第三章 线性方程

第（1）部分：高斯消去法

目录

1. 问题概览
2. 高斯消元法
3. 复杂度
4. 列主元素消去法
5. 高斯消元和 LU 分解法
6. 列主元素消去法和 PLU 分解
7. 总结和附录

解线性方程组

之前介绍的，我们可以通过迭代法来求解方程 $f(x) = 0$。

解线性方程组 $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$ 在大量的复杂计算问题中是一个十分基础的模块。因而，对这类问题的细致研究具有现实的必要性。

我们想要解 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

其中 $\boldsymbol{A}$ 是非奇异矩阵，即 $\det(A) \neq 0$。

- 直接法：针对低阶稠密矩阵

  高斯消元法、LU 分解法等

- 迭代法：针对大型的稀疏矩阵

  Jacobi 迭代法、Gauss-Seidel 迭代法、超松弛迭代法等

目录

1. 问题概览
2. 高斯消元法
3. 复杂度
4. 列主元素消去法
5. 高斯消元和 LU 分解法
6. 列主元素消去法和 PLU 分解
7. 总结和附录

高斯消元法的核心例子

$$\begin{cases} x_1 + x_2 + x_3 = 6 \\ 4x_2 - x_3 = 5 \\ \underbrace{2x_1}_{\text{消去}x_1} -2x_2 + x_3 = 1 \end{cases} \Rightarrow \begin{cases} x_1 + x_2 + x_3 = 6 \\ 4x_2 - x_3 = 5 \\ -4x_2 - x_3 = -11 \end{cases}$$

$$\Rightarrow \text{ 再次消去 } x_2 \text{ 可得 } \begin{cases} x_1 + x_2 + x_3 = 6 \\ 4x_2 - x_3 = 5 \\ -2x_3 = -6 \end{cases} \Rightarrow \begin{cases} x_1 = 1 \\ x_2 = 2 \\ x_3 = 3 \end{cases}$$

整个过程可以简写为:

$$(\boldsymbol{A} \mid \boldsymbol{b}) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 2 & -2 & 1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 0 & -4 & -1 & -11 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 0 & 0 & -2 & -6 \end{array}\right)$$

高斯消元法的一般形式

对于给定的 $(\boldsymbol{A}^{(1)} \mid \boldsymbol{b}^{(1)}) \equiv (\boldsymbol{A} \mid \boldsymbol{b})$，第一次消元后可得

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right) \qquad \begin{array}{c} \text{不变} \\ \text{行 2 减去行 1} \times a_{21}^{(1)}/a_{11}^{(1)} \\ \vdots \\ \text{行 n 减去行 1} \times a_{n1}^{(1)}/a_{11}^{(1)} \end{array}$$

因而，我们需要如下符号 $m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \quad (i \geqslant 2)$。新矩阵的元素是

$$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1} a_{1j}^{(1)} \quad (i \geqslant 2, j \geqslant 2)$$
$$b_i^{(2)} = b_i^{(1)} - m_{i1} b_1^{(1)} \qquad (i \geqslant 2)$$

问题：我们需要多少次加减法、乘法、除法？（包含更新 $\boldsymbol{b}_i^{(2)}$）

高斯消元法的一般形式

对于给定的 $(\boldsymbol{A}^{(1)} \mid \boldsymbol{b}^{(1)}) \equiv (\boldsymbol{A} \mid \boldsymbol{b})$，第一次消元后可得

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right) \quad \begin{array}{c} \text{不变} \\ \text{行 2 减去行 1} \times a_{21}^{(1)}/a_{11}^{(1)} \\ \vdots \\ \text{行 n 减去行 1} \times a_{n1}^{(1)}/a_{11}^{(1)} \end{array}$$

因而，我们需要如下符号 $m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \quad (i \geqslant 2)$。新矩阵的元素是

$$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1} a_{1j}^{(1)} \quad (i \geqslant 2, j \geqslant 2)$$
$$b_i^{(2)} = b_i^{(1)} - m_{i1} b_1^{(1)} \qquad (i \geqslant 2)$$

