5.2 迭代法的收敛性

5.2.1 一般迭代法的收敛性

5.2.2 Jacobi 迭代法和 Gauss-Seidel 迭代法的收敛性

5.2.1 一般迭代法的收敛性

设 x^*是方程组（**5.1.2**）的解，即 $x^* = Bx^* + f$。该式与（**5.1.3**）式相减，并记误差向量 $e^{(k)} = x^{(k)} - x^*$ 则有

$$e^{(k+1)} = Be^{(k)}, k = 0,1,\cdots.$$

由此可推出

$$e^{(k)} = B^k e^{(0)}, \qquad （\mathbf{5.2.1}）$$

其中 $e^{(0)} = x^{(0)} - x^*$ 与 **k** 无关。所以，迭代法（**5.1.3**）收敛就意味着对任意初始向量 $x^{(0)} \in R^n$，都有

$$\lim_{k\to\infty} e^{(k)} = \lim_{k\to\infty} B^K e^{(0)} = 0$$

下面给出迭代法收敛的充分必要条件。

定理 5.1 设距阵 $B\in R^{(n\times n)}$，则 $\lim\limits_{k\to\infty}B^k=0$ 的充分必要条件是 **B** 的谱半径 $\rho(B)<1$。

证 根据距阵 **Jordan** 标准型的结论，对距阵 **B**，存在非奇异距阵 **P**，使得

$$P^{-1}BP=J=diag(J_1,J_2,\cdots,J_r),$$

其中

$$J_i=\begin{pmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{pmatrix}_{\mathrm{n_i}\times n_i}.$$

显然，$B=PJP^{-1},B^K=PJ^KP^{-1}$，

$$J^K=diag(J_1^k,J_2^k,\cdots,J_r^k).$$

因此，$\lim\limits_{k\to\infty}B^k=0$ 的充分必要条件是 $\lim\limits_{k\to\infty}J_i^k=0,i=1,2,\cdots,r.$

记 $J_i = \lambda I + E_i$，则有

$$E_i^k = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & 0 & 1 & \ddots & \vdots \\ & & \ddots & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & \ddots & \ddots & 1 \\ & & & & \ddots & \ddots & 0 \\ & & & & & \ddots & \vdots \\ & & & & & & 0 \end{pmatrix}_{n_i \times n_i},$$

其中第**1**行的第 **k+1** 个元素为**1**。于是有

$$J_i^k = (\lambda_i I + E_i)^k = \sum_{j=0}^{k} C_k^i \lambda_i^{k-j} E_i^j = \sum_{j=0}^{ni-1} C_k^j \lambda_i^{k-j} E_i^j$$

$$= \begin{pmatrix} \lambda_i^k & k\lambda_i^{k-1} & \cdots & C_k^{n_i-1}\lambda_i^{k-n_i+1} \\ & \lambda_i^k & \ddots & \vdots \\ & & \ddots & k\lambda_i^{k-1} \\ & & & \lambda_i^k \end{pmatrix}_{n_i \times n_i},$$

其中，$E_i^0 = I, C_k^j = k!/[j!(k-j)!]$ 。

由于 $\lim\limits_{k\to\infty} k^s \lambda^k = 0(|\lambda| < 1, s \ge 0)$, 所以 $\lim\limits_{k\to\infty} J_i^k = 0$ 的充分必要条件是 $|\lambda_i| < 1(i = 1,2,\cdots,r)$ 。定理得证。

定理 5.2 对于任意的初始向量 $x^{(0)}$和右端向量 **f** ，解方程组（**5.1.2**）的迭代法（**5.1.3**）收敛的充分必要条件是 $\rho(B) < 1$ 。

证　先证充分性。设 $\rho(B) < 1$，则矩阵 $I - B$ 非奇异，方程组（**5.1.2**）有惟一解 x^*，从而的（**5.2.1**）。由定理 **5.1** 知 $\lim\limits_{k\to\infty} B^k = 0$，因此，$\lim\limits_{k\to\infty} e^{(k)} = 0$ ，即 $\lim\limits_{k\to\infty} x^{(k)} = x^*$ 。

