第二讲 线性方程组的直接解法

Gauss 消去法 • 矩阵分解法

目录

2-1 Gauss 消去法

2.1 Gauss 消去法

https://math.ecnu.edu.cn/~jypan/Teaching/NA

Gauss 消去法

- Gauss 消去法的基本思想是消元.
- 早在 2000 年前, 中国古代学者就提出了消元思想（记载在公元初《九章算术》方程章中）, Newton, Lagrange, Gauss, Jacobi 等都对此做过研究.
- 我们目前采用的算法描述方式是十九世纪三十年代后期才形成的.

Gaussian elimination (GE) is the standard method for solving a system of linear equations. As such, it is one of the most ubiquitous numerical algorithms and plays a fundamental role in scientific computation. GE was known to the ancient Chinese[1] and is familiar to many school children as the intuitively natural method of eliminating variables from linear equations.

Nicholas J. Higham

Royal Society
Royal Academy of Engineering
Academia Europaea
National Academy of Engineering
President of SIAM,
ACM Fellow, SIAM Fellow

Nicholas J. Higham, Gaussian elimination, *WIREs Computational Statistics*, 3 (2011), 230-238.

一个例子

例 求解下面的线性方程组

$$\begin{cases} x_1 - 2x_2 + 2x_3 = -2 \\ 2x_1 - 3x_2 - 3x_3 = 4 \\ 4x_1 + x_2 + 6x_3 = 3. \end{cases}$$

Gauss 消去法: 先写出增广矩阵, 然后通过初等变换将其转换为阶梯形 (上三角)

$$\begin{bmatrix} 1 & -2 & 2 & -2 \\ 2 & -3 & -3 & 4 \\ 4 & 1 & 6 & 3 \end{bmatrix} \xrightarrow[③-①\times 4]{②-①\times 2} \begin{bmatrix} 1 & -2 & 2 & -2 \\ 0 & 1 & -7 & 8 \\ 0 & 9 & -2 & 11 \end{bmatrix} \xrightarrow{③-②\times 9} \begin{bmatrix} 1 & -2 & 2 & -2 \\ 0 & 1 & -7 & 8 \\ 0 & 0 & 61 & -61 \end{bmatrix}$$

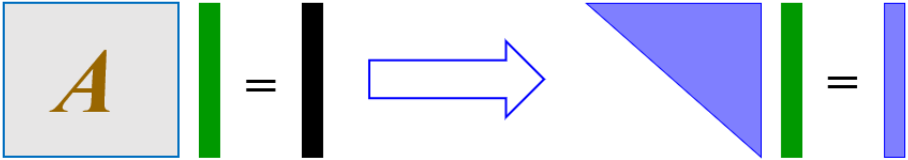
通过回代求解可得

$$x_3 = -1, \quad x_2 = 8 + 7x_3 = 1, \quad x_1 = -2 + 2x_2 - 2x_3 = 2.$$

推广到一般线性方程

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

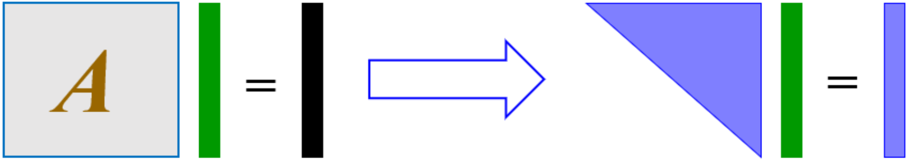
高斯消去法的主要思路: 将系数矩阵 A 化为上三角矩阵, 然后回代求解



推广到一般线性方程

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

高斯消去法的主要思路: 将系数矩阵 A 化为上三角矩阵, 然后回代求解



✍ 高斯消去法是求解线性方程组的经典算法, 是线性代数的重要组成部分, 除了用于线性方程组求解外, 还用于计算行列式、矩阵的秩、矩阵的逆等.

2-1-1 Gauss 消去过程

写出相应算法, 并编程实现

记增广矩阵

$$A^{(1)} = \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{array}\right],$$

其中

$$a_{ij}^{(1)} = a_{ij}, \quad b_i^{(1)} = b_i, \quad i, j = 1, 2, \ldots, n.$$

第 1 步: 消第 1 列.

设 $a_{11}^{(1)} \neq 0$, 计算 $l_{i1} = \dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}}, i = 2, 3, \ldots, n$. 对增广矩阵 $A^{(1)}$ 进行 $n-1$ 次初等变换, 即依次将 $A^{(1)}$ 的第 i 行 $(i > 1)$ 减去第 1 行的 l_{i1} 倍, 将新得到的矩阵记为 $A^{(2)}$, 即

$$A^{(2)} = \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & \vdots & \ddots & \vdots & \vdots \\ & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right],$$

其中

$$a_{ij}^{(2)} = a_{ij}^{(1)} - l_{i1} a_{1j}^{(1)}, \quad b_i^{(2)} = b_i^{(1)} - l_{i1} b_1^{(1)}, \quad i, j = 2, 3, \ldots, n.$$

第 2 步: 消第 2 列.