问题：我们需要多少次加减法、乘法、除法？（包含更新 $\boldsymbol{b}_i^{(2)}$）
答案：$(n-1)n = \mathcal{O}(n^2)$, $(n-1)n = \mathcal{O}(n^2)$, $n-1 = \mathcal{O}(n)$

不断重复后，在 $k-1$ 步后，我们可得

$$(\boldsymbol{A}^{(k)} \mid \boldsymbol{b}^{(k)}) = \left(\begin{array}{cccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \cdots & \cdots & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\ & & & \vdots & \cdots & \vdots & \vdots \\ & & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & b_n^{(k)} \end{array}\right)$$

最后我们可以得到 $\boldsymbol{A}^{(n)}\boldsymbol{x} = \boldsymbol{b}^{(n)}$:

$$
\left(\boldsymbol{A}^{(n)} \mid \boldsymbol{b}^{(n)}\right) = \left(\begin{array}{cccccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\
 & a_{22}^{(2)} & \cdots & \cdots & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\
 & & \ddots & \vdots & \vdots & \vdots & \vdots \\
 & & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\
 & & & & \ddots & \vdots & \vdots \\
 & & & & & a_{nn}^{(n)} & b_n^{(n)}
\end{array}\right) \tag{1}
$$

我们可以直接得到

$$
x_n = b_n^{(n)} / a_{nn}^{(n)} \tag{2}
$$

通过逐步回代的过程:

$$
x_k = \Big(b_k^{(k)} - \sum_{j=k+1}^{n} a_{kj}^{(k)} x_j\Big) / a_{kk}^{(k)} \qquad \text{(回代过程)} \tag{3}
$$

注: 若 $a_{11}^{(1)} = 0$，则通过换行将第一行和第 i_1 行交换（假设 $a_{i_1,1} \neq 0$）。

练习

回代过程需要多少次乘除法运算:

$$x_n = b_n^{(n)} / a_{nn}^{(n)}$$
$$x_k = \left(b_k^{(k)} - \sum_{j=k+1}^{n} a_{kj}^{(k)} x_j\right) / a_{kk}^{(k)}, \qquad k = n-1, \cdots, 1.$$

A $\mathcal{O}(n)$

B $\mathcal{O}(n^2)$

C $\mathcal{O}(n^3)$

练习

回代过程需要多少次乘除法运算：

$$
\begin{aligned}
x_n &= b_n^{(n)} / a_{nn}^{(n)} \\
x_k &= \Big(b_k^{(k)} - \sum_{j=k+1}^{n} a_{kj}^{(k)} x_j\Big) / a_{kk}^{(k)}, \qquad k = n-1, \cdots, 1.
\end{aligned}
$$

A $\mathcal{O}(n)$
B $\mathcal{O}(n^2)$
C $\mathcal{O}(n^3)$

答案： B

总结一下

定理

若 $\boldsymbol{A}$ 为 n 阶非奇异矩阵，则可通过高斯消去法（以及行交换）变成三角方程组。

引理

假设无行交换。则 $a_{ii}^{(i)} \neq 0 (i = 1, 2, \cdots, k)$ 的充要条件是 $D_i \neq 0$ $(i = 1, 2, \cdots, k)$，其中

$$D_i = \begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix} \neq 0 \qquad (\text{这个数量又叫做顺序主子式}).$$

定理

若 $D_k \neq 0$，其中 $k = 1, 2 \ldots, n-1$。则 $\begin{cases} a_{11}^{(1)} = D_1 \\ a_{kk}^{(k)} = D_k / D_{k-1} \quad k = 2, 3, 4 \ldots, n \end{cases}$

核心点就是假设没有行交换，行列式的值在高斯消元过程中不变。

定理 (课本定理 7.2)

如果 n 阶矩阵 $\boldsymbol{A}$ 的所有顺序主子式均不为零，即 $D_i \neq 0 \quad (i = 1, 2, \cdots, n\ ),$ 则可通过 *Gauss* 消去法（不进行交换两行的初等变换），将方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 约化为三角方程组（见讲义内公式(1)）。

