第 7 章 解线性方程组的直接方法

姚遥

南京大学智能科学与技术学院

本章目录

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

直接方法 | Gauss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

1. $\boldsymbol{A} = \boldsymbol{LU}$,　分解 $\boldsymbol{LU}$
2. $\boldsymbol{Ly} = \boldsymbol{b}$, 求 $\boldsymbol{y}$;
3. $\boldsymbol{Ux} = \boldsymbol{y}$, 求 $\boldsymbol{x}$.

直接方法 | 不选主元的三角分解法

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

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

可推导

- L, 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

直接方法 | 不选主元的三角分解法

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

- 考虑 $\boldsymbol{A}$ 的第 1 行元素

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

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

- 考虑 $\boldsymbol{A}$ 的第 1 列元素

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

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

直接方法 | 不选主元的三角分解法

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

- 考虑 $\boldsymbol{A}$ 的第 r 行元素：设已有 $\boldsymbol{U}$ 第 1 到 $r-1$ 行元素与 $\boldsymbol{L}$ 的 1 到 $r-1$ 列元素

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

可得 $\boldsymbol{U}$ 的第 r 行元素

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

直接方法 | 不选主元的三角分解法

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

- 考虑 $\boldsymbol{A}$ 的第 r 列元素：设已有 $\boldsymbol{U}$ 第 1 到 $r-1$ 行元素与 $\boldsymbol{L}$ 的 1 到 $r-1$ 列元素

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

可得 $\boldsymbol{L}$ 的第 r 列元素

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

直接方法 | 直接三角分解法解方程组：算法实现

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

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

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

For $r = 2,3,\cdots,n$，重复步骤2、3

2. 计算 $\boldsymbol{U}$ 的第 r 行元素

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

3. 计算 $\boldsymbol{L}$ 的第 r 列元素

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

直接方法 | 直接三角分解法解方程组：算法实现

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 $\boldsymbol{Ux} = \boldsymbol{y}$

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

直接方法 | 用直接三角分解法解 $\boldsymbol{Ax=b}$

例7.5 用直接三角分解法解

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

- 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

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

例7.5 用直接三角分解法解

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

- 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 5 求解方程

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

得

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

直接方法 | 不选主元的三角分解法：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（省内存）

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

- 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”、“Kahan Summation”等方法提高精度

- 直接分解法大约需要 $n^3/3$ 次乘除运算，和Gauss消去法的计算量基本相同

直接方法 | 不选主元的三角分解法：讨论

- 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（步骤 4、5）

- 矩阵A的分解步骤 2（公式 7.4.2）、步骤 3（公式 7.4.3）又称为Doolittle分解公式

从直接三角分解公式看出：1）当 $u_{rr}=0$ 时计算将中断；2）当 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，可通过交换 $\boldsymbol{A}$ 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

- 设第 $r-1$ 步分解已完成，这时有

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

直接方法 | 选主元素的三角分解法

- 第 r 步分解需用到式(7.4.2)及式(7.4.3)

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

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

- 引进量

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

- 显然

$$u_{rr} = s_r$$

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

直接方法 | 选主元素的三角分解法

- 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

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

直接方法 | 选主元素的三角分解法：算法实现

算法 4：用 $PA = I_{n-1,i_{n-1}} \cdots I_{1i_1}A$ 的三角分解冲掉 A，用 idx(n) 记录主行，解 x 存放在 b 内，r 表示消元次数

For r in $[1,2,\cdots,n]$, do 1, 2, 3, 4

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

2. 选主元： $\left|s_{i_r}\right| = \max\limits_{r \le i \le n} |s_i|$, $\operatorname{idx}(r) \leftarrow i_r$

3. 交换 A 的 r 行与 i_r 行元素： $a_{ri} \leftrightarrow a_{i_r i} \quad (i = 1,2,\cdots,n)$

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

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

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

直接方法 | 选主元素的三角分解法：算法实现

上述计算过程完成后就实现了 $\boldsymbol{PA}$ 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 $\mathrm{idx}(n)$ 最后记录可知

5. For i in $[1,2,\cdots,n]$, do

   1) $t \leftarrow \mathrm{idx}(i)$

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

   3) 继续循环

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

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

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

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

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

直接方法 | 选主元素的三角分解法：矩阵求逆

利用算法4的结果，可计算 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 PA 的三角分解计算 A^{-1} 步骤

- 计算上三角阵的逆矩阵 $\boldsymbol{U}^{-1}$
- 计算 $\boldsymbol{U}^{-1}\boldsymbol{L}^{-1}$
- 交换 $\boldsymbol{U}^{-1}\boldsymbol{L}^{-1}$ 列（利用 $\mathrm{idx}(n)$ 最后记录）
- 大约需要 n^3 次乘法运算

直接方法 | 平方根法

考虑系数矩阵是对称正定的：平方根法是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

先考虑 A 为对称矩阵，且 A 的所有顺序主子式均不为零，由定理7.3知，A 可唯一分解为

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

直接方法 | 对称矩阵的三角分解

为了利用 A 的对称性，将 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： $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{U}_0$
- 根据对称性： $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{U}_0^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}})$
- 根据定理7.3中分解的唯一性，可知

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

直接方法 | 对称矩阵的三角分解

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

直接方法 | 对称正定矩阵的三角分解

重温7.2节引理（需要用到证明中间结论）：

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 $D_i \neq 0\ \ (i = 1, 2, \cdots, k)$

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

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

直接方法 | 对称正定矩阵的三角分解

- 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 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^{T}$ ，当限定 $\boldsymbol{L}$ 的对角元素为正时，这种分解是唯一的

直接方法 | 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 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 $l_{ii}>0\ (i=1,2,\cdots,n)$

