计算方法

李大明

上海交通大学数学科学学院
Email: lidaming@sjtu.edu.cn

September 8, 2020

解线性方程组的迭代法(第6章)

当矩阵的阶数n很大, Gauss消去法的计算量$O(n^3/3)$很大. 当不计舍入误差, Gauss消去法给出精确解.

当A稀疏时, Gauss消去法可能会破坏稀疏结构. (例子)

迭代法:

$$x^{(0)} \to x^{(1)} \to \cdots \to x$$

每次迭代只需要矩阵和向量相乘, 计算量为$O(n^2)$. 若A稀疏时, 每次迭代的计算量更少.

误差$\|x^{(1)} - x\|$!

稀疏矩阵的存贮(按行压缩存贮)

$$A = \begin{pmatrix} a_{11} & 0 & a_{13} & 0 & 0 \\ a_{21} & a_{22} & 0 & a_{24} & 0 \\ 0 & a_{32} & a_{33} & 0 & a_{35} \\ 0 & a_{42} & 0 & a_{44} & 0 \\ 0 & 0 & 0 & a_{54} & a_{55} \end{pmatrix}.$$

$$\begin{aligned} a &= \{a_{11}, a_{13} | a_{21}, a_{22}, a_{24} | a_{32}, a_{33}, a_{35} | a_{42}, a_{44} | a_{54}, a_{55}\} \\ ja &= \{1, 3 |\ 1, 2, 4 |\ 2, 3, 5 |\ 2, 4 |\ 4, 5\} \\ ia &= \{1, 3, 6, 9, 11, 13\} \Longleftrightarrow \{2, 3, 3, 2, 2\} \end{aligned}$$

ja为非零元的列指标, $ia[i]$为A 的第i 行第一个非零元在A 和ja 中的位置

$$(Ax)_i = \sum_{k=ia[i]}^{ia[i+1]-1} a[k] x_{ja[k]}$$

A 的第i行就是a的第$ia[i]$个和第$ia[i+1]-1$个之间的元素.

$$\begin{pmatrix} 7 & -6 \\ -8 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -4 \end{pmatrix} \tag{1}$$

Jacobi 迭代:

$$\begin{cases} x_1^{(k)} = \frac{6}{7} x_2^{(k-1)} + \frac{3}{7} \\ x_2^{(k)} = \frac{8}{9} x_1^{(k-1)} - \frac{4}{9} \end{cases}$$

Gauss-Seidel 迭代:

$$\begin{cases} x_1^{(k)} = \frac{6}{7} x_2^{(k-1)} + \frac{3}{7} \\ x_2^{(k)} = \frac{8}{9} x_1^{(k)} - \frac{4}{9} \end{cases}$$

Jacobi 迭代和Gauss-Seidel 迭代都收敛到方程的解,
Gauss-Seidel 迭代收敛快一些.

给定A的一个分裂

$$A = M - N, \tag{2}$$

$Ax = b$ 可表示为

$$Mx = Nx + b. \tag{3}$$

迭代格式

$$Mx^{(k+1)} = Nx^{(k)} + b, \quad k = 0, 1, \cdots, \tag{4}$$

$Mx = Nx + b$ 等价于

$$x = Bx + f, \quad B = M^{-1}N, \ f = M^{-1}b, \tag{5}$$

迭代格式也可写为

$$x^{(k+1)} = Bx^{(k)} + f, \quad k = 0, 1, \cdots. \tag{6}$$

若对于任意的初始向量$x^{(0)}$, 由(6)得到的序列都收敛, 称迭代格式(6)收敛. 若收敛, 则极限必定是方程$Ax = b$的解.

误差

$$\varepsilon^{(k)} = x^{(k)} - x^*, \tag{7}$$

x^* 为准确解. 误差传播

$$\varepsilon^{(k+1)} = B\varepsilon^{(k)} \Longleftrightarrow \varepsilon^{(k)} = B^k\varepsilon^{(0)}. \tag{8}$$

B为迭代矩阵.

若对于任意$1 \le i, j \le n$,

$$a_{ij}^{(k)} \to a_{ij}, \quad k \to \infty$$

称$A^{(k)} = (a^{(k)})$收敛到$A = (a_{ij})$.

