

## 第六章 解线性方程组的迭代方法

随着计算技术的发展, 计算机的存储量日益增大, 计算速度也迅速提高, 直接法(如Gauss消去法、平方根法等)在计算机上可以求解的线性方程组的规模也越来越大, 但直接法大多数均需对系数矩阵 $A$ 进行分解, 因而一般不能保持 $A$ 的稀疏性. 而实际应用中, 特别是偏微分方程的数值求解时, 常常遇到的恰恰就是大型稀疏线性方程的求解问题. 因此寻求能够保持稀疏性的有效算法就成为数值线性代数中一个非常重要的研究课题.

目前发展起来的求解稀疏线性方程组的方法主要有两类: 一类是迭代法; 另一类是稀疏直接法. 稀疏直接法是直接法与某些稀疏矩阵技巧有机结合的产物, 它充分利用所给线性方程组的系数矩阵的零元素的分布特点(即矩阵的结构), 采用灵活的主元选取策略, 使得分解出的矩阵因子尽可能地保持原有的稀疏性. 限于篇幅, 本书不涉及这类方法.

迭代法是按照某种规则构造一个向量序列 $\{x_k\}$ , 使其极限向量 $x_*$ 是方程组 $Ax = b$ 的精确解. 因此, 对迭代法来说一般下面几个问题:

- (1) 如何构造迭代序列?
- (2) 构造的迭代序列是否收敛? 在什么情况下收敛?
- (3) 如果收敛, 收敛的速度如何? 我们应该给予量的刻画, 用以比较各种迭代方法收敛的快慢.
- (4) 因为计算总是有限次的, 所以总要讨论近似解的误差估计和迭代过程的中断处理等问题, 这又和舍入误差的分析有关.

一个方法是否有效要看得具有某个精确度的近似解而付出的代价如何, 通常以运算量和存储量的要求为标志. 在这个标准下, 直接法在很多情况下比迭代法好, 但是对大型的稀疏方程组来说, 迭代法更适用. 本章我们将主要讨论古典迭代法.

### 1 Jacobi迭代和Gauss-Seidel迭代

#### 1.1 Jacobi 迭代

考虑非奇异线性代数方程组

$$Ax = b. \quad (1.1)$$

令

$$A = D - L - U, \quad (1.2)$$

其中

$$D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}),$$

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

$$U = \begin{bmatrix} 0 & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ & 0 & -a_{23} & \cdots & -a_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & 0 & -a_{n-1,n} \\ & & & & 0 \end{bmatrix}.$$
(1.3)

那么(1.1)可以写成

$$x = Bx + g, \quad (1.4)$$

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

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

并代入(1.4)的右端, 就可以计算出一个新的向量  $x_1$ , 即

$$x_1 = Bx_0 + g;$$

再把  $x_1$  代入(1.4)右端, 又可得到一个向量  $x_2$ ; 依次类推有

$$x_k = Bx_{k-1} + g, \quad k = 1, 2, 3, \dots \quad (1.5)$$

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.

## 1.2 Gauss-Seidel 迭代

注意到Jacobi迭代各分量的计算顺序是没有关系的, 先算哪个分量都一样. 现在, 假设不按Jacobi 格式, 而是在计算  $x_k$  的第一个分量用  $x_{k-1}$  的各个分量计算. 但当计算  $x_k$  的第二个分量  $x_2^{(k)}$  时, 因  $x_1^{(k)}$  已经算出, 用它代替  $x_1^{(k-1)}$ , 其它分量仍用  $x_i^{(k-1)}$ . 类似地, 计算  $x_l^{(k)}$  时, 因  $x_1^{(k)}, \dots, x_{l-1}^{(k)}$  都已算出, 用它们代替  $x_1^{(k-1)}, \dots, x_{l-1}^{(k-1)}$ , 其它分量仍用  $x_{k-1}$  的分量, 于是有

$$x_k = D^{-1}Lx_k + D^{-1}Ux_{k-1} + g, \quad k = 1, 2, \dots \quad (1.6)$$

这种迭代格式称为Gauss-Seidel 迭代, 简称为G-S迭代. 它的一个明显的好处是在编写程序时存储量减少了. 如果  $(D - L)^{-1}$  存在, G-S迭代可以改写成

$$x_k = (D - L)^{-1}Ux_{k-1} + (D - L)^{-1}b. \quad (1.7)$$

我们把  $L_1 = (D - L)^{-1}U$  叫做G-S迭代的迭代矩阵,  $(D - L)^{-1}b$  叫做常数项.

对G-S迭代法来说, 计算分量的次序是不能改变的.

## 2 Jacobi与G-S迭代的收敛性分析

### 2.1 收敛的充分与必要条件

大家已经注意到, 上面介绍的两种算法, 有一个共同的特点, 那就是新的近似解是已知近似解的线性函数, 并且只与 $x_{k-1}$ 有关, 即它们都可以表示成如下形式

$$x_{k+1} = Mx_k + g. \quad (2.1)$$

事实上, 对Jacobi迭代, 有

$$M = D^{-1}(L + U), \quad g = D^{-1}b;$$

对G-S迭代, 有

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

形如(2.1)的迭代称作**单步线性定常迭代**, 其中 $M \in \mathbf{R}^{n \times n}$ 叫做**迭代矩阵**,  $g \in \mathbf{R}^n$ 叫做**常数项**,  $x_0 \in \mathbf{R}^n$ 叫做**初始向量**. 如果对任意的初始向量由(2.1)产生的迭代序列都有极限, 则称该迭代法是**收敛的**; 否则就称它是不收敛的或**发散的**.

如果迭代法(2.1)是收敛的, 并记其极限为 $x_*$ , 则于(2.1)两端取极限, 得到

$$x_* = Mx_* + g. \quad (2.2)$$

这表明迭代序列 $x_k$ 的极限 $x_*$ 恰为(2.2)的解. 如果线性方程组(1.1)与(2.2)等价, 即存在非奇异矩阵 $G \in \mathbf{R}^{n \times n}$ 使得

$$G(I - M) = A, \quad Gg = b, \quad (2.3)$$

则该迭代序列也收敛到非奇异线性方程组(1.1)的唯一解. 因此, 如果迭代序列收敛而且(2.3)成立, 则对充分大的 $k$ ,  $x_k$ 就可以作为方程组(1.1)的近似解. 当条件(2.3)成立时, 就称迭代法(2.1)与线性方程组(1.1)是**相容的**. 显然, Jacobi迭代法和G-S迭代法都是**相容的**. 从实用的角度考虑, 当然我们只对相容的收敛迭代法感兴趣. 因此, 在本章如果没有特别说明, 我们总假定迭代法是相容的.

设 $x_*$ 为(1.1)的解, 向量

$$y_k = x_k - x_* \quad (2.4)$$

称为 $x_k$ 的误差向量, 由(2.1)减去(2.2)得

$$y_{k+1} = My_k, \quad k = 0, 1, 2, \dots \quad (2.5)$$

利用(2.5)即可导出

$$y_k = M^k y_0. \quad (2.6)$$

由(2.6)容易看出, 对任意给定的 $y_0$ 都有 $y_k \rightarrow 0$  (即 $x_k \rightarrow x_*$ ) 的充分与必要条件是 $M^k \rightarrow 0$ . 上述结果可以写成下述引理.

**引理4.2.1** 迭代法(2.1)收敛的充分与必要条件是

$$M^k \rightarrow 0.$$

又据定理2.1.7知,  $M^k \rightarrow 0$  的充分与必要条件为  $\rho(M) < 1$ , 这样我们就得到了单步线性迭代收敛的基本定理.

**定理4.2.1** 解方程组(1.1)的单步线性迭代(2.1)收敛的充分与必要条件是  $M$  的谱半径  $\rho(M) < 1$ .

从上面定理看到, 迭代序列收敛取决于迭代矩阵的谱半径, 而与初始向量的选取和常数项无关.

因为解同一个方程组时, Jacobi 迭代矩阵和G-S 迭代矩阵的谱半径不一定相同, 而且并无包含关系, 因此, 有时Jacobi 迭代收敛, 而G-S 迭代不收敛.当然, 也有Jacobi 迭代不收敛而G-S 迭代收敛的情形. 例如, 若方程组的系数矩阵分别为

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

则可以验证:  $A_1$  对Jacobi 迭代是收敛的, 但G-S 迭代不收敛; 而  $A_2$  则对Jacobi迭代不收敛, 但G-S迭代收敛. 请自己验证.

用谱半径来判别迭代过程是否收敛, 显然是不够方便的. 因为计算迭代矩阵的谱半径相当困难. 因此, 需要给出一些方便的判别条件, 也就是比较容易计算的条件.

