解线性方程组的迭代法

迭代法

对于线性方程组

$$Ax = b$$

其中 A 非异，当 A 为低阶稠密矩阵时，选主元消去法是有效的，但对于大型稀疏矩阵方程，迭代法求解效果较好

迭代法要求 A 有某种性质，保证迭代过程的收敛

迭代法构造原则

对于方程组

$$Ax = b$$

将其改写为等价形式

$$x = Bx + f$$

然后得到迭代格式

$$x^{(k+1)} = Bx^{(k)} + f$$

B 称为**迭代矩阵**，利用上述公式求解即为迭代法（一阶定常迭代法）

迭代法收敛性

若

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

存在则称迭代法收敛。研究序列 $\{x^{(k)}\}$ 的收敛性，引入误差向量

$$\varepsilon^{(k+1)} = x^{(k+1)} - x^*$$

则可得

$$\varepsilon^{(k)} = B\varepsilon^{(k-1)} = \cdots = B^k\varepsilon^{(0)}$$

要求迭代收敛即要求

$$\lim_{k\to\infty} \varepsilon^{(k)} = 0$$

等价于

$$\lim_{k\to\infty} B^k = 0$$

基本迭代法

构造迭代法

对于解线性方程组

$$Ax = b$$

基本思路是将 A 分裂

$$A = M - N$$

其中 M 为可选的非异矩阵，且 $Mx = d$ 易于求解，将 M 称为**分裂矩阵**，于是方程组求解转化为

$$Mx = Nx + b$$

即

$$x = M^{-1}Nx + M^{-1}b$$

可构造迭代法

$$x^{(k+1)} = Bx^{(k)} + f$$

其中

$$B = M^{-1}N = M^{-1}(M - A) = I - M^{-1}A$$
$$f = M^{-1}b$$

称 B 为迭代矩阵，根据分裂矩阵的不同可得到不同的迭代法。

设 A 的主对角元素不为 0，可将 A 写为三部分

$$
\begin{aligned}
A = & \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix} - \begin{pmatrix} 0 & & & \\ -a_{21} & 0 & & \\ \vdots & \vdots & \ddots & \\ -a_{n1} & -a_{n2} & \cdots & 0 \end{pmatrix} \\
& - \begin{pmatrix} 0 & \cdots & -a_{1,n-1} & -a_{1,n} \\ & \ddots & \vdots & \vdots \\ & & 0 & -a_{n-1,n} \\ & & & 0 \end{pmatrix} \\
& = D - L - U
\end{aligned}
$$

D, L, U 分别对应主对角线，下三角，上三角三部分

Jacobi 迭代法

选取 $M = D$，则 $N = M - A = L + U$

则迭代矩阵

$$
J = I - D^{-1}A = D^{-1}(L+U)
$$
$$
f = D^{-1}b
$$

故其形式可写为

$$
x^{(k+1)} = D^{-1}(L+U)x^{(k)} + D^{-1}b
$$

其分量计算公式为

$$
\begin{aligned}
x_i^{(k+1)} &= \frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i\right) \quad (i = 1, 2, \cdots, n) \\
&= \frac{1}{a_{ii}}\left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j^{(k)}\right)
\end{aligned}
$$

Jacobi 迭代法计算简单，且迭代一次只需计算一次矩阵与向量的乘法

Gauss-Seidel 迭代法

基本思路为，在 Jacobi 迭代法计算 $x_i^{(k+1)}$ 时，用已经算出的 $x_j^{(k+1)}$ 代替 $x_j^{(k)}$，$1 \leqslant j \leqslant i-1$

即

$$x_i^{(k+1)}=\frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}+b_i\right)\quad(i=1,2,\cdots,n)$$

以矩阵形式表示，则方程由 Jacobi 的

$$Dx^{(k+1)}=(Lx^{(k)}+Ux^{(k)})+b$$

变为

$$Dx^{(k+1)}=(Lx^{(k+1)}+Ux^{(k)})+b$$

即

$$(D-L)x^{(k+1)}=Ux^{(k)}+b$$
$$x^{(k+1)}=(D-L)^{-1}Ux^{(k)}+(D-L)^{-1}b$$

可得迭代矩阵

$$G=(D-L)^{-1}U$$
$$f=(D-L)^{-1}b$$

等价于分裂矩阵选择了 A 的下三角部分（$D-L$，包括主对角线）

Gauss-Seidel 迭代法是 Jacobi 迭代法的一种改进，计算量不变，但收敛加快（不是绝对的）

逐次超松弛法

考察 Gauss-Seidel 迭代法

$$x_i^{(k+1)}=\frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}+b_i\right)\quad(i=1,2,\cdots,n)$$

可以看作是在第 k 次迭代的基础之上添加了一个修正项

$$\Delta=\frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum_{j=i}^{n}a_{ij}x_j^{(k)}+b_i\right)$$

