迭代法解线性方程组

教师: 胡俊

北京大学数学科学学院

October 18, 2019

目录

目录

$$Ax = b, \quad A = D - L - U$$

其中

$$D = \mathrm{diag}(a_{11}, a_{22}, \cdots, a_{nn})$$

$$L = \begin{bmatrix} 0 & & & & \\ -a_{21} & 0 & & & \\ -a_{31} & -a_{32} & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ -a_{n1} & -a_{n2} & \cdots & -a_{n,n-1} & 0 \end{bmatrix}$$

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

则有

$$x = Bx + g$$

其中$B = D^{-1}(L+U)$, $g = D^{-1}b$. 若给定初始向量

$$x_0 = (x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})^T.$$

$$x_1 = Bx_0 + g$$

一般地, 若x_{k-1}已知, 则

$$x_k = Bx_{k-1} + g, \ k = 1, 2 \cdots$$

这就是Jacobi迭代格式, B称为Jacobi迭代的迭代矩阵, g称为常数项.

$$x_k = D^{-1}Lx_k + D^{-1}Ux_{k-1} + g \quad (g = D^{-1}b)$$

$$(I - D^{-1}L)x_k = D^{-1}Ux_{k-1} + g \Rightarrow x_k = (D-L)^{-1}Ux_{k-1} + (D-L)^{-1}b.$$

$(D-L)^{-1}U \rightarrow$ G-S迭代矩阵

$(D-L)^{-1}b \rightarrow$ G-S迭代常数项

目录

$$x_{k+1} = Mx_k + g$$

Jacobi迭代: $M = D^{-1}(L+U)$, $g = D^{-1}b$
G-S迭代: $M = (D-L)^{-1}U$, $g = (D-L)^{-1}b$

M是迭代矩阵, g是常数项, x_0是初始向量.

$Ax = b$, 若$G(I-M)X = Gg$, 即$A = G(I-M)$, $b = Gg$, 则称迭代法与线性方程组是相容的. Jacobi与G-S迭代都是相容的.

如果迭代法收敛, 并记其极限为x_*, 则有(下面总假定迭代法相容)

$$x_* = Mx_* + g$$

则x_*是$Ax = b$的唯一解(设A非奇异). 令$y_k = x_k - x_*$, 于是有

$$y_{k+1} = My_k \quad k = 0, 1, 2, \cdots \implies y_k = M^k y_0.$$

因此对$\forall y_0$, 都有$y_k \to 0$(即$x_k \to x_*$) 的充分必要条件是$M^k \to 0$.

Theorem 2.1

单步线性迭代法收敛的充分必要条件是$\rho(M) < 1$.

$$A_1 = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

A_1对Jacobi迭代收敛, 但G-S迭代不收敛. A_2对G-S迭代收敛, 但Jacobi迭代不收敛.

设方程组$Ax = b$的右端项为$b = [-3, -1, 4, 7]^T$, 系数矩阵为

$$A = \begin{bmatrix} 6 & -3 & -4 & -2 \\ -6 & 8 & 0 & -3 \\ 0 & -4 & 12 & -2 \\ -2 & -2 & -4 & 15 \end{bmatrix}.$$

证明Gauss-Seidel迭代法是收敛的.

Theorem 2.2

若$\|M\| = q < 1$, 且假定$\|I\| = 1$, 则有

$$\|x_k - x_*\| \leq \frac{q^k}{1-q}\|x_1 - x_0\| \longrightarrow \text{可计算的界}$$

证明:

$$\|y_k\| = \|M^k y_0\| \leq \|M^k\|\|y_0\| = q^k\|y_0\|$$

下面估计y_0

$$\begin{aligned} y_0 &= x_0 - x_* = x_0 - x_1 + x_1 - x_* \\ &= x_0 - x_1 + M(x_0 - x_*) \\ &\Longrightarrow (I - M)(x_0 - x_*) = x_0 - x_1 \end{aligned}$$

迭代法的停止标准

Theorem 2.2

若$\|M\| = q < 1$, 且假定$\|I\| = 1$, 则有

$$\|x_k - x_*\| \leq \frac{q^k}{1-q}\|x_1 - x_0\| \longrightarrow \text{可计算的界}$$

证明:

$$\|y_k\| = \|M^k y_0\| \leq \|M^k\|\|y_0\| = q^k\|y_0\|$$

下面估计y_0

$$\begin{aligned} y_0 &= x_0 - x_* = x_0 - x_1 + x_1 - x_* \\ &= x_0 - x_1 + M(x_0 - x_*) \\ &\Longrightarrow (I - M)(x_0 - x_*) = x_0 - x_1 \end{aligned}$$

由$\|M\| < 1$可知$(I-M)^{-1}$存在. 这样

$$\begin{aligned}\|x_0 - x_*\| &= \|(I-M)^{-1}(x_0 - x_1)\| \le \|(I-M)^{-1}\|\|x_0 - x_1\| \\ &\le \frac{1}{1-\|M\|}\|x_0 - x_1\| = \frac{1}{1-q}\|x_0 - x_1\|.\end{aligned}$$

证明结束.

Theorem 2.3

若$\|M\| = q < 1$且$\|I\| = 1$, 则有

$$\|x_k - x_*\| \le \frac{q}{1-q}\|x_{k-1} - x_k\| \longrightarrow \text{可计算界, 估计两步之间差}$$

证明:

$$\begin{aligned} x_k - x_* &= x_k - x_{k+1} + x_{k+1} - x_* \\ &= x_k - x_{k+1} + Mx_k - Mx_* \end{aligned}$$

$$\begin{aligned} (I-M)(x_k - x_*) = x_k - x_{k+1} \Longrightarrow \|x_k - x_*\| &\le \|(I-M)^{-1}(x_k - x_{k+1})\| \\ &\le \frac{1}{1-q}\|x_k - x_{k+1}\| \\ &\le \frac{q}{1-q}\|x_{k-1} - x_k\|. \end{aligned}$$

Theorem 2.3

若$\|M\| = q < 1$且$\|I\| = 1$, 则有

$$\|x_k - x_*\| \le \frac{q}{1-q}\|x_{k-1} - x_k\| \longrightarrow \text{可计算界, 估计两步之间差}$$

证明:

$$\begin{aligned} x_k - x_* &= x_k - x_{k+1} + x_{k+1} - x_* \\ &= x_k - x_{k+1} + Mx_k - Mx_* \end{aligned}$$

$$\begin{aligned} (I-M)(x_k - x_*) = x_k - x_{k+1} \Longrightarrow \|x_k - x_*\| &\le \|(I-M)^{-1}(x_k - x_{k+1})\| \\ &\le \frac{1}{1-q}\|x_k - x_{k+1}\| \\ &\le \frac{q}{1-q}\|x_{k-1} - x_k\|. \end{aligned}$$

目录

G-S迭代收敛的两个充分条件

Theorem 3.1

设B是*Jacobi*迭代的迭代矩阵. 若$\|B\|_\infty < 1$, 则*G-S*迭代收敛, 且

$$\|x_k - x_*\|_\infty \le \frac{\mu^k}{1-\mu}\|x_1 - x_0\|_\infty$$

其中

$$\mu = \max_i \left(\sum_{j=i+1}^{n} |b_{ij}| / (1 - \sum_{j=1}^{i-1} |b_{ij}|) \right),$$

且有$\mu \le \|B\|_\infty < 1$, 这里b_{ij}是矩阵B的元素.