直接方法 | 对称正定矩阵的直接三角分解

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

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

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

其中第二个等号可根据 $l_{jk} = 0\ (j < k)$ 推导

- 依据上式，可以逐列计算矩阵 $\boldsymbol{L}$ 的元素

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

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

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

直接方法 | 平方根法：算法实现

解对称正定方程组 $Ax = b$ 的平方根法计算公式

首先对 A 进行分解

For j in $[1,2,\cdots,n]$, do 1, 2

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}

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

直接方法 | 平方根法：算法实现

解对称正定方程组 $Ax=b$ 的平方根法计算公式

得到 $A=LL^T$后，继续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 $\boldsymbol{L}^T\boldsymbol{x}=\boldsymbol{y}$

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

直接方法 | 平方根法：数值稳定性

- 根据前的推导

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

- 因此

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

- 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

直接方法 | 平方根法：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 $\boldsymbol{\frac{n^3}{6}}$ 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 $\boldsymbol{\frac{n(n+1)}{2}}$ 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

直接方法 | 改进的平方根法：避免开方

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

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}=\begin{bmatrix}1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1\end{bmatrix}\begin{bmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{bmatrix}\begin{bmatrix}1 & l_{21} & \cdots & l_{n1} \\ & 1 & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1\end{bmatrix}$$

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

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

其中第二个等号可根据 $l_{jk}=0\ (j<k)$、$l_{jj}=1$ 推导

直接方法 | 改进的平方根法：避免开方

- 依据上式，可以逐行计算矩阵 $\boldsymbol{L}$ 的元素

$$\boldsymbol{L}=\begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix}$$

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

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

- 同样可得到 $\boldsymbol{D}$ 的对角元素公式

$$\boldsymbol{D}=\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$$

$$d_i=a_{ii}-\sum_{k=1}^{i-1} l_{ik}^2 d_k$$

直接方法 | 改进的平方根法：算法实现

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 L,T 元素的公式

For i in $[1,2,\cdots,n]$, do 1, 2, 3

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

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

直接方法 | 改进的平方根法：算法实现

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

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

- 对称正定矩阵 $\boldsymbol{A}$ 按 $\boldsymbol{LDL}^T$ 及 $\boldsymbol{LL}^T$ 分解计算量差不多，但 $\boldsymbol{LDL}^T$ 分解无需开方计算

直接方法 | 改进的平方根法：算法实现

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

4. 求解 $\boldsymbol{Ly=b}$

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

5. 求解$\boldsymbol{DL^{T}x=y}$

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

直接方法 | 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

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

- 对角占优条件1： $|b_1| > |c_1| > 0$
- 对角占优条件2： $|b_i| \geq |a_i| + |c_i|, \quad a_i c_i \neq 0 \quad (i = 2, \cdots, n-1)$
- 对角占优条件3： $|b_n| > |a_n| > 0$

直接方法 | 对角占优的三对角矩阵：三角分解

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

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

- $\boldsymbol{L}$ 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

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

直接方法 | 对角占优的三对角矩阵：三角分解

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

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

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

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

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

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

直接方法 | 对角占优的三对角矩阵：三角分解分析

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

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

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

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

直接方法 | 对角占优的三对角矩阵：三角分解分析

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

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

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

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

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

直接方法 | 解对角占优的三对角方程组：追赶法

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

1. 计算 $\{\beta_i\}$ 的递推公式

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

2. 解 $\boldsymbol{Ly} = \boldsymbol{f}$

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

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

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

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

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

直接方法 | 解对角占优的三对角方程组：追赶法

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

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

$$x_n=y_n$$

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

将计算方程组的解 $x_n\rightarrow x_{n-1}\rightarrow\cdots\rightarrow x_1$ 的过程称为赶的过程

定理7.9（具体见教科书描述）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

直接方法 | 解对角占优的三对角方程组：追赶法

追赶法：

- 追赶法公式实际上就是把Gauss消去法/LU分解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 $\{x_i\}$

本章目录

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

直接方法 | 向量和矩阵的范数

为了研究线性方程组近似解的误差估计和迭代法的收敛性，需要对 $\mathbf{R}^n$ 中向量（或 $\mathbf{R}^{n\times n}$ 中矩阵）的“大小”引进某种度量——向量（或矩阵）范数的概念

- 向量范数是三维Euclid空间中向量长度概念的推广，在数值分析中起着重要作用

首先将向量长度概念推广到 $\mathbf{R}^n$

- 注意到本节的内容同样可以拓展到 n 维复向量空间 $\mathbf{C}^n$

直接方法 | 数量积与Euclid范数

定义7.1 设

$$\boldsymbol{x} = (x_1, x_2, \dots, x_n)^{\mathrm{T}}, \ \boldsymbol{y} = (y_1, y_2, \dots, y_n)^{\mathrm{T}} \in \mathbf{R}^n$$

向量 $\boldsymbol{x}$，$\boldsymbol{y}$ 的数量积（内积）定义为：

$$(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} = \sum_{i=1}^{n} x_i y_i$$

向量 $\boldsymbol{x}$ 的Euclid范数定义为：

$$\|\boldsymbol{x}\|_2 = (\boldsymbol{x}, \boldsymbol{x})^{\frac{1}{2}} = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} \geq 0$$

直接方法 | 数量积与Euclid范数

定理7.10 设 $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{R}^n$ ，则

- $(\boldsymbol{x}, \boldsymbol{x}) = 0$ ，当且仅当 $\boldsymbol{x} = \mathbf{0}$ 时成立
- $(\alpha \boldsymbol{x}, \boldsymbol{y}) = \alpha(\boldsymbol{x}, \boldsymbol{y})$ ，α 为实数
- $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})$
- $(\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}) = (\boldsymbol{x}_1, \boldsymbol{y}) + (\boldsymbol{x}_2, \boldsymbol{y})$
- （Cauchy−Schwarz不等式）$|(\boldsymbol{x}, \boldsymbol{y})| \le \|\boldsymbol{x}\|_2 \cdot \|\boldsymbol{y}\|_2$ ，等式当且仅当 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 线性相关时成立
- （三角不等式）$\|\boldsymbol{x} + \boldsymbol{y}\|_2 \le \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2$