计算公式如下*:*

(1) 消元计算 $(k = 1, 2, \cdots, n-1)$

$$
\begin{aligned}
m_{ik} &= a_{ik}^{(k)} / a_{kk}^{(k)} \quad (i = k+1, \cdots, n) \\
a_{ij}^{(k+1)} &= a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)} \quad (i, j = k+1, \cdots, n) \\
b_i^{(k+1)} &= b_i^{(k)} - m_{ik} b_k^{(k)} \quad (i = k+1, \cdots, n)
\end{aligned}
$$

(2) 回代计算求解公式为讲义内公式(2)和(3)。

目录

1. 问题概览
2. 高斯消元法
3. 复杂度
4. 列主元素消去法
5. 高斯消元和 LU 分解法
6. 列主元素消去法和 PLU 分解
7. 总结和附录

高斯消去法需要的运算次数

对于 $n \times n$ 的矩阵，第一层的高斯消元需要加减法 $(n-1)n$，乘法 $(n-1)n$，除法 $n-1$。因而

总的加减法次数是 ________

总的乘法次数是 ________

总的除法次数是 ________

总的乘除法次数是 ________

高斯消去法需要的运算次数

对于 $n \times n$ 的矩阵，第一层的高斯消元需要加减法 $(n-1)n$，乘法 $(n-1)n$，除法 $n-1$。因而

总的加减法次数是 $\sum_{i=2}^{n}(i-1)i = \mathcal{O}(n^3)$

总的乘法次数是 $\sum_{i=2}^{n}(i-1)i = \mathcal{O}(n^3)$

总的除法次数是 $\sum_{i=2}^{n} i-1 = \mathcal{O}(n^2)$

总的乘除法次数是 $\mathcal{O}(n^3) + \mathcal{O}(n^2) = \mathcal{O}(n^3)$

由于回代过程需要的计算量是 $\mathcal{O}(n^2)$，因而对于高斯消去法求解线性方程组需要的乘除法/加减法的计算量主要由消元过程来主导，且该方法来解线性方程组的复杂度是 $\mathcal{O}(n^3)$。

目录

1. 问题概览
2. 高斯消元法
3. 复杂度
4. 列主元素消去法
5. 高斯消元和 LU 分解法
6. 列主元素消去法和 PLU 分解
7. 总结和附录

高斯顺序消去法的问题

问题 1：若 $a_{11}^{(1)} = 0$，我们需要行交换才能使用高斯消元。

问题 2：精度问题

高斯消元过程理论上没有任何的数值方法层面的近似，但是有效数字对误差精度有影响。

例子

$$\begin{pmatrix} 0.001 & 2.000 & 3.000 \\ -1.000 & 3.712 & 4.623 \\ -2.000 & 1.072 & 5.643 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1.000 \\ 2.000 \\ 3.000 \end{pmatrix}$$

如果计算过程使用 4 位有效数字，解离实际值相差很远。

$$(\boldsymbol{A}, \boldsymbol{b}) = \begin{pmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ -1.000 & 3.712 & 4.623 & 2.000 \\ -2.000 & 1.072 & 5.643 & 3.000 \end{pmatrix} \quad \begin{aligned} m_{21} &= -1.000/0.001 = -1000 \\ m_{31} &= -2.000/0.001 = -2000 \end{aligned}$$

$$\to \begin{pmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ 0 & 2004 & 3005 & 1002 \\ 0 & 4001 & 6006 & 2003 \end{pmatrix} m_{32} = 4001/2004 = 1.997$$

$$\to \begin{pmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ 0 & 2004 & 3005 & 1002 \\ 0 & 0 & 5.000 & 2.000 \end{pmatrix} \to \bar{\boldsymbol{x}} = (-0.4000, -0.09980, 0.4000)^{\mathrm{T}}.$$

其中真实值为 $\boldsymbol{x}^* = (-0.4904, -0.05104, 0.3675)^{\mathrm{T}}$.