证明:

先证$\mu < 1$. 令

$$l_i = \sum_{j=1}^{i-1} |b_{ij}|, \; u_i = \sum_{j=i+1}^{n} |b_{ij}|.$$

显然, 对任意$1 \le i \le n$, 都有

$$l_i + u_i \le \|B\|_\infty < 1.$$

由

$$l_i + u_i - \frac{u_i}{1-l_i} = \frac{(l_i+u_i)(1-l_i)-u_i}{1-l_i} = \frac{l_i(1-l_i-u_i)}{1-l_i} \ge 0$$

可推出

$$\frac{u_i}{1-l_i} \le l_i + u_i, \; i = 1, 2, \cdots, n.$$

两边对i取最大值, 得

$$\mu = \max_i \frac{u_i}{1-l_i} \le \max_i (l_i + u_i) = \|B\|_\infty < 1.$$

证明:

先证$\mu < 1$. 令

$$l_i = \sum_{j=1}^{i-1} |b_{ij}|,\ u_i = \sum_{j=i+1}^{n} |b_{ij}|.$$

显然, 对任意$1 \le i \le n$, 都有

$$l_i + u_i \le \|B\|_\infty < 1.$$

由

$$l_i + u_i - \frac{u_i}{1-l_i} = \frac{(l_i+u_i)(1-l_i) - u_i}{1-l_i} = \frac{l_i(1-l_i-u_i)}{1-l_i} \ge 0$$

可推出

$$\frac{u_i}{1-l_i} \le l_i + u_i,\ i = 1, 2, \cdots, n.$$

两边对i取最大值, 得

$$\mu = \max_i \frac{u_i}{1-l_i} \le \max_i (l_i + u_i) = \|B\|_\infty < 1.$$

证明:

先证$\mu < 1$. 令

$$l_i = \sum_{j=1}^{i-1} |b_{ij}|,\ u_i = \sum_{j=i+1}^{n} |b_{ij}|.$$

显然, 对任意$1 \le i \le n$, 都有

$$l_i + u_i \le \|B\|_\infty < 1.$$

由

$$l_i + u_i - \frac{u_i}{1-l_i} = \frac{(l_i+u_i)(1-l_i)-u_i}{1-l_i} = \frac{l_i(1-l_i-u_i)}{1-l_i} \ge 0$$

可推出

$$\frac{u_i}{1-l_i} \le l_i + u_i,\ i = 1, 2, \cdots, n.$$

两边对i取最大值, 得

$$\mu = \max_i \frac{u_i}{1-l_i} \le \max_i (l_i + u_i) = \|B\|_\infty < 1.$$

再证G-S迭代收敛. 也就是要证明$\rho((D-L)^{-1}U)<1$.
设λ是$(D-L)^{-1}U$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则

$$(D-L)^{-1}Ux=\lambda x$$

或者

$$\lambda x=\lambda D^{-1}Lx+D^{-1}Ux$$

注意到$B=D^{-1}L+D^{-1}U$, 即$D^{-1}L$为B的下三角部分,
而$D^{-1}U$为B的上三角部分, 于是上式的第i个方程为

$$\lambda\xi_i=\lambda\sum_{j=1}^{i-1}b_{ij}\xi_j+\sum_{j=n+1}^{n}b_{ij}\xi_j$$

取绝对值, 得

$$|\lambda|\le|\lambda|l_i+u_i\Rightarrow|\lambda|\le\frac{u_i}{1-l_i}\le l_i+u_i<1.$$

即$\rho((D-L)^{-1}U)<1$. 从而G-S迭代收敛.

再证G-S迭代收敛. 也就是要证明$\rho((D-L)^{-1}U)<1$.
设λ是$(D-L)^{-1}U$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则

$$(D-L)^{-1}Ux=\lambda x$$

或者

$$\lambda x=\lambda D^{-1}Lx+D^{-1}Ux$$

注意到$B=D^{-1}L+D^{-1}U$, 即$D^{-1}L$为B的下三角部分, 而$D^{-1}U$为B的上三角部分, 于是上式的第i个方程为

$$\lambda\xi_i=\lambda\sum_{j=1}^{i-1}b_{ij}\xi_j+\sum_{j=n+1}^{n}b_{ij}\xi_j$$

取绝对值, 得

$$|\lambda|\le|\lambda|l_i+u_i\Rightarrow|\lambda|\le\frac{u_i}{1-l_i}\le l_i+u_i<1.$$

即$\rho((D-L)^{-1}U)<1$. 从而G-S迭代收敛.

再证G-S迭代收敛. 也就是要证明$\rho((D-L)^{-1}U)<1$.
设λ是$(D-L)^{-1}U$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则

$$(D-L)^{-1}Ux=\lambda x$$

或者

$$\lambda x=\lambda D^{-1}Lx+D^{-1}Ux$$

注意到$B=D^{-1}L+D^{-1}U$, 即$D^{-1}L$为B的下三角部分,
而$D^{-1}U$为B的上三角部分, 于是上式的第i个方程为

$$\lambda\xi_i=\lambda\sum_{j=1}^{i-1}b_{ij}\xi_j+\sum_{j=n+1}^{n}b_{ij}\xi_j$$

取绝对值, 得

$$|\lambda|\le|\lambda|l_i+u_i\Rightarrow|\lambda|\le\frac{u_i}{1-l_i}\le l_i+u_i<1.$$

即$\rho((D-L)^{-1}U)<1$. 从而G-S迭代收敛.

再证G-S迭代收敛. 也就是要证明$\rho((D-L)^{-1}U)<1$.
设λ是$(D-L)^{-1}U$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则

$$(D-L)^{-1}Ux=\lambda x$$

或者

$$\lambda x=\lambda D^{-1}Lx+D^{-1}Ux$$

注意到$B=D^{-1}L+D^{-1}U$, 即$D^{-1}L$为B的下三角部分,
而$D^{-1}U$为B的上三角部分, 于是上式的第i个方程为

$$\lambda\xi_i=\lambda\sum_{j=1}^{i-1}b_{ij}\xi_j+\sum_{j=n+1}^{n}b_{ij}\xi_j$$

取绝对值, 得

$$|\lambda|\le|\lambda|l_i+u_i\Rightarrow|\lambda|\le\frac{u_i}{1-l_i}\le l_i+u_i<1.$$

即$\rho((D-L)^{-1}U)<1$. 从而G-S迭代收敛.