## 2.2 收敛的充分条件及误差估计

**定理4.2.2** 若迭代矩阵  $M$  的范数  $\|M\| = q < 1$ , 并假定范数满足  $\|I\| = 1$ , 则线性迭代法(2.1)的第  $k$  次迭代向量  $x_k$  与准确解  $x_*$  的误差有估计式:

$$\|x_k - x_*\| \leq \frac{q^k}{1 - q} \|x_1 - x_0\|. \quad (2.7)$$

**证明** 由(2.5)知  $y_k = M^k y_0$ , 两端取范数, 得

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

现在估计  $y_0$ , 根据定义  $\|y_0\| = \|x_0 - x_*\|$ , 由  $\|M\| < 1$  可知  $(I - M)^{-1}$  存在, 而且

$$x_* = (I - M)^{-1}g,$$

于是

$$\begin{aligned} x_0 - x_* &= x_0 - (I - M)^{-1}g = (I - M)^{-1}((I - M)x_0 - g) \\ &= (I - M)^{-1}(x_0 - (Mx_0 + g)) = (I - M)^{-1}(x_0 - x_1), \end{aligned}$$

所以

$$\|y_0\| \leq \|(I - M)^{-1}\| \|x_0 - x_1\|.$$

再注意到当  $\|M\| = q < 1$  时, 有

$$\|(I - M)^{-1}\| \leq \frac{1}{1 - \|M\|} = \frac{1}{1 - q},$$

即知定理成立.  $\square$

从近似解的误差估计可以计算出要得到满足精度要求的近似解需要迭代多少次, 但这种估计往往偏高, 在实际计算时用它控制并不方便, 所以给出下面定理.

**定理4.2.3** 若  $\|M\| = q < 1$  且  $\|I\| = 1$ , 则迭代序列  $\{x_k\}$  的第  $k$  次近似和准确解之差有估计式

$$\|x_k - x_*\| \leq \frac{q}{1-q} \|x_{k-1} - x_k\|. \quad (2.8)$$

**证明** 因为

$$\begin{aligned} x_k - x_* &= Mx_{k-1} + g - (Mx_* + g) = Mx_{k-1} - Mx_* \\ &= Mx_{k-1} - M(I - M)^{-1}g = M(I - M)^{-1}(x_{k-1} - x_k), \end{aligned}$$

所以上式两端取范数, 即得

$$\|x_k - x_*\| \leq \frac{q}{1-q} \|x_{k-1} - x_k\|.$$

于是, 定理得证.  $\square$

上面定理表明, 我们可以从两次相邻近似值的差来判别迭代是否应该终止, 这对实际计算是非常好用的.

用范数来判定迭代过程是否收敛尽管是一个充分条件, 但用起来比较方便. 通常是用矩阵的1范数和 $\infty$ 范数来判定的, 这是因为当矩阵知道以后它们是很容易计算的. 对Jacobi迭代来说, 上述判别法基本上能令人满意了, 这是因为给定方程组后, Jacobi迭代的迭代矩阵是比较容易得到的; 而对G-S迭代来说, 仍有一些困难, 这是因为由方程组的系数矩阵去计算G-S迭代矩阵需要求  $(D - L)^{-1}U$  就不那么方便之故. 为此, 给出下面定理.

**定理4.2.4** 设  $B$  是Jacobi迭代的迭代矩阵. 若  $\|B\|_\infty < 1$ , 则G-S迭代收敛, 而且有估计式

$$\|x_k - x_*\|_\infty \leq \frac{\mu^k}{1-\mu} \|x_1 - x_0\|_\infty, \quad (2.9)$$

其中

$$\mu = \max_i \left( \sum_{j=i+1}^n |b_{ij}| / \left( 1 - \sum_{j=1}^{i-1} |b_{ij}| \right) \right), \quad (2.10)$$

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

**证明** 先证  $\mu < 1$ . 令

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

显然, 对任意  $1 \leq i \leq n$  都有  $l_i + u_i \leq \|B\|_\infty < 1$ . 由

$$\begin{aligned} l_i + u_i - \frac{u_i}{1-l_i} &= \frac{1}{1-l_i} ((l_i + u_i)(1-l_i) - u_i) \\ &= \frac{l_i}{1-l_i} (1-l_i-u_i) \geq 0 \end{aligned}$$

可推出

$$\frac{u_i}{1-l_i} \leq l_i + u_i,$$

两边对 $i$ 取最大值, 得

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

再证G-S迭代收敛, 也就是要证明 $\rho(L_1) < 1$ . 设 $\lambda$  是 $L_1 = (D - L)^{-1}U$  的任一特征值,  $x = (\xi_1, \dots, \xi_n)^T$  为相应的特征向量, 并假定 $|\xi_i| = \|x\|_\infty = 1$ , 则

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

或写成

$$\lambda x = \lambda D^{-1}Lx + D^{-1}Ux. \quad (2.11)$$

注意到 $B = D^{-1}L + D^{-1}U$ , 即知 $D^{-1}L$  为 $B$  的下三角部分, 而 $D^{-1}U$  为 $B$  的上三角部分. 于是, (2.11)的第 $i$  个方程为

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

取绝对值得

$$|\lambda| \leq |\lambda| l_i + u_i.$$

由此推得

$$|\lambda| \leq \frac{u_i}{1-l_i} \leq l_i + u_i < 1,$$

即 $\rho(L_1) < 1$ , 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

$$x_i^{(k)} - x_i^{(k-1)} = \sum_{j=1}^{i-1} b_{ij} (x_j^{(k)} - x_j^{(k-1)}) + \sum_{j=i+1}^n b_{ij} (x_j^{(k-1)} - x_j^{(k-2)}), \quad (2.12)$$

于是, 取绝对值就有

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

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

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

从而有

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

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

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

即(2.9)得证. □

类似地我们可以给出下面定理.

**定理4.2.5** 设  $B = [b_{ij}]$  是Jacobi迭代矩阵. 若  $\|B\|_1 < 1$ , 则G-S迭代收敛, 且有估计式:

$$\|x_k - x_*\|_1 \leq \frac{\tilde{\mu}^k}{(1 - \tilde{\mu})(1 - s)} \|x_1 - x_0\|_1, \quad (2.13)$$

其中

$$s = \max_j \sum_{i=j+1}^n |b_{ij}|, \quad \tilde{\mu} = \max_j \frac{\sum_{i=1}^{j-1} |b_{ij}|}{1 - \sum_{i=j+1}^n |b_{ij}|} \leq \|B\|_1 < 1. \quad (2.14)$$

**证明** 显然,  $\tilde{\mu} \leq \|B\|_1 < 1$ . 下面证明  $\rho(L_1) < 1$ . 因为

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

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  $\lambda$  是  $(D^{-1}U(I - D^{-1}L)^{-1})^T$  的任一特征值,  $x = (\xi_1, \dots, \xi_n)^T$  为相应的特征向量, 并假定  $|\xi_i| = \|x\|_\infty = 1$ , 则由

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

可推出

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

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

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

于是推出

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

由此即知  $\rho(L_1) < 1$ , 所以G-S迭代收敛.

为证明估计式(2.13), 我们把(2.12)两边取绝对值以后, 对  $i$  求和, 得

$$\begin{aligned} \sum_{i=1}^n |x_i^{(k)} - x_i^{(k-1)}| &\leq \sum_{i=1}^n \left( \sum_{j=1}^{i-1} |b_{ij}| |x_j^{(k)} - x_j^{(k-1)}| \right. \\ &\quad \left. + \sum_{j=i+1}^n |b_{ij}| |x_j^{(k-1)} - x_j^{(k-2)}| \right). \end{aligned}$$

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

$$\begin{aligned} \sum_{i=1}^n |x_i^{(k)} - x_i^{(k-1)}| &\leq \sum_{j=1}^n \left( \sum_{i=j+1}^n |b_{ij}| |x_j^{(k)} - x_j^{(k-1)}| \right. \\ &\quad \left. + \sum_{i=1}^{j-1} |b_{ij}| |x_j^{(k-1)} - x_j^{(k-2)}| \right). \end{aligned}$$

我们令

$$\tilde{u}_j = \sum_{i=j+1}^n |b_{ij}|, \quad \tilde{l}_j = \sum_{i=1}^{j-1} |b_{ij}|,$$

则

$$\sum_{i=1}^n |x_i^{(k)} - x_i^{(k-1)}| \leq \sum_{j=1}^n (\tilde{u}_j |x_j^{(k)} - x_j^{(k-1)}| + \tilde{l}_j |x_j^{(k-1)} - x_j^{(k-2)}|),$$

或

$$\begin{aligned} \sum_{j=1}^n (1 - \tilde{u}_j) |x_j^{(k)} - x_j^{(k-1)}| &\leq \sum_{j=1}^n \tilde{l}_j |x_j^{(k-1)} - x_j^{(k-2)}| \\ &\leq \tilde{\mu} \sum_{j=1}^n (1 - \tilde{u}_j) |x_j^{(k-1)} - x_j^{(k-2)}| \\ &\quad \dots\dots\dots \\ &\leq \tilde{\mu}^{k-1} \sum_{j=1}^n (1 - \tilde{u}_j) |x_j^{(1)} - x_j^{(0)}|. \end{aligned}$$

根据  $\tilde{u}_j$  和  $s$  的定义知

$$1 - s \leq 1 - \tilde{u}_j < 1,$$



所以

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

即

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

又因为

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

所以

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

即(2.13)得证. □

如果线性方程组的系数矩阵正定. 我们还能推出一些更好的结论.