即

$$x_i^{(k+1)}=x_i^{(k)}+\Delta$$

而为了加速迭代，对修正项 Δ 乘以调节因子 ω

$$x_i^{(k+1)} = x_i^{(k)} + \omega \cdot \Delta$$

ω 称为松弛因子，该方法称为逐次超松弛迭代法，简称 SOR (Successive Over Relaxation) 迭代法

- $\omega < 1$ 时为低松弛迭代法
- $\omega = 1$ 时为 Gauss-Seidel 迭代法
- $\omega > 1$ 时为 SOR

则令

$$\begin{cases} y_i^{(k+1)} = \frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i\right) \\ x_i^{(k+1)} = x_i^{(k)} + \omega(y_i^{(k+1)} - x_i^{(k)}) \end{cases}$$

可看作

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \omega y_i^{(k+1)}$$

是一个加权平均

从矩阵视角来看

$$\begin{aligned} Dx^{(k+1)} &= \omega(Lx^{(k+1)} + Ux^{(k)} + b) + (1-\omega)Dx^{(k)} \\ (D-\omega L)x^{(k+1)} &= (\omega U + (1-\omega)D)x^{(k)} + \omega b \\ x^{(k+1)} &= (D-\omega L)^{-1}(\omega U + (1-\omega)D)x^{(k)} + \omega(D-\omega L)^{-1}b \end{aligned}$$

其分裂矩阵为带参的下三角矩阵

$$M = \frac{1}{\omega}(D - \omega L)$$

而迭代矩阵为

$$\begin{gathered} L_\omega = (D-\omega L)^{-1}[\omega U + (1-\omega)D] \\ f = \omega(D-\omega L)^{-1}b \end{gathered}$$

其分量公式为

$$x_i^{(k+1)} = x_i^{(k)} + \omega \times \frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)} + b_i\right)$$

迭代法的收敛性

一阶定常迭代法的基本定理

由上文的讨论，一阶定长迭代法若收敛，需要

$$\lim_{k\to\infty} B^k = 0$$

即矩阵序列 $\{B^k\}$ 收敛。考虑矩阵范数

$$\lim_{k\to\infty} A_k = A \iff \lim_{k\to\infty} \|A_k - A\| = 0$$

或是

$$\lim_{k\to\infty} A_k = A \iff \forall x \in \mathbb{R}^n, \lim_{k\to\infty} A_k x = Ax$$

或是

$$\lim_{k\to\infty} B^k = 0 \iff \rho(B) < 1$$

其中 $\rho(B)$ 为矩阵的谱半径。上述办法均可判断收敛性

> 设矩阵 A 有特征值 $\lambda_1, \lambda_2, \ldots, \lambda_n$ ，则 $\rho(A) = \max\{|\lambda_1|, \ldots, |\lambda_n|\}$

迭代法收敛 $\iff$ 其迭代矩阵谱半径 $\rho(B) < 1$

根据谱半径性质，有

$$\rho(B) \leqslant \|B\|$$

可根据矩阵范数得到迭代法收敛的充分条件：任意一种范数小于 1

$\rho(B)$ 可用来衡量收敛速度，其值越小收敛越快

对于一阶定常迭代法，若存在某个范数 $\|B\| = q < 1$ ，则有

$$\begin{gathered}
\left\|x^* - x^{(k)}\right\| \leqslant q^k \left\|x^* - x^{(0)}\right\| \\
\left\|x^* - x^{(k)}\right\| \leqslant \frac{q}{1-q}\left\|x^{(k)} - x^{(k-1)}\right\| \\
\left\|x^* - x^{(k)}\right\| \leqslant \frac{q^k}{1-q}\left\|x^{(1)} - x^{(0)}\right\|
\end{gathered}$$

可用来估计迭代次数或精度

特殊方程迭代法的收敛性

设矩阵 $A = (a_{ij})_{n\times n}$

若满足

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}| \quad (i=1,2,\dots,n)$$

则称 A 为**严格（按行）对角占优阵**

若满足

$$|a_{ii}| \geqslant \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}| \quad (i=1,2,\dots,n)$$

且上式至少有一个大于号成立，则称 A 为**弱（按行）对角占优阵**

当 $n \geqslant 2$ 时，若存在置换阵 P 使得

$$P^T A P = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

其中 A_{11} 为 r 阶方阵，A_{22} 为 $n-r$ 阶方阵，则称 A 为**可约矩阵**，否则为**不可约矩阵**

> 设 $A=(a_{ij})_{n\times n}$ ，$N=\{1,2,\dots,n\}$ ，若存在集合 N 的子集 I,J 满足
>
> $$I \cup J = N \quad I \cap J = \varnothing$$
>
> 则 A 可约 $\iff \forall i\in I, j\in J, a_{ij}=0$

对角占优定理：若 A 为**严格对角占优阵**或**不可约弱对角占优阵**则 A 为非奇异矩阵

设方程组

$$Ax=b$$

- 若 A 为严格对角占优阵，则 Jacobi 和 Gauss-Seidel 均收敛
- 若 A 为弱对角占优阵，且 A 不可约，则 Jacobi 和 Gauss-Seidel 均收敛
- 若 A 为正定矩阵，则 Gauss-Seidel 收敛

SOR 收敛的必要条件：若 SOR 收敛，则 $0<\omega<2$

SOR 收敛的充分条件：若 A 为对称正定矩阵，$A=D-L-L^T$ ，且 $0<\omega<2$ ，则 SOR 收敛