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad A^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix} \to 0$$

$$\begin{aligned} A^{(k)} \to A \quad & \Longleftrightarrow \quad \|A^{(k)} - A\|_\infty \to 0 \\ & \Longleftrightarrow \quad \|A^{(k)} - A\|_p \to 0, \quad 1 \le p \le \infty \\ & \Longleftrightarrow \quad A^{(k)}x \to Ax, \quad \forall x \end{aligned}$$

迭代格式(6)对任意$x^{(0)}$都收敛的充分必要条件为$\rho(B) < 1$.

$$P^{-1}BP = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix}, J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}_{n_i \times n_i}$$

$$\rho(B) < 1 \Longleftrightarrow B^k \to 0 \Longleftrightarrow \varepsilon^{(k)} = B^k \varepsilon^{(0)} \to 0$$

$$
\begin{aligned}
J_i^k &= (\lambda_i I + E_i)^k = \sum_{j=0}^{n_i-1} C_k^j \lambda_i^{k-j} (E_i)^j \\
&= \begin{pmatrix}
\lambda_i^k & C_k^1 \lambda_i^{k-1} & & & C_k^{n_i-1} \lambda_i^{k-(n_i-1)} \\
 & \lambda_i^k & C_k^1 \lambda_i^{k-1} & & C_k^{n_i-2} \lambda_i^{k-(n_i-2)} \\
 & & \ddots & \ddots & \vdots \\
 & & & \ddots & C_k^1 \lambda_i^{k-1} \\
 & & & & \lambda_i^k
\end{pmatrix}
\end{aligned}
$$

设$u_1,\cdots,u_n$ 为B为的特征向量, 对应的特征值为$\lambda_1,\cdots,\lambda_n$.
$\varepsilon^{(0)}=\sum_{i=1}^n a_i u_i$,

$$\varepsilon^{(k)}=B^k\varepsilon^{(0)}=\sum_{i=1}^n a_i B^k u_i=\sum_{i=1}^n a_i\lambda_i^k u_i. \tag{9}$$

若$\rho(B)(<1)$ 越小, 收敛越快.

$$\Big(\rho(B)\Big)^k\leq 10^{-s}\Longleftrightarrow k\geq\frac{s\ln 10}{-\ln\rho(B)}. \tag{10}$$

(渐近)收敛速度

$$R(B)=-\ln\rho(B). \tag{11}$$

若$\rho(B)$ 越小, $R(B)$ 越大, 收敛越快. 平均收敛速度

$$R_k(B)=-\ln\|B^k\|^{1/k}\to R(B),\quad k\to\infty \tag{12}$$

设$q = \|B\| < 1$,

1. 迭代格式(6) 收敛到x^*.
2. 误差传播

$$\|x^{(k)} - x^*\| \le \frac{q}{1-q}\|x^{(k)} - x^{(k-1)}\|,$$

$$\|x^{(k)} - x^*\| \le \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|.$$

迭代终止的判断?

迭代格式的构造

$$A = M - N,$$

迭代格式

$$Mx^{(k+1)} = Nx^{(k)} + b, \quad k = 0, 1, \cdots,$$

容易计算，如M取为A的对角部分，但是当M简单时，收敛往往很慢!

取$M = A$?

给定A的分裂

$$A = D - L - U$$

$$D = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}, \quad L = -\begin{pmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & & \ddots & \\ a_{n1} & a_{n2} & \cdots & & 0 \end{pmatrix}$$

$$U = -\begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 0 & a_{23} & \cdots & a_{2n} \\ & & \ddots & & \vdots \\ & & & & 0 \end{pmatrix}.$$

$$M = D, \quad N = L + U,$$

(6)中

$$B = D^{-1}(L + U), \quad f = D^{-1}b.$$

(6)的分量形式

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

Jacobi 迭代. 例子(1)

$$A = \begin{pmatrix} 7 & -6 \\ -8 & 9 \end{pmatrix}, \quad M = \begin{pmatrix} 7 & 0 \\ 0 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 6/7 \\ 8/9 & 0 \end{pmatrix}$$

