6 解线性代数方程组的迭代法

前一章已经介绍了解线性代数方程组

$$AX = b \tag{6.1}$$

的一些直接法.这里 $\boldsymbol{A}$ 是 $n\times n$ 矩阵;$\boldsymbol{X},\boldsymbol{b}$ 都是 n 维向量.而大多数计算过程均需对系数矩阵进行分解,因而一般不能保持 $\boldsymbol{A}$ 的稀疏性.而在实际问题中,特别是偏微分方程数值求解,常会遇到大型稀疏矩阵.另外,从上一章分析可以看到,高斯消去法、LU 分解法的乘除次数为 $\frac{1}{3}n^3+O(n^2)$,LDL$^{\mathrm{T}}$ 分解法的乘除次数为 $\frac{1}{6}n^3+O(n^2)$,当 n 较大时,计算量相当大.迭代法的基本思想是构造一个向量序列$\{\boldsymbol{X}^{(n)}\}$,使其收敛至某个极限向量 $\boldsymbol{X}^*$,而 $\boldsymbol{X}^*$ 就是要求的方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的准确解.迭代法能充分利用系数矩阵的稀疏性,当 n 较大时,能有效控制计算量.

迭代法要解决的主要问题如下:

(1)如何构造迭代格式?

(2)构造的格式所产生的序列在什么情况下收敛?

(3)如果收敛,收敛的速率如何?

(4)近似解的误差估计.

6.1 几种常用的迭代格式

6.1.1 简单迭代法(Jacobi 迭代)

设有方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \qquad\qquad\qquad\qquad\qquad\quad\vdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases} \tag{6.2}$$

用矩阵表示为

$$\boldsymbol{AX}=\boldsymbol{b}$$

其中 $\boldsymbol{A}$ 是系数矩阵,非奇异;$\boldsymbol{b},\boldsymbol{X}$ 为 n 维向量.

假设 $a_{ii}\neq 0$,将方程组(6.2)中第一个方程的 $a_{11}x_1$ 项保留在左端,其余项移到右端,然后两边再除以 a_{11}.同理将第 k 个方程的 $a_{kk}x_k$ 项保留在左端,其余项移到

右端,然后两边再除以 $a_{kk}(k=2,3,\cdots,n)$. 这样,方程组(6.2)变成下列等价方程组

$$\begin{cases} x_1 = \qquad b_{12}x_2 + b_{12}x_3 + \cdots + b_{1n}x_n + g_1 \\ x_2 = b_{21}x_1 \qquad + b_{23}x_3 + \cdots + b_{2n}x_n + g_2 \\ \quad \vdots \\ x_n = b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{n,n-1}x_{n-1} + g_n \end{cases} \tag{6.3}$$

其中

$$\begin{cases} b_{ij} = \dfrac{-a_{ij}}{a_{ii}} & (i \neq j) \\ b_{ii} = 0 & (i=1,2,\cdots,n) \\ g_i = \dfrac{b_i}{a_{ii}} & (i=1,2,\cdots,n) \end{cases}$$

若令

$$\boldsymbol{B} = \begin{bmatrix} 0 & b_{12} & b_{13} & \cdots & b_{1n} \\ b_{21} & 0 & b_{23} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & b_{n3} & \cdots & 0 \end{bmatrix} \qquad \boldsymbol{D} = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix} \qquad \boldsymbol{g} = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{bmatrix}$$

容易看出

$$\boldsymbol{B} = \boldsymbol{D}^{-1}(\boldsymbol{D}-\boldsymbol{A}) = \boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{A}$$

$$\boldsymbol{g} = \boldsymbol{D}^{-1}\boldsymbol{b}$$

方程(6.3)用矩阵可表示为

$$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{X} + \boldsymbol{g} \tag{6.4}$$

由此便可构造一个迭代格式(Jacobi 迭代)

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{B}\boldsymbol{X}^{(k)} + \boldsymbol{g} \tag{6.5}$$

当取定初始向量 $\boldsymbol{X}^{(0)} \in \mathbf{R}^n$ 后,上式便产生一个向量序列

$$\boldsymbol{X}^{(0)}, \boldsymbol{X}^{(1)}, \cdots, \boldsymbol{X}^{(k)}, \cdots$$

若它收敛于 $\boldsymbol{X}^*$,则 $\boldsymbol{X}^*$ 一定是方程(6.4)的解,当然也是原方程 $\boldsymbol{AX}=\boldsymbol{b}$ 的解. 此时,称 Jacobi 迭代关于初始向量 $\boldsymbol{X}^{(0)}$ 收敛.

