

第8章 解线性方程组的迭代法

张利军

zlj@nju.edu.cn

<https://ai.nju.edu.cn/zlj>





目录

□ 引言

□ Jacobi迭代法

□ Gauss-Seidel迭代法

□ 迭代法的收敛性

□ 解线性方程组的超松弛迭代法



迭代法

□ 考虑线性方程组，其中 **A** 为非奇异矩阵

$$Ax = b \quad (8.1.1)$$

- 当 **A** 为低阶稠密矩阵时，第7章所讨论的选主元素消去法解式(8.1.1)的有效方法
- 但是，对于由工程技术中产生的**大型稀疏矩阵**方程组（ **A** 的阶数 **n** 很大，但零元素较多，例如 $n \geq 10^4$ ），利用迭代法求解式(8.1.1)是合适的
- 在计算机内存和运算两方面，**迭代法**都可利用 **A** 中有大量零元素的特点



迭代法

□ 考察下列形式的方程：

$$x = \varphi(x) \quad (6.2.1)$$

■ 这种方程是隐式的，因而不能直接得出它的根

□ 如果给出根的某个猜测值 x_0 ，将它代入式(6.2.1)的右端，即可求得 $x_1 = \varphi(x_0)$

□ 然后，又可取 x_1 作为猜测值，进一步得到 $x_2 = \varphi(x_1)$



迭代法（续）

□ 如此反复迭代，如果按公式

$$x_{k+1} = \varphi(x_k), \quad k = 0, 1, \dots \quad (6.2.2)$$

确定的数列 $\{x_k\}$ 有极限 $x^* = \lim_{k \rightarrow \infty} x_k$ ，则称迭代过程式(6.2.2)收敛，这时极限值 x^* 显然就是方程(6.2.1)的根

□ 上述迭代法是一种逐次逼近法，基本思想是将隐式方程 $x = \varphi(x)$ 归结为一组显式的计算公式(6.2.2)

■ 迭代过程是逐步显式化的过程



例8.1

□ 求解方程组

$$\begin{cases} 8x_1 - 3x_2 + 2x_3 = 20 \\ 4x_1 + 11x_2 - x_3 = 33 \\ 6x_1 + 3x_2 + 12x_3 = 36 \end{cases} \quad (8.1.2)$$

■ 记作

$$Ax = b$$

其中

$$A = \begin{bmatrix} 8 & -3 & 2 \\ 4 & 11 & -1 \\ 6 & 3 & 12 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 20 \\ 33 \\ 36 \end{bmatrix}$$

■ 方程组的精确解是

$$x^* = (3, 2, 1)^T$$



例8.1（续）

■ 现将式(8.1.2)

$$\begin{cases} 8x_1 - 3x_2 + 2x_3 = 20 \\ 4x_1 + 11x_2 - x_3 = 33 \\ 6x_1 + 3x_2 + 12x_3 = 36 \end{cases} \quad (8.1.2)$$

改写成

$$\begin{cases} x_1 = \frac{1}{8}(3x_2 - 2x_3 + 20) \\ x_2 = \frac{1}{11}(-4x_1 + x_3 + 33) \\ x_3 = \frac{1}{12}(-6x_1 - 3x_2 + 36) \end{cases} \quad (8.1.3)$$



例8.1 (续)

■ 记为

$$x = B_0 x + f$$

其中

$$B_0 = \begin{bmatrix} 0 & \frac{3}{8} & -\frac{2}{8} \\ -\frac{4}{11} & 0 & \frac{1}{11} \\ -\frac{6}{12} & -\frac{3}{12} & 0 \end{bmatrix} = I - D^{-1}A, \quad f = \begin{bmatrix} \frac{20}{8} \\ \frac{33}{11} \\ \frac{36}{12} \end{bmatrix} = D^{-1}b$$

■ 一般推导:

$$Dx = (D - A)x + Ax = (D - A)x + b$$

$$x = D^{-1}(D - A)x + D^{-1}b = (I - D^{-1}A)x + D^{-1}b$$



例8.1（续）

- 任取初始值，例如取 $\mathbf{x}^{(0)} = (0, 0, 0)^T$ ，将这些值代入式(8.1.3)右端，得到新的值

$$\mathbf{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)})^T = (3.5, 3, 3)^T$$

✓ 若式(8.1.3)为等式即求得方程组的解

- 再将 $\mathbf{x}^{(1)}$ 分量代入式(8.1.3)右端，得到 $\mathbf{x}^{(2)}$

- 反复利用这个计算程序，得到一向量序列

$$\mathbf{x}^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix}, \quad \mathbf{x}^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix}, \dots, \quad \mathbf{x}^{(k)} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{bmatrix}, \dots$$



$$\mathbf{x}^* = (3, 2, 1)^T$$

例8.1 (续)