直接方法 | 向量范数

除了Euclid范数之外，还可以用其他办法来度量 R^n 中向量的“大小”

- 例如对于$\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}\in\mathbf{R}^n$，用一个 $\boldsymbol{x}$ 的函数 $N(\boldsymbol{x})=\max\limits_{i=1,2}|x_i|$ 来度量 $\boldsymbol{x}$ 的“大小”，而且这种度量方法计算起来比Euclid范数更方便
- 在许多应用中，对度量 $\boldsymbol{x}$ 的“大小”的函数 $N(\boldsymbol{x})$ 都要求是正定的、齐次的且满足三角不等式

向量范数的一般定义？

直接方法 | 向量范数

定义7.2（向量的范数）如果向量 $x \in R^n$ 的某个实值函数 $N(x) = \|x\|$，满足条件：

- 正定性：$\|\boldsymbol{x}\| \geq 0$， $\|\boldsymbol{x}\| = 0$ 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$
- 齐次性：$\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$， $\forall\alpha \in \mathbf{R}$
- 三角不等式：$\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

则称 $N(\boldsymbol{x})$ 是 $\mathbf{R}^n$ 上的一个向量范数（或模）

- 由条件3可以推出不等式：$\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$

直接方法 | 常用范数

向量的 ∞-范数（最大范数）

$$\|\boldsymbol{x}\|_\infty = \max_{1\le i\le n} |x_i|$$

向量的 1-范数

$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

向量的 2-范数

$$\|\boldsymbol{x}\|_2 = (\boldsymbol{x},\boldsymbol{x})^{\frac{1}{2}} = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{\frac{1}{2}}$$

向量的 p-范数，$p \in [1,\infty)$

$$\|\boldsymbol{x}\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

直接方法 | 范数的连续性

定义7.3 设 $\{\boldsymbol{x}^{(k)}\}$ 为 $\mathbf{R}^n$ 中一向量序列，$\boldsymbol{x}^* \in \mathbf{R}^n$，记$\boldsymbol{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^T$，$\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_n^*)^T$。如果$\lim\limits_{k\to\infty} x_i^{(k)} = x_i^*$，$i = 1,2,\cdots,n$，则称 $\boldsymbol{x}^{(k)}$ 收敛于向量 $\boldsymbol{x}^*$，记作

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

定理7.11（$N(\boldsymbol{x})$ 的连续性） 设非负函数 $N(\boldsymbol{x}) = \|\boldsymbol{x}\|$ 为 $\mathbf{R}^n$ 上任一向量范数，则 $N(\boldsymbol{x})$ 是 $\boldsymbol{x}$ 分量 $x_1, x_2, \ldots, x_n$ 的连续函数

直接方法 | 范数的等价性

定理7.12 （向量范数的等价性）设 $\|\boldsymbol{x}\|_s$, $\|\boldsymbol{x}\|_t$ 为 R^n 上向量的任意两种范数，则存在常数 $c_1, c_2 > 0$，使得

$$c_1\|\boldsymbol{x}\|_s \le \|\boldsymbol{x}\|_t \le c_2\|\boldsymbol{x}\|_s, \qquad \forall \boldsymbol{x} \in \mathbf{R}^n$$

- 不能推广到无穷维空间
- 从上述定理可知：如果在一种范数意义下向量序列收敛，则在任何一种范数意义下该向量序列亦收敛

定理7.13 $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$ 当且仅当 $k \to \infty$ 时，$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\| \to 0$，其中 $\|\cdot\|$ 为向量的任一种范数

直接方法 | 矩阵范数

将向量范数的概念推广到矩阵上

- 由向量的 2-范数可以得到 $\mathbf{R}^{n\times n}$ 中矩阵的一种范数

$$F(\boldsymbol{A}) = \|\boldsymbol{A}\|_F = \left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}$$

称为 $\boldsymbol{A}$ 的Frobenius范数

- $\|\boldsymbol{A}\|_F$ 显然满足正定性、齐次性及三角不等式

直接方法 | 矩阵范数

定义7.4（矩阵范数）如果矩阵 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$ 的某个实值函数 $N(\boldsymbol{A}) = \|\boldsymbol{A}\|$ 满足条件：

- 正定性：$\|\boldsymbol{A}\| \geq 0$，$\|\boldsymbol{A}\| = 0$ 当且仅当 $\boldsymbol{A} = \mathbf{0}$
- 齐次性：$\|c\boldsymbol{A}\| = |c|\|\boldsymbol{A}\|$，$c$ 为实数
- 三角不等式：$\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$
- $\|\boldsymbol{AB}\| \leq \|\boldsymbol{A}\| \, \|\boldsymbol{B}\|$

则称 $N(\boldsymbol{A})$ 是 $\mathbf{R}^{n\times n}$ 上的一个矩阵范数（或模）

- 最后的条件未必需要
- 上一页定义的 $F(\boldsymbol{A}) = \|\boldsymbol{A}\|_F$ 就是 $\mathbf{R}^{n\times n}$ 上的一个矩阵范数

直接方法 | 矩阵的算子范数

在大多数与估计有关的问题中，矩阵和向量会同时参与讨论，所以希望引进一种矩阵的范数，它是和向量范数相联系而且相容，即

$$\|\boldsymbol{A}\boldsymbol{x}\| \le \|\boldsymbol{A}\| \, \|\boldsymbol{x}\|$$

对于任意向量 $\boldsymbol{x} \in \mathbf{R}^n$ 及 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$ 都成立