设 $a_{22}^{(2)} \neq 0$, 计算 $l_{i2} = \dfrac{a_{i2}^{(2)}}{a_{22}^{(2)}}, i = 3, 4, \ldots, n$. 依次将 $A^{(2)}$ 的第 i 行 $(i > 2)$ 减去第 2 行的 l_{i2} 倍, 将新得到的矩阵记为 $A^{(3)}$, 即

$$A^{(3)} = \left[\begin{array}{cccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & & b_1^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & & b_2^{(2)} \\ & & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & & b_3^{(3)} \\ & & \vdots & \ddots & \vdots & & \vdots \\ & & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & & b_n^{(3)} \end{array}\right],$$

其中

$$a_{ij}^{(3)} = a_{ij}^{(2)} - l_{i2}a_{2j}^{(2)}, \quad b_i^{(3)} = b_i^{(2)} - l_{i2}b_2^{(2)}, \quad i, j = 3, 4, \ldots, n.$$

依此类推, 经过 $k-1$ 步后, 可得新矩阵 $A^{(k)}$:

$$A^{(k)} = \left[\begin{array}{ccccc|c} a_{11}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & \ddots & \vdots & & \vdots & \vdots \\ & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\ & & \vdots & \ddots & \vdots & \vdots \\ & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & b_n^{(k)} \end{array}\right], \tag{2.1}$$

第 k 步: 消第 k 列.

设 $a_{kk}^{(k)} \neq 0$, 计算 $l_{ik} = \dfrac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$, $i = k+1, \ldots, n$. 依次将 $A^{(k)}$ 的第 i 行 $(i > k)$ 减去第 k 行的 l_{ik} 倍, 将新得到的矩阵记为 $A^{(k+1)}$, 矩阵元素的更新公式为

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik} a_{kj}^{(k)}, \quad b_i^{(k+1)} = b_i^{(k)} - l_{ik} b_k^{(k)}, \quad i, j = k+1, k+2, \ldots, n. \tag{2.2}$$

这样, **经过 $n-1$ 步后**, 即可得到一个上三角矩阵 $A^{(n)}$:

$$A^{(n)} = \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & \vdots & \vdots \\ & & & a_{nn}^{(n)} & b_n^{(n)} \end{array}\right].$$

最后, **回代求解**

$$x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}}, \quad x_i = \frac{1}{a_{ii}^{(i)}} \left(b_i^{(i)} - \sum_{j=i+1}^{n} a_{ij}^{(i)} x_j \right), \quad i = n-1, n-2, \ldots, 1.$$

以上就是 Gauss 消去法的整个计算过程.

> 由上面的计算过程可知, Gauss 消去法能顺利进行下去的充要条件是 $a_{kk}^{(k)} \neq 0$, $k = 1, 2, \ldots, n$, 这些元素被称为 **主元**.

主元

定理 设 $A \in \mathbb{R}^{n\times n}$, 则所有主元 $a_{kk}^{(k)}$ 都不为零的充要条件是 A 的所有顺序主子式都不为零, 即

$$D_1 \triangleq a_{11} \neq 0, \quad D_k \triangleq \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} \neq 0, \quad k = 2, 3, \ldots, n.$$

✍ 事实上, 如果 A 的所有顺序主子式都不为零, 则主元为

$$a_{11}^{(1)} = D_1, \quad a_{kk}^{(k)} = \frac{D_k}{D_{k-1}}, \quad k = 2, 3, \ldots, n.$$

推论 Gauss 消去法能顺利完成的充要条件是 A 的所有顺序主子式都不为零.

2-1-2 | Gauss 消去法的运算量

在第 k 步中, 我们需要计算

$$
\begin{aligned}
l_{ik} &= a_{ik}^{(k)}/a_{kk}^{(k)}, i = k+1,\ldots,n \quad \rightarrow \quad n-k \text{ 次除法} \\
a_{ij}^{(k+1)} &= a_{ij}^{(k)} - l_{ik}a_{kj}^{(k)}, i,j = k+1,\ldots,n \quad \rightarrow \quad (n-k)^2 \text{ 次乘法} \\
b_i^{(k+1)} &= b_i^{(k)} - l_{ik}b_k^{(k)}, i = k+1,\ldots,n \quad \rightarrow \quad n-k \text{ 次乘法}
\end{aligned}
$$

所以整个消去过程的**乘除**运算为

$$
\sum_{k=1}^{n-1} 2(n-k) + (n-k)^2 = \sum_{\ell=1}^{n-1} 2\ell + \ell^2 = n(n-1) + \frac{n(n-1)(2n-3)}{6}.
$$

易知, 回代求解过程的乘除运算为

$$
1 + \sum_{i=1}^{n-1} n-i+1 = \frac{n(n+1)}{2}.
$$

计算复杂度

所以整个 Gauss 消去法的乘除运算为

$$\frac{n^3}{3} + n^2 - \frac{n}{3}.$$

同理, 也可统计出, Gauss 消去法中的加减运算次数为

$$\frac{n^3}{3} + \frac{n^2}{2} - \frac{5n}{6}.$$

所以总运算量为

$$\boxed{\frac{2n^3}{3} + \frac{3n^2}{2} - \frac{7n}{6} = \frac{2n^3}{3} + \mathcal{O}(n^2)}$$

算法评价

✍ 评价算法的一个重要指标是 **执行时间**, 但这依赖于计算机硬件和编程技巧等, 因此直接给出算法执行时间是不太现实的. 所以我们通常是统计算法中算术运算 (加减乘除) 的次数.