和一般计算公式 (迭代公式)

$$\begin{cases} x_1^{(k+1)} = (3x_2^{(k)} - 2x_3^{(k)} + 20) / 8 \\ x_2^{(k+1)} = (-4x_1^{(k)} + x_3^{(k)} + 33) / 11 \\ x_3^{(k+1)} = (-6x_1^{(k)} - 3x_2^{(k)} + 36) / 12 \end{cases} \quad (8.1.4)$$

■ 简写为 $\mathbf{x}^{(k+1)} = \mathbf{B}_0 \mathbf{x}^{(k)} + \mathbf{f}$

其中 k ($k = 0, 1, 2, \dots$) 表示迭代次数

■ 迭代到第10次有

逐步逼近方程组的精确解 $\mathbf{x}^*$

$$\mathbf{x}^{(10)} = (3.000032, 1.999838, 0.9998813)^T$$

$$\|\boldsymbol{\varepsilon}^{(10)}\|_{\infty} = 0.000187 \quad (\boldsymbol{\varepsilon}^{(10)} = \mathbf{x}^{(10)} - \mathbf{x}^*)$$



迭代法的收敛性

□ 对于任何一个方程组： $\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{f}$ ，按迭代法作出的向量序列 $\mathbf{x}^{(k)}$ 是否一定逐步逼近方程组的解 $\mathbf{x}^*$ 呢？

■ $\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{f}$ 是由 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 变形得到的等价方程组

□ 回答是**不一定**

■ 考虑用迭代法解下述方程组

$$\begin{cases} x_1 = 2x_2 + 5 \\ x_2 = 3x_1 + 5 \end{cases}$$



迭代法的收敛性（续）

- 对于给定方程组 $\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{f}$ ，设有唯一解 $\mathbf{x}^*$ ，则

$$\mathbf{x}^* = \mathbf{B}\mathbf{x}^* + \mathbf{f} \quad (8.1.5)$$

- 又设 $\mathbf{x}^{(0)}$ 为任取的初始向量，按下述公式构造向量序列

$$\begin{cases} \mathbf{x}^{(1)} = \mathbf{B}\mathbf{x}^{(0)} + \mathbf{f} \\ \mathbf{x}^{(2)} = \mathbf{B}\mathbf{x}^{(1)} + \mathbf{f} \\ \dots \\ \mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f} \end{cases} \quad (8.1.6)$$

- k 表示迭代次数



迭代法的收敛性（续）

□ 定义8.1

1. 对于给定的方程组 $\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{f}$ ，用式(8.1.6)逐步代入求近似解的方法称为**迭代法**（或称一阶定常迭代法，这里 $\mathbf{B}$ 与 k 无关）

$$\begin{cases} \mathbf{x}^{(1)} = \mathbf{B}\mathbf{x}^{(0)} + \mathbf{f} \\ \mathbf{x}^{(2)} = \mathbf{B}\mathbf{x}^{(1)} + \mathbf{f} \\ \dots \\ \mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f} \end{cases} \quad (8.1.6)$$

2. 如果 $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)}$ 存在（记作 $\mathbf{x}^*$ ），称此迭代法**收敛**，显然 $\mathbf{x}^*$ 就是方程组的解，否则称此迭代法**发散**



迭代法的收敛性（续）

□ 为了研究 $\{\mathbf{x}^{(k)}\}$ 的收敛性，引进误差向量

$$\boldsymbol{\varepsilon}^{(k+1)} = \mathbf{x}^{(k+1)} - \mathbf{x}^*$$

■ 由式(8.1.6)减去(8.1.5)，得

$$\left. \begin{array}{l} \mathbf{x}^* = \mathbf{B}\mathbf{x}^* + \mathbf{f} \\ \mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f} \end{array} \right\} \Rightarrow \boldsymbol{\varepsilon}^{(k+1)} = \mathbf{B}\boldsymbol{\varepsilon}^{(k)} \quad (k = 0, 1, 2 \dots)$$

■ 递推得

$$\boldsymbol{\varepsilon}^{(k)} = \mathbf{B}\boldsymbol{\varepsilon}^{(k-1)} = \dots = \mathbf{B}^k \boldsymbol{\varepsilon}^{(0)}$$

■ 要考察 $\{\mathbf{x}^{(k)}\}$ 的收敛性，就要研究 $\mathbf{B}$ 在什么条件下有 $\boldsymbol{\varepsilon}^{(k)} \rightarrow \mathbf{0} \ (k \rightarrow \infty)$ ，亦即要研究 $\mathbf{B}$ 满足什么条件时有 $\mathbf{B}^k \rightarrow \mathbf{0}$ （零矩阵） $(k \rightarrow \infty)$