Jacobi迭代收敛, B的特征值

$$\pm\sqrt{\frac{8}{9} \times \frac{6}{7}}$$

$$M = D - L, \quad N = U,$$

(6)中

$$B = (D-L)^{-1}U, \quad f = (D-L)^{-1}b.$$

(6)的分量形式($(D-L)x^{(k+1)} = Ux^{(k)} + b$)

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

Gauss-Seidel 迭代. 例子(1)

$$A = \begin{pmatrix} 7 & -6 \\ -8 & 9 \end{pmatrix}, \quad M = \begin{pmatrix} 7 & 0 \\ -8 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 6/7 \\ 0 & 6*8/(7*9) \end{pmatrix}$$

Gauss-Seidel格式收敛, B的特征值

$$0, \quad 6*8/(7*9) < \sqrt{\frac{8}{9} \times \frac{6}{7}}$$

$$M = \frac{1}{\omega}D - L, \quad N = M - A = \frac{1-\omega}{\omega}D + U,$$

(6)中

$$B = (D - \omega L)^{-1}\Big((1-\omega)D + \omega U\Big), \quad f = \omega(D - \omega L)^{-1}b.$$

(6)的分量形式

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \frac{\omega}{a_{ii}}\Big(b_i - \sum_{j\le i-1} a_{ij}x_j^{(k+1)} - \sum_{j\ge i+1} a_{ij}x_j^{(k)}\Big)$$

松弛因子为ω的SOR 迭代.

三个迭代法的具体执行

和Gauss-Seidel, SOR迭代不同, Jacobi迭代法中分量的更新和次序无关.

SOR迭代为

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

交换L 和U, 得到另一个迭代. SSOR iteration

$$(D-\omega L)x^{(k+\frac{1}{2})} = ((1-\omega)D+\omega U)x^{(k)}+\omega b,$$

$$(D-\omega U)x^{(k+1)} = ((1-\omega)D+\omega L)x^{(k+\frac{1}{2})}+\omega b.$$

消去$x^{(k+\frac{1}{2})}$, 得到(6),

$$M = \frac{1}{\omega(2-\omega)}(D-\omega L)D^{-1}(D-\omega U),$$

$$N = \frac{1}{\omega(2-\omega)}\Big((1-\omega)D+\omega L\Big)D^{-1}\Big((1-\omega)D+\omega U\Big).$$

迭代矩阵

$$B = M^{-1}N = \bar{B}_\omega B_\omega,$$

B_ω 是SOR迭代的迭代矩阵, B_ω中L 和U互换得到$\bar{B}_\omega$.

给定$n \times n$矩阵A. 若

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad i = 1, 2, \cdots n,$$

称A为严格对角占优. 若

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|, \quad i = 1, 2, \cdots n,$$

并且至少有一个等号成立, 称A为弱对角占优.

若存在$n \times n$ 的置换矩阵P 使得

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

A_{11} 为$r \times r$矩阵, $1 \le r < n$, 称A 为可约的; 否则称A不可约.

A 可约的充分必要条件为存在一个非空子集$J \subsetneq \{1, 2, \cdots, n\}$, 使得

$$a_{ij} = 0, \quad i \in J, j \notin J.$$

$P = (\delta_{p_i,j})$, 其中$(p_1, \cdots, p_n)$ 为$(1, 2, \cdots, n)$的排列.

$$(P^TAP)_{ij} = \sum_{k,l}(P^T)_{ik}A_{kl}P_{lj} = \sum_{k,l}\delta_{p_k,i}a_{kl}\delta_{p_l,j} = a_{p_i^{-1},p_j^{-1}}$$

$$p_i^{-1} = k \Longleftrightarrow p_k = i$$

A 可约当且仅当存在$r(1 \le r < n)$使得

$$a_{p_i^{-1},p_j^{-1}} = 0, \quad r+1 \le i \le n, \ 1 \le j \le r$$

取$J = \{p_i^{-1}\}_{i=r+1}^{n}$. 若非空真子集$J$存在, 可取相应的$(p_1, \cdots, p_n)$.