定义7.5 （矩阵的算子范数）设 $\boldsymbol{x} \in \mathbf{R}^n$ 及 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$，给出一种向量范数 $\|\boldsymbol{x}\|_v$（如 $v = 1, 2$ 或 ∞），定义 $\mathbf{R}^{n\times n}$ 上的矩阵范数

$$\|\boldsymbol{A}\|_v = \max_{\boldsymbol{x}\ne\mathbf{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_v}{\|\boldsymbol{x}\|_v}$$

直接方法 | 矩阵的算子范数

$$\|\boldsymbol{A}\|_v = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_v}{\|\boldsymbol{x}\|_v}$$

- 可验证 $\|\boldsymbol{A}\|_v$ 满足定义7.4，所以 $\|\boldsymbol{A}\|_v$ 是 $\mathbf{R}^{n\times n}$ 上的一个矩阵范数，且有如下定理

定理7.14 设 $\|\boldsymbol{x}\|_v$ 是 $\mathbf{R}^n$上一个向量范数，则 $\|\boldsymbol{A}\|_v$ 是 $\mathbf{R}^{n\times n}$ 上的矩阵范数，且满足相容条件

$$\|\boldsymbol{A}\boldsymbol{x}\|_v \leq \|\boldsymbol{A}\|_v \|\boldsymbol{x}\|_v$$

- 矩阵范数 $\|\boldsymbol{A}\|_v$ 依赖于向量范数 $\|\boldsymbol{x}\|_v$ 的具体含义

直接方法 | 常用的矩阵范数

定理7.15 设 $\boldsymbol{x} \in \mathbf{R}^n$，$\boldsymbol{A} \in \mathbf{R}^{n\times n}$，则

- 行范数：$\|\boldsymbol{A}\|_\infty = \max\limits_{1\le i\le n} \sum_{j=1}^{n}|a_{ij}|$
- 列范数：$\|\boldsymbol{A}\|_1 = \max\limits_{1\le j\le n} \sum_{i=1}^{n}|a_{ij}|$
- 2−范数：$\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}$，其中 $\lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$表示 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的最大特征值

矩阵的 2-范数 $\|\boldsymbol{A}\|_2$ 在计算上不方便，但是矩阵的2-范数具有许多好的性质，它在理论上是有用的

直接方法 | 特征值与范数

定义7.6 设 $\boldsymbol{A} \in \boldsymbol{R}^{n\times n}$ 的特征值为 λ_i, $i = 1,2, \ldots, n$，称 $\rho(\boldsymbol{A}) = \max\limits_{1\le i\le n}|\lambda_i|$ 为 $\boldsymbol{A}$ 的谱半径

定理7.16（特征值上界） 设 $\boldsymbol{A} \in R^{n\times n}$，则 $\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|$，即 $\boldsymbol{A}$ 的谱半径不超过 $\boldsymbol{A}$ 的任何一种算子范数（对于 $\|\boldsymbol{A}\|_F$ 亦成立）

定理7.17 如果 $A \in R^{n\times n}$ 为对称矩阵，则 $\|A\|_2 = \rho(A)$

直接方法 | 特征值与范数

定理7.18 如果 $\|\boldsymbol{B}\| < 1$，则 $\boldsymbol{I} \pm \boldsymbol{B}$ 为非奇异矩阵，且 $\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\| \leq \frac{1}{1-\|\boldsymbol{B}\|}$，$\|\cdot\|$ 为算子范数

反证法

- 若 $\boldsymbol{I} \pm \boldsymbol{B}$ 是奇异矩阵，则行线性相关 $(\boldsymbol{I} \pm \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$ 有非零解，即存在 $\boldsymbol{x}_0 \neq \boldsymbol{0}$ 使

$$\boldsymbol{B}\boldsymbol{x}_0 = \boldsymbol{x}_0$$

- 注意到

$$\|\boldsymbol{B}\| = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{B}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \geq \frac{\|\boldsymbol{B}\boldsymbol{x}_0\|}{\|\boldsymbol{x}_0\|} = 1$$

与假设矛盾，因此 $\boldsymbol{I} \pm \boldsymbol{B}$ 为非奇异矩阵

直接方法 | 特征值与范数

定理7.18 如果 $\|\boldsymbol{B}\| < 1$，则 $\boldsymbol{I} \pm \boldsymbol{B}$ 为非奇异矩阵，且 $\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\| \le \frac{1}{1-\|\boldsymbol{B}\|}$，$\|\cdot\|$ 为算子范数

- 由 $(\boldsymbol{I} \pm \boldsymbol{B})(\boldsymbol{I} \pm \boldsymbol{B})^{-1} = \boldsymbol{I}$，有

$$(\boldsymbol{I} \pm \boldsymbol{B})^{-1} = \boldsymbol{I} \mp \boldsymbol{B}(\boldsymbol{I} \pm \boldsymbol{B})^{-1}$$

- 可得

$$\Rightarrow \|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\| \le 1 + \|\boldsymbol{B}\|\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\| \quad \Rightarrow \|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\| \le \frac{1}{1-\|\boldsymbol{B}\|}$$

本章目录

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

直接方法 | 微小误差对解的影响

考虑线性方程组 $Ax = b$，其中 A 为非奇异矩阵，x 为方程组的精确解

- 由于 $\boldsymbol{A}$ 或 $\boldsymbol{b}$ 中的元素是测量得到的，或者是计算的结果，所以 $\boldsymbol{A}$ 或 $\boldsymbol{b}$ 常带有某些观测误差或者包含有舍入误差

处理实际问题 $(A + \delta A)x = b + \delta b$ 时，通常要考虑 A 或 b 的微小误差对解的影响

- 考虑估计 $\boldsymbol{x} - \boldsymbol{y}$，其中

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

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

直接方法 | 微小误差对解的影响

例7.8 设有方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，其精确解为 $\boldsymbol{x}=(2,0)^T$

$$\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}2\\ 2\end{bmatrix}$$

- 现在考虑常数项的微小变化对方程组解的影响，即考察

$$\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}\begin{bmatrix}y_1\\ y_2\end{bmatrix}=\begin{bmatrix}2\\ 2.0001\end{bmatrix}$$