Jacobi 迭代算法

①输入 $\boldsymbol{A}, \boldsymbol{b}$,初始向量 $\boldsymbol{Y}$,容许误差 ε,容许最大迭代次数 M.
②置 $k=1$.
③形成迭代矩阵 $\boldsymbol{B}$(存放在 $\boldsymbol{A}$ 中).
　对 $i=1,2,\cdots,n$,循环
　　若 $|a_{ii}| < \varepsilon$,则打印"求解失败",停机;否则

$T = a_{ii}$

对 $j = 1, 2, \cdots, n$,计算

$$a_{ij} = -\frac{a_{ij}}{T}$$

$$a_{ii} = 0, g_i = \frac{b_i}{T}$$

④迭代:

对 $i = 1, 2, \cdots, n$,计算

$$x_i = g_i + \sum_{j \neq i} a_{ij} y_j$$

⑤若 $\| \boldsymbol{X} - \boldsymbol{Y} \| < \varepsilon$,输出 $\boldsymbol{X}, k$,停机;否则

⑥若 $k < M$,则 $k = k + 1$,将 $\boldsymbol{X}$ 赋值给 $\boldsymbol{Y}$,转④;否则,输出求解失败信息,停机.

6.1.2 Seidel 迭代法

设有简单迭代法

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{B}\boldsymbol{X}^{(k)} + \boldsymbol{g}$$

将迭代矩阵 $\boldsymbol{B} = (b_{ij})_{n \times n}$ 分解为 $\boldsymbol{B} = \boldsymbol{L} + \boldsymbol{U}$,其中

$$\boldsymbol{L} = \begin{bmatrix} 0 & & & \\ b_{21} & 0 & & \\ \vdots & \vdots & \ddots & \\ b_{n1} & b_{n2} & \cdots & 0 \end{bmatrix} \qquad \boldsymbol{U} = \begin{bmatrix} 0 & b_{12} & \cdots & b_{1n} \\ & 0 & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & 0 \end{bmatrix}$$

则
$$\boldsymbol{X}^{(k+1)} = \boldsymbol{L}\boldsymbol{X}^{(k)} + \boldsymbol{U}\boldsymbol{X}^{(k)} + \boldsymbol{g} \quad (k = 0, 1, 2, \cdots)$$

将其修改为

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{L}\boldsymbol{X}^{(k+1)} + \boldsymbol{U}\boldsymbol{X}^{(k)} + \boldsymbol{g} \quad (k = 0, 1, 2, \cdots) \tag{6.6}$$

上式称为 Seidel 迭代,其分量形式为

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} b_{ij} x_j^{(k+1)} + \sum_{j=i+1}^{n} b_{ij} x_j^{(k)} + g_i \quad (i = 1, 2, \cdots, n) \tag{6.7}$$

它的特点在于,计算第 i 个分量 $x_i^{(k+1)}$ 时,前边的 $i-1$ 个分量用的是最新算出的 $x_1^{(k+1)}, \cdots, x_{i-1}^{(k+1)}$,而不是旧值 $x_1^{(k)}, \cdots, x_n^{(k)}$,这样有可能提高收敛速度.

因 $(\boldsymbol{I} - \boldsymbol{L})^{-1}$ 存在,迭代格式(6.6)可以改写成

$$\boldsymbol{X}^{(k+1)} = (\boldsymbol{I} - \boldsymbol{L})^{-1} \boldsymbol{U}\boldsymbol{X}^{(k)} + (\boldsymbol{I} - \boldsymbol{L})^{-1} \boldsymbol{g} \tag{6.8}$$

称 $\boldsymbol{B}_1 = (\boldsymbol{I} - \boldsymbol{L})^{-1} \boldsymbol{U}$ 为 Seidel 迭代法的迭代矩阵.

Seidel 迭代算法

①输入 $\boldsymbol{A}$,$\boldsymbol{b}$,初始向量 $\boldsymbol{Y}$,容许误差 ε,容许最大迭代次数 M

②置 $k=1$,对 $i=1,2,\cdots,n$,$x_i=y_i$

③形成迭代矩阵 $\boldsymbol{B}$(存放在 $\boldsymbol{A}$ 中)

对 $i=1,2,\cdots,n$,循环

若 $|a_{ii}|<\varepsilon$,则打印"求解失败",停机;否则

$T=a_{ii}$

对 $j=1,2,\cdots,n$,计算

$$a_{ij}=\frac{-a_{ij}}{T}$$

$$a_{ii}=0 \qquad g_i=\frac{b_i}{T}$$

④迭代:

对 $i=1,2,\cdots,n$,计算

$$x_i=g_i+\sum_{j\neq i}a_{ij}x_j$$

⑤若 $\|\boldsymbol{X}-\boldsymbol{Y}\|<\varepsilon$,输出 $\boldsymbol{X}$,k,停机;否则

⑥若 $k<M$,则 $k=k+1$,将 $\boldsymbol{X}$ 赋值给 $\boldsymbol{Y}$,转④;否则,输出求解失败信息,停机

6.1.3 松弛法(SOR 迭代)