该问题的来源：$|m_{k,1}|$ 太大了。

所以问题的解决变成了如何使得 $a_{11}^{(1)}$ 足够大，且不改变问题本身 $\Rightarrow$ 行交换

[方法 2] 交换行，避免绝对值小的主元作除数.

$$(A,b)\xrightarrow{r_1\leftrightarrow r_3}\left(\begin{array}{ccc:c}-2.000 & 1.072 & 5.643 & 3.000\\ -1.000 & 3.712 & 4.623 & 2.000\\ 0.001 & 2.000 & 3.000 & 1.000\end{array}\right)\begin{array}{l}m_{21}=0.5000\\ m_{31}=-0.0005\end{array}$$

$$\rightarrow\left(\begin{array}{ccc:c}-2.000 & 1.072 & 5.643 & 3.000\\ 0 & 3.176 & 1.801 & 0.5000\\ 0 & 2.001 & 3.003 & 1.002\end{array}\right)m_{32}=0.6300$$

$$\rightarrow\left(\begin{array}{ccc:c}-2.000 & 1.072 & 5.643 & 3.000\\ 0 & 3.176 & 1.801 & 0.5000\\ 0 & 0 & 1.868 & 0.6870\end{array}\right),$$

得计算解为 $x=(-0.4900,\ -0.05113,\ 0.3678)^{\mathrm{T}}\approx x^*$.

（同步可参考代码里的实验。）

列主元素消去法

比如在第 k 步，

$$\left(\boldsymbol{A}^{(k)} \mid \boldsymbol{b}^{(k)}\right) = \left(\begin{array}{cccccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\
 & a_{22}^{(2)} & \cdots & \cdots & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\
 & & \ddots & \vdots & \vdots & \vdots & \vdots \\
 & & & \boxed{a_{kk}^{(k)}} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\
 & & & \boxed{\vdots} & \cdots & \vdots & \vdots \\
 & & & \boxed{a_{nk}^{(k)}} & \cdots & a_{nn}^{(k)} & b_n^{(k)}
\end{array}\right)$$

找到 $a_{kk}^{(k)}, a_{(k+1)k}^{(k)}, a_{(k+2)k}^{(k)}, \cdots, a_{nk}^{(k)}$（即带框的部分）中绝对值最大的，

$$\left|a_{i_k,k}^{(k)}\right| = \max_{k \leq i \leq n} \left|a_{ik}^{(k)}\right|.$$

然后将第 k 行和第 i_k 行交换。

练习:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 8 \\ 17 \\ 28 \end{pmatrix}.$$

目录

1. 问题概览
2. 高斯消元法
3. 复杂度
4. 列主元素消去法
5. 高斯消元和 LU 分解法
6. 列主元素消去法和 PLU 分解
7. 总结和附录

矩阵的三角分解

假设各顺度主子式不为 0

我们可以验证一下，高斯消元其实就等价于左乘某个矩阵 $\boldsymbol{L}_1$：

$$\boldsymbol{L}_1\boldsymbol{A}^{(1)} = \boldsymbol{A}^{(2)} \qquad \boldsymbol{L}_1 = \begin{pmatrix} 1 & & & & \\ -m_{21} & 1 & & & \\ -m_{31} & & 1 & & \\ \vdots & & & \ddots & \\ -m_{n1} & & & & 1 \end{pmatrix}$$

重复这个过程

$$\boldsymbol{L}_k \boldsymbol{A}^{(k)} = \boldsymbol{A}^{(k+1)} \qquad \boldsymbol{L}_k \boldsymbol{b}^{(k)} = \boldsymbol{b}^{(k+1)}$$

其中

$$\boldsymbol{L}_k = \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & -m_{k+1,k} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & -m_{n,k} & & & 1 \end{pmatrix}$$

因而，高斯消去法的过程可以写成如下的矩阵形式:

$$\boldsymbol{L}_{n-1}\boldsymbol{L}_{n-2}\cdots\cdots\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{A}^{(1)} = \underbrace{\boldsymbol{A}^{(n)}}_{\text{上三角矩阵}} \quad \text{标记为}\,\boldsymbol{U}$$

$$\boldsymbol{L}_{n-1}\boldsymbol{L}_{n-2}\cdots\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{b}^{(1)} = \boldsymbol{b}^{(n)}$$

$$\Rightarrow \boldsymbol{A} \equiv \boldsymbol{A}^{(1)} = \underbrace{\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\cdots\boldsymbol{L}_{n-1}^{-1}}\boldsymbol{U} = \boldsymbol{L}\boldsymbol{U}$$

