\times n}$，$\forall \alpha \in \boldsymbol{R}$有$\|\alpha \boldsymbol{A}\| = |\alpha| \cdot \|\boldsymbol{A}\|$
3. 三角不等式：$\forall \boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{R}^{n\times n}$有$\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$
4. 乘法不等式：$\forall \boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{R}^{n\times n}$有$\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|$

则$\|\boldsymbol{A}\|$称为矩阵$\boldsymbol{A}$的范数.

定义（矩阵的算子范数）：设$X \in R^n, A \in \mathrm{R}^{n\times n}$，给出一种向量范数$\|X\|_v$(如$v=1,2$或者$\infty$)，相应地定义一个矩阵的非负函数，称其为由向量$\|\cdot\|$诱导出的矩阵范数.

$$\|A\|_v = \max_{X\neq 0}\frac{\|AX\|_v}{\|X\|_v} = \max_{\|X\|=1}\|AX\|_v$$

定义：对于给定的向量范数$\|\cdot\|$和矩阵范数$\|\cdot\|$，如果$\forall A \in R^{n\times n}$和$\forall X \in R^n$，并满足

$$\|AX\| \leq \|A\| \cdot \|X\|$$

则称此矩阵范数与向量范数是相容的.

定理： 诱导矩阵范数$\|A\|$是相容的.

定理： 设A为$n \times n$阶矩阵，$\|\cdot\|$是$\boldsymbol{R}^n$中的向量范数，则

$$\|A\| = \max_{\|X\|=1} \|AX\|$$

是一种矩阵范数.

定理： 设$\{\boldsymbol{A}^{(k)}\}$是$\boldsymbol{R}^{n\times n}$中矩阵系列，$\boldsymbol{A}$是$\boldsymbol{R}^{n\times n}$中矩阵，则

$$\lim_{k\to\infty}\left\|\boldsymbol{A}^{(k)}-\boldsymbol{A}\right\|=0$$

当且仅当$\lim_{k\to\infty} a_{ij}^{(k)}=a_{ij}$，其中$a_{ij}^{(k)}$和$a_{ij}$分别表示$\boldsymbol{A}^{(k)}$和$\boldsymbol{A}$的分量.

定理：设$A = (a_{ij}) \in R^{n\times n}$，则有：

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$$

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\|_2 = [\lambda_{\max} \quad (A^H A)]^{\frac{1}{2}}$$

其中A^H是A的共轭.

- Frobenius范数：$\|A\|_F = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}}$

□ 谱半径

定义：$\boldsymbol{A} \in \boldsymbol{R}^{n\times n}$，$\boldsymbol{A}$的特征值为$\lambda_i(i=1,2,\cdots,n)$，称$\rho(\boldsymbol{A}) = \max\limits_{1\le i\le n}|\lambda_i|$为$\boldsymbol{A}$的谱半径.

定理：$\boldsymbol{A}$的谱半径不超过$\boldsymbol{A}$的任何一个矩阵范数，即 $\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|$

定理：设$\boldsymbol{A}$是任意$n\times n$矩阵，由$\boldsymbol{A}$的各次幂组成的矩阵序列 $\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2, \cdots, \boldsymbol{A}^k, \cdots$

收敛于零，即$\lim\limits_{k\to\infty} \boldsymbol{A}^k = 0$的充要条件是$\rho(\boldsymbol{A}) < 1$

方程右端误差对解的影响

- 设右端项有误差δb，相应方程组的解为$X+\delta X$

$$AX=b$$
$$A(X+\delta X)=b+\delta b$$
$$A\delta X=\delta b \text{ 或者 } \delta X=A^{-1}\delta b$$

- 假定A非奇异

$$\|\delta X\|=\|A^{-1}\delta b\|\le\|A^{-1}\|\cdot\|\delta b\|$$
$$\|AX\|=\|b\|\le\|A\|\cdot\|X\|$$
$$\frac{\|\delta X\|}{\|X\|}\le\|A\|\cdot\|A^{-1}\|\cdot\frac{\|\delta b\|}{\|b\|}$$