松弛法可以看作是 Seidel 迭代法的加速,Seidel 迭代法是松弛法的特例. Seidel 迭代格式为

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{L}\boldsymbol{X}^{(k+1)}+\boldsymbol{U}\boldsymbol{X}^{(k)}+\boldsymbol{g}$$

现令
$$\Delta\boldsymbol{X}=\boldsymbol{X}^{(k+1)}-\boldsymbol{X}^{(k)}=\boldsymbol{L}\boldsymbol{X}^{(k+1)}+\boldsymbol{U}\boldsymbol{X}^{(k)}+\boldsymbol{g}-\boldsymbol{X}^{(k)}$$

于是 $\boldsymbol{X}^{(k+1)}=\boldsymbol{X}^{(k)}+\Delta\boldsymbol{X}$.

对于 Seidel 迭代法,$\boldsymbol{X}^{(k+1)}$ 可以看成向量 $\boldsymbol{X}^{(k)}$ 加上修正项 $\Delta\boldsymbol{X}$ 得到. 现在,若在修正项的前面加上一个参数 ω,便得到松弛法的计算格式

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{X}^{(k)}+\omega\Delta\boldsymbol{X}=(1-\omega)\boldsymbol{X}^{(k)}+\omega(\boldsymbol{L}\boldsymbol{X}^{(k+1)}+\boldsymbol{U}\boldsymbol{X}^{(k)}+\boldsymbol{g})\quad(k=0,1,\cdots)\tag{6.9}$$

用分量形式即为

$$x_i^{(k+1)}=(1-\omega)x_i^k+\omega\left(\sum_{j=1}^{i-1}b_{ij}x_j^{(k+1)}+\sum_{j=i+1}^{n}b_{ij}x_j^{(k)}+g_i\right)$$
$$(i=1,2,\cdots,n;\ k=0,1,2,\cdots)\tag{6.10}$$

其中,ω 称为松弛因子,当 $\omega>1$ 时叫超松弛;$\omega<1$ 时称低松弛;$\omega=1$ 时就是 Seidel 迭代法.

迭代格式(6.9)还可改写成

$$X^{(k+1)} = (I - \omega L)^{-1}[(1-\omega)I + \omega U]X^{(k)} + \omega(I - \omega L)^{-1}g$$

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

称为松弛法的迭代矩阵.

松弛迭代法算法

①输入 $\boldsymbol{A}, \boldsymbol{b}$,初始向量 $\boldsymbol{Y}$,松弛因子 ω,容许误差 ε,容许最大迭代次数 M.

②置 $k=1$,对 $i=1,2,\cdots,n, x_i = y_i$.

③形成迭代矩阵 $\boldsymbol{B}$(仍存放在 $\boldsymbol{A}$ 中):

对 $i=1,2,\cdots,n$,循环

若 $|a_{ii}| < \varepsilon$,则打印"求解失败",停机;否则

$T = a_{ii}$

对 $j=1,2,\cdots,n$,计算

$$a_{ij} = -\omega * \frac{a_{ij}}{T}$$

$$a_{ii} = 1 - \omega,\ g_i = \omega * \frac{b_i}{T}$$

④迭代:

对 $i=1,2,\cdots,n$,计算

$$x_i = g_i + \sum_{i=1}^{n} a_{ij} x_j$$

⑤若 $\|\boldsymbol{X} - \boldsymbol{Y}\| < \varepsilon$,输出 $\boldsymbol{X}, k$;停机;否则

⑥若 $k < M$,则 $k = k+1$,将 $\boldsymbol{X}$ 赋值给 $\boldsymbol{Y}$,转④;否则,输出求解失败信息,停机

6.2 迭代法收敛性理论

任意选取初始向量,利用迭代格式(6.5)构造向量序列$\{\boldsymbol{X}^{(k)}\}$,向量序列是否一定收敛呢?先看两个例子.

例 6-1 用 Jacobi 迭代法解方程组

$$\begin{cases} 10x_1 - x_2 - 2x_3 = 7.2 \\ -x_1 + 10x_2 - 2x_3 = 8.3 \\ -x_1 - x_2 + 5x_3 = 4.2 \end{cases}$$

解 上述方程组的准确解是 $x_1 = 1.1, x_2 = 1.2, x_3 = 1.3$.