再证必要性。设对任意初始向量$x^{(0)}$和右端项 **f**，均有 $\lim\limits_{k\to\infty} x^{(k)} = x^*$，则得 $x^* = Bx^* + f, x^{(k)} - x^* = B^k(x^{(0)} - x^*)$。因此，对任意 $x^{(0)}$ 都有 $\lim\limits_{k\to\infty} B^k(x^{(0)} - x^*) = 0$，由此推出 $\lim\limits_{k\to\infty} B^k = 0$，即得 $\rho(B) < 1$。定理得证。

例 **5.2** 判断用 **J** 法和 **GS** 法解方程组 $Ax = b$ 的收敛性，其中

$$(1) A = \begin{pmatrix} 1 & -9 & -10 \\ -9 & 1 & 5 \\ 8 & 7 & 1 \end{pmatrix}, \quad (2) A = \begin{pmatrix} 10 & 4 & 5 \\ 4 & 10 & 7 \\ 5 & 7 & 10 \end{pmatrix}.$$

解 （**1**）按（**5.1.6**）和（**5.1.11**）有

$$B_J = \begin{pmatrix} 0 & 9 & 10 \\ 9 & 0 & -5 \\ -8 & -7 & 0 \end{pmatrix}, \quad B_{GS} = \begin{pmatrix} 0 & 9 & 10 \\ 0 & 81 & 85 \\ 0 & -639 & -675 \end{pmatrix}.$$

B_J 的特征值为：$\lambda_1 = 4.1412 + 3.9306i, \lambda_2 = 4.1412 - 3.9306i, \lambda_3 = -8.2825.$

$\rho(B_J) = 8.2825 > 1.$ B_{GS} 的特征值为:$\lambda_1 = 0, \lambda_2 = 0.6054, \lambda_3 = -594.6054.$

$\rho(B_{GS}) = 594.6054 > 1.$ 因此，两种迭代法均发散。

（2） 按（5.1.6）和（5.1.11）求得

$$B_J = \begin{pmatrix} 0 & -0.4 & -0.5 \\ -0.4 & 0 & -0.7 \\ -0.5 & -0.7 & 0 \end{pmatrix}, B_{GS} = \begin{pmatrix} 0 & -0.4 & -0.5 \\ 0 & 0.16 & -0.5 \\ 0 & 0.088 & 0.6 \end{pmatrix}.$$

B_J 的特征值为：$\lambda_1 = 0.3653, \lambda_2 = 0.7108, \lambda_3 = -1.0770. \rho(B_J) = 1.0770 > 1.$

B_{GS} 的特征值为:$\lambda_1 = 0, \lambda_2 = 0.3137, \lambda_3 = 0.4463. \rho(B_{GS}) = 0.4463 < 1.$ 因此，**J** 法发散，而 **GS** 法收敛。

有时实际判别一个迭代法是否收敛，条件$\rho(B) < 1$ 是很难检验的。而一些矩阵范数 $\|B\|$ 可以用 **B** 的元素表示，所以用 $\|B\| < 1$作为收敛的充分条件较为方便。

定理 5.3　对某种算子范数，若 $\|B\|<1$ 则迭代法（**5.1.3**）式产生的向量序列 $\{x^{(k)}\}$ 收敛于（**5.1.2**）的精确解 x^* ，且有误差估计式

$$\left\|x^{(k)}-x^*\right\| \le \frac{\|B\|}{1-\|B\|}\left\|x^{(k)}-x^{(k-1)}\right\|, \tag{5.2.2}$$

$$\left\|x^{(k)}-x^*\right\| \le \frac{\|B\|^k}{1-\|B\|}\left\|x^{(1)}-x^{(0)}\right\|. \tag{5.2.3}$$