目录

□ 引言

□ Jacobi迭代法

□ Gauss-Seidel迭代法

□ 迭代法的收敛性

□ 解线性方程组的超松弛迭代法



Jacobi (雅可比) 迭代法

□ 设有方程组

$$\sum_{j=1}^n a_{ij}x_j = b_i \quad (i = 1, 2, \dots, n)$$

■ 记作 $Ax = b$ (8.2.1)

A 为非奇异矩阵且 $a_{ii} \neq 0$ ($i = 1, 2, \dots, n$)

□ 将 A 分裂为 $A = D - L - U$, 其中

$$D = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & \ddots \\ & & & & a_{nn} \end{bmatrix}$$



Jacobi迭代法（续）

$$L = - \begin{bmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \quad U = - \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 0 & a_{23} & \cdots & a_{2n} \\ & & 0 & \ddots & \vdots \\ & & & \ddots & a_{n-1,n} \\ & & & & 0 \end{bmatrix}$$

□ 将 $Ax = b$ 第 i ($i = 1, 2, \dots, n$)个方程用 a_{ii} 去除再移项，得到等价方程组

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j \right) \quad (i = 1, 2, \dots, n) \quad (8.2.2)$$



Jacobi迭代法（续）

$$A = D - L - U$$

□ 简记作

$$\mathbf{x} = B_0 \mathbf{x} + \mathbf{f}$$

■ 其中 $B_0 = I - D^{-1}A = D^{-1}(D - A) = D^{-1}(L + U)$

$$\mathbf{f} = D^{-1}\mathbf{b}$$

□ 对方程组(8.2.2)应用迭代法，得到解式(8.2.1)的Jacobi迭代公式

$$\begin{cases} \mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T \text{ (初始向量)} \\ x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right) \end{cases} \quad (8.2.3)$$

■ 其中 $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})^T$ 为第 k 次迭代向量



Jacobi迭代法（续）

□ 显然迭代公式(8.2.3)的矩阵形式为

$$\begin{cases} \mathbf{x}^{(0)} & \text{(初始向量)} \\ \mathbf{x}^{(k+1)} = \mathbf{B}_0 \mathbf{x}^{(k)} + \mathbf{f} \end{cases} \quad (8.2.4)$$

■ 其中 $\mathbf{B}_0$ 称为Jacobi方法迭代矩阵

□ Jacobi迭代法公式简单

- 每迭代一次只需计算一次矩阵和向量乘法
- 在用计算机计算时，需要两组工作单元，以存储 $\mathbf{x}^{(k)}$ 及 $\mathbf{x}^{(k+1)}$



目录

- 引言
- Jacobi迭代法
- Gauss-Seidel迭代法
- 迭代法的收敛性
- 解线性方程组的超松弛迭代法

Gauss-Seidel (高斯-赛德尔) 迭代法



□ 由Jacobi方法迭代公式(8.2.4)可知

$$\begin{cases} \mathbf{x}^{(0)} & \text{(初始向量)} \\ \mathbf{x}^{(k+1)} = \mathbf{B}_0 \mathbf{x}^{(k)} + \mathbf{f} \end{cases} \quad (8.2.4)$$

- 迭代的每一步计算过程，都是用 $\mathbf{x}^{(k)}$ 的全部分量来计算 $\mathbf{x}^{(k+1)}$ 的所有分量
- 显然在计算第 i 个分量时，已经计算出的最新分量 $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ 没有被利用
- 从直观上看，最新计算出的分量可能比旧的分量要好些。因此，对这些最新计算出来的 $\mathbf{x}^{(k+1)}$ 的分量 $x_j^{(k+1)}$ 加以利用，就得到所谓解方程组的

Gauss-Seidel迭代法 (简称G-S方法)



Gauss-Seidel迭代法（续）

□ Gauss-Seidel迭代法

$$\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T \quad (\text{初始向量})$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) \\ (k = 0, 1, 2, \dots; i = 1, 2, \dots, n), \quad (8.2.5)$$

□ Jacobi迭代公式

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



Gauss-Seidel迭代法（续）

□ Gauss-Seidel迭代法

$$\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)} \right)^T \quad (\text{初始向量})$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) \\ (k = 0, 1, 2, \dots; i = 1, 2, \dots, n), \quad (8.2.5)$$

- 第2个式子利用了最新计算出的分量 $x_1^{(k+1)}$
- 第 i 个式子利用了计算出的最新分量 $x_j^{(k+1)}$ ($j = 1, 2, \dots, i - 1$)



Gauss-Seidel迭代法（续）

□ Gauss-Seidel迭代法

$$\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)} \right)^T \quad (\text{初始向量})$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) \quad (k = 0, 1, 2, \dots; i = 1, 2, \dots, n), \quad (8.2.5)$$

■ 可写成

$$\begin{cases} x_i^{(k+1)} = x_i^{(k)} + \Delta x_i & (k = 0, 1, 2, \dots; i = 1, 2, \dots, n) \\ \Delta x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right) \end{cases}$$