先把方程改写成

$$\begin{cases} x_1 = \qquad\quad 0.1x_2 + 0.2x_3 + 0.72 \\ x_2 = 0.1x_1 + \qquad\quad 0.2x_3 + 0.83 \\ x_3 = 0.2x_1 + 0.2x_2 + \qquad\quad 0.84 \end{cases}$$

取 $\boldsymbol{X}^{(0)} = (0,0,0)^{\mathrm{T}}$,利用 Jacobi 迭代,计算结果如表 6-1 所示.

表 6-1 计算结果

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0	5	1.095 1	1.195 1	1.294 1
1	0.72	0.83	0.84	6	1.098 3	1.198 3	1.298 0
2	0.971	1.070	1.150	7	1.099 4	1.199 8	1.299 3
3	1.057	1.157 1	1.248 2	8	1.099 8	1.199 8	1.299 7
4	1.085 3	1.185 3	1.282 8	9	1.099 9	1.199 9	1.299 9

从表 6-1 中可以看到,近似解向量序列收敛,且收敛到准确解.

例 6-2 用 Jacobi 迭代法解方程组

$$\begin{cases} x_1 - 10x_2 + 20x_3 = 11 \\ -10x_1 + x_2 - 5x_3 = -14 \\ 5x_1 - x_2 - x_3 = 3 \end{cases}$$

上述方程的准确解是 $x_1 = x_2 = x_3 = 1$.

同样取 $\boldsymbol{X}^{(0)} = (0,0,0)^{\mathrm{T}}$,利用 Jacobi 迭代,计算结果如表 6-2 所示.

表 6-2 计算结果

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	11	-14	-3
2	-69	81	66
3	-499	-374	-429

从表 6-2 中可以看到,此迭代发散. 可以证明,除 $\boldsymbol{X}^{(0)} = (1,1,1)^{\mathrm{T}}$ 外,无论选什么初值都不会收敛.

从上面两个例子可以看出,迭代序列收敛是有条件的. 下面给出迭代法收敛性基本理论.

定理 6.1 对任何初始向量 $\boldsymbol{X}^{(0)}$ 和常数项 $\boldsymbol{f}$,由迭代格式

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{M}\boldsymbol{X}^{(k)} + \boldsymbol{f} \quad (k = 0,1,2,\cdots)$$

产生的向量序列 $\{\boldsymbol{X}^{(k)}\}$ 收敛的充要条件是迭代矩阵的谱半径

$$\rho(\boldsymbol{M}) < 1$$

证明 先证必要性. 假设 $\{\boldsymbol{X}^{(k)}\}$ 收敛到 $\boldsymbol{X}^*$,即 $\lim \boldsymbol{X}^{(k)} = \boldsymbol{X}^*$,则有

$$\boldsymbol{X}^* = \boldsymbol{M}\boldsymbol{X}^* + \boldsymbol{f}$$

令 $\boldsymbol{\varepsilon}_k = \boldsymbol{X}^{(k)} - \boldsymbol{X}^*$,则

$$\boldsymbol{\varepsilon}_{k+1}=\boldsymbol{X}^{(k+1)}-\boldsymbol{X}^{*}=\boldsymbol{M}\boldsymbol{X}^{(k)}-\boldsymbol{M}\boldsymbol{X}^{*}=\boldsymbol{M}(\boldsymbol{X}^{(k)}-\boldsymbol{X}^{*})$$

所以有

$$\boldsymbol{\varepsilon}_{k+1}=\boldsymbol{M}\boldsymbol{\varepsilon}_{k}=\boldsymbol{M}^{2}\boldsymbol{\varepsilon}_{k-1}=\cdots=\boldsymbol{M}^{k+1}\boldsymbol{\varepsilon}_{0} \tag{6.11}$$

对任意初始向量 $\boldsymbol{\varepsilon}_0$,要使向量序列$\{\boldsymbol{M}^k\boldsymbol{\varepsilon}_0\}$收敛于零向量,必须

$$\lim \boldsymbol{M}^{k}=0 \tag{6.12}$$

由定理 5.10 可知

$$\rho(\boldsymbol{M})<1$$

再证充分性. 假设 $\rho(\boldsymbol{M})<1$,则 $\boldsymbol{I}-\boldsymbol{M}$ 非奇异,从而方程组$(\boldsymbol{I}-\boldsymbol{M})\boldsymbol{X}=\boldsymbol{f}$ 有惟一解,并记为 $\boldsymbol{X}^*$,于是(6.11)式仍成立. 由(6.12)推出$\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}=\boldsymbol{X}^*$. 证毕.

在具体问题中,谱半径 $\rho(\boldsymbol{M})$往往很难计算,但由于 $\rho(\boldsymbol{M})\leqslant\|\boldsymbol{M}\|$,所以有时可以用$\|\boldsymbol{M}\|$作为 $\rho(\boldsymbol{M})$的一种估计. 当$\|\boldsymbol{M}\|<1$ 时迭代一定收敛,不过这只是充分条件.