✍ 为了尽可能地减少运算量, 在实际计算中, 数, 向量和矩阵做乘法运算时的先后执行次序为: 先计算数与向量的乘法, 然后计算矩阵与向量的乘法, 最后才计算矩阵与矩阵的乘法.

2-2 矩阵分解法

2.2 矩阵分解法

2-2-1 矩阵 LU 分解

换个角度看 Gauss 消去过程

每次都是做矩阵初等变换, 因此也可理解为不断地左乘初等矩阵. 将所有这些初等矩阵的乘积记为 $\tilde{L}$, 则可得 $\tilde{L}A = U$, 其中 U 是一个上三角矩阵. 记 $L \triangleq \tilde{L}^{-1}$, 则

$$A = LU,$$

这就是著名的矩阵 **LU 分解**.

2-2-1 矩阵 LU 分解

换个角度看 Gauss 消去过程

每次都是做矩阵初等变换, 因此也可理解为不断地左乘初等矩阵. 将所有这些初等矩阵的乘积记为 $\tilde{L}$, 则可得 $\tilde{L}A = U$, 其中 U 是一个上三角矩阵. 记 $L \triangleq \tilde{L}^{-1}$, 则

$$A = LU,$$

这就是著名的矩阵 **LU 分解**.

矩阵分解

将矩阵分解成若干具有简单结构的矩阵的乘积, 是矩阵计算中一个很重要的技术.

2-2-1 矩阵 LU 分解

换个角度看 Gauss 消去过程

每次都是做矩阵初等变换, 因此也可理解为不断地左乘初等矩阵. 将所有这些初等矩阵的乘积记为 $\tilde{L}$, 则可得 $\tilde{L}A = U$, 其中 U 是一个上三角矩阵. 记 $L \triangleq \tilde{L}^{-1}$, 则

$$A = LU,$$

这就是著名的矩阵 **LU 分解**.

矩阵分解

将矩阵分解成若干具有简单结构的矩阵的乘积, 是矩阵计算中一个很重要的技术.

假定 Gauss 消去过程能顺利进行, 则 U 一定是非奇异上三角矩阵, L 有什么特殊结构?

L 的结构

考察第 k 步的情形, 即 $A^{(k+1)}$ 与 $A^{(k)}$ 之间的关系式. 由 Gauss 消去过程可知

$$A^{(k+1)} = L_k A^{(k)},$$

其中 $L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{n,k} & & & 1 \end{bmatrix}, \quad l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}},\ i = k+1, \ldots, n.$

L 的结构

考察第 k 步的情形, 即 $A^{(k+1)}$ 与 $A^{(k)}$ 之间的关系式. 由 Gauss 消去过程可知

$$A^{(k+1)} = L_k A^{(k)},$$

其中 $$L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{n,k} & & & 1 \end{bmatrix}, \quad l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}},\ i = k+1, \ldots, n.$$

令 $k = 1, 2, \ldots, n-1$, 并将所有 Gauss 消去过程结合在一起即可得

$$\boxed{A^{(n)} = L_{n-1} L_{n-2} \cdots L_1 A} \iff \boxed{A = (L_{n-1} L_{n-2} \cdots L_1)^{-1} A^{(n)} \triangleq LU}$$

引理

$$L_k^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{n,k} & & & 1 \end{bmatrix}, \quad L_1^{-1} \cdots L_{n-1}^{-1} = \begin{bmatrix} 1 & & & & & \\ l_{21} & 1 & & & & \\ l_{31} & l_{32} & 1 & & & \\ l_{41} & l_{42} & l_{4,3} & 1 & & \\ \vdots & \vdots & \vdots & & \ddots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{n,n-1} & 1 \end{bmatrix}$$

(留作练习)

LU 分解的存在性和唯一性

定理 (LU 分解的存在性和唯一性)

设 $A \in \mathbb{R}^{n\times n}$, 则存在唯一的单位下三角矩阵 L 和非奇异上三角矩阵 U, 使得 $A = LU$ 的充要条件是 A 的所有顺序主子矩阵 $A_k = A(1:k, 1:k)$ 都非奇异, $k = 1, 2, \ldots, n$.

(板书)

LU 分解的存在性和唯一性

定理 (LU 分解的存在性和唯一性)

设 $A \in \mathbb{R}^{n\times n}$, 则存在唯一的单位下三角矩阵 L 和非奇异上三角矩阵 U, 使得 $A = LU$ 的充要条件是 A 的所有顺序主子矩阵 $A_k = A(1:k, 1:k)$ 都非奇异, $k = 1, 2, \ldots, n$.