Gauss-Seidel迭代法（续）

□ 矩阵形式

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

$$L = - \begin{bmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \quad U = - \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 0 & a_{23} & \cdots & a_{2n} \\ & & 0 & \ddots & \vdots \\ & & & \ddots & a_{n-1,n} \\ & & & & 0 \end{bmatrix}$$

□ Jacobi迭代公式

$$\left. \begin{aligned} \mathbf{x}^{(k+1)} &= B_0 \mathbf{x}^{(k)} + \mathbf{f} \\ B_0 &= D^{-1}(L + U) \\ \mathbf{f} &= D^{-1}\mathbf{b} \end{aligned} \right\} \Rightarrow D\mathbf{x}^{(k+1)} = (L + U)\mathbf{x}^{(k)} + \mathbf{b}$$



Gauss-Seidel迭代法（续）

□ 矩阵形式

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

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

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

■ 假设 $(D - L)^{-1}$ 存在

□ 可改写为

$$\mathbf{x}^{(k+1)} = G\mathbf{x}^{(k)} + \mathbf{f} \quad (8.2.6)$$

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

□ Jacobi迭代公式

$$\mathbf{x}^{(k+1)} = B_0\mathbf{x}^{(k)} + \mathbf{f}$$



Gauss-Seidel迭代法（续）

- 应用Gauss-Seidel迭代法解式(8.2.1)就是对方程组 $\mathbf{x} = \mathbf{G}\mathbf{x} + \mathbf{f}$ 应用迭代法
 - $\mathbf{G}$ 称为Gauss-Seidel迭代法的迭代矩阵
- Gauss-Seidel迭代法的一个明显的优点是，在用计算机计算时，只需一组工作单元，以便存放近似解，此外，每迭代一步只需计算一次矩阵与向量的乘法

$$\Delta x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right)$$



例8.2

□ 用Gauss-Seidel迭代法解例8.1

$$\begin{cases} 8x_1 - 3x_2 + 2x_3 = 20 \\ 4x_1 + 11x_2 - x_3 = 33 \\ 6x_1 + 3x_2 + 12x_3 = 36 \end{cases} \quad (8.1.2)$$

■ 取 $\mathbf{x}^{(0)} = \mathbf{0}$ ，迭代公式为

$$\begin{cases} x_1^{(k+1)} = (20 + 3x_2^{(k)} - 2x_3^{(k)})/8 \\ x_2^{(k+1)} = (33 - 4x_1^{(k+1)} + x_3^{(k)})/11 \\ x_3^{(k+1)} = (36 - 6x_1^{(k+1)} - 3x_2^{(k+1)})/12 \end{cases}$$



$$\mathbf{x}^* = (3, 2, 1)^T$$

例8.2（续）

- 迭代到第5次时，有

$$\mathbf{x}^{(5)} = (2.999843, 2.000072, 1.000061)^T$$

$$\|\boldsymbol{\varepsilon}^{(5)}\|_{\infty} = 0.00157$$

- 从此例看出，Gauss-Seidel迭代法比Jacobi迭代法收敛快（达到同样的精度所需的迭代次数较少）
- 但这个结论在**一定条件下才是对的**，甚至有这样的方程组，Jacobi方法收敛，而Gauss-Seidel迭代法却是发散的



例8.3

□ 方程组

$$\begin{cases} x_1 + 2x_2 - 2x_3 = 1 \\ x_1 + x_2 + x_3 = 1 \\ 2x_1 + 2x_2 + x_3 = 1 \end{cases}$$

- 能够说明解此方程组的Jacobi迭代法收敛而Gauss-Seidel迭代法发散



目录

- 引言
- Jacobi迭代法
- Gauss-Seidel迭代法
- 迭代法的收敛性
- 解线性方程组的超松弛迭代法



矩阵收敛

□ 定义**8.2** 设有矩阵序列 $\mathbf{A}_k = (a_{ij}^{(k)})_{n \times n} (k = 1, 2, \dots)$ 及 $\mathbf{A} = (a_{ij})_{n \times n}$, 如果

$$\lim_{k \rightarrow \infty} a_{ij}^{(k)} = a_{ij} \quad (i, j = 1, 2, \dots, n)$$

成立, 则称 $\{\mathbf{A}_k\}$ 收敛于 $\mathbf{A}$, 记作 $\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{A}$

□ 例**8.4** 矩阵序列

$$\mathbf{A} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \mathbf{A}^2 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}, \dots, \mathbf{A}^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}, \dots,$$

当 $|\lambda| < 1$ 时, $\mathbf{A}^k \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ (当 $k \rightarrow \infty$ 时)



矩阵收敛（续）

□ 矩阵序列极限的概念可以用任何矩阵范数来描述

□ 定理**8.1** $\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{A}$ 的充要条件是 $\|\mathbf{A}_k - \mathbf{A}\| \rightarrow 0 \ (k \rightarrow \infty)$