证　由 $\|B\|<1$ 知迭代法是收敛的，且 $\lim\limits_{k\to\infty} x^{(k)}=x^*$ 。由（**5.1.3**）和 $x^*=Bx^*+f$ 易得 $x^{(k+1)}-x^*=B(x^{(k)}-x^*)$，$x^{(k+1)}-x^{(k)}=B(x^{(k)}-x^{(k-1)})$

于是有

$$\left\|x^{(k+1)}-x^*\right\| \le \|B\|\left\|x^{(k)}-x^*\right\|,$$

$$\left\|x^{(k+1)}-x^{(k)}\right\| \le \|B\|\left\|x^{(k)}-x^{(k-1)}\right\|.$$

由此可得

$$\left\|x^{(k)}-x^{*}\right\|=\left\|x^{(k)}-x^{(k+1)}+x^{(k+1)}-x^{*}\right\|$$
$$\leq\|B\|\left\|x^{(k)}-x^{(k-1)}\right\|+\|B\|\left\|x^{(k)}-x^{*}\right\|.$$

因 $1-\|B\|>0$，由上式即得（**5.2.2**），反复运用

$$\left\|x^{(k)}-x^{(k-1)}\right\|=\left\|B(x^{(k-1)}-x^{(k-2)})\right\|$$
$$\leq\|B\|\left\|(x^{(k-1)}-x^{(k-2)})\right\|,$$

即可得（**5.2.3**），定理得证。

式（**5.2.2**）说明，若 $\|B\|<1$ 但不接近 **1**，则当相邻两次迭代向量 $x^{(k-1)}$ 和 $x^{(k)}$ 很接近时，$x^{(k)}$ 与精确解很靠近。因此，在实际计算中，用 $\left\|x^{(k+1)}-x^{(k)}\right\|\leq\varepsilon$ 作为迭代终止条件是合理的。

对给定的精度要求，由（**5.2.3**）可以得到需要迭代的次数，并且，由

（**5.2.3**）可见，$\|B\|$ 越小，序列 $\{x^{(k)}\}$ 收敛越快。由于 $\|B\|$ 依赖于所选择的范数，而且 $\rho(B)\le\|B\|$ ，我们以 $\rho(B)$给出收敛速度的概念。

定义 5.2　称 $R(B)=-\ln\rho(B)$为迭代法（**5.1.3**）的渐进收敛速度。

由此定义可以看出，$\rho(B)<1$ 越小，**R**（**B**）就越大。

例 **5.3**　证明用 **J** 法和 **GS** 法解下列方程组

$$\begin{cases} 10x_1-2x_2-2x_3=1, \\ -2x_1+10x_2-x_3=0.5, \\ -x_1-2x_2+3x_3=1. \end{cases}$$

必收敛，并比较满足$\left\|x^{(k)}-x^{(k-1)}\right\|_\infty\le 10^{-5}$ 的迭代次数。

解　按（**5.1.6**）和（**5.1.11**）有

$$B_J=\begin{pmatrix} 0 & 0.2 & 0.2 \\ 0.2 & 0 & 0.1 \\ 1/3 & 2/3 & 0 \end{pmatrix},B_{GS}=\frac{1}{150}\begin{pmatrix} 0 & 30 & 30 \\ 0 & 6 & 21 \\ 0 & 14 & 24 \end{pmatrix}.$$

由于$\|B_J\|_1 = 13/15 < 1, \|B_{GS}\|_1 = 1/2 < 1$，所以，**J 法和 GS 法必收敛，**
并且，$\|B_{GS}\|_1 < \|B_J\|_1$，**GS 法比 J 法收敛快。**

取 $x^{(0)} = (0,0,0)^T$，**J 法的计算结果如表 5－1，GS 法的计算结果如表 5-2** 。对 **J 法有** $\|x^{(15)} - x^{(14)}\|_\infty = 10^{-5}$，对 **GS 法有** $\|x^{(9)} - x^{(8)}\| = 0.4 \times 10^{-5}$