下面给出估计式. 因为

$$x_k - x_{k-1} = D^{-1}L(x_k - x_{k-1}) + D^{-1}U(x_{k-1} - x_{k-2}).$$

用分量表示即为

$$x_i^{(k)} - x_i^{(k-1)} = \sum_{j=1}^{i-1} b_{ij}(x_j^{(k)} - x_j^{(k-1)}) + \sum_{j=i+1}^{n} (x_j^{(k-1)} - x_j^{(k-2)}).$$

于是

$$|x_i^{(k)} - x_i^{(k-1)}| \le l_i \max_j |x_i^{(k)} - x_i^{(k-1)}| + u_i \max_j |x_j^{(k-1)} - x_j^{(k-2)}|$$

假定分量中在某个i_0达到极大, 即

$$|x_{i_0}^{(k)} - x_{i_0}^{(k-1)}| = \max_i |x_i^{(k)} - x_i^{(k-1)}|.$$

下面给出估计式. 因为

$$x_k - x_{k-1} = D^{-1}L(x_k - x_{k-1}) + D^{-1}U(x_{k-1} - x_{k-2}).$$

用分量表示即为

$$x_i^{(k)} - x_i^{(k-1)} = \sum_{j=1}^{i-1} b_{ij}(x_j^{(k)} - x_j^{(k-1)}) + \sum_{j=i+1}^{n} (x_j^{(k-1)} - x_j^{(k-2)}).$$

于是

$$|x_i^{(k)} - x_i^{(k-1)}| \le l_i \max_j |x_i^{(k)} - x_i^{(k-1)}| + u_i \max_j |x_j^{(k-1)} - x_j^{(k-2)}|$$

假定分量中在某个i_0达到极大, 即

$$|x_{i_0}^{(k)} - x_{i_0}^{(k-1)}| = \max_i |x_i^{(k)} - x_i^{(k-1)}|.$$

下面给出估计式. 因为

$$x_k - x_{k-1} = D^{-1}L(x_k - x_{k-1}) + D^{-1}U(x_{k-1} - x_{k-2}).$$

用分量表示即为

$$x_i^{(k)} - x_i^{(k-1)} = \sum_{j=1}^{i-1} b_{ij}(x_j^{(k)} - x_j^{(k-1)}) + \sum_{j=i+1}^{n} (x_j^{(k-1)} - x_j^{(k-2)}).$$

于是

$$|x_i^{(k)} - x_i^{(k-1)}| \le l_i \max_j |x_i^{(k)} - x_i^{(k-1)}| + u_i \max_j |x_j^{(k-1)} - x_j^{(k-2)}|$$

假定分量中在某个i_0达到极大, 即

$$|x_{i_0}^{(k)} - x_{i_0}^{(k-1)}| = \max_i |x_i^{(k)} - x_i^{(k-1)}|.$$

下面给出估计式. 因为

$$x_k - x_{k-1} = D^{-1}L(x_k - x_{k-1}) + D^{-1}U(x_{k-1} - x_{k-2}).$$

用分量表示即为

$$x_i^{(k)} - x_i^{(k-1)} = \sum_{j=1}^{i-1} b_{ij}(x_j^{(k)} - x_j^{(k-1)}) + \sum_{j=i+1}^{n} (x_j^{(k-1)} - x_j^{(k-2)}).$$

于是

$$|x_i^{(k)} - x_i^{(k-1)}| \leq l_i \max_j |x_i^{(k)} - x_i^{(k-1)}| + u_i \max_j |x_j^{(k-1)} - x_j^{(k-2)}|$$

假定分量中在某个i_0达到极大, 即

$$|x_{i_0}^{(k)} - x_{i_0}^{(k-1)}| = \max_i |x_i^{(k)} - x_i^{(k-1)}|.$$

那么就有

$$\|x_k - x_{k-1}\|_\infty \le l_{i_0}\|x_k - x_{k-1}\|_\infty + u_{i_0}\|x_{k-1} - x_{k-2}\|_\infty.$$

从而有

$$\begin{aligned}\|x_k - x_{k-1}\|_\infty \le & \frac{u_{i_0}}{1-l_{i_0}}\|x_{k-1} - x_{k-2}\|_\infty \le \mu\|x_{k-1} - x_{k-2}\|_\infty \\ & \cdots \\ \le & \mu^{k-1}\|x_1 - x_0\|_\infty.\end{aligned}$$

由$x_k - x_* = \sum\limits_{i=k}^{\infty}(x_i - x_{i+1})$得

$$\begin{aligned}\|x_k - x_*\|_\infty \le \sum_{i=k}^{\infty}\|x_i - x_{i+1}\|_\infty \le \sum_{i=k}^{\infty}\mu^i\|x_1 - x_0\|_\infty \\ = \frac{\mu^k}{1-\mu}\|x_1 - x_0\|_\infty\end{aligned}$$

证明结束.

那么就有

$$\|x_k - x_{k-1}\|_\infty \le l_{i_0}\|x_k - x_{k-1}\|_\infty + u_{i_0}\|x_{k-1} - x_{k-2}\|_\infty.$$

从而有

$$\begin{aligned}\|x_k - x_{k-1}\|_\infty \le & \frac{u_{i_0}}{1-l_{i_0}}\|x_{k-1} - x_{k-2}\|_\infty \le \mu\|x_{k-1} - x_{k-2}\|_\infty \\ & \cdots \\ \le & \mu^{k-1}\|x_1 - x_0\|_\infty.\end{aligned}$$

由$x_k - x_* = \sum\limits_{i=k}^{\infty}(x_i - x_{i+1})$得

$$\begin{aligned}\|x_k - x_*\|_\infty \le \sum_{i=k}^{\infty}\|x_i - x_{i+1}\|_\infty \le \sum_{i=k}^{\infty}\mu^i\|x_1 - x_0\|_\infty \\ = \frac{\mu^k}{1-\mu}\|x_1 - x_0\|_\infty\end{aligned}$$

证明结束.

那么就有

$$\|x_k - x_{k-1}\|_\infty \le l_{i_0}\|x_k - x_{k-1}\|_\infty + u_{i_0}\|x_{k-1} - x_{k-2}\|_\infty.$$

从而有

$$\begin{aligned}\|x_k - x_{k-1}\|_\infty \le & \frac{u_{i_0}}{1-l_{i_0}}\|x_{k-1} - x_{k-2}\|_\infty \le \mu\|x_{k-1} - x_{k-2}\|_\infty \\ & \cdots \\ \le & \mu^{k-1}\|x_1 - x_0\|_\infty.\end{aligned}$$

由$x_k - x_* = \sum_{i=k}^{\infty}(x_i - x_{i+1})$得

$$\begin{aligned}\|x_k - x_*\|_\infty \le \sum_{i=k}^{\infty}\|x_i - x_{i+1}\|_\infty \le \sum_{i=k}^{\infty}\mu^i\|x_1 - x_0\|_\infty \\ = \frac{\mu^k}{1-\mu}\|x_1 - x_0\|_\infty\end{aligned}$$

证明结束.

类似地我们可以给出下面的定理.