(板书)

✍ 如果 A 存在 LU 分解, 则 $Ax = b$ 就等价于求解下面两个方程组

$$\begin{cases} Ly = b, \\ Ux = y. \end{cases}$$

算法 LU 分解

Set $L = I, U = 0$ % 将 L 设为单位矩阵, U 设为零矩阵

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

  for $i = k+1$ to n **do**

    $l_{ik} = a_{ik}/a_{kk}$ % 计算 L 的第 k 列

  end for

  for $i = k$ to n **do**

    $u_{ki} = a_{ki}$ % 计算 U 的第 k 行

  end for

  for $i = k+1$ to n **do**

    for $j = k+1$ to n **do**

      $a_{ij} = a_{ij} - l_{ik}u_{kj}$ % 更新 $A(k+1:n, k+1:n)$

    end for

  end for

end for

实施细节: 矩阵 L 和 U 的存储

- 当 A 的第 i 列被用于计算 L 的第 i 列后, 后面的计算中不再被使用
- 同样, A 的第 i 行被用于计算 U 的第 i 行后, 后面的计算中也不再使用

因此, 可以将 L 的第 i 列存放在 A 的第 i 列, 将 U 的第 i 行存放在 A 的第 i 行.

实施细节: 矩阵 L 和 U 的存储

- 当 A 的第 i 列被用于计算 L 的第 i 列后, 后面的计算中不再被使用
- 同样, A 的第 i 行被用于计算 U 的第 i 行后, 后面的计算中也不再使用

因此, 可以将 L 的第 i 列存放在 A 的第 i 列, 将 U 的第 i 行存放在 A 的第 i 行.

算法 LU 分解

```
for k = 1 to n − 1 do
    for i = k + 1 to n do
        a_ik = a_ik / a_kk
        for j = k + 1 to n do
            a_ij = a_ij − a_ik a_kj
        end for
    end for
end for
```

2-2-2 列主元 Gauss 消去法与 PLU 分解

为什么选主元

- A 非奇异, 则线性方程组就存在唯一解, 但 Gauss 消去法却不一定有效.
- 实际计算中, 即使主元都不为零, 但如果主元的值很小, 由于舍入误差的原因, 也可能会给计算结果带来很大的误差.

2-2-2 列主元 Gauss 消去法与 PLU 分解

为什么选主元

- A 非奇异, 则线性方程组就存在唯一解, 但 Gauss 消去法却不一定有效.
- 实际计算中, 即使主元都不为零, 但如果主元的值很小, 由于舍入误差的原因, 也可能会给计算结果带来很大的误差.

例 求解线性方程组 $Ax = b$, 其中 $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

由于主元 $a_{11} = 0$, 因此 Gauss 消去法无法顺利进行下去.

例 求解线性方程组 $Ax = b$, 其中 $A = \begin{bmatrix} 0.02 & 61.3 \\ 3.43 & -8.5 \end{bmatrix}$, $b = \begin{bmatrix} 61.5 \\ 25.8 \end{bmatrix}$,

要求在运算过程中保留 3 位有效数字.

解. 根据 LU 分解算法可得

$$
\begin{aligned}
& l_{11} = 1.00,\ l_{21} = a_{21}/a_{11} \approx 1.72 \times 10^2,\ l_{22} = 1.00, \\
& u_{11} = a_{11} = 2.00 \times 10^{-2},\ u_{12} = a_{12} = 6.13 \times 10, \\
& u_{22} = a_{22} - l_{21}u_{12} \approx -8.5 - 1.05 \times 10^4 \approx -1.05 \times 10^4,
\end{aligned}
$$

即

$$
A \approx \begin{bmatrix} 1.00 & 0 \\ 1.72 \times 10^2 & 1.00 \end{bmatrix} \begin{bmatrix} 2.00 \times 10^{-2} & 6.12 \times 10 \\ 0 & -1.05 \times 10^4 \end{bmatrix}.
$$

解方程组 $Ly = b$ 得 $y_1 = 6.15 \times 10,\quad y_2 = b_2 - l_{21}y_1 \approx -1.06 \times 10^4$.

解方程组 $Ux = y$ 得 $x_2 = y_2/u_{22} \approx 1.01,\quad x_1 = (y_1 - u_{12}x_2)/u_{11} \approx -20.7$ □

✍ 实际上, 方程的精确解为 $x_1 = 10.0$ 和 $x_2 = 1.00$.

选主元

在执行 Gauss 消去过程的第 k 步之前, 插入下面的选主元过程.

① 选取 **列主元** : $\left|a_{i_k,k}^{(k)}\right| = \max\limits_{k\leq i\leq n}\left\{\left|a_{i,k}^{(k)}\right|\right\}$

② 交换: 如果 $i_k \neq k$, 则交换第 k 行与第 i_k 行

即第 k 步时, 先在 $A^{(k)}$ 中第 k 列的第 k 至 n 的元素中选取绝对值最大的元素:

$$\begin{bmatrix} a_{11}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} \\ & \ddots & \vdots & & \vdots \\ & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ & & \vdots & \ddots & \vdots \\ & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{bmatrix}.$$

然后根据需要判断是否需要更换两行.