其中矩阵 $\boldsymbol{L}$ 是一个下三角矩阵:

$$\boldsymbol{L} = \begin{pmatrix} 1 & & & & \\ m_{21} & 1 & & & \\ m_{31} & m_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ m_{n1} & m_{n2} & \cdots & \cdots & 1 \end{pmatrix}$$

问题：不做任何的计算的前提下，请解释为什么 $\boldsymbol{L}_{n-1}\boldsymbol{L}_{n-2}\cdots\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{L} = \boldsymbol{I}$?

结论：LU 分解

定理 (矩阵 LU 分解)

假设 $\boldsymbol{A}$ 为 n 阶可逆矩阵，若 $D_i \neq 0$ $(i = 1, 2, \ldots, n-1)$，则 $\boldsymbol{A}$ 可分解为一个单位下三角矩阵 $\boldsymbol{L}$ 和一个上三角矩阵 $\boldsymbol{U}$ 的乘积，且分解唯一。

该分解的存在性其实通过高斯消去法已经得到了。

引理 (证明可见附录)

- 若 $\boldsymbol{U}$ 为上三角矩阵，且 $\boldsymbol{U}$ 可逆，则 $\boldsymbol{U}^{-1}$ 为上三角矩阵。
- 若 $\boldsymbol{U}$ 为单位上三角矩阵，且 $\boldsymbol{U}$ 可逆，则 $\boldsymbol{U}^{-1}$ 为单位上三角矩阵。

引理

- 若 $\boldsymbol{L}$ 为下三角矩阵，且 $\boldsymbol{L}$ 可逆，则 $\boldsymbol{L}^{-1}$ 为下三角矩阵。
- 若 $\boldsymbol{L}$ 为单位下三角矩阵，且 $\boldsymbol{L}$ 可逆，则 $\boldsymbol{L}^{-1}$ 为单位下三角矩阵。

引理

- 若 $\boldsymbol{U}_1$，$\boldsymbol{U}_2$ 为（单位）上三角，则 $\boldsymbol{U}_1\boldsymbol{U}_2$ 为（单位）上三角。
- 若 $\boldsymbol{L}_1$，$\boldsymbol{L}_2$ 为（单位）下三角，则 $\boldsymbol{L}_1\boldsymbol{L}_2$ 为（单位）下三角。

证明：矩阵 LU 分解的唯一性.

若不唯一，则存在 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} = \boldsymbol{L}_1\boldsymbol{U}_1$，进而我们可知

$$\underbrace{\boldsymbol{L}^{-1}\boldsymbol{L}_1}_{\text{单位下三角}} = \underbrace{\boldsymbol{U}\boldsymbol{U}_1^{-1}}_{\text{上三角}}$$

因而，$\boldsymbol{L}^{-1}\boldsymbol{L}_1 = \boldsymbol{U}\boldsymbol{U}_1^{-1} = \boldsymbol{I} \to \boldsymbol{L} = \boldsymbol{L}_1$, $\boldsymbol{U} = \boldsymbol{U}_1$，故而假设不成立。通过反证法可知唯一性。 $\square$

例子

对于讲义一开始的 3×3 矩阵，

$$(\boldsymbol{A} \mid \boldsymbol{b}) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 2 & -2 & 1 & 1 \end{array}\right)$$