- 可描述为 $\boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{b}+\delta\boldsymbol{b}$，$\delta\boldsymbol{b}=(0,0.0001)^{\mathrm{T}}$，易得

$$\boldsymbol{x}+\delta\boldsymbol{x}=(1,1)^{\mathrm{T}}$$

- 可以看到，常数项$\boldsymbol{b}$的第二个分量只有1/10000的 微小变化，方程组的解却变化很大，这样的方程组称为病态方程组

定义7.7 设如果矩阵 $\boldsymbol{A}$ 或常数项 $\boldsymbol{b}$ 的微小变化，引起方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 解的巨大变化，则称此方程组为病态方程组，矩阵 $\boldsymbol{A}$ 称为病态矩阵（相对于方程组而言），否则方程组为良态方程组，矩阵 $\boldsymbol{A}$ 称为良态矩阵

应该注意，矩阵的病态性质是矩阵本身的特性，我们希望可以定量地刻画矩阵的病态程度，且无需通过解具体的方程组

直接方法 | 微小误差影响的定量分析

设方程组 $Ax=b$，其中 A 为非奇异阵，x 为方程组的精确解

1. 设 $\boldsymbol{A}$ 是精确的，$\boldsymbol{b}$ 有微小误差 $\delta\boldsymbol{b}$，解为 $\boldsymbol{x}+\boldsymbol{\delta x}$

- 由 $\boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{b}+\delta\boldsymbol{b}$

$$\Rightarrow \boldsymbol{A}\delta\boldsymbol{x}=\delta\boldsymbol{b}$$

$$\Rightarrow \delta\boldsymbol{x}=\boldsymbol{A}^{-1}\delta\boldsymbol{b}$$

$$\Rightarrow \|\delta\boldsymbol{x}\|\le\|\boldsymbol{A}^{-1}\delta\boldsymbol{b}\|\le\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{b}\|$$

- 由 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$

$$\Rightarrow \|\boldsymbol{b}\|=\|\boldsymbol{A}\boldsymbol{x}\|\le\|\boldsymbol{A}\|\|\boldsymbol{x}\|$$

$$\Rightarrow \frac{1}{\|\boldsymbol{x}\|}\le\frac{\|\boldsymbol{A}\|}{\|\boldsymbol{b}\|}$$ （假设$\boldsymbol{b}\neq\boldsymbol{0}$） $$\Rightarrow \frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\le\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}$$

直接方法 | 微小误差影响的定量分析

定理7.19 设 A 为非奇异阵，$\boldsymbol{Ax}=\boldsymbol{b}\neq 0$，且 $A(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{b}+\delta\boldsymbol{b}$，则

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\leq \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}$$

- 常数项 $\boldsymbol{b}$ 的相对误差在解中可能放大 $\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$ 倍

2. 设 $\boldsymbol{b}$ 是精确的，$\boldsymbol{A}$ 有微小误差 $\delta\boldsymbol{A}$，解为 $\boldsymbol{x}+\delta\boldsymbol{x}$

- 由 $(\boldsymbol{A}+\delta\boldsymbol{A})(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{b}$， $\boldsymbol{Ax}=\boldsymbol{b}$ 可得

$$\Rightarrow \boldsymbol{Ax}+(\delta\boldsymbol{A})\boldsymbol{x}+(\boldsymbol{A}+\delta\boldsymbol{A})\delta\boldsymbol{x}=\boldsymbol{b}$$

$$\Rightarrow (\boldsymbol{A}+\delta\boldsymbol{A})\delta\boldsymbol{x}=-(\delta\boldsymbol{A})\boldsymbol{x}$$

直接方法 | 微小误差影响的定量分析

- 如果 δA 不受限制，$\boldsymbol{A}+\delta\boldsymbol{A}$ 可能奇异，注意到

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

- 由定理7.18知，当 $\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}\|<1$ 时，$(\boldsymbol{I}+\boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}$ 存在

- 可得 $\delta\boldsymbol{x}=-(\boldsymbol{I}+\boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\boldsymbol{A}^{-1}(\delta\boldsymbol{A})\boldsymbol{x}$

$$\Rightarrow \|\delta\boldsymbol{x}\|\le\|(\boldsymbol{I}+\boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\|\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\|\|\boldsymbol{x}\|$$

$$\le\frac{\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\|\|\boldsymbol{x}\|}{1-\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}\|}\le\frac{\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\|\|\boldsymbol{x}\|}{1-\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\|}\qquad(\ \|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\|<1\ )$$

直接方法 | 微小误差影响的定量分析：条件数

定理7.20 设 $\boldsymbol{A}$ 为非奇异阵，$\boldsymbol{Ax} = \boldsymbol{b} \neq 0$，且 $(\boldsymbol{A} + \boldsymbol{\delta A})(\boldsymbol{x} + \boldsymbol{\delta x}) = \boldsymbol{b}$，如果 $\|\boldsymbol{A}^{-1}\|\|\boldsymbol{\delta A}\| < 1$

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \frac{\|\boldsymbol{A}^{-1}\|\|\delta \boldsymbol{A}\|}{1 - \|\boldsymbol{A}^{-1}\|\|\delta \boldsymbol{A}\|} = \frac{\|\boldsymbol{A}^{-1}\|\,\|\boldsymbol{A}\|\,\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}}{1 - \|\boldsymbol{A}^{-1}\|\,\|\boldsymbol{A}\|\,\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}}$$

- 若 $\delta\boldsymbol{A}$ 充分小且 $\|\boldsymbol{A}^{-1}\|\|\delta \boldsymbol{A}\| < 1$，则矩阵 $\boldsymbol{A}$ 的相对误差 $\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}$ 在解中（$\boldsymbol{x}$ 的相对误差 ）可能放大$\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$倍
- 量 $\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$ 实际上刻画了解对原始数据变化的灵敏度，即刻画了方程组的病态程度

