

线性方程组的直接解法

线性方程组直接解法

线性方程组可使用矩阵表示为

$$Ax = b$$

若矩阵非奇异，即

$$\det(A) \neq 0$$

则方程组有唯一解。克莱姆法则计算量过大 $((n^2 - 1)n!)$ ，需要构造出适合计算机使用的求解线性方程组的方法

解线性方程组的直接法即不计舍入误差，通过有限次算术运算求得精确解，适用于低阶稠密方程组。除此之外还有迭代法，通过设计迭代公式，产生逼近精确解的近似序列，迭代法适用于大型稀疏方程组

Gauss 消去

基本思想

Gauss 消去的基本思想即是通过逐次消元将线性方程组转化成等价的上三角方程组。由消元和回代两个过程组成。即将原方程消去成如下形式

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\&\vdots \\a_{nn}x_n &= b_n\end{aligned}$$

然后再回代求解，要求主对角元素均不为 0

顺序 Gauss 消去

第一步，设 $a_{11} \neq 0$ ，矩阵为 $[A^{(1)}, b^{(1)}]$ ，将第一列中 $a_{11}^{(1)}$ 之下的元素均消为 0，即依次用

$$-l_{i1} = -\frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \quad i = 2, 3, \dots, n$$

乘以第 1 行再加到第 i 行, 得到矩阵 $[A^{(2)}, b^{(2)}]$

之后每一步以此类推, 对于第 k 步, 将第 k 列中 $a_{kk}^{(k)}$ 以下的元素均消为 0, 即

$$\begin{aligned} a_{ij}^{(k+1)} &= a_{ij}^{(k)} - l_{ik} a_{kj}^{(k)} & i, j = k+1, k+2, \dots, n \\ b_i^{(k+1)} &= b_i^{(k)} - l_{ik} b_k^{(k)} & i = k+1, k+2, \dots, n \end{aligned}$$

其中

$$l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$$

最终得到 $[A^{(n)}, b^{(n)}]$, 然后进行回代求解

$$\begin{aligned} x_n &= \frac{b_n^{(n)}}{a_{nn}^{(n)}} \\ x_i &= \frac{1}{a_{ii}^{(i)}} \left(b_i^{(i)} - \sum_{j=i+1}^n a_{ij}^{(j)} x_j \right) & i = n-1, n-2, \dots, 1 \end{aligned}$$

顺序 Gauss 消去要求所有主对角元不为 0, 即 $a_{ii}^{(i)} \neq 0, i = 1, 2, \dots, n$

约化的主元素 $a_{ii}^{(i)} \neq 0 \iff D_i \neq 0, i = 1, 2, \dots, n$, 其中 D_i 为顺序主子式, 并且有

$$\begin{cases} a_{11}^{(1)} = D_1 \\ a_{kk}^{(k)} = D_k / D_{k-1} & (k = 2, 3, \dots, n) \end{cases}$$

顺序 Gauss 的运算量大约为 $\frac{1}{3}(n^3 + 3n^2 - n)$ 。

若 $a_{kk}^{(k)} \neq 0$, 可通过 Gauss 消去法将原方程化为等价的三角方程

若 $\det(A) \neq 0$, 可通过 Gauss 消去和**行交换的初等变换**将原方程化为等价的三角方程

列主元的 Gauss 消去

顺序 Gauss 消去中 $a_{kk}^{(k)}$ 称为主元素, 在第 k 步消元时主元素作为分母, 此时可能会有以下几种情况

- $a_{kk}^{(k)} = 0$, 消元过程不能继续
- $|a_{kk}^{(k)}|$ 过小, 作为分母会导致舍入误差过大, 精度下降

为了避免这种情况, 可将原方程等价变换, 再用顺序消去法求解, 即是列主元 Gauss 消去

列主元的步骤为在第 i 步消元时, 选择主元

$$|a_{ki}^{(i)}| = \max_{i \leq j \leq n} |a_{ji}^{(i)}|$$

然后交换第 k 行和第 i 行

$$a_{kj}^{(i)} \leftrightarrow a_{ij}^{(i)} \quad j = i, i+1, \dots, n$$

$$b_k^{(i)} \leftrightarrow b_i^{(i)}$$

之后同正常的消元过程一样消元即可

只要 $\det(A) \neq 0$ ，列主元 Gauss 消去即可完成

Gauss-Jordan 消去

Gauss-Jordan 的消去过程中，第 k 步消去时，将第 k 列除 $a_{kk}^{(k)}$ 的元素均变为 0，即

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik} a_{kj}^{(k)} \quad \begin{cases} i = 1, 2, \dots, k-1, k+1, \dots, n \\ j = k+1, k+2, \dots, n \end{cases}$$

$$b_i^{(k+1)} = b_i^{(k)} - l_{ik} b_k^{(k)} \quad i = k+1, k+2, \dots, n$$

Gauss-Jordan 消去的代价高于 Gauss 消去，最终得到一个对角阵，这种方法便于计算多个系数相同的方程组，也可以用来求矩阵的逆

直接三角分解法

Gauss 消去的矩阵运算

Gauss 消去中每一步运算等价于左乘一个矩阵，如第一步消去等价于左乘矩阵

$$L_1 = \begin{bmatrix} 1 & & & & \\ -l_{21} & 1 & & & \\ -l_{31} & 0 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ -l_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