几点注记

- ✍ 上面选出的 $a_{i_k,k}^{(k)}$ 就称为 **列主元**.
- ✍ 加入这个选主元过程后, 就不会出现主元为零的情形 (除非 A 是奇异的).
- ✍ 带这种选主元方法的 Gauss 消去法就称为 **列主元 Gauss 消去法** 或 **部分选主元 Gauss 消去法** (Gaussian Elimination with Partial Pivoting, GEPP).

算法 列主元 Gauss 消去法

for $k = 1$ to $n-1$ **do**
  $a_{i_k,k} = \max_{k \le i \le n} |a_{i,k}|$ % 选列主元
  if $i_k \neq k$ **then**
    for $j = 1$ to n **do**
      $a_{tmp} = a_{i_k,j},\ a_{i_k,j} = a_{k,j},\ a_{k,j} = a_{tmp}$ % 交换 A 的第 i_k 行与第 k 行
    end for
    $b_{tmp} = b_{i_k},\ b_{i_k} = b_k,\ b_k = b_{tmp}$ % 交换 b 的第 i_k 与第 k 个分量
  end if
  for $i = k+1$ to n **do**
    $a_{ik} = a_{ik}/a_{kk}$ % 计算 L 的第 i 列
    for $j = k+1$ to n **do**
      $a_{ij} = a_{ij} - a_{ik} * a_{ij}$ % 更新 $A(k+1:n, k+1:n)$
    end for
    $b_i = b_i - a_{ik} b_k$
  end for
end for
$x_n = b_n / a_{nn}$ % 向后回代求解 $Ux = y$
for $i = n-1$ to 1 **do**
  for $j = i+1$ to n **do**
    $b_i = b_i - a_{ij} x_j$
  end for
  $x_i = b_i / a_{ii}$
end for

列主元 LU 分解的存在性

定理 (列主元 LU 分解) 若矩阵 $A \in \mathbb{R}^{n\times n}$ 非奇异, 则存在置换矩阵 P, 使得

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

其中 L 为单位下三角矩阵, U 为上三角矩阵.

(板书)

$$\boxed{Ax = b} \Longleftrightarrow \boxed{Ly = Pb, \quad Ux = y}$$

例 用部分选主元 LU 分解求解线性方程组 $Ax=b$, 其中 $A=\begin{bmatrix}0.02 & 61.3\\ 3.43 & -8.5\end{bmatrix}$, $b=\begin{bmatrix}61.5\\ 25.8\end{bmatrix}$, 要求在运算过程中保留 3 位有效数字.

解. 由于 $|a_{21}|>|a_{11}|$, 需要选主元, 并交换第 1 行与第 2 行, 即取 $P=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$, 然后计算 $\tilde{A}=PA=\begin{bmatrix}3.43 & -8.5\\ 0.02 & 61.3\end{bmatrix}$ 的 LU 分解, 即令

$$\tilde{A}=LU=\begin{bmatrix}1 & 0\\ l_{21} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12}\\ 0 & u_{22}\end{bmatrix}.$$

直接比较等式两边可得

$$u_{11}=\tilde{a}_{11}=3.43,\ u_{12}=\tilde{a}_{12}=-8.5,$$
$$l_{21}=\tilde{a}_{21}/u_{11}\approx 5.83\times 10^{-3},$$
$$u_{22}=\tilde{a}_{22}-l_{21}u_{12}\approx 61.3+0.0496\approx 61.3,$$

即

$$PA \approx \begin{bmatrix} 1.00 & 0 \\ 5.83\times 10^{-3} & 1.00 \end{bmatrix} \begin{bmatrix} 3.43 & -8.50 \\ 0 & 61.3 \end{bmatrix}.$$

解方程组 $Ly = Pb$ 可得

$$y_1 = 25.8, \quad y_2 \approx 61.2.$$

解方程组 $Ux = y$ 可得

$$x_2 = y_2/u_{22} \approx 0.998, \quad x_1 = (y_1 - u_{12}x_2)/u_{11} \approx 10.0.$$

所以, 数值解具有 3 位有效数字. (精确解为 $x_1 = 10, x_2 = 1$) $\square$

几点注记

✍ 列主元 Gauss 消去法要多做一些比较运算, 但

(1) 对系数矩阵要求低, 只需非奇异即可;

(2) 比普通 Gauss 消去法更稳定.

列主元 Gauss 消去法是当前求解线性方程组的直接法中的首选算法.