**定理4.2.6** 若系数矩阵 $A$ 对称, 对角线元素 $a_{ii} > 0$  ( $i = 1, 2, \dots, n$ ), 则Jacobi迭代收敛的充分与必要条件是 $A$ 和 $2D - A$ 都正定.

**证明** 因为 $B = D^{-1}(L + U) = D^{-1}(D - A) = I - D^{-1}A$ , 而 $D = \text{diag}(a_{ii})$ 的对角元素均大于零, 故

$$B = I - D^{-1}A = D^{-\frac{1}{2}}(I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}$$

由 $A$ 的对称性推出 $I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 也是对称的, 而且它与 $B$ 相似, 有相同的特征值, 从而 $B$ 的特征值均为实数.

先证必要性. 假设Jacobi迭代收敛, 即 $\rho(B) < 1$ , 则矩阵 $I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值绝对值也都小于1, 亦即 $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值必在区间 $(0, 2)$ 内, 从而 $A$ 是正定矩阵. 另外, 矩阵 $2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值也全为正数, 所以 $2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 也正定, 而

$$D^{-\frac{1}{2}}(2D - A)D^{-\frac{1}{2}} = 2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}},$$

因此 $2D - A$ 正定.

再证充分性. 因为

$$D^{\frac{1}{2}}(I - B)D^{-\frac{1}{2}} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}},$$

所以由 $A$ 正定推出 $I - B$ 的特征值均大于零, 即 $B$ 的特征值均小于1. 由 $2D - A$ 正定, 再利用

$$D^{-\frac{1}{2}}(2D - A)D^{-\frac{1}{2}} = D^{\frac{1}{2}}(I + B)D^{-\frac{1}{2}},$$

即可推出 $I + B$ 的特征值全大于0, 即 $B$ 的特征值均大于-1, 这样就证明了 $\rho(B) < 1$ , 从而Jacobi迭代收敛. □

**定理4.2.7** 若系数矩阵 $A$ 正定, 则G-S迭代收敛.

**证明** 设 $\lambda$ 是G-S迭代矩阵任一特征值,  $v$ 是相应的特征向量.则有

$$(D - L)^{-1}Uv = \lambda v.$$

因为 $A$ 正定, 所以 $U = L^T$ , 因此,

$$\lambda(D - L)v = L^T v.$$

用 $v^*$ 左乘上式两边, 得

$$\lambda v^* Dv - \lambda v^* Lv = v^* L^T v.$$

令

$$v^* Dv = \delta, \quad v^* Lv = \alpha + i\beta,$$

则有

$$v^* L^T v = (Lv)^* v = \overline{v^* Lv} = \alpha - i\beta.$$

所以

$$\lambda(\delta - (\alpha + i\beta)) = \alpha - i\beta,$$

取模得

$$|\lambda|^2((\delta - \alpha)^2 + \beta^2) = \alpha^2 + \beta^2.$$

而

$$0 < v^* Av = v^*(D - L - L^T)v = \delta - 2\alpha,$$

所以

$$\begin{aligned} (\delta - \alpha)^2 + \beta^2 &= \delta^2 + \alpha^2 + \beta^2 - 2\delta\alpha \\ &= \delta(\delta - 2\alpha) + \alpha^2 + \beta^2 \\ &> \alpha^2 + \beta^2, \end{aligned}$$

由此即知 $|\lambda| < 1$ , 从而推出G-S迭代收敛.  $\square$

从上面的两个定理看到, 对Jacobi迭代和G-S迭代收敛性的判别, 已从对迭代矩阵性质的研究转到对系数矩阵性质的研究上来了, 这应该说是很大的进展, 然而矩阵正定的判别, 并不很直观. 所以要把判别条件再深入一步, 需引进两个定义.

**定义4.2.1** 设矩阵 $A = [a_{ij}] \in \mathbf{R}^{n \times n}$ . 若对所有的 $i$  ( $1 \leq i \leq n$ )都有

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad (2.15)$$

并且(2.15)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; 如果(2.15)式对所有 $i$ 都有严格不等号成立, 则称 $A$ 为**严格对角占优**的.

**定义4.2.2** 如果存在 $n$ 阶排列矩阵 $P$ , 使

$$PAP^T = \begin{bmatrix} A_{11} & 0 \\ A_{12} & A_{22} \end{bmatrix}, \quad (2.16)$$

其中 $A_{11}$ 是 $r$ 阶方阵,  $A_{22}$ 是 $n-r$ 阶方阵, 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**).

如果 $A$ 可分, 则可将方程组 $Ax = b$ 化为

$$PAP^T Px = Pb.$$

记 $Px = y, Pb = f$ , 则有

$$\begin{bmatrix} A_{11} & 0 \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix},$$

其中 $y_1$ 和 $f_1$ 是 $r$ 维向量,  $y_2$ 和 $f_2$ 是 $n-r$ 维的向量. 我们可以先解 $r$ 阶方程组

$$A_{11}y_1 = f_1,$$

求出 $y_1$ , 再代入

$$A_{12}y_1 + A_{22}y_2 = f_2,$$

解出 $y_2$ . 这样, 就把一个求解 $n$ 阶方程组的问题化为求解两个低阶方程组问题了. 这也正是可分这一概念的由来.

定义4.2.2的一个等价说法是: 设 $A$ 为 $n$ 阶方阵( $n \geq 2$ ),  $\mathcal{W} = \{1, 2, \dots, n\}$ . 如果存在 $\mathcal{W}$ 两个非空的子集 $\mathcal{S}$ 和 $\mathcal{T}$ 满足

$$\mathcal{S} \cup \mathcal{T} = \mathcal{W}, \quad \mathcal{S} \cap \mathcal{T} = \emptyset,$$

使得

$$a_{ij} = 0 \quad (i \in \mathcal{S}, j \in \mathcal{T}),$$

则称矩阵 $A$ 为可约; 否则, 称 $A$ 为不可约.

关于它们的等价性, 我们就不证明了. 我们主要是利用矩阵所具有的这些特征, 来研究判别收敛的一些条件.

如果一个矩阵不可约, 并且是弱严格对角占优的, 则称该矩阵为**不可约对角占优的**. 例如, 三对角矩阵

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

是不可约对角占优的.

**定理4.2.8** 若矩阵 $A$ 是严格对角占优的或不可约对角占优的, 则 $A$ 非奇异.

**证明** 先证 $A$ 为严格对角占优的情形. 假设 $A$ 奇异, 则齐次方程组 $Ax = 0$ 有非零解 $x$ . 不妨假设 $|x_i| = \|x\|_\infty = 1$ , 则有

$$|a_{ii}| = |a_{ii}x_i| = \left| \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j \right| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|.$$

这与 $A$ 严格对角占优矛盾.

再证 $A$ 为不可约对角占优时,  $A$ 亦非奇异. 仍用反证法. 设 $x$ 满足 $\|x\|_\infty = 1$ 使得 $Ax = 0$ , 并定义

$$\mathcal{S} = \{i : |x_i| = 1\}, \\ \mathcal{T} = \{k : |x_k| < 1\}.$$

显然  $\mathcal{S} \cup \mathcal{T} = \mathcal{W}$ ,  $\mathcal{S} \cap \mathcal{T} = \emptyset$ , 而且 $\mathcal{T}$ 为非空. 这是因为假设 $\mathcal{T}$ 为空集, 则 $x$ 的各个分量的绝对值均为1, 那么不论 $i$ 为 $\mathcal{S}$ 中的何值均有

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

这与 $A$ 弱严格对角占优矛盾. 另外, 因为 $A$ 不可约, 必定存在 $i, k$ 使得

$$a_{ik} \neq 0, \quad i \in \mathcal{S}, k \in \mathcal{T}.$$

于是,  $|a_{ik}x_k| < |a_{ik}|$ , 并且

$$\begin{aligned} |a_{ii}| &\leq \sum_{j \in \mathcal{S}, j \neq i} |a_{ij}||x_j| + \sum_{j \in \mathcal{T}} |a_{ij}||x_j| \\ &< \sum_{j \in \mathcal{S}, j \neq i} |a_{ij}| + \sum_{j \in \mathcal{T}} |a_{ij}| \\ &= \sum_{j \neq i} |a_{ij}|, \end{aligned}$$

这又与 $A$ 弱严格对角占优矛盾. 因此, 定理得证.  $\square$

从上面定理可以直接得到下面推论.