□ 由8.1节知，要考虑序列 $\{\mathbf{x}^{(k)}\}$ 的收敛性，就要研究 $\mathbf{B}$ 在什么条件下误差向量趋于零向量，即

$$\boldsymbol{\varepsilon}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^* = \mathbf{B}^k \boldsymbol{\varepsilon}^{(0)} \rightarrow \mathbf{0} \quad (k \rightarrow \infty)$$



迭代法收敛的充要条件

□ 定理8.2 设 $B = (b_{ij})_{n \times n}$, 则 $B^k \rightarrow O (k \rightarrow \infty)$ 的充要条件是 $\rho(B) < 1$

■ 由矩阵 B 的Jordan标准形, 存在非奇异矩阵 P , 使

$$P^{-1}BP = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_r \end{bmatrix} \equiv J \quad J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}_{n_i \times n_i}$$

其中 λ_i 为第 i 个特征值, $\sum_{i=1}^r n_i = n$

https://en.wikipedia.org/wiki/Jordan_normal_form



定理8.2证明

■ 显然 $\mathbf{B} = \mathbf{PJP}^{-1}$, 故

$$\mathbf{B}^k = (\mathbf{PJP}^{-1})^k = \mathbf{PJ}^k \mathbf{P}^{-1}$$

其中

$$\mathbf{J}^k = \begin{bmatrix} J_1^k & & & \\ & J_2^k & & \\ & & \ddots & \\ & & & J_r^k \end{bmatrix} (k = 1, 2, \dots)$$

■ 于是

$$\mathbf{B}^k \rightarrow \mathbf{O} \quad (k \rightarrow \infty) \Leftrightarrow \mathbf{J}^k \rightarrow \mathbf{O} \quad (k \rightarrow \infty)$$

$$\mathbf{J}^k \rightarrow \mathbf{O} (k \rightarrow \infty) \Leftrightarrow J_i^k \rightarrow \mathbf{O} (k \rightarrow \infty) \quad (i = 1, 2, \dots, r)$$



定理8.2证明（续）

■ 引进记号

$$E_{tk} = \begin{bmatrix} \overbrace{0 \quad \cdots \quad 0}^{k \text{ 个}} & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ & & & & 0 \\ & & & & \vdots \\ & & & & 0 \end{bmatrix}_{t \times t} \quad (\text{其中 } t = n_i)$$

■ 可以证明当 $k < t$ 时, $(E_{t1})^k = E_{tk}$; 当 $k \geq t$ 时 $(E_{t1})^k = \mathbf{0}$



定理8.2证明（续）

■ 显然

$$J_i = \begin{bmatrix} \lambda_i & & & \\ & \lambda_i & & \\ & & \ddots & \\ & & & \ddots \\ & & & & \lambda_i \end{bmatrix} + \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$$
$$= \lambda_i \mathbf{I} + \mathbf{E}_{t1}$$

■ 可得

$$J_i^k = (\lambda_i \mathbf{I} + \mathbf{E}_{t1})^k = \sum_{j=0}^k \binom{k}{j} \lambda_i^{k-j} (\mathbf{E}_{t1})^j = \sum_{j=0}^k \binom{k}{j} \lambda_i^{k-j} \mathbf{E}_{tj}$$
$$(\text{当 } k \geq t \text{ 时 } \mathbf{E}_{tk} = \mathbf{O}) = \sum_{j=0}^{t-1} \binom{k}{j} \lambda_i^{k-j} \mathbf{E}_{tj}$$



定理8.2证明（续）

■ 进而

$$J_i^k = (\lambda_i I + E_{t1})^k = \sum_{j=0}^{t-1} \binom{k}{j} \lambda_i^{k-j} E_{tj}$$
$$= \begin{bmatrix} \lambda_i^k & \binom{k}{1} \lambda_i^{k-1} & \dots & \dots & \binom{k}{t-1} \lambda_i^{k-(t-1)} \\ & \lambda_i^k & \binom{k}{1} \lambda_i^{k-1} & \dots & \binom{k}{t-2} \lambda_i^{k-(t-2)} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \binom{k}{1} \lambda_i^{k-1} \\ & & & & \lambda_i^k \end{bmatrix}$$

$$\text{其中 } \binom{k}{j} = \frac{k!}{j!(k-j)!} = \frac{k(k-1)\cdots(k-j+1)}{j!}$$



定理8.2证明（续）

- 利用极限 $\lim_{k \rightarrow \infty} k^r c^k = 0$ ($0 < c < 1, r \geq 0$), 得到

$$\binom{k}{j} \lambda^{k-j} \rightarrow 0 \ (k \rightarrow \infty) \Leftrightarrow |\lambda| < 1$$