全主元 Gauss 消去法

为了获得更好的数值稳定性, 我们可以在剩余的子矩阵中选取主元, 即在 $A^{(k)}(k:n,k:n)$ 中选取绝对值最大的元素作为主元.

$$
\begin{bmatrix}
a_{11}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} \\
 & \ddots & \vdots & & \vdots \\
 & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\
 & & \vdots & \ddots & \vdots \\
 & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)}
\end{bmatrix}.
$$

全主元 Gauss 消去法

为了获得更好的数值稳定性, 我们可以在剩余的子矩阵中选取主元, 即在 $A^{(k)}(k:n,k:n)$ 中选取绝对值最大的元素作为主元.

$$\begin{bmatrix} a_{11}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} \\ & \ddots & \vdots & & \vdots \\ & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ & & \vdots & \ddots & \vdots \\ & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{bmatrix}.$$

全主元

① 选取 **全主元** : $|a_{i_k,j_k}^{(k)}| = \max_{k\leq i,j\leq n} \left\{ |a_{i,j}^{(k)}| \right\}$

② 行交换: 如果 $i_k \neq k$, 则交换第 k 行与第 i_k 行

③ 列交换: 如果 $j_k \neq k$, 则交换第 k 列与第 j_k 列

如果有列交换, 则会改变 x_i 的顺序, 因此需要记录每次的列交换次序.

全主元 LU 分解

定理 (全主元 LU 分解)

设矩阵 A 非奇异, 则存在置换矩阵 P_l, P_r, 以及单位下三角矩阵 L 和非奇异上三角矩阵 U, 使得

$$P_lAP_r = LU.$$

✍ 全主元高斯消去法具有更好的稳定性, 但很费时, 一般情况下很少使用.

其他矩阵分解

Crout 分解

$$\boxed{A = \tilde{L}\tilde{U}}, \quad \tilde{L} \text{ 非奇异下三角}, \tilde{U} \text{ 单位上三角}$$

LDR 分解

$$\boxed{A = LDR}, \quad L \text{ 单位下三角}, U \text{ 单位上三角}, D \text{ 对角}$$

显然, 这两种分解在本质上与 LU 分解没有任何区别, 在实际计算中可以根据需要选择其中的一种.

2-2-3 Cholesky 分解与平方根法

特殊线性方程组

如果 A 是**对称正定**的, 则可以得到更加简洁高效的方法.

定理 (Cholesky 分解) 设 $A \in \mathbb{R}^{n\times n}$ 对称正定, 则存在唯一的对角线元素全为正的下三角矩阵 L, 使得

$$A = LL^{\mathsf{T}}.$$

该分解称为 **Cholesky 分解**.

(板书)

如何计算 Cholesky 分解: 待定系数法

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

直接比较等式两边可得 $a_{ij} = \sum_{k=1}^{n} l_{ik}l_{jk} = \sum_{k=1}^{j-1} l_{ik}l_{jk} + l_{jj}l_{ij}$, 所以有

$$\boxed{l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2\right)^{\frac{1}{2}}, \quad l_{ij} = \frac{1}{l_{jj}}\left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}l_{jk}\right)}, j = 1, 2, \ldots, n, \ i = j, j{+}1, \ldots, n$$

根据这个公式即可得下面的算法描述.

for $j = 1$ to n **do**

$$l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2 \right)^{\frac{1}{2}} \qquad \text{\% 先计算对角线元素}$$

for $i = j+1$ to n **do**

$$l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right) / l_{jj} \qquad \text{\% 计算 } L \text{ 的第 } j \text{ 列}$$

end for

end for

算法 Cholesky 分解 (L 存放在 A 的下三角部分)

```
for j = 1 to n do
    for k = 1 to j − 1 do
        a_jj = a_jj − a_jk^2
    end for
    a_jj = sqrt(a_jj)
    for i = j + 1 to n do
        for k = 1 to j − 1 do
            a_ij = a_ij − a_ik a_jk
        end for
        a_ij = a_ij / a_jj
    end for
end for
```

- ✍ Cholesky 分解算法的乘除运算量为 $n^3/3 + \mathcal{O}(n^2)$.
- ✍ Cholesky 分解算法是稳定的 (全主元 Gauss 消去法相当), 不需要选主元.

平方根法

算法 Cholesky 分解求解线性方程组

计算 Cholesky 分解 (此处省略)
$y_1 = b_1/a_{11}$ % 向前回代求解 $Ly = b$
for $i = 2$ to n **do**
  for $j = 1$ to $i - 1$ **do**
    $b_i = b_i - a_{ij}y_j$
  end for
  $y_i = b_i/a_{ii}$
end for
$x_n = b_n/a_{nn}$ % 向后回代求解 $L^{\mathsf{T}}x = y$
for $i = n - 1$ to 1 **do**
  for $j = i + 1$ to n **do**
    $y_i = y_i - a_{ji}x_j$
  end for
  $x_i = y_i/a_{ii}$
end for

LDL$^\intercal$ 分解

$$\boxed{A = LDL^\intercal}, \qquad L \text{ 单位下三角},\ D \text{ 对角}.$$

- 目的: 避免在 Cholesky 分解计算平方根
- 计算方法: 待定系数法

$$A = LDL^\intercal = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{bmatrix} \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & \ddots & \vdots \\ & & \ddots & l_{n,n-1} \\ & & & 1 \end{bmatrix}.$$

$$\Longrightarrow \boxed{a_{ij} = d_j l_{ij} + \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}}, \quad j = 1, \ldots, n, \quad i = j+1, \ldots, n.$$

$$\boxed{a_{ij} = d_j l_{ij} + \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}}, \quad j = 1, \ldots, n, \quad i = j+1, \ldots, n.$$