直接方法 | 微小误差影响的定量分析：条件数

定义7.8 设 $\boldsymbol{A}$ 为非奇异矩阵，称 $cond(\boldsymbol{A})_v = \|\boldsymbol{A}^{-1}\|_v \|\boldsymbol{A}\|_v$（$v = 1,2$或$\infty$）为矩阵 $\boldsymbol{A}$ 的条件数

- 矩阵的条件数与范数有关
- 当 $\boldsymbol{A}$ 的条件数相对地大时，即 $\mathrm{cond}(\boldsymbol{A}) \gg 1$ 时，式 $\boldsymbol{Ax} = \boldsymbol{b}$ 是病态的（即 $\boldsymbol{A}$ 是病态矩阵，或者说 $\boldsymbol{A}$ 是坏条件的）
- 当 $\boldsymbol{A}$ 的条件数相对地小时，式 $\boldsymbol{Ax} = \boldsymbol{b}$ 是良态的（或者说 $\boldsymbol{A}$ 是好条件的）
- $\boldsymbol{A}$ 的条件数愈大，方程组病态程度愈严重，也就愈难得到方程组比较准确的解

直接方法丨微小误差影响的定量分析：常用条件数

1. $\mathrm{cond}(\boldsymbol{A})_{\infty} = \|\boldsymbol{A}^{-1}\|_{\infty}\|\boldsymbol{A}\|_{\infty}$

2. $\boldsymbol{A}$ 的谱条件数

$$\mathrm{cond}(\boldsymbol{A})_{2} = \|\boldsymbol{A}^{-1}\|_{2}\|\boldsymbol{A}\|_{2} = \sqrt{\frac{\lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}{\lambda_{\min}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}}$$

- 当 $\boldsymbol{A}$ 为对称阵时，$\mathrm{cond}(\boldsymbol{A})_{2} = \frac{|\lambda_1|}{|\lambda_n|}$，其中 λ_1 和 λ_n 为 $\boldsymbol{A}$ 的绝对值最大的和绝对值最小的特征值
- 这是最常见的条件数

直接方法 | 微小误差影响的定量分析：条件数性质

- 任何非奇异矩阵 $\boldsymbol{A}$，都有 $\mathrm{cond}(\boldsymbol{A})_v \geq 1$

$$\mathrm{cond}(\boldsymbol{A})_v = \|\boldsymbol{A}^{-1}\|_v \|\boldsymbol{A}\|_v \geq \|\boldsymbol{A}^{-1}\boldsymbol{A}\|_v = \|\boldsymbol{I}\|_v = 1$$

- $\boldsymbol{A}$ 为非奇异矩阵，c 为不等于零的常数，则

$$\mathrm{cond}(c\boldsymbol{A})_v = \mathrm{cond}(\boldsymbol{A})_v$$

- 如果 $\boldsymbol{A}$ 为正交矩阵，则 $\mathrm{cond}(\boldsymbol{A})_2 = 1$；
- 如果 $\boldsymbol{A}$ 为非奇异矩阵，$\boldsymbol{R}$ 为正交矩阵，则

$$\mathrm{cond}(\boldsymbol{R}\boldsymbol{A})_2 = \mathrm{cond}(\boldsymbol{A}\boldsymbol{R})_2 = \mathrm{cond}(\boldsymbol{A})_2$$

直接方法 | 微小误差影响的定量分析：例子

例7.9 已知Hilbert矩阵 $\boldsymbol{H}_n$，计算 $\boldsymbol{H}_3$ 的条件数

$$\boldsymbol{H}_n=\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}$$

- 先求逆

$$\boldsymbol{H}_3=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}, \qquad \boldsymbol{H}_3^{-1}=\begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$$

直接方法 | 微小误差影响的定量分析：例子

例7.9 已知Hilbert矩阵 $\boldsymbol{H}_n$ ，计算 $\boldsymbol{H}_3$ 的条件数

- 计算 $\boldsymbol{H}_3$ 条件数 $\mathrm{cond}(\boldsymbol{H}_3)_\infty$

$$\|\boldsymbol{H}_3\|_\infty = \frac{11}{6}, \qquad \left\|\boldsymbol{H}_3^{-1}\right\|_\infty = 408, \qquad \mathrm{cond}(\boldsymbol{H}_3)_\infty = 748$$

- 同样可计算 $\mathrm{cond}(\boldsymbol{H}_6)_\infty = 2.9\times 10^6$ ，一般当 n 越大时，$\boldsymbol{H}_n$ 的病态越严重

考虑方程组

$$\boldsymbol{H}_3\boldsymbol{x} = \boldsymbol{b} = \left(\frac{11}{6}, \frac{13}{12}, \frac{47}{60}\right)^{\mathrm{T}}$$

- 精确解为 $\boldsymbol{x} = (1,1,1)^{\mathrm{T}}$

直接方法 | 微小误差影响的定量分析：例子

例7.9 已知Hilbert矩阵 $\boldsymbol{H}_n$，计算 $\boldsymbol{H}_3$ 的条件数

- 设 $\boldsymbol{H}_3$ 及 $\boldsymbol{b}$ 有微小误差（取三位有效数字）

$$\begin{bmatrix} 1.00 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{bmatrix} \begin{bmatrix} x_1 + \delta x_1 \\ x_2 + \delta x_2 \\ x_3 + \delta x_3 \end{bmatrix} = \begin{bmatrix} 1.83 \\ 1.08 \\ 0.783 \end{bmatrix} \qquad (7.6.8)$$

简记作 $(\boldsymbol{H}_3 + \delta \boldsymbol{H}_3)(\boldsymbol{x} + \delta \boldsymbol{x}) = \boldsymbol{b} + \delta \boldsymbol{b}$