若A 严格对角占优或不可约弱对角占优, 则A非奇异. 对任意$x^{(0)}$, Jacobi 迭代, Gauss-Seidel 迭代和松弛因子为$0 < \omega (0 < \omega \leq 1)$的SOR 迭代都收敛.

例子(1)中

$$A = \begin{pmatrix} 7 & -6 \\ -8 & 9 \end{pmatrix}$$

它的Jacobi 迭代, Gauss-Seidel 迭代和松弛因子为$0 < \omega (0 < \omega \leq 1)$ 的SOR 迭代都收敛.

收敛性

若A 奇异, 存在非零向量x使得$Ax = 0$. 记$|x_k| = \max_j |x_j| \neq 0$. $Ax = 0$的第k个方程

$$|a_{kk}x_k| \leq \sum_{j\neq k} |a_{kj}x_j| \leq |x_k| \sum_{j\neq k} |a_{kj}| \tag{13}$$

和A为严格对角占优矛盾.

若A为弱对角占优，要证A非奇异. 否则，存在非零向量x使得$Ax = 0$. 由于A为弱对角占优，存在m使得

$$|a_{mm}| > \sum_{j \neq m} |a_{mj}| \tag{14}$$

定义

$$J = \{k \big| |x_k| \geq |x_i|, i = 1, \cdots, n; |x_k| > |x_j| \ for\ some\ j\}$$

显然，J非空. 否则，$|x_1| = \cdots = |x_n|$ $(x \neq 0)$ 和(13) 中$k = m$和(14)矛盾. 当然，$J \neq \{1, \cdots, n\}$.

任取$k \in J$, $x_k \neq 0$,

$$|a_{kk}| \leq \sum_{j \neq k} |a_{kj}|\,|x_j|/|x_k| \tag{15}$$

若$|x_k| > |x_j|$, 则$a_{kj} = 0$. 否则,

$$|a_{kk}| \leq \sum_{j' \neq k} |a_{kj'}|\,|x_{j'}|/|x_k| = |a_{kj}|\,|x_j|/|x_k| + \sum_{j' \neq k,j} |a_{kj'}|\,|x_{j'}|/|x_k| < \sum_{j' \neq k} |a_{kj'}|$$

这和A为弱对角占优矛盾. 对任意$k \in J$, $j \notin J$, $|x_k| > |x_j|$,

$$a_{kj} = 0, \quad k \in J, j \notin J$$

这表明A 为可约, 矛盾.

设A 为不可约弱对角占优, 证明Gauss-Seidel 迭代收敛. 它的迭代矩阵

$$B = (D-L)^{-1}U$$

$$\det(\lambda I - B) = 0 \iff \det(\lambda(D-L) - U) = 0 \tag{16}$$

要证明B的所有特征值的模都小于1.

$$H = \lambda(D-L) - U = \begin{pmatrix} \lambda a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & a_{23} & \dots & a_{2n} \\ \dots & & & & \\ \lambda a_{n1} & \lambda a_{n2} & \lambda a_{n3} & \dots & \lambda a_{nn} \end{pmatrix} \tag{17}$$

若$|\lambda| \geq 1$, H为不可约弱对角占优, 所以$\det(H) \neq 0$. 和(16)矛盾, 所以, $|\lambda| < 1$, Gauss-Seidel 迭代收敛.

若SOR 对任意初始向量都收敛, 则

$$0 < \omega < 2.$$

SOR 迭代矩阵B 满足

$$|\det(B)| = |\lambda_1\lambda_2\dots\lambda_n| \leq (\rho(B))^n,$$

$$\det(B) = \det((D-\omega L)^{-1})\det((1-\omega)D + \omega U) = (1-\omega)^n$$

由$\rho(B) < 1$得到

$$|1-\omega| < 1$$

若A 为正定的Hermite矩阵, 若$0 < \omega < 2$, 松弛因子为ω的SOR 迭代法对任意初始向量都收敛.

二维Poisson 方程的离散, 以及各种迭代法构造(参见教材P. 197)

最速下降法和共轭梯度法.

作业

李庆杨等，数值分析，第5版，P.210, 1,2(1),3,4,7,9,