实际计算结果也表明 **GS 法比 J 法收敛快。**

表 5-1

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	**0**	**0**	**0**
1	**0.100000**	**0.050000**	**0.333333**
2	**0.176667**	**0.103333**	**0.400000**
…	…	…	…
13	**0.231069**	**0.147041**	**0.508362**
14	**0.231081**	**0.147050**	**0.508383**
15	**0.231087**	**0.147055**	**0.508393**

表 5-2

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	0.100000	0.070000	0.413333
2	0.196667	0.130667	0.486000
⋮	⋮	⋮	⋮
7	0.231071	0.147048	0.508389
8	0.231087	0.147056	0.508399
9	0.231091	0.147058	0.508402

5.2.2 Jacobi 迭代法和 Gauss-Seidel 迭代法的收敛性

显然可以利用定理**5.2**和定理**5.3**判别 **J** 法和 **GS** 法的收敛性，但其中只有定理**5.3**对 **J** 法使用比较方便。对于大型方程组，要求出迭代矩阵 B_{GS} 和

谱半径 $\rho(B_J)$以及 $\rho(B_{GS})$都是不容易的。下面给出一些容易验证收敛性的充分条件，先讨论对角占优矩阵的性质。

定义 5.3 若 $A=(a_{ij})\in R^{n\times n}$ 满足

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

则称 **A** 为严格对角占优矩阵。若满足

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

且其中至少有一个严格不等式成立，则称 **A** 为弱对角占优矩阵。

定义 5.4 设 $A\in R^{n\times n}$，若存在一个排列阵 **P**，使得

$$P^T AP = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \qquad \textbf{(5.2.4)}$$

其中 A_{11} 和 A_{12} 均为方阵，则称 **A** 为可约的。如果不存在排列阵 **P** 使（**5.2.4**）成立，则称 **A** 为不可约的。

如下矩阵 **A** 是可约的，**B** 是不可约的：

$$A=\begin{pmatrix}5&3&1&2\\0&1&0&3\\3&2&1&4\\0&2&0&3\end{pmatrix},B=\begin{pmatrix}4&-1&&\\-1&4&-1&\\&-1&4&-1\\&&-1&4\end{pmatrix}.$$

因为，对于矩阵 **A** 有

$$P=\begin{pmatrix}1&&&\\&0&1&\\&1&0&\\&&&1\end{pmatrix},P^TAP=\begin{pmatrix}5&1&3&2\\3&1&2&4\\0&0&1&3\\0&0&2&3\end{pmatrix}.$$

而对于矩阵 **B**，不存在一个排列阵使（**5.2.4**）成立。

定理 5.4 若 $A=(a_{ij})$ 严格对角占优，则 $a_{ii}\neq 0, i=1,2,\cdots,n$，且 **A** 非奇异。

证 由严格占优矩阵的定义可知 $a_{ii}\neq 0, i=1,2,\cdots,n$，若 **A** 奇异，则有 $x=(x_1,x_2,\cdots,x_n)^T\neq 0$，使 **Ax=0**。设 $|x_k|=\|x\|_\infty>0$，则 **Ax=0**的第**k**个方程有

$$a_{kk}x_k=-\sum_{j=1,j\neq k}^{n}a_{kj}x_j,$$

由此得到

$$|a_{kk}|\leq \sum_{j=1,j\neq k}^{n}|a_{kj}|\frac{|x_j|}{|x_k|}\leq \sum_{j=1,j\neq k}^{n}|a_{kj}|,$$

这与严格对角占优矛盾，定理得证。

定理 5.5 若 $A=(a_{ij})$ 为不可约弱对角占优阵，则 $a_{ii}\neq 0, i=1,2,\cdots,n,$ 且 **A** 非奇异。

证 若有某个 $a_{kk}=0$，由 **A** 的弱对角占优性质可知 **A** 的第**k**行元素均为零。交换 **A** 的第**k**行和第**n**行，并交换 **A** 的第**k**列和第**n**列，就得到（**5.2.4**）的形式，这与 **A** 的不可约性质矛盾，故 $a_{ii}\neq 0, i=1,2,\cdots,n.$