✓ 注意到

$$\binom{k}{j} = \frac{k!}{j! (k-j)!} = \frac{k(k-1) \cdots (k-j+1)}{j!}$$

- 所以 $\mathbf{J}_i^k \rightarrow \mathbf{O}$ ($k \rightarrow \infty$) 充要条件是 $|\lambda_i| < 1$ ($i = 1, 2, \dots, r$), 即 $\rho(\mathbf{B}) < 1$



迭代法基本定理

□ 定理8.3 设有方程组

$$\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f} \quad (8.3.1)$$

对于任意初始向量 $\boldsymbol{x}^{(0)}$ 及任意 $\boldsymbol{f}$ ，解此方程组的迭代法

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f}$$

收敛的充要条件是 $\rho(\boldsymbol{B}) < 1$

□ 充分性

■ 设 $\rho(\boldsymbol{B}) < 1$ ，由定理7.18， $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{B}$ 为非奇异矩阵。因此， $\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{f}$ 有唯一解，记作

$$\boldsymbol{x}^* = \boldsymbol{B}\boldsymbol{x}^* + \boldsymbol{f} \quad (8.3.2)$$



定理8.3证明

■ 误差向量

$$\boldsymbol{\varepsilon}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}^* = \boldsymbol{B}^k \boldsymbol{\varepsilon}^{(0)}, \quad \boldsymbol{\varepsilon}^{(0)} = \boldsymbol{x}^{(0)} - \boldsymbol{x}^*$$

■ 由于 $\rho(\boldsymbol{B}) < 1$, 应用定理8.2, 有 $\boldsymbol{B}^k \rightarrow \boldsymbol{0} \ (k \rightarrow \infty)$.

■ 于是对于任意 $\boldsymbol{x}^{(0)}$ 及 $\boldsymbol{f}$, 有 $\boldsymbol{\varepsilon}^{(k)} \rightarrow \boldsymbol{0} \ (k \rightarrow \infty)$, 即 $\boldsymbol{x}^{(k)} \rightarrow \boldsymbol{x}^* \ (k \rightarrow \infty)$

□ 必要性

■ 设对于任意 $\boldsymbol{x}^{(0)}$ 及任意 $\boldsymbol{f}$, 皆有 $\lim_{k \rightarrow \infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$, 其中 $\boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f}$



定理8.3证明（续）

■ 显然，

1. 极限 $\boldsymbol{x}^*$ 是方程组 $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f}$ 的解

2. 对于任意 $\boldsymbol{x}^{(0)}$ 及任意 $\boldsymbol{f}$ ，有

$$\boldsymbol{\varepsilon}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}^* = \boldsymbol{B}^k \boldsymbol{\varepsilon}^{(0)} \rightarrow \mathbf{0} \quad (k \rightarrow \infty)$$

■ 由第2点可知， $\boldsymbol{B}^k \rightarrow \mathbf{0} \quad (k \rightarrow \infty)$

■ 再次应用定理8.2，即得 $\rho(\boldsymbol{B}) < 1$



例8.5

□ 考察用Jacobi方法解例8.1方程组的收敛性

$$\begin{cases} 8x_1 - 3x_2 + 2x_3 = 20 \\ 4x_1 + 11x_2 - x_3 = 33 \\ 6x_1 + 3x_2 + 12x_3 = 36 \end{cases} \quad (8.1.2)$$

■ Jacobi方法

$$x = B_0 x + f$$

其中

$$B_0 = \begin{bmatrix} 0 & \frac{3}{8} & -\frac{2}{8} \\ -\frac{4}{11} & 0 & \frac{1}{11} \\ -\frac{6}{12} & -\frac{3}{12} & 0 \end{bmatrix}$$



例8.5（续）

■ 迭代矩阵 $\mathbf{B}_0$ 的特征方程为

$$\det(\lambda \mathbf{I} - \mathbf{B}_0) = \lambda^3 + 0.034090909\lambda + 0.039772727 = 0$$

■ 解得 $\lambda_1 = -0.3082$

$$\lambda_2 = 0.1541 + i0.3245$$

$$\lambda_3 = 0.1541 - i0.3245,$$

■ 可知

$$|\lambda_1| < 1, \quad |\lambda_2| = |\lambda_3| = 0.3592 < 1$$

即 $\rho(\mathbf{B}_0) < 1$ ，所以用Jacobi迭代法是收敛的



例8.6

□ 考察用迭代过程的收敛性。

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f} \quad (k = 0, 1, 2, \dots)$$

□ 其中 $\mathbf{B} = \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$

■ 特征方程为

$$\det(\lambda \mathbf{I} - \mathbf{B}) = \lambda^2 - 6 = 0$$

■ 特征根 $\lambda_{1,2} = \pm\sqrt{6}$, 即 $\rho(\mathbf{B}) > 1$

■ 说明此迭代过程不收敛



迭代法的收敛速度

□ 考察误差向量 $\boldsymbol{\varepsilon}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}^* = \boldsymbol{B}^k \boldsymbol{\varepsilon}^{(0)}$