同理，第二步得到

$$L_2 = \begin{bmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & -l_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & -l_{n2} & 0 & \cdots & 1 \end{bmatrix}$$

依此类推，即

$$A^{(n)} = L_{n-1} \dots L_2 L_1 A$$

$$A = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1} A^{(n)}$$

令

$$L = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1}$$

则

$$L = \begin{bmatrix} 1 & & & & \\ -l_{21} & 1 & & & \\ -l_{31} & -l_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ -l_{n1} & -l_{n2} & -l_{n3} & \dots & 1 \end{bmatrix}$$

则系数矩阵 A 可分解为一个单位下三角矩阵 L 和一个上三角矩阵 $U(A^{(n)})$ 的乘积

$$U = A^{(n)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n)} \end{bmatrix}$$

将 $A = LU$ 称为矩阵的三角分解，此时方程组为

$$Ax = b \Rightarrow LUx = b$$

令

$$Ux = y$$

则有

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

两个方程均易求解。

对于 $Ly = b$

$$y_1 = b_1$$

$$y_k = b_k - \sum_{i=1}^{k-1} l_{ki} y_i \quad k = 2, 3, \dots, n$$

对于 $Ux = y$

$$x_n = \frac{y_n}{u_{nn}}$$

$$x_i = \frac{1}{u_{ii}} \left(y_i - \sum_{j=i+1}^n u_{ij} x_j \right) \quad i = n-1, n-2, \dots, 1$$

系数矩阵的三角分解很关键

Doolittle 分解

若顺序 Gauss 消去法每一步主元均不为 0，则矩阵 A 可分解为 $A = LU$ ，且 L 为**单位下三角矩阵**， U 为上三角矩阵，这样的分解称为 Doolittle 分解

若 A 的各阶顺序主子式不为 0，则存在唯一的 Doolittle 分解，可按如下公式进行

$$u_{1j} = a_{1j} \quad j = 1, 2, \dots, n$$

$$l_{i1} = \frac{a_{i1}}{u_{11}} \quad i = 2, 3, \dots, n$$

对于 $i = 2, 3, \dots, n$

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \quad j = i, i+1, \dots, n$$

对于 $j = 2, 3, \dots, n$

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

这样的分解过程叫做 Doolittle 分解。

存储时为了方便可以将 L, U 存储在同一矩阵。分解后再分别求解，运算量与 Gauss 消去相近

平方根法

当线性方程组的系数矩阵为对称正定矩阵时，将得出更有效的三角分解法。

若对称正定矩阵有如下分解

$$A = LL^T$$

其中 L 为下三角矩阵，则称其为对称正定矩阵的 Cholesky 分解，原方程等价于 $LL^T x = b$

则可以得到计算公式为

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

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

若将方程组化为 $Ly = b, L^T x = y$, 则解为

$$y_j = \left(b_j - \sum_{k=1}^{j-1} l_{jk} y_k \right) / l_{jj}, \quad j = 1, 2, \dots, n$$

$$x_j = \left(y_j - \sum_{k=j+1}^n l_{kj} y_k \right) / l_{ij}, \quad j = n, n-1, \dots, 1$$

追赶法

追赶法一般用来解三对角方程组，其系数矩阵有如下形式

$$A = \begin{bmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & a_3 & b_3 & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ & & & a_n & b_n \end{bmatrix}$$

对于一个三对角矩阵，其顺序主子式不为 0 的条件为

$$|b_1| > |c_1| > 0$$

$$|b_n| > |a_n| > 0$$

$$|b_i| \geq |c_i| + |a_i|, c_i a_i \neq 0$$

在此条件下可进行三角分解

$$A = \begin{bmatrix} \alpha_1 & & & & \\ \gamma_2 & \alpha_2 & & & \\ & \gamma_3 & \alpha_3 & & \\ & & \ddots & \ddots & \\ & & & \gamma_n & \alpha_n \end{bmatrix} \begin{bmatrix} 1 & \beta_1 & & & \\ & 1 & \beta_2 & & \\ & & 1 & \ddots & \\ & & & \ddots & \beta_{n-1} \\ & & & & 1 \end{bmatrix}$$

根据矩阵乘法可得

$$\begin{aligned}\gamma_i &= a_i, i = 2, 3, \dots, n \\ \alpha_1 &= b_1 \\ \alpha_i &= b_i - \gamma_i \beta_{i-1}, i = 2, 3, \dots, n \\ \beta_i &= \frac{c_i}{\alpha_i}, i = 1, 2, \dots, n-1\end{aligned}$$