Theorem 3.2

设$B = [b_{ij}]$是*Jacobi*迭代的迭代矩阵. 若$\|B\|_1 < 1$, 则*G-S*迭代收敛, 且有估计式

$$\|x_k - x_*\|_1 \le \frac{\tilde{\mu}^k}{(1-\tilde{\mu})(1-s)}\|x_1 - x_0\|_1$$

其中

$$s = \max_j \sum_{i=j+1}^{n} |b_{ij}|,\ \tilde{\mu} = \max_j \frac{\sum_{i=1}^{j-1} |b_{ij}|}{1 - \sum_{i=j+1}^{n} |b_{ij}|} \le \|B\|_1 < 1.,$$

证明:

令$v_j = \sum\limits_{i=1}^{j-1} |b_{ij}|$, $s_j = \sum\limits_{i=j+1}^{n} |b_{ij}|$, 则有

$$s_j + v_j < 1, \quad \forall 1 \le j \le n.$$

比较

$$s_j + v_j - \frac{v_j}{1 - s_j} = \frac{(1 - s_j)(s_j + v_j) - v_j}{1 - s_j} = \frac{s_j(1 - s_j - v_j)}{1 - s_j} \ge 0,$$

故

$$\frac{v_j}{1 - s_j} \le s_j + v_j.$$

于是

$$\tilde{\mu} = \max_j \frac{v_j}{1 - s_j} \le \max_j (s_j + v_j) \le \|B\|_1 < 1.$$

下面证明$\rho((D-L)^{-1}U)<1$. 因为

$$
\begin{aligned}
(D-L)^{-1}U &= (D(I-D^{-1}L))^{-1}U = (I-D^{-1}L)^{-1}D^{-1}U \\
&= (I-D^{-1}L)^{-1}D^{-1}U(I-D^{-1}L)^{-1}(I-D^{-1}L),
\end{aligned}
$$

因此$(D-L)^{-1}U$与$D^{-1}U(I-D^{-1}L)^{-1}$相似, 有相同的特征值. 现在假定λ 是$(D^{-1}U(I-D^{-1}L)^{-1})^T$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则有

$$(D^{-1}U(I-D^{-1}L)^{-1})^T x = \lambda x.$$

于是

$$\lambda x = \lambda (D^{-1}L)^T x + (D^{-1}U)^T x.$$

下面证明$\rho((D-L)^{-1}U)<1$. 因为

$$
\begin{aligned}
(D-L)^{-1}U &= (D(I-D^{-1}L))^{-1}U = (I-D^{-1}L)^{-1}D^{-1}U \\
&= (I-D^{-1}L)^{-1}D^{-1}U(I-D^{-1}L)^{-1}(I-D^{-1}L),
\end{aligned}
$$

因此$(D-L)^{-1}U$与$D^{-1}U(I-D^{-1}L)^{-1}$相似, 有相同的特征值. 现在假定λ 是$(D^{-1}U(I-D^{-1}L)^{-1})^T$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则有

$$(D^{-1}U(I-D^{-1}L)^{-1})^T x = \lambda x.$$

于是

$$\lambda x = \lambda(D^{-1}L)^T x + (D^{-1}U)^T x.$$

下面证明$\rho((D-L)^{-1}U)<1$. 因为

$$
\begin{aligned}
(D-L)^{-1}U &= (D(I-D^{-1}L))^{-1}U = (I-D^{-1}L)^{-1}D^{-1}U \\
&= (I-D^{-1}L)^{-1}D^{-1}U(I-D^{-1}L)^{-1}(I-D^{-1}L),
\end{aligned}
$$

因此$(D-L)^{-1}U$与$D^{-1}U(I-D^{-1}L)^{-1}$相似, 有相同的特征值. 现在假定λ 是$(D^{-1}U(I-D^{-1}L)^{-1})^T$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则有

$$(D^{-1}U(I-D^{-1}L)^{-1})^T x = \lambda x.$$

于是

$$\lambda x = \lambda(D^{-1}L)^T x + (D^{-1}U)^T x.$$

下面证明$\rho((D-L)^{-1}U)<1$. 因为

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

因此$(D-L)^{-1}U$与$D^{-1}U(I-D^{-1}L)^{-1}$相似, 有相同的特征值. 现在假定λ 是$(D^{-1}U(I-D^{-1}L)^{-1})^T$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则有

$$(D^{-1}U(I-D^{-1}L)^{-1})^T x=\lambda x.$$

于是

$$\lambda x=\lambda(D^{-1}L)^T x+(D^{-1}U)^T x.$$

下面证明$\rho((D-L)^{-1}U)<1$. 因为

$$
\begin{aligned}
(D-L)^{-1}U &= (D(I-D^{-1}L))^{-1}U = (I-D^{-1}L)^{-1}D^{-1}U \\
&= (I-D^{-1}L)^{-1}D^{-1}U(I-D^{-1}L)^{-1}(I-D^{-1}L),
\end{aligned}
$$

因此$(D-L)^{-1}U$与$D^{-1}U(I-D^{-1}L)^{-1}$相似, 有相同的特征值. 现在假定λ是$(D^{-1}U(I-D^{-1}L)^{-1})^T$的任一特征值, $x=(\xi_1,\cdots,\xi_n)^T$为相应的特征向量, 并假定$|\xi_i|=\|x\|_\infty=1$, 则有

$$(D^{-1}U(I-D^{-1}L)^{-1})^T x = \lambda x.$$

于是

$$\lambda x = \lambda(D^{-1}L)^T x + (D^{-1}U)^T x.$$

比较两边向量的第i个分量, 得

$$\lambda\xi_i = \sum_{j=1}^{i-1} b_{ji}\xi_j + \lambda \sum_{j=i+1}^{n} b_{ji}\xi_j,$$

两边取绝对值

$$|\lambda| \le \sum_{j=1}^{i-1} |b_{ji}| + \lambda \sum_{j=i+1}^{n} |b_{ji}|.$$

于是

$$\lambda \le \frac{\sum\limits_{j=1}^{i-1} |b_{ji}|}{1 - \sum\limits_{j=i+1}^{n} |b_{ji}|} \le \tilde{\mu} < 1.$$

故G-S迭代收敛.

比较两边向量的第i个分量, 得

$$\lambda\xi_i = \sum_{j=1}^{i-1} b_{ji}\xi_j + \lambda \sum_{j=i+1}^{n} b_{ji}\xi_j,$$

两边取绝对值

$$|\lambda| \leq \sum_{j=1}^{i-1} |b_{ji}| + \lambda \sum_{j=i+1}^{n} |b_{ji}|.$$

于是

$$\lambda \leq \frac{\sum\limits_{j=1}^{i-1} |b_{ji}|}{1 - \sum\limits_{j=i+1}^{n} |b_{ji}|} \leq \tilde{\mu} < 1.$$

故G-S迭代收敛.

比较两边向量的第i个分量, 得

$$\lambda \xi_i = \sum_{j=1}^{i-1} b_{ji}\xi_j + \lambda \sum_{j=i+1}^{n} b_{ji}\xi_j,$$

两边取绝对值

$$|\lambda| \leq \sum_{j=1}^{i-1} |b_{ji}| + \lambda \sum_{j=i+1}^{n} |b_{ji}|.$$

于是

$$\lambda \leq \frac{\sum\limits_{j=1}^{i-1} |b_{ji}|}{1 - \sum\limits_{j=i+1}^{n} |b_{ji}|} \leq \tilde{\mu} < 1.$$

故G-S迭代收敛.

$$\sum_{i=1}^{n}|x_i^{(k)}-x_i^{(k-1)}| \le \sum_{i=1}^{n}\Big(\sum_{j=1}^{i-1}|b_{ij}||x_j^{(k)}-x_j^{(k-1)}|+\sum_{j=i+1}^{n}|b_{ij}||x_j^{(k-1)}-x_j^{(k-2)}|\Big)$$