- $\|A\|\cdot\|A^{-1}\|$是误差的放大倍数，称为矩阵A的条件数，记作 **Cond(A)**

系数矩阵误差对解的影响

- 设系数矩阵$\boldsymbol{A}$有误差$\delta\boldsymbol{A}$，$\boldsymbol{X}+\delta\boldsymbol{X}$是系数矩阵为$\boldsymbol{A}+\delta\boldsymbol{A}$时的准确解，即

$$(\boldsymbol{A}+\delta\boldsymbol{A})(\boldsymbol{X}+\delta\boldsymbol{X})=\boldsymbol{b}$$

  - 即使矩阵$\boldsymbol{A}$非奇异，其元素有$\delta\boldsymbol{A}$扰动后，矩阵$\boldsymbol{A}+\delta\boldsymbol{A}$有可能变为奇异
  - 若$\delta\boldsymbol{A}$很小，矩阵$\boldsymbol{A}+\delta\boldsymbol{A}$仍然可以保持它的非奇异性.

引理： 设矩阵$\boldsymbol{F}$的范数小于1，即$\|\boldsymbol{F}\| < 1$，则矩阵$(\boldsymbol{I}+\boldsymbol{F})$非奇异，且

$$\|(\boldsymbol{I}+\boldsymbol{F})^{-1}\| \leq \frac{1}{1-\|\boldsymbol{F}\|}$$

- $\|(I+F)X\| = \|X+FX\| \geq \|X\| - \|FX\| \geq \|X\| - \|F\|\|X\| \geq (1 - \|F\|)\|X\|$
- 因为$1 - \|F\| > 0$，所以对$X \neq 0, (1 - \|F\|)\|X\| > 0$，$\Rightarrow \forall\ X \neq 0$,有 $\|(I+F)X\| > 0$
- 从而，方程组$(I + F)X = 0$只有零解，因此矩阵$(I + F)$非奇异。
- 又 $1= \|I\| = \|(I+F)(I+F)^{-1}\| = \|(I+F)^{-1} + F(I+F)^{-1}\|$
- $\geq \|(I+F)^{-1}\| - \|F\|\|(I+F)^{-1}\| = \|(I+F)^{-1}\|(1-\|F\|) > 0$
- 从而必有

$$\|(I+F)^{-1}\| \leq \frac{1}{1-\|F\|}$$

- 考虑矩阵$A+\delta A$非奇异条件
  - $A+\delta A=A(I+A^{-1}\delta A)$
  - 如果$\|A^{-1}\delta A\|<1$，则矩阵$A+\delta A$非奇异

$$\left\|(I+A^{-1}\delta A)^{-1}\right\|\le\frac{1}{1-\|A^{-1}\delta A\|}$$

  - 由$(A+\delta A)(X+\delta X)=b$可得

$$\delta X=(A+\delta A)^{-1}(-\delta A)X$$

  - $(A+\delta A)(X+\delta X) = b \Rightarrow (X+\delta X) = (A+\delta A)^{-1}b$
  - $\Rightarrow \delta X = (A+\delta A)^{-1}b - X \Rightarrow \delta X = (A+\delta A)^{-1}(b - (A+\delta A)X)$
  - 因为$AX = b$，所以 $\delta X = (A+\delta A)^{-1}(-\delta A)X$

$$\frac{||\delta X||}{||X||} \le \left\| (A+\delta A)^{-1}\delta A \right\| = ||(A(I+A^{-1}\delta A))^{-1}\delta A||$$

$$= ||(I+A^{-1}\delta A)^{-1}A^{-1}\delta A||$$

$$\le ||(I+A^{-1}\delta A)^{-1}||\,||A^{-1}\delta A||$$

$$\le \frac{||A^{-1}\delta A||}{1-||A^{-1}\delta A||} \le \frac{||A^{-1}||\,||\delta A||}{1-||A^{-1}||\,||\delta A||}$$

$$= \frac{cond(A)}{1-cond(A)\dfrac{||\delta A||}{||A||}}\frac{||\delta A||}{||A||}$$

- 若$Cond(\boldsymbol{A})\ \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}$很小，则$Cond(\boldsymbol{A})$ 仍表示相对误差的近似放大率.
- $Cond(\boldsymbol{A})$反映线性方程组的解对初始数据误差的灵敏度
- 如果值很大，即使对很小的初始误差$\delta\boldsymbol{b}$或$\delta\boldsymbol{A}$，也会产生较大的相对误差，破坏解的精确度.

作业

- 课后第1、2题