- 设 $\boldsymbol{B}$ 有 n 个线性无关的特征向量 $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$, 相应的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$
- 由于线性无关假设, $\boldsymbol{\varepsilon}^{(0)} = \sum_{i=1}^n a_i \boldsymbol{u}_i$, 得

$$\boldsymbol{\varepsilon}^{(k)} = \boldsymbol{B}^k \boldsymbol{\varepsilon}^{(0)} = \sum_{i=1}^n a_i \boldsymbol{B}^k \boldsymbol{u}_i = \sum_{i=1}^n a_i \lambda_i^k \boldsymbol{u}_i$$

- 可以看出, 当 $\rho(\boldsymbol{B}) < 1$ 愈小时, $\lambda_i^k \rightarrow 0$ ($i = 1, 2, \dots, n; k \rightarrow \infty$) 愈快, 即 $\boldsymbol{\varepsilon}^{(k)} \rightarrow \mathbf{0}$ 愈快, 故可用量 $\rho(\boldsymbol{B})$ 来刻画迭代法的收敛快慢



迭代法的收敛速度（续）

- 现在依据给定精度要求来确定迭代次数 k ，即使

$$[\rho(\mathbf{B})]^k \leq 10^{-s} \quad (8.3.3)$$

- 取对数得

$$k \geq \frac{s \ln 10}{-\ln \rho(\mathbf{B})}$$

- 定义**8.3** 称 $R(\mathbf{B}) = -\ln \rho(\mathbf{B})$ 为迭代法的收敛速度

- 可以看出 $\rho(\mathbf{B}) < 1$ 愈小， $-\ln \rho(\mathbf{B})$ 就愈大，式(8.3.3)成立所需迭代次数就愈少



迭代矩阵的特征值

□ 矩阵特征值的重要性

- 定理8.3表明迭代法收敛的充要条件是 $\rho(\mathbf{B}) < 1$
- 收敛速度由 $R(\mathbf{B}) = -\ln \rho(\mathbf{B})$ 决定

□ 一般当 n 较大时，矩阵特征值计算比较困难

- 定理8.3的条件比较难验证，所以最好建立与矩阵元素直接有关的条件来判别迭代法的收敛性
- 由于 $\rho(\mathbf{B}) \leq \|\mathbf{B}\|_v (v = 1, 2, \infty \text{ 或 } F)$ ，所以可用 $\|\mathbf{B}\|_v$ 来作为 $\rho(\mathbf{B})$ 上界的一种估计



迭代法收敛的充分条件

□ 定理**8.4** 如果方程组 $\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{f}$ 的迭代公式为

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}$$

其中 $\mathbf{x}^0$ 为任意初始向量，且迭代矩阵的某一范数 $\|\mathbf{B}\|_v = q < 1$ ，则：

1. 迭代法收敛

2. $\|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v \leq \frac{q}{1-q} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_v$

3. $\|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v \leq \frac{q^k}{1-q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_v$



定理8.4证明

■ 根据定理8.3和 $\rho(\mathbf{B}) \leq \|\mathbf{B}\|_v$ ，第1点显然成立

■ 由 $\mathbf{x}^* = \mathbf{B}\mathbf{x}^* + \mathbf{f}$

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}$$

可得

$$\mathbf{x}^* - \mathbf{x}^{(k+1)} = \mathbf{B}(\mathbf{x}^* - \mathbf{x}^{(k)}) \quad (8.3.4)$$

■ 根据范数性质，可得

$$\begin{aligned} \|\mathbf{x}^* - \mathbf{x}^{(k+1)}\|_v &= \|\mathbf{B}(\mathbf{x}^* - \mathbf{x}^{(k)})\|_v \\ &\leq \|\mathbf{B}\|_v \|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v = q \|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v \\ &\quad (k = 1, 2, \dots) \end{aligned} \quad (8.3.6)$$

$$\|\mathbf{x}^* - \mathbf{x}^{(k+1)}\|_v \leq q \|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v$$



定理8.4证明（续）

(8.3.6)

■ 类似可得,

$$\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} = \mathbf{B}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)})$$

■ 因此,

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_v \leq q \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_v \quad (8.3.5)$$

■ 另一方面,

$$\begin{aligned} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_v &= \|\mathbf{x}^* - \mathbf{x}^{(k)} - (\mathbf{x}^* - \mathbf{x}^{(k+1)})\|_v \\ &\geq \|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v - \|\mathbf{x}^* - \mathbf{x}^{(k+1)}\|_v \\ &\geq (1 - q) \|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v \end{aligned}$$



$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_v \leq q \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_v$$

定理8.4证明（续）

(8.3.5)