把不等式右边的两个求和号交换, 并注意到B矩阵的特点, 就有

$$\sum_{i=1}^{n}|x_i^{(k)}-x_i^{(k-1)}| \le \sum_{j=1}^{n}\Big(\sum_{i=j+1}^{n}|b_{ij}||x_j^{(k)}-x_j^{(k-1)}|+\sum_{i=1}^{j-1}|b_{ij}||x_j^{(k-1)}-x_j^{(k-2)}|\Big).$$

注意到

$$s_j=\sum_{i=j+1}^{n}|b_{ij}|,\ v_j=\sum_{i=1}^{j-1}|b_{ij}|,$$

$$\sum_{i=1}^{n}|x_i^{(k)}-x_i^{(k-1)}| \le \sum_{i=1}^{n}\Big(\sum_{j=1}^{i-1}|b_{ij}||x_j^{(k)}-x_j^{(k-1)}| + \sum_{j=i+1}^{n}|b_{ij}||x_j^{(k-1)}-x_j^{(k-2)}|\Big)$$

把不等式右边的两个求和号交换, 并注意到B矩阵的特点, 就有

$$\sum_{i=1}^{n}|x_i^{(k)}-x_i^{(k-1)}| \le \sum_{j=1}^{n}\Big(\sum_{i=j+1}^{n}|b_{ij}||x_j^{(k)}-x_j^{(k-1)}| + \sum_{i=1}^{j-1}|b_{ij}||x_j^{(k-1)}-x_j^{(k-2)}|\Big).$$

注意到

$$s_j = \sum_{i=j+1}^{n}|b_{ij}|,\ v_j = \sum_{i=1}^{j-1}|b_{ij}|,$$

$$\sum_{i=1}^{n} |x_i^{(k)} - x_i^{(k-1)}| \le \sum_{i=1}^{n} \Big(\sum_{j=1}^{i-1} |b_{ij}||x_j^{(k)} - x_j^{(k-1)}| + \sum_{j=i+1}^{n} |b_{ij}||x_j^{(k-1)} - x_j^{(k-2)}| \Big)$$

把不等式右边的两个求和号交换, 并注意到B矩阵的特点, 就有

$$\sum_{i=1}^{n} |x_i^{(k)} - x_i^{(k-1)}| \le \sum_{j=1}^{n} \Big(\sum_{i=j+1}^{n} |b_{ij}||x_j^{(k)} - x_j^{(k-1)}| + \sum_{i=1}^{j-1} |b_{ij}||x_j^{(k-1)} - x_j^{(k-2)}| \Big).$$

注意到

$$s_j = \sum_{i=j+1}^{n} |b_{ij}|, \ v_j = \sum_{i=1}^{j-1} |b_{ij}|,$$

则有

$$\sum_{i=1}^{n}|x_i^{(k)}-x_i^{(k-1)}|\le\sum_{j=1}^{n}\left(s_j|x_j^{(k)}-x_j^{(k-1)}|+v_j|x_j^{(k-1)}-x_j^{(k-2)}|\right),$$

或者

$$\begin{aligned}\sum_{j=1}^{n}(1-s_j)|x_j^{(k)}-x_j^{(k-1)}| &\le \sum_{j=1}^{n}v_j|x_j^{(k-1)}-x_j^{(k-2)}|\\ &\le \tilde{\mu}\sum_{j=1}^{n}(1-s_j)|x_j^{(k-1)}-x_j^{(k-2)}|\\ &\cdots\\ &\le \tilde{\mu}^{k-1}\sum_{j=1}^{n}(1-s_j)|x_j^{(1)}-x_j^{(0)}|.\end{aligned}$$

则有

$$\sum_{i=1}^{n}|x_i^{(k)}-x_i^{(k-1)}|\leq\sum_{j=1}^{n}\left(s_j|x_j^{(k)}-x_j^{(k-1)}|+v_j|x_j^{(k-1)}-x_j^{(k-2)}|\right),$$

或者

$$\begin{aligned}\sum_{j=1}^{n}(1-s_j)|x_j^{(k)}-x_j^{(k-1)}| &\leq \sum_{j=1}^{n}v_j|x_j^{(k-1)}-x_j^{(k-2)}| \\ &\leq \tilde{\mu}\sum_{j=1}^{n}(1-s_j)|x_j^{(k-1)}-x_j^{(k-2)}| \\ &\cdots \\ &\leq \tilde{\mu}^{k-1}\sum_{j=1}^{n}(1-s_j)|x_j^{(1)}-x_j^{(0)}|.\end{aligned}$$

由定义

$$1 - s \le 1 - s_j < 1.$$

所以

$$(1-s)\sum_{j=1}^{n}|x_j^{(k)} - x_j^{(k-1)}| \le \tilde{\mu}^{k-1}\sum_{j=1}^{n}(1-s_j)|x_j^{(1)} - x_j^{(0)}|,$$

即

$$(1-s)\|x_k - x_{k-1}\|_1 \le \tilde{\mu}^{k-1}\|x_1 - x_0\|_1.$$

又因为

$$x_k - x_* = \sum_{i=k}^{\infty}(x_i - x_{i+1})$$

由定义

$$1 - s \leq 1 - s_j < 1.$$

所以

$$(1 - s)\sum_{j=1}^{n} |x_j^{(k)} - x_j^{(k-1)}| \leq \tilde{\mu}^{k-1} \sum_{j=1}^{n} (1 - s_j)|x_j^{(1)} - x_j^{(0)}|,$$

即

$$(1 - s)\|x_k - x_{k-1}\|_1 \leq \tilde{\mu}^{k-1}\|x_1 - x_0\|_1.$$

又因为

$$x_k - x_* = \sum_{i=k}^{\infty} (x_i - x_{i+1})$$

由定义

$$1 - s \leq 1 - s_j < 1.$$

所以

$$(1-s)\sum_{j=1}^{n}|x_j^{(k)} - x_j^{(k-1)}| \leq \tilde{\mu}^{k-1}\sum_{j=1}^{n}(1-s_j)|x_j^{(1)} - x_j^{(0)}|,$$

即

$$(1-s)\|x_k - x_{k-1}\|_1 \leq \tilde{\mu}^{k-1}\|x_1 - x_0\|_1.$$

又因为

$$x_k - x_* = \sum_{i=k}^{\infty}(x_i - x_{i+1})$$

由定义

$$1 - s \le 1 - s_j < 1.$$

所以

$$(1-s)\sum_{j=1}^{n}|x_j^{(k)} - x_j^{(k-1)}| \le \tilde{\mu}^{k-1}\sum_{j=1}^{n}(1-s_j)|x_j^{(1)} - x_j^{(0)}|,$$

即

$$(1-s)\|x_k - x_{k-1}\|_1 \le \tilde{\mu}^{k-1}\|x_1 - x_0\|_1.$$

又因为

$$x_k - x_* = \sum_{i=k}^{\infty}(x_i - x_{i+1})$$

所以

$$
\begin{aligned}
\|x_k - x_*\|_1 &\le \sum_{i=k}^{\infty} \|x_i - x_{i+1}\|_1 \\
&\le \frac{1}{1-s} \sum_{i=k}^{\infty} \tilde{\mu}^i \|x_1 - x_0\|_1 \\
&\le \frac{\tilde{\mu}^k}{(1-\tilde{\mu})(1-s)} \|x_1 - x_0\|_1.
\end{aligned}
$$

证明结束.