$m_{21} = 0$, $m_{31} = 2$, $m_{32} = -1$，因此

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

目录

1. 问题概览
2. 高斯消元法
3. 复杂度
4. 列主元素消去法
5. 高斯消元和 LU 分解法
6. 列主元素消去法和 PLU 分解
7. 总结和附录

对于列主元素消去法，我们也可以类似写成矩阵的形式

新的概念是排列矩阵

定义

排列矩阵是单位矩阵通过行交换得到的；$\boldsymbol{I}_{j,k}$ 是指将单位矩阵的第 j 行和第 k 行交换。

例子：我们考虑 4×4 的矩阵

$$\boldsymbol{I}_{2,3}=\begin{pmatrix}1&&&\\&&1&\\&1&&\\&&&1\end{pmatrix},\qquad \boldsymbol{I}_{2,4}=\begin{pmatrix}1&&&\\&&&1\\&&1&\\&1&&\end{pmatrix}$$

消元的过程可以写成

$$
\begin{aligned}
&\boldsymbol{L}_1 \boldsymbol{I}_{1,i_1} \boldsymbol{A}^{(1)} = \boldsymbol{A}^{(2)} && \boldsymbol{L}_1 \boldsymbol{I}_{1,i_1} \boldsymbol{b}^{(1)} = \boldsymbol{b}^{(2)} \\
&\boldsymbol{L}_2 \boldsymbol{I}_{2,i_2} \boldsymbol{A}^{(2)} = \boldsymbol{A}^{(3)} && \boldsymbol{L}_2 \boldsymbol{I}_{2,i_2} \boldsymbol{b}^{(2)} = \boldsymbol{b}^{(3)} \\
&\vdots \\
&\boldsymbol{L}_k \boldsymbol{I}_{k,i_k} \boldsymbol{A}^{(k)} = \boldsymbol{A}^{(k+1)} && \boldsymbol{L}_k \boldsymbol{I}_{k,i_k} \boldsymbol{b}^{(k)} = \boldsymbol{b}^{(k+1)} \\
&\vdots
\end{aligned}
$$

其中 $\boldsymbol{L}_k$ 中的元素满足 $|m_{i,k}| \le 1$ $(i = k+1, \cdots, n)$。

问题：为什么能满足小于等于 1?

合并

$$\underbrace{\boldsymbol{L}_{n-1}\boldsymbol{I}_{n-1,i_{n-1}}\cdots\boldsymbol{L}_2\boldsymbol{I}_{2,i_2}\boldsymbol{L}_1\boldsymbol{I}_{1,i_1}}_{=:\widetilde{\boldsymbol{P}}}\boldsymbol{A}=\boldsymbol{A}^{(n)}=\boldsymbol{U}$$

所以

$$\widetilde{\boldsymbol{P}}\boldsymbol{A}=\boldsymbol{U}\qquad\widetilde{\boldsymbol{P}}\boldsymbol{b}=\boldsymbol{b}^{(n)}$$

问题： $\widetilde{\boldsymbol{P}}$ 怎么能分离出下三角矩阵的结构呢?

例子：$n = 4$

$$
\begin{aligned}
\boldsymbol{U} = \boldsymbol{A}^{(4)} &= \boldsymbol{L}_3 \boldsymbol{I}_{3,i_3} \boldsymbol{L}_2 \boldsymbol{I}_{2,i_2} \boldsymbol{L}_1 \boldsymbol{I}_{1,i_1} \boldsymbol{A} \\
&= \underbrace{\boldsymbol{L}_3}_{\widetilde{\boldsymbol{L}}_3} \underbrace{\boldsymbol{I}_{3,i_3} \boldsymbol{L}_2 \boldsymbol{I}_{3,i_3}}_{\widetilde{\boldsymbol{L}}_2} \underbrace{\boldsymbol{I}_{3,i_3} \boldsymbol{I}_{2,i_2} \boldsymbol{L}_1 \boldsymbol{I}_{2,i_2} \boldsymbol{I}_{3,i_3}}_{\widetilde{\boldsymbol{L}}_1} \underbrace{\boldsymbol{I}_{3,i_3} \boldsymbol{I}_{2,i_2} \boldsymbol{I}_{1,i_1}}_{\boldsymbol{P}} \boldsymbol{A} = \boldsymbol{L}^{-1} \boldsymbol{P} \boldsymbol{A}.
\end{aligned}
$$