**推论4.2.1** 若 $A$ 是严格对角占优的或不可约对角占优的对称矩阵, 且对角线元素皆为正, 则 $A$ 正定.

**证明** 设 $\lambda \leq 0$ , 我们考虑矩阵 $A - \lambda I$ . 因为 $A - \lambda I$ 只在 $A$ 的对角线元素上增加了一些, 所以 $A - \lambda I$ 和 $A$ 一样是严格对角占优的或不可约对角占优的. 因此,  $A - \lambda I$ 非奇异. 注意到 $\lambda \leq 0$ 的任意性, 即知 $A$ 的特征值皆大于零, 从而 $A$ 正定.  $\square$

**定理4.2.9** 若 $A$ 是严格对角占优的或不可约对角占优的, 则Jacobi迭代和G-S迭代收敛.

**证明** 若 $A$ 是严格对角占优的或不可约对角占优的, 则必有 $|a_{ii}| > 0$ , 因此 $D$ 可逆. 现在假设Jacobi迭代矩阵 $B$ 的某个特征值 $|\lambda| \geq 1$ , 考察矩阵 $\lambda D - L - U$ , 显然,  $\lambda D - L - U$ 也是严格对角占优的或不可约对角占优的, 因此 $\lambda D - L - U$ 非奇异. 再由

$$\lambda I - B = \lambda I - D^{-1}(L + U) = D^{-1}(\lambda D - L - U)$$

可推出

$$\det(\lambda I - B) = \det(D^{-1}) \det(\lambda D - L - U) \neq 0,$$

这与 $\lambda$ 是 $B$ 的特征值矛盾. 这也就是说 $B$ 的特征值的模均小于1, 即Jacobi迭代收敛.

关于G-S迭代, 我们只要考虑矩阵 $\lambda D - \lambda L - U$ , 用同样的方法可以证明 $L_1$ 的特征值的模均小于1, 即G-S迭代收敛. 详细的证明留作练习.  $\square$

### 3 收敛速度

#### 3.1 平均收敛速度和渐近收敛速度

任何一种迭代格式都是根据某种特定的规则, 构造一个向量序列 $\{x_k\}$ , 使其极限为线性方程组的解. 一般可以表示成

$$x_k = M_k x_{k-1} + g_k. \quad (3.1)$$

当迭代矩阵和常数项与迭代次数无关时, 即为单步线性定常迭代

$$x_k = M x_{k-1} + g.$$

若单步线性定常迭代方法收敛, 则其极限 $x_*$ 必满足

$$x_* = M x_* + g.$$

因此, 误差 $y_k = x_k - x_*$ 就满足

$$y_k = M y_{k-1} = \dots = M^k y_0,$$

从而

$$\|y_k\| \leq \|M^k\| \|y_0\|.$$

初始误差一般是不知道的, 所以通常用 $\|M^k\|$ 的大小来刻画迭代方法收敛的速度. 因此, 定义

$$R_k(M) = \frac{-\ln \|M^k\|}{k}, \quad (3.2)$$

称其为 $k$ 次迭代的平均收敛速度.

注意, 如果迭代过程收敛, 则当 $k \rightarrow \infty$ 时,  $\|M^k\| \rightarrow 0$ , 故当 $k$ 充分大时, 总有 $R_k(M) > 0$ . 设 $R_k = R_k(M) > 0$ , 量

$$\sigma = \left( \frac{\|y_k\|}{\|y_0\|} \right)^{\frac{1}{k}} \approx \|M^k\|^{\frac{1}{k}} = e^{-R_k}$$

就表示误差范数在 $k$ 次迭代中平均每次迭代所缩减的比例因子.

如果有两个迭代矩阵 $G$ 和 $H$ , 而 $R_k(H) > R_k(G) > 0$ , 我们就说对于 $k$ 次迭代来说, 对应于 $H$ 的迭代法比对应于 $G$ 的迭代法快.

如果 $M$ 是对称矩阵(或Hermite矩阵, 或正规矩阵), 则显然有

$$\|M^k\|_2 = (\rho(M))^k,$$

所以

$$R_k(M) = -\ln \rho(M), \quad (3.3)$$

它与 $k$ 无关. 但一般情况下 $R_k$ 是依赖于 $k$ 的, 这时计算 $R_k$ 也就很复杂. 另外, 在比较两个迭代矩阵 $G$ 和 $H$ 时, 可能对某些 $k$ 有 $R_k(G) < R_k(H)$ , 而对另一些 $k$ 又有 $R_k(G) > R_k(H)$ . 好在我们是研究迭代过程的收敛性, 最终是通过 $k \rightarrow \infty$ 来实现的. 因此, 为了刻画整个迭代过程的收敛速度, 我们自然考虑

$$R_\infty(M) = \lim_{k \rightarrow \infty} R_k(M),$$

这就是渐近收敛速度.

**定理4.3.1** 设单步线性定常迭代矩阵为 $M$ , 则有

$$R_\infty(M) = -\ln \rho(M). \quad (3.4)$$

**证明** 只需证明  $\lim_{k \rightarrow \infty} \|M^k\|^{\frac{1}{k}} = \rho(M)$  即可. 因为

$$(\rho(M))^k = \rho(M^k) \leq \|M^k\|,$$

所以

$$\rho(M) \leq \|M^k\|^{\frac{1}{k}}.$$

另一方面, 对任给的 $\varepsilon > 0$ , 考虑矩阵

$$B_\varepsilon = \frac{1}{\rho(M) + \varepsilon} M.$$

显然,  $\rho(B_\varepsilon) < 1$ . 于是, 由定理2.1.7知,  $\lim_{k \rightarrow \infty} B_\varepsilon^k = 0$ . 因此, 存在自然数 $K$ , 使当 $k \geq K$ 时, 有

$$\|B_\varepsilon^k\| \leq 1,$$

即

$$\|M^k\| \leq (\rho(M) + \varepsilon)^k.$$

从而我们已经证明了: 对任给的 $\varepsilon > 0$ , 存在自然数 $K$ , 使当 $k \geq K$ 时, 有

$$\rho(M) \leq \|M^k\|^{\frac{1}{k}} \leq \rho(M) + \varepsilon,$$

即

$$\lim_{k \rightarrow \infty} \|M^k\|^{\frac{1}{k}} = \rho(M)$$

□

从上面定理看到, 渐近收敛速度是迭代矩阵谱半径的负对数. 为了更具体地说明迭代法的收敛速度, 我们来考虑模型问题.

### 3.2 模型问题

模型问题是指用五点差分格式求解单位正方形区域的Poisson方程第一边值问题：

$$\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -f(x, y), & 0 < x, y < 1, \\ u|_{\Gamma} = \phi, \end{cases} \quad (3.5)$$

其中 $\Gamma$ 为正方形区域的边界.