对角占优矩阵迭代法的收敛性

Definition 4.1

设矩阵$A = [a_{ij}] \in \mathbb{R}^{n\times n}$. 若对所有的$i\,(1 \le i \le n)$都有

$$|a_{ii}| \ge \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|,$$

并且上式中至少对一个i有严格不等号成立, 则称A为弱严格对角占优的; 如上式对所有i 都有严格不等式成立, 则称A为严格对角占优的.

Definition 4.2

若存在n阶排列矩阵P, 使

$$PAP^T = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$$

其中A_{11}是r阶方阵, A_{22}是$n-r$阶方阵$(0<r<n)$, 则称A是可约的(或可分的), 反之, 如果不存在这样的排列矩阵, 则称A是不可约的(或不可分的).

Theorem 4.3

若A严格对角占优或不可约对角占优, 则*Jacobi*和*G-S*迭代收敛.

证明:

由条件$|a_{ii}| > 0$, 因此D可逆. 现假设Jacobi迭代矩阵B的某个特征值$|\lambda| \geq 1$. 则$\lambda D - L - U$ 是严格对角占优或者不可约对角占优, 因此$\lambda D - L - U$非奇异. 即$\lambda D - L - U$ 不可能有零特征值, 这和λ为$D^{-1}(L + U)$的特征值矛盾, 故$\rho(D^{-1}(L + U)) < 1$, 即Jacobi迭代收敛.

关于G-S迭代, 只要考虑矩阵$\lambda D - \lambda L - U$, 用同样的方法即可. 证毕

G-S迭代格式为

$$x_{k+1} = D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b$$

令$\Delta x = x_{k+1} - x_k$, 则有

$$x_{k+1} = x_k + \Delta x.$$

$$\begin{aligned} x_{k+1} &= x_k + \omega\Delta x \\ &= x_k + \omega(x_{k+1} - x_k) = (1-\omega)x_k + \omega x_{k+1} \\ &= (1-\omega)x_k + \omega(D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b) \end{aligned}$$

G-S迭代格式为

$$x_{k+1} = D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b$$

令$\Delta x = x_{k+1} - x_k$, 则有

$$x_{k+1} = x_k + \Delta x.$$

$$\begin{aligned} x_{k+1} &= x_k + \omega\Delta x \\ &= x_k + \omega(x_{k+1} - x_k) = (1-\omega)x_k + \omega x_{k+1} \\ &= (1-\omega)x_k + \omega(D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b) \end{aligned}$$

G-S迭代格式为

$$x_{k+1} = D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b$$

令$\Delta x = x_{k+1} - x_k$, 则有

$$x_{k+1} = x_k + \Delta x.$$

$$\begin{aligned} x_{k+1} &= x_k + \omega\Delta x \\ &= x_k + \omega(x_{k+1} - x_k) = (1-\omega)x_k + \omega x_{k+1} \\ &= (1-\omega)x_k + \omega(D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b) \end{aligned}$$

分量形式

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \omega\Big(\sum_{j=1}^{i-1} b_{ij}x_j^{(k+1)} + \sum_{j=i+1}^{n} b_{ij}x_j^{(k)} + g_i\Big)$$

其中ω叫做松弛因子. 当$\omega > 1$时叫超松弛; 当$\omega < 1$时叫低松弛; $\omega = 1$ 时就是G-S 迭代. 将超松弛迭代简称为SOR (Successive Over Relaxation). 因为$(I - \omega D^{-1}L)^{-1}$存在, 故有矩阵形式:

$$x_{k+1} = L_\omega x_k + \omega(D-\omega L)^{-1}b$$

其中

$$L_\omega = (D-\omega L)^{-1}[(1-\omega)D + \omega U]$$

Theorem 5.1

*SOR*收敛的必要条件是$0 < \omega < 2$.

证明: 因为SOR收敛, 得$\rho(L_\omega) < 1$, 从而

$$|\det(L_\omega)| = |\lambda_1 \cdots \lambda_n| < 1,$$

其中$\lambda_1, \cdots, \lambda_n$是$L_\omega$的$n$个特征值. 再由

$$L_\omega = (I - \omega D^{-1}L)^{-1}[(1-\omega)I + \omega D^{-1}U],$$

$$\det[(1-\omega)I + \omega D^{-1}U] = (1-\omega)^n,$$

$$\det(I - \omega D^{-1}L)^{-1} = 1,$$

知

$$
\begin{aligned}
|\det(L_\omega)| =& |\det(I-\omega D^{-1}L)^{-1}||\det[(1-\omega)I+\omega D^{-1}U]| \\
=& |(1-\omega)^n| = |\lambda_1\cdots\lambda_n| < 1,
\end{aligned}
$$

从而有$|1-\omega|<1$, 即$0<\omega<2$. 证明结束.

Theorem 5.2

若系数矩阵A是严格对角占优的或某不可约对角占优的, 且松弛因子$\omega\in(0,1)$, 则*SOR* 收敛.

设矩阵A有如下形式：

$$A = \begin{bmatrix} D_1 & K_1 & & & & \\ H_1 & D_2 & K_2 & & & \\ & H_2 & D_3 & \ddots & & \\ & & \ddots & \ddots & & K_{t-1} \\ & & & & H_{t-1} & D_t \end{bmatrix}$$

其中$D_i(i = 1, \cdots, t)$是非奇异的对角阵. 假设Jabobi迭代的矩阵B的特征值都是实的，则最佳松弛因子为

$$\omega = \omega_{\text{opt}} = \frac{2}{1 + \sqrt{1 - \rho(B)^2}}$$

相应的谱半径为

$$\rho(L_{\omega_{\text{opt}}}) = \frac{1 - \sqrt{1 - \rho(B)^2}}{1 + \sqrt{1 - \rho(B)^2}}$$

自适应松弛因子选取方法(姚嘉豪2016)

这是对于 Jacob 迭代进行参数化的修正，从而引发出如下的算法。

松弛因子参数为 $\omega > 0$,

$$x_{k+1} = x_k - \omega D^{-1}(Ax_k - b)$$

或者

$$x_{k+1} = (I - wD^{-1}A)x_k + wD^{-1}b$$

分量的形式是

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

定义残差函数 $r(\omega) = ||b - Ax||_2^2 = x^T A^T Ax - 2b^T Ax + b^T b$

$$\frac{dr(\omega)^{(k+1)}}{d\omega} = \frac{\partial r(\omega)^{(k+1)}}{\partial x^{(k+1)}} \frac{dx^{(k+1)}}{dw}$$

其中

$$\frac{\partial r(\omega)^{(k+1)}}{\partial x^{(k+1)}} = 2x^{(k+1)T} A^T A - 2b^T A$$

$$\frac{d^2 r(\omega)^{(k+1)}}{d\omega^2} = 2\frac{dx^{(k+1)}}{dw}^T A^T A \frac{dx^{(k+1)}}{dw} > 0$$

这里的二阶导数为正数，因为 $A^T A$ 正定，故 $r(\omega)$ 是 ω 的凸函数，通过牛顿法求得极值点一定是该问题的最小值点，其收敛的情况也是如是得证！