1: **for** $j = 1$ to n **do**

2: $\quad d_j = a_{jj} - \sum_{k=1}^{j-1} d_k l_{jk}^2$ % 先计算 D 的对角线元素

3: $\quad$ **for** $i = j+1$ to n **do**

4: $\quad\quad l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} d_k l_{ik} l_{jk} \right) / d_j$ % 计算 L 的第 j 列

5: $\quad$ **end for**

6: **end for**

改进的平方根法 I

算法 改进的平方根法 (L 存放在 A 的下三角, D 存放在 A 的对角)

```
for j = 1 to n do
    for k = 1 to j − 1 do
        a_jj = a_jj − a_jk^2 a_kk
    end for
    for i = j + 1 to n do
        for k = 1 to j − 1 do
            a_ij = a_ij − a_ik a_kk a_jk
        end for
        a_ij = a_ij / a_jj
    end for
end for
y_1 = b_1      % 向前回代求解 Ly = b
```

改进的平方根法 II

for $i = 2$ to n **do**
  for $j = 1$ to $i - 1$ **do**
    $b_i = b_i - a_{ij} y_j$
  end for
  $y_i = b_i$
end for
$x_n = b_n / a_{nn}$ % 向后回代求解 $DL^{\intercal} x = y$
for $i = n - 1$ to 1 **do**
  $x_i = y_i / a_{ii}$
  for $j = i + 1$ to n **do**
    $x_i = x_i - a_{ji} x_j$
  end for
end for

LDLT 分解

- 若 A 对称正定, 则 A 存在 $\mathrm{LDL}^{\intercal}$ 分解, 且 D 的对角线元素都是正的.
- 但并非只有对称正定矩阵才存在 $\mathrm{LDL}^{\intercal}$ 分解.

LDLT 分解

- 若 A 对称正定, 则 A 存在 $\mathrm{LDL}^{\intercal}$ 分解, 且 D 的对角线元素都是正的.
- 但并非只有对称正定矩阵才存在 $\mathrm{LDL}^{\intercal}$ 分解.

定理 若 A 对称, 且所有顺序主子式都不为 0, 则 A 可唯一分解为

$$A = LDL^{\intercal},$$

其中 L 为单位下三角矩阵, D 为对角矩阵. (留作课外自习)

LDLT 分解

若 A 对称正定, 则 A 存在 $\mathrm{LDL}^{\mathsf{T}}$ 分解, 且 D 的对角线元素都是正的.

但并非只有对称正定矩阵才存在 $\mathrm{LDL}^{\mathsf{T}}$ 分解.

定理 若 A 对称, 且所有顺序主子式都不为 0, 则 A 可唯一分解为

$$A = LDL^{\mathsf{T}},$$

其中 L 为单位下三角矩阵, D 为对角矩阵. (留作课外自习)

思考: 若 $A \in \mathbb{R}^{n\times n}$ 存在 $\mathrm{LDL}^{\mathsf{T}}$ 分解, 则 A 一定是对称的. 反之, 如果 A 对称, 则 A 一定存在 $\mathrm{LDL}^{\mathsf{T}}$ 分解吗?

2-2-4 三对角线性方程组

特殊线性方程组: 不可约对角占优三对角线性方程组

$$\begin{bmatrix} b_1 & c_1 & & \\ a_1 & \ddots & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}.$$

我们假定

$$|b_1| > |c_1| > 0, \quad |b_n| > |a_{n-1}| > 0, \tag{2.3}$$

且

$$|b_i| \geq |a_{i-1}| + |c_i|, \quad a_i c_i \neq 0, \quad i = 1, \ldots, n-1. \tag{2.4}$$

即 A 是**不可约 (行) 弱对角占优**的.

计算 A 的 Crout 分解:

$$A = \begin{bmatrix} b_1 & c_1 & & \\ a_1 & \ddots & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix} = \begin{bmatrix} \alpha_1 & & & \\ a_1 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & a_{n-1} & \alpha_n \end{bmatrix} \begin{bmatrix} 1 & \beta_1 & & \\ & 1 & \ddots & \\ & & \ddots & \beta_{n-1} \\ & & & 1 \end{bmatrix} \triangleq LU. \quad (2.5)$$

由待定系数法, 我们可以得到递推计算公式:

$$\alpha_1 = b_1, \quad \beta_1 = c_1/\alpha_1 = c_1/b_1,$$

$$\begin{cases} \alpha_i = b_i - a_{i-1}\beta_{i-1}, \\ \beta_i = c_i/\alpha_i = c_i/(b_i - a_{i-1}\beta_{i-1}), \quad i = 2, 3, \ldots, n-1 \end{cases}$$

$$\alpha_n = b_n - a_{n-1}\beta_{n-1}.$$

为了使得算法能够顺利进行下去, 我们需要证明 $\alpha_i \neq 0$.