定理 6.2 若迭代矩阵 $\boldsymbol{M}$ 的范数$\|\boldsymbol{M}\|=q<1$,则迭代格式 $\boldsymbol{X}^{(k+1)}=\boldsymbol{M}\boldsymbol{X}^{(k)}+\boldsymbol{f}$ 对任何初始向量 $\boldsymbol{X}^{(0)}$一定收敛,且

$$\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^{*}\|\leqslant\frac{q}{1-q}\|\boldsymbol{X}^{(k-1)}-\boldsymbol{X}^{(k)}\|\leqslant\frac{q^{k}}{1-q}\|\boldsymbol{X}^{(1)}-\boldsymbol{X}^{(0)}\|$$

证 由迭代格式得

$$\boldsymbol{X}^{(k+1)}-\boldsymbol{X}^{(k)}=\boldsymbol{M}(\boldsymbol{X}^{(k)}-\boldsymbol{X}^{(k-1)})=\cdots=\boldsymbol{M}^{k}(\boldsymbol{X}^{(1)}-\boldsymbol{X}^{(0)})$$

$$\|\boldsymbol{X}^{(k+1)}-\boldsymbol{X}^{(k)}\|\leqslant q\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^{(k-1)}\|\leqslant q^{k}\|\boldsymbol{X}^{(1)}-\boldsymbol{X}^{(0)}\|$$

注意到

$$\boldsymbol{X}^{(k)}-\boldsymbol{X}^{*}=\boldsymbol{X}^{(k)}-\boldsymbol{X}^{(k+1)}+\boldsymbol{X}^{(k+1)}-\boldsymbol{X}^{*}=\boldsymbol{X}^{(k)}-\boldsymbol{X}^{(k+1)}+\boldsymbol{M}(\boldsymbol{X}^{(k)}-\boldsymbol{X}^{*})$$

从而

$$\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^{*}\|\leqslant\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^{(k+1)}\|+q\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^{*}\|$$

故

$$\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^{*}\|\leqslant\frac{1}{1-q}\|\boldsymbol{X}^{(k+1)}-\boldsymbol{X}^{(k)}\|\leqslant\frac{q}{1-q}\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^{(k-1)}\|$$

$$\leqslant\frac{q^{k}}{1-q}\|\boldsymbol{X}^{(1)}-\boldsymbol{X}^{(0)}\|$$

6.2.1 三种迭代法迭代矩阵的谱半径与系数矩阵 A 的关系

6.2.1.1 Jacobi 迭代

Jacobi 迭代的迭代矩阵 $\boldsymbol{B}=\boldsymbol{D}^{-1}(\boldsymbol{D}-\boldsymbol{A})$. 由

$$|\lambda\boldsymbol{I}-\boldsymbol{B}|=0$$

得

$$|\lambda\boldsymbol{I}-\boldsymbol{D}^{-1}(\boldsymbol{D}-\boldsymbol{A})|=0$$

即

$$|\boldsymbol{D}^{-1}|\cdot|\lambda\boldsymbol{D}-(\boldsymbol{D}-\boldsymbol{A})|=0$$

所以

$$|\lambda\boldsymbol{D}-(\boldsymbol{D}-\boldsymbol{A})|=0$$

上式写成分量形式为

$$\begin{vmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \lambda a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & \lambda a_{nn} \end{vmatrix} = 0 \tag{6.13}$$

由定理6.1可知：

定理6.3 Jacobi迭代收敛的充要条件是行列式(6.13)的所有根 $\lambda_i(i=1,2,\cdots,n)$ 的绝对值(复数理解为模)小于1.

6.2.1.2 Seidel迭代

对于Seidel迭代，$\boldsymbol{B}=(\boldsymbol{I}-\boldsymbol{L})^{-1}\boldsymbol{U}$. 由

$$|\lambda\boldsymbol{I}-\boldsymbol{B}|=0$$

得
$$|\lambda\boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{L})^{-1}\boldsymbol{U}|=0$$

即
$$|\lambda(\boldsymbol{I}-\boldsymbol{L})^{-1}(\boldsymbol{I}-\boldsymbol{L})-(\boldsymbol{I}-\boldsymbol{L})^{-1}\boldsymbol{U}|=0$$

所以
$$|\lambda(\boldsymbol{I}-\boldsymbol{L})-\boldsymbol{U}|=0 \qquad |\lambda\boldsymbol{D}-\lambda\boldsymbol{DL}-\boldsymbol{DU}|=0$$

写成分量的形式为

$$\begin{vmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn} \end{vmatrix} = 0 \tag{6.14}$$

定理6.4 Seidel迭代收敛的充要条件是行列式(6.14)的所有根 $\lambda_i(i=1,2,\cdots,n)$ 的绝对值小于1.