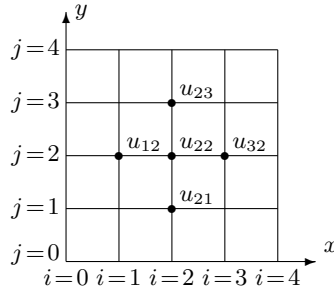


图 4.1

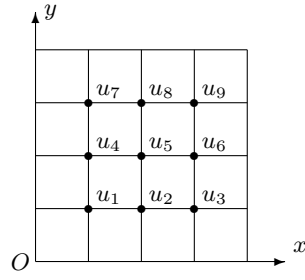


图 4.2

为把微分方程离散化，把正方形的每边 $n$ 等分，并令 $h = \frac{1}{n}$ ，用组两平行于坐标轴的等距直线

$$\begin{cases} x = ih, & i = 1, 2, \dots, n-1, \\ y = jh, & j = 1, 2, \dots, n-1, \end{cases}$$

把单位正方形分割成 $n^2$ 个小正方形，小正方形的顶点为 $(x_i, y_j)$ ,  $i, j = 0, 1, 2, \dots, n$ .  $n = 4$ 的情形如图4.1所示. 用二阶差商

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} \Big|_{(x_i, y_j)} &= \frac{1}{h^2} [u(x_{i-1}, y_j) - 2u(x_i, y_j) + u(x_{i+1}, y_j)] + O(h^2), \\ \frac{\partial^2 u}{\partial y^2} \Big|_{(x_i, y_j)} &= \frac{1}{h^2} [u(x_i, y_{j-1}) - 2u(x_i, y_j) + u(x_i, y_{j+1})] + O(h^2), \end{aligned}$$

代入(3.5)后，再用 $u_{ij}$ 代替 $u(x_i, y_j)$ ，并略去误差项得

$$\begin{aligned} 4u_{ij} - (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) &= h^2 f_{ij}, \\ i, j &= 1, 2, \dots, n-1, \\ u_{i,0} = \phi_{i,0}, \quad u_{i,n} = \phi_{i,n}, \quad i &= 0, 1, 2, \dots, n, \\ u_{0,j} = \phi_{0,j}, \quad u_{n,j} = \phi_{n,j}, \quad j &= 0, 1, 2, \dots, n. \end{aligned} \quad (3.6)$$

现假定 $\phi \equiv 0$ ，将(3.6)写成矩阵形式即为

$$T_{n-1}U + UT_{n-1} = h^2 F, \quad (3.7)$$

其中

$$U = [u_{ij}], \quad F = [f_{ij}],$$

$$T_{n-1} = \begin{bmatrix} 2 & -1 & & & \\ -1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & -1 & 2 \end{bmatrix} \in \mathbf{R}^{(n-1) \times (n-1)}.$$

若把正方形的顶点按如图4.2所示的次序排列，即先按 $j$ 由小到大， $j$ 相同的按 $i$ 由小到大，这种排列方式叫“自然顺序排列”。用自然顺序排列后得到的线性方程组具有下面形状

$$Au = h^2 f,$$

其中

$$A = \begin{bmatrix} T_{n-1} + 2I_{n-1} & -I_{n-1} & & & \\ & -I_{n-1} & \ddots & & \\ & & \ddots & \ddots & \\ & & & -I_{n-1} & \\ & & & -I_{n-1} & T_{n-1} + 2I_{n-1} \end{bmatrix}.$$

易知， $A$ 是 $(n-1)^2$ 阶矩阵。如果结点的排列次序改变了， $A$ 的形状要改变，但可以证明 $A$ 的特征值仍旧不变。

上面的系数矩阵 $A$ 具有这样几个特点：

- (1)  $A$ 是块三对角矩阵，共有五条对角线上有非零元素；
- (2)  $A$ 是不可约对角占优的；
- (3)  $A$ 是正定对称的，而且是稀疏的。

容易验证 $T_{n-1}$ 的特征值为 $\lambda_j = 2\left(1 - \cos \frac{j\pi}{n}\right)$ ，对应的单位特征向量为

$$z_j = \left( \sqrt{\frac{2}{n}} \sin \frac{j\pi}{n}, \sqrt{\frac{2}{n}} \sin \frac{2j\pi}{n}, \dots, \sqrt{\frac{2}{n}} \sin \frac{(n-1)j\pi}{n} \right)^T.$$

利用(3.7)，由此即可推出 $A$ 的特征值为

$$\lambda_{pq} = \lambda_p + \lambda_q = 2 \left( 2 - \cos \frac{p\pi}{n} - \cos \frac{q\pi}{n} \right), \quad (3.8)$$

对应的特征向量 $v_{pq}$ 为矩阵 $z_p z_q^T$ 按列“拉直”得到的 $(n-1)^2$ 维向量，即

$$v_{pq} = \sqrt{\frac{2}{n}} \left( \sin \frac{q\pi}{n} z_p^T, \sin \frac{2q\pi}{n} z_p^T, \dots, \sin \frac{(n-1)q\pi}{n} z_p^T \right)^T. \quad (3.9)$$

### 3.3 Jacobi和G-S迭代的渐近收敛速度.

容易导出，用Jacobi迭代计算模型问题的迭代矩阵为 $B = D^{-1}(D - A) = I - \frac{1}{4}A$ ，其计算公式为

$$u_{ij}^{(k)} = \frac{1}{4} (u_{i+1,j}^{(k-1)} + u_{i-1,j}^{(k-1)} + u_{i,j+1}^{(k-1)} + u_{i,j-1}^{(k-1)}) + \frac{h^2}{4} f_{ij}, \quad (3.10)$$

$$i, j = 1, 2, \dots, n-1.$$



而对应的Jacobi迭代矩阵 $B$ 的特征值问题为

$$\begin{cases} \mu V_{ij} = \frac{1}{4}(V_{i-1,j} + V_{i+1,j} + V_{i,j-1} + V_{i,j+1}), \\ V_{i0} = V_{in} = V_{0j} = V_{nj} = 0. \end{cases} \quad (3.11)$$

因 $B = I - \frac{1}{4}A$ , 所以 $B$ 矩阵特征值为

$$\begin{aligned} \mu_{pq} &= 1 - \frac{1}{4}\lambda_{pq} \\ &= \frac{1}{2} \left( \cos \frac{p\pi}{n} + \cos \frac{q\pi}{n} \right), \quad p, q = 1, 2, \dots, n-1. \end{aligned}$$

于是 $\rho(B) = \cos \frac{\pi}{n} = \cos h\pi$ . 这样我们就得到用Jacobi迭代计算模型问题时的渐近收敛速度

$$R_\infty(B) = -\ln \rho(B) = -\ln \cos h\pi \sim \frac{1}{2}\pi^2 h^2, \quad h \rightarrow 0.$$

把G-S迭代应用到模型问题上, 迭代矩阵 $L_1 = (D - L)^{-1}U$ 的特征值问题是

$$\begin{cases} \lambda \xi_{ij} = \frac{1}{4}(\lambda \xi_{i-1,j} + \lambda \xi_{i,j-1} + \xi_{i+1,j} + \xi_{i,j+1}), \\ \xi_{i0} = \xi_{in} = \xi_{0j} = \xi_{nj} = 0. \end{cases} \quad (3.12)$$

设 $\lambda \neq 0$ , 作变换 $\xi_{ij} = \lambda^{\frac{i+j}{2}} V_{ij}$  得

$$\begin{aligned} \lambda^{1+\frac{i+j}{2}} V_{ij} &= \frac{1}{4}(\lambda^{1+\frac{i+j-1}{2}} V_{i-1,j} + \lambda^{1+\frac{i+j-1}{2}} V_{i,j-1} \\ &\quad + \lambda^{\frac{i+j+1}{2}} V_{i+1,j} + \lambda^{\frac{i+j+1}{2}} V_{i,j+1}) \\ &= \frac{1}{4}\lambda^{\frac{i+j+1}{2}} (V_{i-1,j} + V_{i,j-1} + V_{i+1,j} + V_{i,j+1}), \end{aligned}$$

即

$$\lambda^{\frac{1}{2}} V_{ij} = \frac{1}{4}(V_{i-1,j} + V_{i,j-1} + V_{i+1,j} + V_{i,j+1}).$$

由边值为零, 并对照(3.11), 可以看出 $\lambda$  是 $L_1$ 的非零特征值当且仅当 $\lambda^{1/2}$ 是 $B$ 的非零特征值. 于是, 用G-S迭代计算模型问题时的渐收敛速度为

$$R_\infty(L_1) = -\ln \rho(L_1) = -2 \ln \rho(B) \sim \pi^2 h^2, \quad h \rightarrow 0.$$

这说明G-S迭代比Jacobi迭代快一倍.

此外, 需注意的是运算量也是作为比较各种迭代法以及直接法的优劣的重要标志.

## 4 超松弛迭代法

本节介绍一种新的迭代算法, 它是G-S迭代的引深和推广, 也可看作是G-S迭代的加速.

## 4.1 迭代格式

大家知道, G-S迭代格式为

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

现在令  $\Delta x = x_{k+1} - x_k$ , 则有

$$x_{k+1} = x_k + \Delta x. \quad (4.1)$$

这就是说, 对G-S迭代法来说,  $x_{k+1}$ 可以看作在向量 $x_k$ 上加上修正项 $\Delta x$ 而得到的. 若修正项的前面加上一个参数 $\omega$ , 便得到松弛法的计算公式

$$\begin{aligned} x_{k+1} &= x_k + \omega \Delta x \\ &= (1 - \omega)x_k + \omega(D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b). \end{aligned} \quad (4.2)$$

用分量形式表示即为

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \omega\left(\sum_{j=1}^{i-1} b_{ij}x_j^{(k+1)} + \sum_{j=i+1}^n b_{ij}x_j^{(k)} + g_i\right), \quad (4.3)$$

其中 $\omega$ 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. 我们把超松弛迭代简称为SOR. 因为 $(I - \omega D^{-1}L)^{-1}$ 存在, 所以(4.2)可以改写为

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

其中

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

称作是松弛法的迭代矩阵. 下面我们来研究SOR的收敛性判别和松弛参数 $\omega$ 的选取范围.