牛顿法的更新:

$$\omega^{(k+1)} = \omega^{(k)} - \frac{\frac{dr(\omega)^{(k)}}{d\omega}}{\frac{d^2 r(\omega)^{(k)}}{d\omega^2}}$$

算例

$$A = \begin{pmatrix} 1+\epsilon & -1 & -1 \\ -1 & 2+\epsilon & -1 \\ -1 & -1 & 1+\epsilon \end{pmatrix}$$

其中 $\epsilon = 0.01, b = \begin{pmatrix} 1 \\ 10 \\ 23 \end{pmatrix}$

一个 3 乘 3 线性方程组的简单例子	
算法	迭代次数
AdJOR	25
Jacob	234

Poisson型方程

解偏微分方程	
算法	迭代次数
AdJOR	363
AdSOR	427
Gauss-Seidel	8437
Jacob	14912

Figure: N=80即h=1/80, $\|x_{k+1} - x_k\|_2 \leq 10^{-7}$

目录

系数矩阵对称正定时迭代法的收敛性

定理： 已知$Ax=b$, A是实对称的正定矩阵, B为A的近似逆, 则当$B^{-t}+B^{-1}-A$正定时, 迭代法

$$x_k = x_{k-1} + B(b - Ax_{k-1}), \quad k = 1, 2, \cdots$$

收敛.

证明： 考虑矩阵$I-\bar{B}A := (I-B^tA)(I-BA)$, 则它关于能量内积$(\cdot,\cdot)_A$对称：

$$\begin{aligned}
&(A(I-B^tA)(I-BA)x, y)\\
&= (x, (I-A^tB^t)(I-A^tB)Ay)\\
&= (x, (I-AB^t)(I-AB)Ay))\\
&= (x, (I-AB^t)A(I-BA)y))\\
&= (x, A(I-B^tA)(I-BA)y)\\
&= (Ax, (I-B^tA)(I-BA)y)
\end{aligned}$$

这样

$$\rho(1-\bar{B}A) \le \|I-\bar{B}A\|_A.$$

设r和μ分别为$I-\bar{B}A$和$\bar{B}A$的特征值, 有$r=1-\mu$.

$$\begin{aligned}
&(A(I-B^tA)(I-BA)x,x) \\
&=((I-BA)x,(I-AB)Ax) \\
&=((I-BA)x,A(I-BA)x)\ge 0 \quad \forall x\in\mathbb{R}^n
\end{aligned}$$

因此$I-\bar{B}A$非负定($(I-\bar{B}A)x=\lambda x, (A(I-\bar{B}A)x,x)=\lambda(Ax,x)$ 这样得$\lambda\ge 0$). 因此, 如$\mu>0$, 则$r<1$.

$$\begin{aligned}
&(I-B^tA)(I-BA) \\
&=I-BA-B^tA-B^tABA \\
&=I-(B+B^t-B^tAB)A
\end{aligned}$$

这样
$$\rho(1-\bar{B}A) \le \|I-\bar{B}A\|_A.$$
设r和μ分别为$I-\bar{B}A$和$\bar{B}A$的特征值, 有$r=1-\mu$.

$$\begin{aligned}&(A(I-B^tA)(I-BA)x,x)\\&=((I-BA)x,(I-AB)Ax)\\&=((I-BA)x,A(I-BA)x)\ge 0 \quad \forall x\in\mathbb{R}^n\end{aligned}$$

因此$I-\bar{B}A$非负定($(I-\bar{B}A)x=\lambda x, (A(I-\bar{B}A)x,x)=\lambda(Ax,x)$ 这样得$\lambda\ge 0$). 因此, 如$\mu>0$, 则$r<1$.

$$\begin{aligned}&(I-B^tA)(I-BA)\\&=I-BA-B^tA-B^tABA\\&=I-(B+B^t-B^tAB)A\end{aligned}$$

这样

$$\rho(1-\bar{B}A) \le \|I-\bar{B}A\|_A.$$

设r和μ分别为$I-\bar{B}A$和$\bar{B}A$的特征值, 有$r=1-\mu$.

$$\begin{aligned}
&(A(I-B^tA)(I-BA)x,x)\\
&=((I-BA)x,(I-AB)Ax)\\
&=((I-BA)x,A(I-BA)x)\ge 0 \quad \forall x\in\mathbb{R}^n
\end{aligned}$$

因此$I-\bar{B}A$非负定($(I-\bar{B}A)x=\lambda x, (A(I-\bar{B}A)x,x)=\lambda(Ax,x)$ 这样得$\lambda\ge 0$). 因此, 如$\mu>0$, 则$r<1$.

$$\begin{aligned}
&(I-B^tA)(I-BA)\\
&=I-BA-B^tA-B^tABA\\
&=I-(B+B^t-B^tAB)A
\end{aligned}$$

这样

$$\rho(1-\bar{B}A) \leq \|I-\bar{B}A\|_A.$$

设r和μ分别为$I-\bar{B}A$和$\bar{B}A$的特征值, 有$r=1-\mu$.

$$\begin{aligned}
&(A(I-B^tA)(I-BA)x,x)\\
&=((I-BA)x,(I-AB)Ax)\\
&=((I-BA)x,A(I-BA)x)\geq 0 \quad \forall x\in\mathbb{R}^n
\end{aligned}$$

因此$I-\bar{B}A$非负定($(I-\bar{B}A)x=\lambda x, (A(I-\bar{B}A)x,x)=\lambda(Ax,x)$ 这样得$\lambda\geq 0$). 因此, 如$\mu>0$, 则$r<1$.

$$\begin{aligned}
&(I-B^tA)(I-BA)\\
&=I-BA-B^tA-B^tABA\\
&=I-(B+B^t-B^tAB)A
\end{aligned}$$

$$\bar{B} = B + B^t - B^t AB = B^t(B^{-t} + B^{-1} - A)B$$

另一方面$\mu > 0$等价于$\bar{B}A$关于能量内积$(\cdot,\cdot)_A$正定$((\bar{B}Ax, Ax))$. 这又等价于$\bar{B}$正定, 后者又等价于$B^{-t} + B^{-1} - A$正定, 因此有$\rho(I - \bar{B}A) < 1$. 最后, 有

$$\rho(I - BA) \leq \|I - BA\|_A = \|I - \bar{B}A\|_A^{\frac{1}{2}} < 1$$

因此迭代法收敛. 证明结束.

$$\bar{B} = B + B^t - B^tAB = B^t(B^{-t} + B^{-1} - A)B$$

另一方面$\mu > 0$等价于$\bar{B}A$关于能量内积$(\cdot,\cdot)_A$正定$((\bar{B}Ax, Ax))$. 这又等价于$\bar{B}$正定, 后者又等价于$B^{-t} + B^{-1} - A$正定, 因此有$\rho(I - \bar{B}A) < 1$. 最后, 有

$$\rho(I - BA) \le \|I - BA\|_A = \|I - \bar{B}A\|_A^{\frac{1}{2}} < 1$$

因此迭代法收敛. 证明结束.