6.2.1.3 松弛迭代(SOR迭代)

对于松弛迭代，

$$\boldsymbol{B}_\omega=(\boldsymbol{I}-\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{I}+\omega\boldsymbol{U}]$$

由 $|\lambda\boldsymbol{I}-\boldsymbol{B}_\omega|=0$ 得

$$\left|\lambda\boldsymbol{I}-(\boldsymbol{I}-\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{I}+\omega\boldsymbol{U})]\right|=0$$

化简即得

$$|\lambda(\boldsymbol{I}-\omega\boldsymbol{L})-[(1-\omega)\boldsymbol{I}+\omega\boldsymbol{U}]|=0$$

注意到 $|\boldsymbol{D}|\neq 0$，得

$$|\lambda\boldsymbol{D}-\omega\lambda\boldsymbol{DL}-[(1-\omega)\boldsymbol{D}+\omega\boldsymbol{DU}]|=0$$

写成分量形式为

$$\begin{vmatrix} (\lambda+\omega-1)a_{11} & \omega a_{12} & \cdots & \omega a_{1n} \\ \omega\lambda a_{21} & (\lambda+\omega-1)a_{22} & \cdots & \omega a_{2n} \\ \vdots & \vdots & & \vdots \\ \omega\lambda a_{n1} & \omega\lambda a_{n2} & \cdots & (\lambda+\omega-1)a_{nn} \end{vmatrix} = 0 \tag{6.15}$$

定理 6.5 松弛迭代收敛的充要条件是行列式(6.15)的所有根 $\lambda_i(i=1,2,\cdots,n)$的绝对值小于1.

对于一些特定的系数矩阵 $\boldsymbol{A}$,有一些特定的判别方法. 为了说明方便起见,在此先引进一些概念.

定义 6.1 如果矩阵 $\boldsymbol{A}$ 不能通过行交换和相应的列交换变成

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ 0 & \boldsymbol{A}_{22} \end{bmatrix}$$

其中 $\boldsymbol{A}_{11}$,$\boldsymbol{A}_{22}$为方阵,则称 $\boldsymbol{A}$ 为不可约.

定义 6.2 若矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$满足

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

且至少有一个 i 值使上式中不等式严格成立,则称矩阵 $\boldsymbol{A}$ 具有对角优势(又称为弱对角占优). 特别地,若所有的 $\boldsymbol{i}$ 值对上式中不等式都严格成立,则称矩阵 $\boldsymbol{A}$ 具有强对角占优.

定理 6.6 若 $\boldsymbol{A}$ 是强对角占优,或者 $\boldsymbol{A}$ 弱对角占优且不可约,则 $\det\boldsymbol{A}\neq 0$.

矩阵的不可约与矩阵对应的邻接图有一个必然的联系.

定理 6.7 $\boldsymbol{A}$ 矩阵不可约的充要条件是 $\boldsymbol{A}$ 矩阵对应的邻接图是一个强连通图.

定理 6.8 若 $\boldsymbol{A}$ 是强对角占优,或者 $\boldsymbol{A}$ 弱对角占优且不可约,则

(1)Jacobi 迭代、Seidel 迭代一定收敛.

(2)若松弛因子 ω 满足 $0<\omega\leqslant 1$,则松弛迭代一定收敛.

证明 现以 Jacobi 迭代为例予以证明,Seidel 迭代同样适用.

采用反证法. 若 Jacobi 迭代不收敛,即存在一个 λ_0,$|\lambda_0|\geqslant 1$,使

$$\begin{vmatrix} \lambda_0 a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \lambda_0 a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & \lambda_0 a_{nn} \end{vmatrix} = 0 \tag{6.16}$$

注意到 $$|\lambda_0|\cdot|a_{ii}| \geqslant |a_{ii}| \geqslant \sum_{j\neq i} |a_{ij}| \quad (i=1,2,\cdots,n)$$

因此(6.16)式构成的矩阵的对角占优性与 A 矩阵相同,因此(6.16)式不可能成立. 说明 Jacobi 迭代收敛.

定理 6.9 松弛法收敛的必要条件是 $0<\omega<2$.