## 4.2 收敛性分析

**定理4.4.1** SOR收敛的充分和必要条件是 $\rho(L_\omega) < 1$ .

另外, 也要给出一些判别SOR收敛的其他条件, 首先有

**定理4.4.2** SOR收敛的必要条件是 $0 < \omega < 2$ .

**证明** 因为SOR收敛, 推出 $\rho(L_\omega) < 1$ , 从而

$$|\det(L_\omega)| = |\lambda_1 \cdot \lambda_2 \cdots \lambda_n| < 1,$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 $L_\omega$ 的 $n$ 个特征值. 再由

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

易知

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

$$= |(1-\omega)^n| = |\lambda_1 \lambda_2 \cdots \lambda_n| < 1,$$

从而有 $|1-\omega| < 1$ , 即 $0 < \omega < 2$ . 定理得证.  $\square$

这个定理说明, 对任何系数矩阵, 若要SOR收敛, 必须选取松弛因子 $\omega \in (0, 2)$ . 下面再给出一个充分条件.

**定理4.4.3** 若系数矩阵 $A$ 是严格对角占优的或不可约对角占优的, 且松弛因子 $\omega \in (0, 1)$ , 则SOR收敛.

**证明** 与定理4.2.9的证明类似, 请读者自己补出.  $\square$

对正定矩阵也有很好的结果.

**定理4.4.4** 若系数矩阵是实对称的正定矩阵, 则当 $0 < \omega < 2$ 时SOR收敛.

**证明** 设 $\lambda$ 是 $L_\omega$ 的任一特征值,  $x$ 为对应的特征向量, 则有

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

或

$$[(1 - \omega)D + \omega L^T]x = \lambda(D - \omega L)x.$$

用 $x^*$ 左乘上式两边, 得

$$x^*[(1 - \omega)D + \omega L^T]x = \lambda x^*(D - \omega L)x.$$

现在令 $x^*Dx = \delta$ ,  $x^*Lx = \alpha + i\beta$ , 则有 $x^*L^Tx = \alpha - i\beta$ , 因此可推出

$$(1 - \omega)\delta + \omega(\alpha - i\beta) = \lambda(\delta - \omega(\alpha + i\beta)).$$

两边取模得

$$|\lambda|^2 = \frac{[(1 - \omega)\delta + \omega\alpha]^2 + \omega^2\beta^2}{(\delta - \omega\alpha)^2 + \omega^2\beta^2}.$$

因为

$$\begin{aligned} & [(1 - \omega)\delta + \omega\alpha]^2 + \omega^2\beta^2 - (\delta - \omega\alpha)^2 - \omega^2\beta^2 \\ &= [\delta - \omega(\delta - \alpha)]^2 - (\delta - \omega\alpha)^2 \\ &= \omega\delta(\delta - 2\alpha)(\omega - 2), \end{aligned}$$

并注意到当 $A$ 正定时有 $\delta$ 和 $\delta - 2\alpha$ 都大于零, 因此, 当 $0 < \omega < 2$ 时, 就有 $|\lambda|^2 < 1$ , 也就是说SOR收敛.  $\square$

### 4.3 最佳松弛因子

因为SOR的谱半径依赖于 $\omega$ , 当然会问能否适当选取 $\omega$ 使收敛速度最快? 这就是选择最佳松弛因子的问题.

要研究收敛速度在 $\omega$ 选取何值时为最快, 我们考虑特征值问题

$$L_\omega x = \lambda x,$$

即

$$[(\lambda - 1 + \omega)D - \lambda\omega L - \omega U]x = 0. \quad (4.4)$$

此外, 由于 $L_\omega$ 的特征值 $\lambda_1, \dots, \lambda_n$ 满足

$$\lambda_1 \cdot \lambda_2 \cdots \lambda_n = (1 - \omega)^n,$$

故当 $\omega \neq 1$ 时,  $L_\omega$ 无零特征值.

对于模型问题, 在自然次序排列下, (4.4)可写成

$$\begin{cases} (\lambda + \omega - 1)u_{ij} - \frac{\omega}{4}(\lambda u_{i-1,j} + \lambda u_{i,j-1} + u_{i+1,j} + u_{i,j+1}) = 0, \\ u_{i0} = u_{in} = u_{0j} = u_{nj} = 0, \quad 0 < i, j < n. \end{cases}$$

下面分两步来讨论

(1) 找出Jacobi迭代矩阵 $B$ 的特征值与 $L_\omega$ 的特征值之间的关系. 对 $\lambda \neq 0$ 作变换

$$u_{ij} = (\pm\lambda^{\frac{1}{2}})^{i+j} V_{ij},$$

则得

$$\mu V_{ij} - \frac{1}{4}(V_{i-1,j} + V_{i,j-1} + V_{i+1,j} + V_{i,j+1}) = 0,$$

其中

$$\mu = \pm \frac{\lambda + \omega - 1}{\omega \lambda^{1/2}}.$$

由此可知, 当 $\omega \neq 1$ 时, 若 $\lambda$ 是 $L_\omega$ 的特征值, 则由

$$(\lambda + \omega - 1)^2 = \mu^2 \omega^2 \lambda \quad (4.5)$$

或

$$(\lambda + \omega - 1) = \pm \mu \omega \lambda^{\frac{1}{2}}$$

所确定的两个 $\mu$ 都是矩阵 $B$ 的特征值.

反过来,  $\omega \neq 1$ , 若 $\mu$ 是 $B$ 的特征值, 将上述过程逆推, 不难得知, 由(4.5)确定的两个非零 $\lambda$ 必是 $L_\omega$ 的特征值.

前面关于 $\omega \neq 1$ 的限制可以去掉的. 事实上,  $\omega = 1$ 时(4.5)退化为

$$\lambda^2 = \mu^2 \lambda,$$

可以看成是 $\omega \rightarrow 1$ 时的极限情况,  $B$ 的特征值为 $\pm\mu_i$ , 对应的 $L_1$ 的特征值为0,  $\mu_i^2$ .

因此, 我们可以归纳出:  $L_\omega$ 的特征值 $\lambda$ 和 $B$ 的特征值 $\mu$ 之间有关系式(4.5), 即若 $\lambda$ 是 $L_\omega$ 的一个特征值, 则由(4.5)确定的 $\mu$ 是 $B$ 的特征值; 反之, 若 $\mu$ 是 $B$ 的特征值, 则由(4.5)确定的 $\lambda$ 是 $L_\omega$ 的特征值. 而且, 若 $\mu$ 是 $B$ 的特征值, 则 $-\mu$ 也是 $B$ 的特征值.

(2) 确定 $\rho(L_\omega)$ 随 $\omega$ 的变化情况.

从前面的讨论已经知道,  $L_\omega$ 的特征值完全由松弛参数 $\omega$ 和Jacobi迭代矩阵 $B$ 的特征值 $\mu$ 确定. 设 $0 \leq \mu < 1$ 是 $B$ 的一个特征值,  $0 < \omega < 2$ , 则由方程

$$\lambda + \omega - 1 = \pm \mu \omega \lambda^{\frac{1}{2}}, \quad (4.6)$$

可得 $L_\omega$ 的两个特征值分别为

$$\begin{aligned}\lambda_+(\omega, \mu) &= \left( \frac{\mu\omega}{2} + \sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)} \right)^2, \\ \lambda_-(\omega, \mu) &= \left( \frac{\mu\omega}{2} - \sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)} \right)^2.\end{aligned}$$

因为我们关心的是 $L_\omega$ 的谱半径随 $\omega$ 的变化情况, 因此我们感兴趣的是

$$M(\omega, \mu) = \max\{|\lambda_+(\omega, \mu)|, |\lambda_-(\omega, \mu)|\}. \quad (4.7)$$

由二次方程(4.6)的判别式为零, 即

$$\Delta = (\mu\omega)^2 - 4(\omega - 1) = 0. \quad (4.8)$$

得位于 $(0, 2)$ 之内的唯一根为

$$\omega_\mu = \frac{2}{1 + \sqrt{1 - \mu^2}}, \quad 0 < \omega_\mu < 2. \quad (4.9)$$