- 式(7.6.8)的解为

$$\boldsymbol{x} + \delta \boldsymbol{x} = (1.089512538, 0.487967062, 1.491002798)^{\mathrm{T}}$$

- 可得

$$\delta \boldsymbol{x} = (0.0895, -0.5120, 0.4910)^{\mathrm{T}}$$

例7.9 已知Hilbert矩阵 $\boldsymbol{H}_n$，计算 $\boldsymbol{H}_3$ 的条件数

- 继续可得

$$\frac{\|\delta \boldsymbol{H}_3\|_\infty}{\|\boldsymbol{H}_3\|_\infty} \approx 0.18\times 10^{-3} < 0.02\%,$$

$$\frac{\|\delta \boldsymbol{b}\|_\infty}{\|\boldsymbol{b}\|_\infty} \approx 0.182\%,$$

$$\frac{\|\delta \boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty} \approx 51.2\%$$

- $\boldsymbol{H}_3$ 与 $\boldsymbol{b}$ 相对误差不超过 0.2%，而引起解的误差超过 50%

直接方法 | 判断矩阵是否病态

要判别一个矩阵是否病态需计算条件数 $\mathrm{cond}(\boldsymbol{A}) = \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$，而计算 $\|\boldsymbol{A}^{-1}\|$ 比较困难，在实际计算中如何发现病态情况？

- 如果在 $\boldsymbol{A}$ 的三角约化时（尤其是用主元素消去法解方程组时）出现小主元，那么对大多数矩阵来说， $\boldsymbol{A}$ 是病态矩阵
  - 例如用选主元的直接三角分解法解(7.6.8)

$$\begin{bmatrix} 1.00 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{bmatrix} \begin{bmatrix} x_1 + \delta x_1 \\ x_2 + \delta x_2 \\ x_3 + \delta x_3 \end{bmatrix} = \begin{bmatrix} 1.83 \\ 1.08 \\ 0.783 \end{bmatrix} \qquad (7.6.8)$$

  - 得到

$$\boldsymbol{I}_{23}(\boldsymbol{H}_3 + \delta \boldsymbol{H}_3) = \begin{bmatrix} 1 & & \\ 0.333 & 1 & \\ 0.500 & 0.994 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0.5000 & 0.3330 \\ & 0.0835 & 0.0891 \\ & & -0.00507 \end{bmatrix}$$

直接方法 | 判断矩阵是否病态

要判别一个矩阵是否病态需计算条件数 $\mathrm{cond}(\boldsymbol{A})=\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$，而计算 $\|\boldsymbol{A}^{-1}\|$ 比较困难，在实际计算中如何发现病态情况？

- 如果 $\boldsymbol{A}$ 的最大特征值和最小特征值之比（按绝对值）是大的，则 $\boldsymbol{A}$ 是病态的，即 $\frac{|\lambda_1|}{|\lambda_n|}$ 是大的，其中 λ_1 和 λ_n 为 $\boldsymbol{A}$ 的绝对值最大和绝对值最小的特征值

$$|\lambda_1|\geq|\lambda_2|\geq\cdots\geq|\lambda_n|>0$$

➢ 由定理7.16 可推导 $\Rightarrow|\lambda_1|\leq\|\boldsymbol{A}\|,\frac{1}{|\lambda_n|}\leq\|\boldsymbol{A}^{-1}\|\quad\Rightarrow\mathrm{cond}(\boldsymbol{A})\geq\frac{|\lambda_1|}{|\lambda_n|}\gg1$

直接方法 | 判断矩阵是否病态

要判别一个矩阵是否病态需计算条件数 $\mathrm{cond}(\boldsymbol{A}) = \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$，而计算 $\|\boldsymbol{A}^{-1}\|$ 比较困难，在实际计算中如何发现病态情况？

- 如果系数矩阵的行列式值相对很小，或系数矩阵某些行近似线性相关，则 $\boldsymbol{A}$ 可能是病态的

- 如果系数矩阵 $\boldsymbol{A}$ 元素间数量级相差很大，并且无一定规则，则$\boldsymbol{A}$可能是病态的

病态问题通常不能用选主元素的消去法来解决

直接方法 | 病态问题的求解

- 一般采用高精度的算术运算（采用双倍字长进行运算）
- 预处理方法，即将求解 $\boldsymbol{Ax=b}$ 的问题转化为求解等价方程组

$$\begin{cases} \boldsymbol{PAQy=Pb} \\ \boldsymbol{y=Q^{-1}x} \end{cases}$$

  - 通过选择非奇异矩阵 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$，使 $\mathrm{cond}(\boldsymbol{PAQ}) < \mathrm{cond}(\boldsymbol{A})$
  - 一般选择 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$ 为对角阵或者三角阵

- 当矩阵 $\boldsymbol{A}$ 的元素大小不均时，在 $\boldsymbol{A}$ 的行（或列）中引进适当的比例因子（使矩阵 $\boldsymbol{A}$ 的所有行或列按∞ –范数大体上有相同的长度，使$\boldsymbol{A}$的系数均衡）
  - 对 $\boldsymbol{A}$ 的条件数是有影响的，但不能保证 $\boldsymbol{A}$ 的条件数一定得到改善

例7.10 设 $A=\begin{bmatrix}1 & 10^4\\ 1 & 1\end{bmatrix}$，方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 为，$\begin{bmatrix}1 & 10^4\\ 1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}10^4\\ 2\end{bmatrix}$，计算 $\mathrm{cond}(\boldsymbol{A})_\infty$ 和解

- 先计算cond

$$\boldsymbol{A}=\begin{bmatrix}1 & 10^4\\ 1 & 1\end{bmatrix}\quad \boldsymbol{A}^{-1}=\frac{1}{10^4-1}\begin{bmatrix}-1 & 10^4\\ 1 & -1\end{bmatrix}\quad \mathrm{cond}(\boldsymbol{A})_\infty=\frac{(1+10^4)^2}{10^4-1}\approx 10^4$$