证明 因为松弛法收敛,故有 $\rho(\boldsymbol{B}_\omega)<1$. 由矩阵 $\boldsymbol{B}_\omega$ 的特征值的性质可知

$$|\det\boldsymbol{B}_\omega| = |\lambda_1\lambda_2\cdots\lambda_n| < 1$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是矩阵 $\boldsymbol{B}_\omega$ 的 n 个特征值. 而

$$\boldsymbol{B}_\omega = (\boldsymbol{I}-\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{I}+\omega\boldsymbol{U}]$$

所以 $$|\det \boldsymbol{B}_\omega| = |(1-\omega)^n| = |\lambda_1\lambda_2\cdots\lambda_n| < 1$$
即 $$|1-\omega| < 1,\ 0<\omega<2.$$

上述定理说明,对于任何系数矩阵 $\boldsymbol{A}$,若要松弛法收敛,其松弛因子 $\omega \in (0,2)$. 然而,当松弛因子满足条件 $0<\omega<2$ 时,并不是对所有系数矩阵 $\boldsymbol{A}$ 松弛法均收敛.

定理 6.10 若矩阵 $\boldsymbol{A}$ 对称且对角线元素均为正实数,则当 $0<\omega<2$ 时,松弛法收敛的充要条件是 $\boldsymbol{A}$ 正定.

证明 松弛法的迭代矩阵

$$\boldsymbol{B}_\omega = (\boldsymbol{I}-\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{I}+\omega\boldsymbol{U}]$$

设其特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$,相应的特征向量为 $\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_n$,则

$$[(1-\omega)\boldsymbol{I}+\omega\boldsymbol{U}]\boldsymbol{Y}_k = \lambda_k(\boldsymbol{I}-\omega\boldsymbol{L})\boldsymbol{Y}_k$$

两边用 $\boldsymbol{Y}_k$ 作内积得

$$\lambda_k = \frac{(1-\omega)(\boldsymbol{Y}_k,\boldsymbol{Y}_k)+\omega(\boldsymbol{U}\boldsymbol{Y}_k,\boldsymbol{Y}_k)}{(\boldsymbol{Y}_k,\boldsymbol{Y}_k)-\omega(\boldsymbol{L}\boldsymbol{Y}_k,\boldsymbol{Y}_k)} \tag{6.17}$$

设 $(\boldsymbol{L}\boldsymbol{Y}_k,\boldsymbol{Y}_k)=\alpha_k+\beta_k i$,$(\boldsymbol{Y}_k,\boldsymbol{Y}_k)=a$. 则

$$(\boldsymbol{U}\boldsymbol{Y}_k,\boldsymbol{Y}_k)=(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{Y}_k,\boldsymbol{Y}_k)=(\boldsymbol{Y}_k,\boldsymbol{L}\boldsymbol{Y}_k)=\alpha_k-\beta_k i$$

将上面结果代入(6.17)式得

$$\lambda_k = \frac{(1-\omega)a+\omega\alpha_k-\omega\beta_k i}{(a-\omega\alpha_k)-i\omega\beta_k}$$

$$|\lambda_k|^2 = \frac{[(1-\omega)a+\omega\alpha_k]^2+\omega^2\beta_k^2}{(a-\omega\alpha_k)^2+\omega^2\beta_k^2} \tag{6.18}$$

由于 $\boldsymbol{A}$ 正定,则 $\boldsymbol{A}_1=\boldsymbol{D}^{-1}\boldsymbol{A}$ 对称正定.

$$(\boldsymbol{A}_1\boldsymbol{Y}_k,\boldsymbol{Y}_k)=((\boldsymbol{I}-\boldsymbol{L}-\boldsymbol{U})\boldsymbol{Y}_k,\boldsymbol{Y}_k)=a-2\alpha_k>0$$

所以 $$[(1-\omega)a+\omega\alpha_k]^2-(a-\omega\alpha_k)^2=(2-\omega)\omega a(2\alpha_k-a)<0$$

即(6.18)式中的分子 $[(1-\omega)a+\omega\alpha_k]^2+\omega^2\beta_k^2$ 小于分母 $(a-\omega\alpha_k)^2+\omega^2\beta_k^2$,于是 $|\lambda_k|^2<1$,$\rho(\boldsymbol{B}_\omega)<1$,松弛迭代法收敛.

松弛迭代法的收敛速度与松弛因子 ω 有关. 下面举一个例子.

例 6-3 取初始向量 $\boldsymbol{X}^{(0)}=(1,1,1)^{\mathrm{T}}$,用 SOR 法求解方程组

$$\begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 24 \\ 30 \\ -24 \end{bmatrix}$$

使 $|\boldsymbol{X}^{(k+1)}-\boldsymbol{X}^{(k)}|<10^{-5}$. 该方程组的精确解为 $\boldsymbol{X}^*=(3,4,-5)$.