这样, 当 $0 < \omega < \omega_\mu$ 时,  $\Delta > 0$ , 从而有

$$\lambda_+(\omega, \mu) > \lambda_-(\omega, \mu) > 0;$$

而当 $\omega_\mu < \omega < 2$ 时,  $\Delta < 0$ , 故 $\lambda_+(\omega, \mu)$ 与 $\lambda_-(\omega, \mu)$ 互为共轭复数, 而且有 $|\lambda_+(\omega, \mu)| = |\lambda_-(\omega, \mu)| = \omega - 1$ . 综合上面的讨论即有

$$M(\omega, \mu) = \begin{cases} \lambda_+(\omega, \mu), & 0 < \omega \leq \omega_\mu, \\ \omega - 1, & \omega_\mu < \omega < 2. \end{cases}$$

先固定 $\mu$ , 令 $\omega$ 从0变到2, 看 $M(\omega, \mu)$ 的变化情况. 首先, 易证

$$M(\omega, \mu) < 1. \quad (4.10)$$

事实上, 当 $\omega_\mu < \omega < 2$ 时上式显然成立; 而当 $0 < \omega \leq \omega_\mu$ 时, 有

$$M(\omega, \mu) = \lambda_+(\omega, \mu) < \left( \frac{\mu\omega}{2} + \left(1 - \frac{\mu\omega}{2}\right) \right)^2 = 1.$$

其次, 当 $0 < \omega \leq \omega_\mu$ 时, 我们有

$$\begin{aligned}\frac{dM(\omega, \mu)}{d\omega} &= 2M(\omega, \mu)^{\frac{1}{2}} \left( \frac{\mu}{2} + \frac{\frac{1}{2}\omega\mu^2 - 1}{2\sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)}} \right) \\ &= M(\omega, \mu)^{\frac{1}{2}} \frac{\mu M(\omega, \mu)^{\frac{1}{2}} - 1}{\sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)}} < 0.\end{aligned} \quad (4.11)$$

于是, 当  $\omega \rightarrow \omega_\mu - 0$  时, 有

$$\frac{dM(\omega, \mu)}{d\omega} \rightarrow -\infty, \quad M(\omega, \mu) \rightarrow \frac{1 - \sqrt{1 - \mu^2}}{1 + \sqrt{1 - \mu^2}} = \omega_\mu - 1. \quad (4.12)$$

综合上面的讨论可知, 对固定的  $\mu$ ,  $M(\omega, \mu)$  在  $0 < \omega < \omega_\mu$  内是严格单调下降的, 而且在  $\omega_\mu$  达到极小, 在极小点的左侧导数为  $-\infty$ .

再固定  $\omega$ , 考虑  $M(\omega, \mu)$  随  $\mu$  的变化状态. 若  $0 < \mu_1 < \mu_2 < 1$  是  $B$  的两个特征值, 则有

$$M(\omega, \mu_1) \leq M(\omega, \mu_2),$$

即  $M(\omega, \mu)$  是  $\mu$  的增函数. 事实上, 当  $\omega_{\mu_2} \leq \omega < 2$  时, 上式显然成立; 当  $0 < \omega \leq \omega_{\mu_1}$  时, 有

$$M(\omega, \mu_2) - M(\omega, \mu_1) = \left( M(\omega, \mu_2)^{\frac{1}{2}} + M(\omega, \mu_1)^{\frac{1}{2}} \right) \left( \frac{\omega(\mu_2 - \mu_1)}{2} + \frac{\frac{\omega^2}{4}(\mu_2^2 - \mu_1^2)}{\sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)} + \sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)}} \right) > 0;$$

而当  $\omega_{\mu_1} < \omega \leq \omega_{\mu_2}$  时, 有

$$M(\omega, \mu_2) \geq \omega_{\mu_2} - 1 \geq \omega - 1 = M(\omega, \mu_1).$$

这样, 对  $B$  的每一个给定的特征值  $\mu$ , 我们都可以画出一条表明  $M(\omega, \mu)$  随  $\omega$  变化的曲线; 因为对确定的  $\omega$ ,  $M(\omega, \mu)$  又是  $\mu$  的增函数, 也就是说, 对  $B$  的较大的特征值所画出的曲线应在  $B$  的较小特征值画出的曲线的上边, 而  $\rho(B)$  所对应的则是  $\rho(L_\omega)$ , 在最上边; 参见图 4.3.

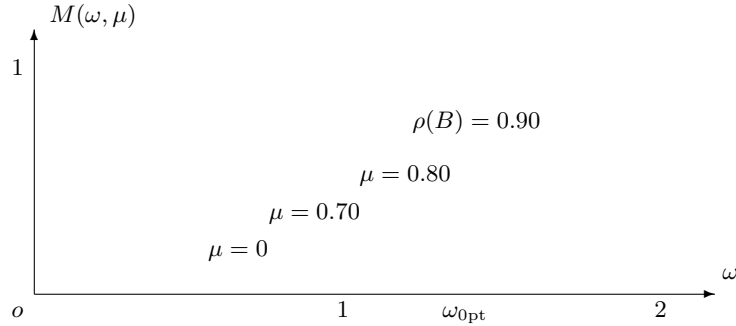


图 4.3

综合上面的讨论, 最后得到结论: 随着  $\omega$  从 0 增加,  $\rho(L_\omega)$  减少, 直至

$$\omega = \omega_{\text{opt}} = \frac{2}{1 + \sqrt{1 - \rho(B)^2}} \quad (4.13)$$

时,  $\rho(L_\omega)$  达到极小

$$\rho(L_{\omega_{\text{opt}}}) = \frac{1 - \sqrt{1 - \rho(B)^2}}{1 + \sqrt{1 - \rho(B)^2}}, \quad (4.14)$$

$\omega$  再增加时,  $\rho(L_\omega)$  开始增加. 因此,  $\omega_{\text{opt}}$  称为最佳松弛因子.

#### 4.4 渐近收敛速度

由(4.14)可得

$$\begin{aligned} R_\infty(L_{\omega_{\text{opt}}}) &= -\ln \rho(L_{\omega_{\text{opt}}}) = -\ln \frac{1 - \sqrt{1 - \rho(B)^2}}{1 + \sqrt{1 - \rho(B)^2}} \\ &= -\ln \frac{1 - \sin(h\pi)}{1 + \sin(h\pi)} \sim 2h\pi, \quad h \rightarrow 0. \end{aligned} \quad (4.15)$$

从上面结果看出, SOR收敛速度要比Jacobi迭代和G-S迭代的收敛速度快得多.

为了对Jacobi迭代法、G-S迭代法和SOR迭代法的收敛速度有一个更清晰的了解, 我们来看一个具体的例子. 现假定模型问题中的 $f(x, y) \equiv 0$ , 于是其真解为 $u(x, y) \equiv 0$ . 取 $h = 0.05$ , 初始向量 $x_0$ 的分量都取为1, 精度要求为 $\|x_k - x_*\|_\infty \leq 10^{-6}$ , 则Jacobi方法需迭代1154次, G-S方法需迭代578次, 而用 $\omega = 1.737$ 作为松弛因子的SOR方法仅需迭代54次. 由此可见, 对一些特殊的问题用带最佳松弛因子的SOR方法是很有有效的.

上面针对模型问题所推导出的结论是Franke给出的, 后来Young进行了推广.

#### 4.5 超松弛理论的推广

前面的讨论中, 我们是对模型问题来分析SOR迭代矩阵的谱半径与Jacobi 迭代矩阵谱半径之间的关系, 事实上我们可以把这些理论推广到更广泛的情形.