- 用列主元素消去法解 $\boldsymbol{Ax}=\boldsymbol{b}$ 时（计算到三位数字），得

$$(\boldsymbol{A},\boldsymbol{b})\rightarrow\begin{bmatrix}1 & 10^4 & 10^4\\ 0 & -10^4 & -10^4\end{bmatrix}$$

得到明显错误的结果：$x_1=0,\ x_2=1$

例7.10 设 $A = \begin{bmatrix} 1 & 10^4 \\ 1 & 1 \end{bmatrix}$，方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 为，$\begin{bmatrix} 1 & 10^4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10^4 \\ 2 \end{bmatrix}$，计算 $\mathrm{cond}(A)_\infty$ 和解

- 在 $\boldsymbol{A}$ 的第一行引进比例因子，如用 $s_1 = \max\limits_{1\le i\le 2} |a_{1i}| = 10^4$ 除第一个方程式，得 $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$，即

$$\begin{bmatrix} 10^{-4} & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad (\boldsymbol{A}')^{-1} = \frac{1}{1-10^{-4}}\begin{bmatrix} -1 & 1 \\ 1 & -10^{-4} \end{bmatrix} \quad \mathrm{cond}(\boldsymbol{A}')_\infty = \frac{1}{1-10^{-4}} \approx 1$$

- 用列主元素消去法解 $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$，得

$$(\boldsymbol{A}', \boldsymbol{b}') \to \begin{bmatrix} 1 & 1 & 2 \\ 10^{-4} & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

从而得到较好的结果：$x_1 = 1,\ x_2 = 1$

设 $\bar{x}$ 为方程组 $Ax = b$ 的近似解，于是可计算 $\bar{x}$ 的剩余向量 $r = b - A\bar{x}$ 衡量误差

定理7.21（事后误差估计） 设 $\boldsymbol{A}$ 为非奇异阵，$\boldsymbol{Ax} = \boldsymbol{b} \neq 0$，$\bar{\boldsymbol{x}}$ 是方程组的近似解，$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\bar{\boldsymbol{x}}$，则

$$\frac{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|}{\|\boldsymbol{x}\|} \leq \operatorname{cond}(\boldsymbol{A}) \frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}$$

- 由

$$\boldsymbol{A}(\boldsymbol{x} - \bar{\boldsymbol{x}}) = \boldsymbol{b} - \boldsymbol{A}\bar{\boldsymbol{x}} = \boldsymbol{r}$$

  得

$$\boldsymbol{x} - \bar{\boldsymbol{x}} = \boldsymbol{A}^{-1}\boldsymbol{r} \quad \Rightarrow \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \leq \|\boldsymbol{A}^{-1}\|\|\boldsymbol{r}\|$$

直接方法 | 事后误差估计

- 又有

$$\|\boldsymbol{b}\| = \|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\,\|\boldsymbol{x}\|$$

得

$$\frac{\mathbf{1}}{\|\boldsymbol{x}\|} \leq \frac{\|\boldsymbol{A}\|}{\|\boldsymbol{b}\|}$$

- 因此

$$\frac{\|\boldsymbol{x}-\overline{\boldsymbol{x}}\|}{\|\boldsymbol{x}\|} \leq \frac{\|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|} = \mathrm{cond}(\boldsymbol{A})\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}$$

直接方法 | 舍入误差

- 在复杂的计算中，由浮点运算而引进的舍入误差可能积累而影响答案，因此对任何算法都需要进行舍入误差分析，看其是否过度影响所得的结果
- 设 $\bar{\boldsymbol{x}}$ 为选主元素Gauss消去法解 $\boldsymbol{Ax} = \boldsymbol{b}$ 的计算解，$\boldsymbol{x}$ 为精确解
  - 若要直接计算每一步舍入误差对解的影响来获得界的估计$\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|$，是非常困难的
- Wilkinson等人提出了“向后误差分析方法”，基本思想是把舍入误差对解的影响归为原始数据变化对解的影响，即 $\bar{\boldsymbol{x}}$ 是扰动方程组 $(\boldsymbol{A} + \delta\boldsymbol{A})\boldsymbol{x} = \boldsymbol{b}$ 的精确解
  - 具体见教材
- 计算解 $\bar{\boldsymbol{x}}$ 的相对误差限依赖于 $\mathrm{cond}(\boldsymbol{A})_\infty$，元素的增长因子、方程组阶数、计算机字长等

直接方法 | 总结

7.1 Gauss消去法

- 消元过程、回代过程、能够成功的充分条件
- Gauss消去法的矩阵描述、矩阵的LU分解
- Gauss消去法的计算量

7.2 Guass主元素消去法

- 完全主元素消去法、列主元素消去法
- 列主元素的三角分解定理
- Gauss-Jordan消去法、矩阵求逆

直接方法 | 总结

7.3 Gauss消去法的变形

- 不选主元的三角分解法、选主元的三角分解法
- 平方根法、对称（正定）矩阵的三角分解、避免开方的分解
- 追赶法、对角占优的矩阵三角分解、分解的性质

7.4 向量和矩阵的范数

- 数量积与Euclid范数、范数、连续性和等价性
- 矩阵范数、算子范数、特征值、谱半径

7.5 误差分析

- 病态方程组、误差分析、条件数、舍入误差

方程求根 | 作业

作业信息

- **编程作业**：线性方程组求解与Attention机制优化，占比 20%，3 周（DDL：6.5）
- **课后作业**
  - 第七章 7、15、21 及第八章 1，5，8
  - 两章合并为 1 次作业，共 6 题，占比 3%，2 周（DDL：5.29）

成绩构成更新：

- **平时作业 15%**（共 7 次作业，6 * 2% + 3%）
- **编程作业 30%**（共 2 次，10% + 20%）
- **期末考试 55%**（半开卷考试，5% 小抄成绩 + 50% 卷面成绩）