解 SOR 法的迭代公式为

$$\begin{cases} x_1^{(k+1)} = (1-\omega)x_1^{(k)} + \dfrac{\omega}{4}[24 - 3x_2^{(k)}] \\ x_2^{(k+1)} = (1-\omega)x_2^{(k)} + \dfrac{\omega}{4}[30 - 3x_1^{(k+1)} + x_3^{(k)}] \\ x_3^{(k+1)} = (1-\omega)x_3^{(k)} + \dfrac{\omega}{4}[-24 + x_2^{(k+1)}] \end{cases}$$

分别取 $\omega=1.8$,$\omega=1.22$,迭代结果如表 6-3 和表 6-4 所示.

表 6-3 $\omega=1.8$ 时的迭代结果

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	1	1	1
1	8.65	1.472 4	-10.937 37
2	1.892 126	4.845 811	0.130 513 5
3	2.744 454	5.977 064	-8.214 29
4	0.535 393 2	4.298 935	-2.293 696
⋮	⋮	⋮	⋮
64	3.000 001	3.999 999	-4.999 996
65	3.000 001	4.000 001	-5.000 002

表 6-4 $\omega=1.22$ 时的迭代结果

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	1	1	1
1	6.185	3.575 732	-6.449 404
2	2.687 512	3.937 199	-4.700 285
3	3.126 210	3.989 747	-5.069 064
4	2.981 615	3.998 013	-4.985 412
⋮	⋮	⋮	⋮
10	3.000 000	3.999 998	-4.999 999
11	3.000 002	4.000 000	-5.000 000

从表 6-3、表 6-4 可以看到,在相同的初始条件下,$\omega=1.8$ 时 SOR 法迭代了 65 步,而 $\omega=1.22$ 时 SOR 法仅迭代了 11 次.

使松弛法收敛最快的松弛因子叫最优松弛因子,记为 ω_{opt}. 对于某些特殊类型

的矩阵,可以证明最优松弛因子为:

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

其中 $\rho(\boldsymbol{B})$是迭代矩阵的谱半径.对一般的矩阵(即使是正定对称矩阵),目前尚无法确定 ω_{opt}的理论值.实际计算时,大部分由经验或通过试算来确定 ω_{opt}的一个近似值.

定理 6.11　设

①$\boldsymbol{A}$ 为分块三对角阵,且 $a_{ii}\neq 0$;

②Jacobi 迭代的迭代矩阵 $\boldsymbol{B}$ 的特征值为实值,且 $0<\rho(\boldsymbol{B})<1$.

则

(1)当 $0<\omega<2$ 时,SOR 法迭代收敛;

(2)SOR 法最优松弛因子

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

练习与思考

1. 用简单迭代法,Seidel 迭代法和取 $\omega=1.46$ 的超松弛迭代法解方程组

$$\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

并写出相应的迭代矩阵.

2. 已知方程组

$$\begin{bmatrix} 10 & -2 & -2 \\ -2 & 10 & -1 \\ -1 & -2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 1 \\ 0.5 \\ 1 \end{bmatrix}$$

(1)构造简单迭代法和 Seidel 迭代法的迭代格式;

(2)讨论这些迭代格式的收敛性.

3. 设有系数矩阵

$$\boldsymbol{A}=\begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}\qquad \boldsymbol{B}=\begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

证明:

(1)对系数矩阵 $\boldsymbol{A}$,简单迭代法收敛,而 Seidel 迭代法不收敛;

(2)对系数矩阵 $\boldsymbol{B}$,简单迭代法收敛,而 Seidel 迭代法收敛.

4. 设常数 $a\neq 0$,试求 a 的取值范围,使得方程组

$$\begin{bmatrix} a & 1 & 3 \\ 1 & a & 2 \\ -3 & 2 & a \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

的简单迭代法关于任何初始向量均收敛.

5. 设 $\boldsymbol{B}\in\mathbf{R}^{n\times n}$,$\rho(\boldsymbol{B})>1$,但 $\boldsymbol{B}$ 有一个特征值满足 $|\lambda|<1$. 试证明存在初始向量 $x^{(0)}$,使得简单迭代

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{B}\boldsymbol{X}^{(k)}+\boldsymbol{f}\quad (k=0,1,2,\cdots)$$

关于此初始向量收敛.

6. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$,且对称正定,其最小特征值和最大特征值分别是 λ_1,λ_n. 试证迭代法

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{X}^{(k)}+\alpha(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{X}^{(k)})$$

收敛的充要条件是 $0<\alpha<2/\lambda_n$. 又问参数取何值时迭代矩阵的谱半径最小?

7. 对方程组

$$\begin{bmatrix} 1 & 2 \\ 0.3 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

拟用迭代法

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{X}^{(k)}+\alpha(\boldsymbol{A}\boldsymbol{X}^{(k)}-\boldsymbol{b})$$

求解. 试确定一个 α 的取值范围,使得上述迭代公式收敛.