$$\bar{B} = B + B^t - B^tAB = B^t(B^{-t} + B^{-1} - A)B$$

另一方面$\mu > 0$等价于$\bar{B}A$关于能量内积$(\cdot,\cdot)_A$正定($(\bar{B}Ax, Ax)$). 这又等价于$\bar{B}$正定, 后者又等价于$B^{-t} + B^{-1} - A$正定, 因此有$\rho(I - \bar{B}A) < 1$. 最后, 有

$$\rho(I - BA) \le \|I - BA\|_A = \|I - \bar{B}A\|_A^{\frac{1}{2}} < 1$$

因此迭代法收敛. 证明结束.

$$\bar{B} = B + B^t - B^tAB = B^t(B^{-t} + B^{-1} - A)B$$

另一方面$\mu > 0$等价于$\bar{B}A$关于能量内积$(\cdot,\cdot)_A$正定($(\bar{B}Ax, Ax)$). 这又等价于$\bar{B}$正定, 后者又等价于$B^{-t} + B^{-1} - A$正定, 因此有$\rho(I - \bar{B}A) < 1$. 最后, 有

$$\rho(I - BA) \le \|I - BA\|_A = \|I - \bar{B}A\|_A^{\frac{1}{2}} < 1$$

因此迭代法收敛. 证明结束.

引理：$\|I-BA\|_A=\|I-\bar{B}A\|_A^{1/2}$.
证明：

$$
\begin{aligned}
&(I-B^tA)(I-BA)\\
&=A^{-\frac{1}{2}}(I-A^{\frac{1}{2}}B^tA^{\frac{1}{2}})A^{\frac{1}{2}}A^{-\frac{1}{2}}(I-A^{\frac{1}{2}}BA^{\frac{1}{2}})A^{\frac{1}{2}}\\
&\sim(I-A^{\frac{1}{2}}B^tA^{\frac{1}{2}})(I-A^{\frac{1}{2}}BA^{\frac{1}{2}})
\end{aligned}
$$

而后者对称, 可以对角化

$$(I-B^tA)(I-BA)x=\lambda x$$

$$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$$

设$x_i, i=1,2,\cdots,n$为λ_i对应的特征值, 且$(x_i,x_j)_A=\delta_{ij}$ 于是

$$\begin{aligned}
&(A(I-BA)x,(I-BA)x)\\
&=(A(I-B^tA)(I-BA)x,x)\\
&=(A\sum_{i=1}^{n}c_i\lambda_i x_i,\sum_{i=1}^{n}c_i x_i)\\
&=\sum_{i=1}^{n}c_i^2\lambda_i \quad (因为(x_i,x_j)_A=\delta_{ij}=\begin{cases}1 & i=j\\ & 0i\neq j\end{cases})\\
&\leq \lambda_1 \quad (\sum_{i=1}^{n}c_i^2=(A\sum_{i=1}^{n}c_i x_i,\sum_{i=1}^{n}c_i x_i)=1)
\end{aligned}$$

取$x = x_1$, 得$(A(I - BA)x, (I - BA)x) = \lambda_1$, 因此

$$\|I - BA\|_A = \sqrt{\lambda_{\max}(I - \bar{B}A)}$$

另一方面

$$\begin{aligned}
&(A(I - \bar{B}A)x, (I - \bar{B}A)x) \\
=&(A(I - B^t A)(I - BA)x, (I - B^t A)(I - BA)x) \\
&\qquad\qquad (x = \sum_{i=1}^{n} c_i x_i,\ 且 (Ax, x) = \sum_{i=1}^{n} c_i^2 = 1) \\
=&(A \sum_{i=1}^{n} c_i \lambda_i x_i, \sum_{i=1}^{n} c_i \lambda_i x_i) \\
=&\sum_{i=1}^{n} c_i^2 \lambda_i^2 \leq \lambda_1^2
\end{aligned}$$

取$x = x_1$, 有$((Ax_1, x_1) = 1)$

$$(A(I - \bar{B}A)x, (I - \bar{B}A)x) = \lambda_1^2$$

故

$$\|I - \bar{B}A\|_A = \lambda_1.$$

这样就得

$$\|I - BA\|_A = \|I - \bar{B}A\|_A^{1/2}$$

因为$I - \bar{B}A$非负定, 故$\rho(I - \bar{B}A) < 1$当且仅当$\lambda(\bar{B}A) > 0$, 即$\bar{B}$正定.证明结束。

Jacobi

$$B = D^{-1}$$
$$B^{-t} + B^{-1} - A = D + D - A = 2D - A \text{ 正定}$$

SOR=successive over relaxation

$$B = \omega(D - \omega L)^{-1} \quad B^t = \omega(D - \omega U)^{-1}$$

$$B^{-1} = \frac{D - \omega L}{\omega} \quad B^{-t} = \frac{D - \omega U}{\omega}$$

$$\begin{aligned} B^{-1} + B^{-t} - A &= \frac{2D - \omega(L + U)}{\omega} - D + L + U \\ &= \frac{2D}{\omega} - D = (\frac{2}{\omega} - 1)D > 0. \end{aligned}$$

目录

求解线性方程组：

$$Ax = b$$

其中

$$A = \begin{pmatrix} 1+\epsilon & -1 & 0 \\ -1 & 2+\epsilon & -1 \\ 0 & -1 & 1+\epsilon \end{pmatrix},$$

$$b = (1 \quad 10 \quad 23)^T.$$

$$D - L = \begin{pmatrix} 1+\epsilon & 0 & 0 \\ -1 & 2+\epsilon & 0 \\ 0 & -1 & 1+\epsilon \end{pmatrix}$$

$$(D-L)^{-1} = \frac{1}{(1+\epsilon)^2(2+\epsilon)} \times \begin{pmatrix} (1+\epsilon)(2+\epsilon) & 0 & 0 \\ 1+\epsilon & (1+\epsilon)^2 & 0 \\ 1 & 1+\epsilon & (1+\epsilon)(2+\epsilon) \end{pmatrix}$$

$$-U = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

G-S迭代矩阵

$$\begin{aligned} M &= (D-L)^{-1}U \\ &= \frac{1}{(1+\epsilon)^2(2+\epsilon)} \begin{pmatrix} 0 & (1+\epsilon)(2+\epsilon) & 0 \\ 0 & 1+\epsilon & (1+\epsilon)^2 \\ 0 & 1 & 1+\epsilon \end{pmatrix} \end{aligned}$$

$$x^{(k)} - x_* = M^{(k-1)}(x^{(0)} - x_*).$$

假设

$$x^{(0)} - x_* = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

则

$$M(x^{(0)} - x_*) = \begin{pmatrix} 0 & \frac{1}{(1+\epsilon)} & 0 \\ 0 & \frac{1}{(1+\epsilon)(2+\epsilon)} & \frac{1}{(2+\epsilon)} \\ 0 & \frac{1}{(1+\epsilon)^2(2+\epsilon)} & \frac{1}{(1+\epsilon)(2+\epsilon)} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{(1+\epsilon)} \\ \frac{1}{(1+\epsilon)} \\ \frac{1}{(1+\epsilon)^2} \end{pmatrix} = \frac{1}{(1+\epsilon)} \begin{pmatrix} 1 \\ 1 \\ \frac{1}{(1+\epsilon)} \end{pmatrix}$$

谢谢！