于是可得到解为

$$\begin{cases} y_1 = b_1 / \alpha_1 \\ y_i = (b_i - \gamma_i y_{i-1}) / \alpha_i \\ x_n = y_n \\ x_i = y_i - \beta_i x_{i+1} \end{cases}$$

当三对角矩阵满足对角占优条件时，追赶法是数值稳定的。追赶法计算简单，存储要求少

应用

求矩阵行列式

对矩阵 A 做三角分解 $A = LU$

可求得行列式为 $\det(A) = \det(L) \cdot \det(U) = u_{11} u_{22} \dots u_{nn}$

求逆矩阵

求解方程组 $AA^{-1} = E$ 即可 (Gauss-Jordan)

向量与矩阵范数

向量范数

范数是三维欧氏空间中对向量长度的推广

设 $x = (x_1, x_2, \dots, x_n)^T, y = (y_1, y_2, \dots, y_n)^T \in R^n$, 将实数

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

或复数

$$(x, y) = y^H x = \sum_{i=1}^n x_i \bar{y}_i$$

称为向量 x, y 的数量积，将非负实数

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

称为向量 x 的欧氏范数, 满足

- $(x, x) = 0 \iff x = 0$
- $(\alpha x, y) = \alpha(x, y)$
- $(x, y) = (y, x)$
- $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$
- $|(x, y)| \leq \|x\|_2 \cdot \|y\|_2$
- $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$

R^n 空间的向量范数 $\|\cdot\|$ 对任意 $x \in R_n$ 满足

- 正定性: $\|x\| \geq 0, \|x\| = 0 \iff x = 0$
- 齐次性: $\|\alpha x\| = |\alpha| \cdot \|x\|$
- 三角不等式: $\|x + y\| \leq \|x\| + \|y\|$

常用的范数有

- 2-范数: $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$
- 1-范数: $\|x\|_1 = \sum_{i=1}^n |x_i|$
- 无穷大范数: $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$
- p 范数: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$

有性质

- $\|-x\| = \|x\|$
- $||x\| - \|y|| \leq \|x - y\|$
- 向量范数是 R^n 上的连续函数

设 $\|\cdot\|_A, \|\cdot\|_B$ 是两种范数, 若存在 $c_1, c_2 > 0$ 使得

$$c_1 \|x\|_B \leq \|x\|_A \leq c_2 \|x\|_B$$

则称两个范数等价。 R^n 上一切范数均等价

矩阵范数

设 $\|\cdot\|$ 是以 n 阶方阵为变量的实值函数, 且满足

- 非负性: $\|A\| \geq 0, \|A\| = 0 \iff A = 0$
- 齐次性: $\|\alpha A\| = |\alpha| \cdot \|A\|$
- 三角不等式: $\|A + B\| \leq \|A\| + \|B\|$
- 相容性: $\|AB\| \leq \|A\| \|B\|$

则称其为矩阵的范数

若 $\|\cdot\|$ 是一种向量范数，且

$$\|A\| = \sup_{x \in R^n, x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \in R^n, \|x\|=1} \|Ax\|$$

称之为由向量范数派生的**矩阵算子范数**

常用矩阵范数有

- 列范数: $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$
- 行范数: $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$
- 2-范数: $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$, 其中 λ_{\max} 是特征值的绝对值的最大值
- Frobenius 范数: $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2}$

设矩阵 $A \in R^{n \times n}$ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则称

$$\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$$

为矩阵 A 的谱半径。显然

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)} = \sqrt{\rho(A^T A)}$$

对任意算子范数有

$$\rho(A) \leq \|A\|$$

若 A 对称，则有

$$\|A\|_2 = \rho(A)$$

误差分析

定义

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

为矩阵 A 的条件数，当 $\text{cond}(A) \gg 1$ 时称矩阵是病态的，当 $\text{cond}(A)$ 较小时称矩阵是良态的。病态的矩阵在计算时结果收到误差的影响较大

通常使用的条件数有

$$\begin{aligned} \text{cond}(A)_\infty &= \|A\|_\infty \cdot \|A^{-1}\|_\infty \\ \text{cond}(A)_2 &= \|A\|_2 \cdot \|A^{-1}\|_2 = \sqrt{\lambda_{\max}(A^T A) / \lambda_{\min}(A^T A)} \end{aligned}$$

特别的，若 A 对称，则

$$\text{cond}(A)_2 = \frac{\max |\lambda|}{\min |\lambda|}$$

一般判断矩阵是否病态，根据经验得出

- 行列式很大或很小（如某些行、列近似相关）
- 元素间相差大数量级，且无规则
- 主元消去过程中出现小主元
- 特征值相差大数量