**定理4.4.5** 设方程组 $Ax = b$ 的系数矩阵 $A$ 有如下形式

$$A = \begin{bmatrix} D_1 & K_1 & & & \\ H_1 & D_2 & K_2 & & \\ & H_2 & D_3 & \ddots & \\ & & \ddots & \ddots & K_{t-1} \\ & & & H_{t-1} & D_t \end{bmatrix}, \quad (4.16)$$

其中 $D_i (i = 1, 2, \dots, t)$ 是非奇异的对角矩阵. 若 $B$ ,  $L_1$ 和 $L_\omega$ 分别是Jacobi 迭代, G-S迭代和SOR的迭代矩阵, 并假定 $B$ 的特征值均为实数, 那么

(1) 若 $\mu \neq 0$ 是矩阵 $B$ 的一个特征值, 则由

$$(\lambda + \omega - 1)^2 = \omega^2 \mu^2 \lambda$$

或

$$(\lambda + \omega - 1) = \pm \omega \mu \lambda^{\frac{1}{2}}$$

所确定的两个 $\lambda$ 是 $L_\omega$ 的特征值; 反之, 若 $\lambda \neq 0$ 是 $L_\omega$ 的特征值, 则由上式确定的两个 $\mu$ 都是 $B$ 的特征值;

(2)  $\rho(L_1) = (\rho(B))^2$ ,  $R_\infty(L_1) = 2R_\infty(B)$ ;

(3) 若 $\rho(B) < 1$ ,  $0 < \omega < 2$ , 则SOR收敛, 且最佳松弛因子为

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

相应的谱半径为

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

$$(4) \ R_{\infty}(L_{\omega_{\text{opt}}}) \sim 2\sqrt{2R_{\infty}(B)}, \quad \rho(B) \rightarrow 1 - 0.$$

**证明** 我们只需指出系数矩阵为(4.16)形状的Jacobi迭代矩阵 $B$ 的特征值和G-S迭代矩阵 $L_{\omega}$ 的特征值具有和模型问题相同的关系即可. 此时 $B$ 的特征值问题是求复数 $\mu$ 和非零向量 $\{V_i\}$ 满足(假定 $H_0 = 0, K_t = 0$ )

$$\begin{cases} H_{i-1}V_{i-1} + \mu D_i V_i + K_i V_{i+1} = 0, \\ V_0 = 0, \quad V_{t+1} = 0; \end{cases} \quad (4.17)$$

而 $L_{\omega}$ 的特征值问题则是计算复数 $\lambda$ 和非零向量 $\{u_i\}$ 满足

$$\begin{cases} \lambda \omega H_{i-1} u_{i-1} + (\lambda + \omega - 1) D_i u_i + \omega K_i u_{i+1} = 0, \\ u_0 = 0, \quad u_{t+1} = 0. \end{cases} \quad (4.18)$$

设 $\lambda \neq 0$ , 作变换 $u_i = \lambda^{\frac{i}{2}} V_i$ , 则有

$$\begin{cases} \lambda^{\frac{i+1}{2}} \omega H_{i-1} V_{i-1} + (\lambda + \omega - 1) D_i \lambda^{\frac{i}{2}} V_i + \lambda^{\frac{i+1}{2}} \omega K_i V_{i+1} = 0, \\ V_0 = 0, \quad V_{t+1} = 0. \end{cases}$$

两边消去 $\lambda^{\frac{i+1}{2}}$ 后再除以 $\omega$ , 得

$$\begin{cases} H_{i-1} V_{i-1} + \frac{\lambda + \omega - 1}{\omega \lambda^{1/2}} D_i V_i + K_i V_{i+1} = 0, \\ V_0 = 0, \quad V_{t+1} = 0. \end{cases} \quad (4.19)$$

把(4.19)与(4.17)相比较, 就可以得到 $B$ 的特征值 $\mu$ 与 $L_{\omega}$ 的特征值 $\lambda$ 之间与(4.5)完全一样的关系式. 这样, 只需把前面的证明平移过来即可证明结论(1), (2)和(3)成立.

至于(4)的证明, 可以假设

$$\rho = \rho(B) = 1 - \alpha, \quad \alpha \rightarrow 0 +.$$

则有

$$R_{\infty}(L_{\omega_{\text{opt}}}) = -\ln \frac{1 - \sqrt{1 - \rho^2}}{1 + \sqrt{1 - \rho^2}} \sim 2\sqrt{2\alpha}, \quad \alpha \rightarrow 0 +.$$

再注意到

$$R_{\infty}(B) = -\ln \rho \sim \alpha, \quad \alpha \rightarrow 0+,$$

便得到结论(4).  $\square$

上述定理要求线性方程组的系数矩阵是块三对角矩阵, 而且其对角块为对角矩阵. 对角块是对角阵这一要求太苛刻了些. 我们上面介绍的模型问题, 按自然次序排列, 它的系数矩阵虽是块三对角矩阵, 但其对角块并不是对角阵. 然而我们却已经对它建立了超松弛理论. 实际上超松弛理论可以适用于一大类更为广泛的矩阵. 为此, 我们引进相容性概念:



**定义4.4.1** 设 $A$ 是 $n$ 阶方阵, 记 $\mathcal{W} = \{1, 2, \dots, n\}$ . 称 $A$ 是具有**相容次序**的矩阵, 如果对某个 $t$ , 存在 $\mathcal{W}$ 的 $t$ 个互不相交的子集 $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_t$ 满足 $\bigcup_{i=1}^t \mathcal{S}_i = \mathcal{W}$ , 使得每个非零的非对角线元素 $a_{ij} \neq 0$  ( $i \neq j$ )的两个下标 $i, j$ 满足: 若 $i \in \mathcal{S}_r$ , 则当 $j < i$ 时,  $j \in \mathcal{S}_{r-1}$ ; 而当 $j > i$ 时,  $j \in \mathcal{S}_{r+1}$ .

作为一个例子, 考虑模型问题的系数矩阵 $A_{9 \times 9}$ . 若我们取

$$\begin{aligned}\mathcal{S}_1 &= \{1\}, & \mathcal{S}_2 &= \{2, 4\}, & \mathcal{S}_3 &= \{3, 5, 7\}, \\ \mathcal{S}_4 &= \{6, 8\}, & \mathcal{S}_5 &= \{9\},\end{aligned}$$

可以验证这些子集是满足定义的, 也就是说 $A$ 是具有相容性次序的矩阵.

对具有相容性次序的矩阵, 亦可建立矩阵 $L_\omega$ 和 $\omega$ 之间特征值的关系.

**定理4.4.6** 设矩阵 $A$ 具有相容次序, 且对角元全不为零, 并假定Jacobi迭代矩阵 $B = I - D^{-1}A$ 的特征值均为实数, 那么

- (1) 若 $\mu \neq 0$ 是 $B$ 的特征值, 则 $-\mu$ 也是 $B$ 的特征值;
- (2) 若 $\mu \neq 0$ 是 $B$ 矩阵的一个特征值, 则由(4.5)所确定的两个 $\lambda$ 是 $L_\omega$ 的特征值; 反之, 若 $\lambda \neq 0$ 是 $L_\omega$ 的特征值, 则由(4.5)确定的两个 $\mu$ 都是 $B$ 矩阵的特征值;
- (3)  $\rho(L_1) = (\rho(B))^2$ ;  $R_\infty(L_1) = 2R_\infty(B)$ ;
- (4) 若 $\rho(B) < 1$ ,  $0 < \omega < 2$ , 则SOR收敛, 且最佳松弛因子 $\omega_{\text{opt}}$ 由(4.13)确定, 相应的谱半径 $\rho(L_{\omega_{\text{opt}}})$ 由(4.14)确定;
- (5)  $R_\infty(L_{\omega_{\text{opt}}}) \sim 2\sqrt{2R_\infty(B)}$ ,  $\rho(B) \rightarrow 1 - 0$ .

因为对角元非零, 且具有相容次序的矩阵, 按照 $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_t$ 重排次序就得到一个形如(4.16)的矩阵, 并且将SOR应用于它们二者所得到的分量表示的迭代公式是完全一样的, 相应的迭代矩阵相似, 因此超松弛法的这些结论自然成立.

最后我们再给出如下定理.

**定理4.4.7** 如果矩阵 $A$ 对称正定, 并且具有相容次序,  $B = I - D^{-1}A$ , 则 $B$ 的特征值全是实数, 并且 $\rho(B) < 1$ .

**证明** 因为 $D = \text{diag}(a_{ii})$ 为正定对角矩阵, 故有

$$B = I - D^{-1}A = D^{-\frac{1}{2}}(I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}},$$

即 $B$ 与对称矩阵 $I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 相似, 因而 $B$ 的特征值全为实数.

因为对于正定对称矩阵来说, G-S迭代是收敛的, 再由定理4.4.6即可推出 $\rho(B) < 1$ . □

SOR方法收敛的速度比G-S方法和Jacobi方法在量级上有改进. 是一个很有效的方法. 还以模型问题为例, 若取 $h = 0.1$ , 这时 $R_\infty(B) \approx 0.05$ , 而 $R_\infty(L_{\omega_{\text{opt}}}) \approx 0.63$ .

松弛因子的选择, 对计算也是很有影响的. 在实际计算中, 因为 $\rho(B)$ 不一定知道, 所以 $\omega_{\text{opt}}$ 也不知道, 如何求近似的 $\omega_{\text{opt}}$ 呢? 通常是选用不同的 $\omega$ 的值, 然后用相同的初始向量, 进行试算迭代相同次数, 比较它们的残向量, 选择使残向量最小的 $\omega$ 作为松弛因子. 另外, 根据 $\rho(L_\omega)$ 随 $\omega$ 变化的曲线看, 我们在不能得到准确的最佳松弛因子时, 宁可取得稍大一些.