如果 **A** 是奇异的，则存在 $x=(x_1,x_2,\cdots,x_n)^T\neq 0$, 使 **Ax=0**，下面分两种情况考虑。

若 $|x_1|=|x_2|=\cdots=|x_n|\neq 0$, 由 **Ax=0**的第**k**个方程有

$$|a_{kk}|\leq \sum_{j=1, j\neq k}^{n}|a_{kj}|, k=1,2,\cdots,n.$$

这与 **A** 的弱对角占优性相矛盾。

若$|x_i|(i=1,2,\cdots,n)$不全相等，记 $J=\{k:|x_k|\geq|x_i|,i=1,2,\cdots,n\}$，显然 **J** 非空，**J** 的补集也非空。若有$k\in J$ 和$m\notin J$ ，使得$a_{km}\neq 0$ ，则由$|x_m/x_k|<1$ 得知

$$|a_{kk}|\leq \sum_{j=1,j\neq k}^{n}|a_{kj}|\frac{|x_j|}{|x_k|}\leq \sum_{j=1,j\neq k}^{n}|a_{kj}|.$$

这与 **A** 的弱对角占优性相矛盾，因此

$$a_{km}=0,\forall k\in J,\forall m\notin J.$$

这又导致与 **A** 的不可约性相矛盾。

故在以上两种情况下，齐次方程组 **Ax=0** 只有零解，所以 **A** 非奇异，定理得证。

以上两个定理说明，若A为严格对角占优或不可约弱对角占优阵，则 J 法和GS法都可以计算。在这种情况下迭代法的收敛性有如下定理。

定理5.6 若A为严格对角占优矩阵，或不可约的弱对角占优矩阵，则解方程组 $Ax=b$ 的 J 法和GS法均收敛。

证 设$A=D-L-U$，这里只给出A为严格对角占优阵时的证明。

对 J法，迭代矩阵$B_J=D^{-1}(L-U)$，易得

$$\|B_J\|_\infty = \max_{1\le i\le n} \sum_{j=1,\, j\ne i} \frac{|a_{ij}|}{|a_{ii}|}。$$

由A的严格对角占优性，得到 $\|B_J\|_\infty < 1$，所以 J 法收敛。

对GS法，迭代矩阵$B_{GS}=(D-L)^{-1}U$，这里$\det(D-L)^{-1}=\prod_{i=1}^{n} a_{ii}^{-1}\ne 0$。

由于

$$\det(\lambda I-B_{GS})=\det(\lambda I-(D-L)^{-1})$$

$$=\det(D-L)^{-1}\det(\lambda(D-L)-U)$$

我们只需要证明 $\det(\lambda(D-L)-U)=0$ 的根 λ，满足$|\lambda|<1$。

用反证法，假设 $|\lambda| \geq 1$ ，则由A的严格对角占优性有

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

$$\geq \sum_{j=1}^{i-1} |\lambda||a_{ij}| + \sum_{j=i+1}^{n} |a_{ij}| ,\quad i = 1,2,\cdots,n 。$$

这说明矩阵

$$\lambda(D - L) - U = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn} \end{pmatrix}$$

是严格对角占优矩阵，因此det($\lambda(D-L)-U) \neq 0$ 。这说明只有当 $|\lambda|<1$ 时，才能使det $(\lambda(D-L)-U)=0$。从而有 $\rho(B_{GS})<1$，**GS**法收敛，定理得证。

由定理**5. 6**的证明可见，矩阵**A**严格对角占优等价于 $\|B_J\|_\infty<1$。因此，由定理**5. 6**又可知，若 $\|B_J\|_\infty<$,1则相应的**GS**法也收敛。

由例**5. 1**所给的系数矩阵是严格对角占优的，由例**5. 3**所给的系数矩阵是不可约弱对角占优的，所以，用**J**法和**GS**法解对应的方程组都收敛。