我们验证为什么 $\boldsymbol{I}_{3,i_3} \boldsymbol{L}_2 \boldsymbol{I}_{3,i_3}$ 是一个单位下三角矩阵：通过分块矩阵的形式

$$
\boldsymbol{L}_2 = \begin{pmatrix} I & 0 \\ C & I \end{pmatrix}, \qquad \boldsymbol{I}_{3,i_3} = \begin{pmatrix} I & 0 \\ 0 & Q \end{pmatrix}
$$

其中 Q 是一个维数小一点的排列矩阵。

$$
\boldsymbol{I}_{3,i_3} \boldsymbol{L}_2 \boldsymbol{I}_{3,i_3} = \begin{pmatrix} I & 0 \\ QC & Q^2 \end{pmatrix} \overset{\text{为什么}}{=} \begin{pmatrix} I & 0 \\ QC & I \end{pmatrix} = \text{单位下三角矩阵}
$$

并且该下三角矩阵的元素的 绝对值不超过 1。

定理 (列主元素的三角分解定理)

如果 $\boldsymbol{A}$ 为非奇异矩阵, 则存在排列矩阵 $\boldsymbol{P}$，使

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

其中 $\boldsymbol{L}$ 为单位下三角阵, $\boldsymbol{U}$ 为上三角阵。

目录

1. 问题概览
2. 高斯消元法
3. 复杂度
4. 列主元素消去法
5. 高斯消元和 LU 分解法
6. 列主元素消去法和 PLU 分解
7. 总结和附录

总结

- 需要掌握高斯消去法和列主元素消去法，以及能进行计算
- 通过高斯消去法来计算线性方程组需要 $\mathcal{O}(n^3)$ 的计算量
- 通过列主元素消去法来计算线性方程组需要 $\mathcal{O}(n^3)$ 的计算量（附加 $\mathcal{O}(n^2)$ 次元素交换和比大小）
- 了解高斯消去法可以帮我们计算 *LU* 分解，且能应用到一个具体的矩阵，如 3 维或 4 维矩阵。
- 了解列主元素消去法和 *PLU* 分解的关系，且能应用到一个具体的矩阵，如 3 维或 4 维矩阵。

附录：对于上三角矩阵性质的证明

对于 $\boldsymbol{U}$ 和其逆，我们可以标记各个部分为 $\boldsymbol{U} = \begin{bmatrix} v & \boldsymbol{u}^T \\ 0 & \boldsymbol{D} \end{bmatrix}$ 和 $\boldsymbol{U}^{-1} = \begin{bmatrix} a & \boldsymbol{c}^T \\ \boldsymbol{b} & \boldsymbol{d} \end{bmatrix}$ (分块矩阵内的符号随便起的)。由于 $\boldsymbol{U}\boldsymbol{U}^{-1} = \boldsymbol{I}$，我们可得

$$\begin{bmatrix} va + \boldsymbol{u}^T\boldsymbol{b} & v\boldsymbol{c}^T + \boldsymbol{u}^T\boldsymbol{d} \\ \boldsymbol{D}\boldsymbol{b} & \boldsymbol{D}\boldsymbol{d} \end{bmatrix} = \begin{bmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$$

由于 $\boldsymbol{D}$ 可逆，$\boldsymbol{b} = \boldsymbol{0}$，且 $\boldsymbol{d} = \boldsymbol{D}^{-1}$, $a = v^{-1}$。进而

$$\boldsymbol{U}^{-1} = \begin{bmatrix} v^{-1} & -v^{-1}\boldsymbol{u}^T\boldsymbol{d} \\ \boldsymbol{0} & \boldsymbol{D}^{-1} \end{bmatrix}$$

若 $\boldsymbol{D}$ 是矩阵，则重复使用上述的方法；最终 $\boldsymbol{U}^{-1}$ 是一个上三角矩阵。

对于单位上三角的例子，由于 $v = 1$，我们可知 $\boldsymbol{U}^{-1}$ 最左上角的值也是 1。通过不断继续使用上述方法，最终可知 $\boldsymbol{U}^{-1}$ 是一个单位上三角矩阵。