定理 设三对角矩阵 A 满足条件 (2.3) 和 (2.4), 则 A 非奇异, 且

(1) $|\alpha_1| = |b_1| > 0$;

(2) $0 < |\beta_i| < 1,\ i = 1, 2, \ldots, n-1$;

(3) $0 < |c_i| \leq |b_i| - |a_{i-1}| < |\alpha_i| < |b_i| + |a_{i-1}|, \quad i = 2, 3, \ldots, n$;

(板书)

追赶法 / Thomas 算法

算法 追赶法 / Thomas 算法 (矩阵分解与方程求解交叉进行)

$\alpha_1 = b_1$
$\beta_1 = c_1/b_1$
$y_1 = f_1/b_1$
for $i = 2$ to $n-1$ **do**
  $\alpha_i = b_i - a_{i-1}\beta_{i-1}$
  $\beta_i = c_i/\alpha_i$
  $y_i = (f_i - a_{i-1}y_{i-1})/\alpha_i$ % 求解 $Ly = f$
end for
$\alpha_n = b_n - a_{n-1}\beta_{n-1}$
$y_n = (f_n - a_{n-1}y_{n-1})/\alpha_n$
$x_n = y_n$
for $i = n-1$ to 1 **do**
  $x_i = y_i - \beta_i x_{i+1}$
end for

几点注记

- ✍ 具体计算时, 由于求解 $Ly = f$ 与矩阵 LU 分解是同时进行的, 因此, α_i 可以不用存储. 但 β_i 需要存储.
- ✍ 由于 $|\beta_i| < 1$, 因此在回代求解 x_i 时, 误差可以得到有效控制.
- ✍ 乘除运算量大约为 $5n$, 加减运算大约为 $3n$.

列对角占优情形

我们也可以考虑下面的分解

$$A=\begin{bmatrix} b_1 & c_1 & & \\ a_1 & \ddots & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix}=\begin{bmatrix} 1 & & & \\ \gamma_1 & 1 & & \\ & \ddots & \ddots & \\ & & \gamma_{n-1} & 1 \end{bmatrix}\begin{bmatrix} \alpha_1 & c_1 & & \\ & \alpha_2 & \ddots & \\ & & \ddots & c_{n-1} \\ & & & \alpha_n \end{bmatrix}. \tag{2.6}$$

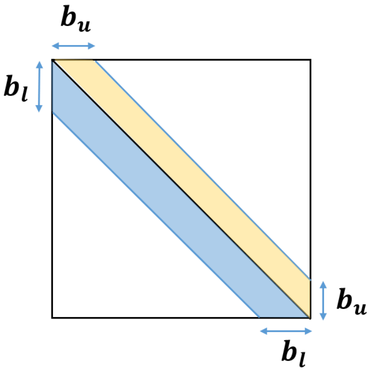
但此时 $|\gamma_i|$ 可能大于 1. 比如 $\gamma_1=a_1/b_1$, 因此当 $|b_1|<|a_1|$ 时, $|\gamma_1|>1$. 所以在回代求解时, 误差可能得不到有效控制. 另外一方面, 计算 γ_i 时也可能会产生较大的舍入误差 (大数除以小数). 但如果 A 是列对角占优, 则可以保证 $|\gamma_i|<1$.

✍ 如果 A 是 (行) 对角占优, 则采用分解 (2.5);
如果 A 是列对角占优, 则采用分解 (2.6).

2-2-5 带状线性方程组

特殊线性方程组: 带状线性方程组

带状矩阵: 下带宽为 b_l, 上带宽为 b_u ⟷ $a_{ij} = 0 \quad \text{for} \quad i > j + b_l \text{ or } i < j - b_u$



带状矩阵的 LU 分解

定理 设 $A \in \mathbb{R}^{n\times n}$ 是带状矩阵, 其下带宽为 b_l, 上带宽为 b_u. 若 $A = LU$ 是不选主元的 LU 分解, 则 L 为下带宽为 b_l 的带状矩阵, U 为上带宽为 b_u 的带状矩阵.

带状矩阵的 LU 分解

定理 设 $A \in \mathbb{R}^{n\times n}$ 是带状矩阵, 其下带宽为 b_l, 上带宽为 b_u. 若 $A = LU$ 是不选主元的 LU 分解, 则 L 为下带宽为 b_l 的带状矩阵, U 为上带宽为 b_u 的带状矩阵.

部分选主元 LU 分解的性质

定理 设 $A \in \mathbb{R}^{n\times n}$ 是带状矩阵, 其下带宽为 b_l, 上带宽为 b_u. 若 $PA = LU$ 是部分选主元的 LU 分解, 则 U 为上带宽不超过 $b_l + b_u$ 的带状矩阵, L 为下带宽 b_l 的 "基本带状矩阵", 即 L 每列的非零元素不超过 $b_l + 1$ 个. (非零元素的个数有限制, 但位置不限)

谢谢

THANK YOU