■ 结合 $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_v \leq q \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_v$

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_v \geq (1 - q) \|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v$$

可证明第2点成立：

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v \leq \frac{q}{1 - q} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_v$$

■ 对上式重复利用(8.3.5)，可得第3点成立：

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v \leq \frac{q^k}{1 - q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_v$$



例8.7

□ 任然考察例8.1, Jacobi方法

$$\boldsymbol{x} = \boldsymbol{B}_0 \boldsymbol{x} + \boldsymbol{f} \quad \boldsymbol{B}_0 = \begin{bmatrix} 0 & \frac{3}{8} & -\frac{2}{8} \\ -\frac{4}{11} & 0 & \frac{1}{11} \\ -\frac{6}{12} & -\frac{3}{12} & 0 \end{bmatrix}$$

■ 迭代矩阵 $\boldsymbol{B}_0$ 的 ∞ -范数为

$$\|\boldsymbol{B}\|_{\infty} = \max_{1 \leq i \leq 3} \sum_{j=1}^3 |b_{ij}| = \max \left\{ \frac{5}{8}, \frac{5}{11}, \frac{9}{12} \right\} = \frac{9}{12} < 1$$

■ 所以对例8.1应用Jacobi迭代方法是收敛的



例8.8

□ 设有迭代过程 $\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}$ ($k = 0, 1, 2, \dots$), 其中

$$\mathbf{B} = \begin{pmatrix} 0.9 & 0 \\ 0.3 & 0.8 \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

■ 可得

$$\|\mathbf{B}\|_{\infty} = 1.1, \quad \|\mathbf{B}\|_1 = 1.2$$

$$\|\mathbf{B}\|_2 = 1.021, \quad \|\mathbf{B}\|_F = \sqrt{1.54}$$

■ 虽然 $\mathbf{B}$ 的这些范数都大于1, 但 $\mathbf{B}$ 的特征值为 $\lambda_1 = 0.9, \lambda_2 = 0.8$, 由定理8.3知, 此迭代过程还是收敛的



定理8.4的应用

□ 由定理8.4可知, $\|B\| = q < 1$ 愈小, 迭代法收敛愈快

□ 当 B 的某一种范数 $\|B\| < 1$ 时, 如果相邻两次迭代

$$\|x^{(k)} - x^{(k-1)}\| < \varepsilon_0$$

ε_0 为给定的精度要求, 则

$$\|x^* - x^{(k)}\| \leq \frac{q}{1-q} \|x^{(k)} - x^{(k-1)}\|_v \leq \frac{q}{1-q} \varepsilon_0$$

所以在用计算机计算时通常利用 $\|x^{(k)} - x^{(k-1)}\| < \varepsilon_0$ 来作为控制迭代的终止条件



定理8.4的应用（续）

- 不过要注意，当 $q \approx 1$ 时， $\frac{q}{1-q}$ 大，尽管 $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$ 已非常小，但误差向量的模 $\|\boldsymbol{\varepsilon}^{(k)}\| = \|\mathbf{x}^* - \mathbf{x}^{(k)}\|$ 可能较大，迭代法收敛将是缓慢的
- 定理8.4中结论3的估计可以用来事先确定需要迭代多少次才能保证 $\|\boldsymbol{\varepsilon}^{(k)}\| < \varepsilon_0$

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\|_v \leq \frac{q^k}{1-q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_v$$



Gauss-Seidel迭代法的收敛性

□ 定理**8.5** 解方程组 $Ax = b$ 的Gauss-Seidel迭代法收敛的充要条件是 $\rho(G) < 1$ ，其中 G 为Gauss-Seidel迭代法的迭代矩阵

■ Gauss-Seidel迭代法

$$x^{(k+1)} = Gx^{(k)} + f \quad (8.2.6)$$

■ 对式(8.2.6)应用定理8.3



特殊系数矩阵

- 在实际应用中常遇到一些线性代数方程组，其系数矩阵具有某些性质
 - 系数矩阵的对角元素占优
 - 系数矩阵可约、不可约
 - 系数矩阵为对称正定等

- 充分利用这些性质往往可使判定迭代法收敛的问题变得简单



对角占优矩阵

□ 定义8.4 （对角占优阵） 设 $\mathbf{A} = (a_{ij})_{n \times n} \in \mathbf{R}^{n \times n}$

1. 如果矩阵 $\mathbf{A}$ 满足条件

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

即 $\mathbf{A}$ 的每一行对角元素的绝对值都严格大于同行其他元素绝对值之和，则称 $\mathbf{A}$ 为严格对角占优阵

2. 如果 $|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}| \quad (i = 1, 2, \dots, n)$ 且至少有一个不等式严格成立，称 $\mathbf{A}$ 为弱对角占优阵



例8.9

□ 显然

■ $A = \begin{bmatrix} -4 & 1 & 0 & 0 \\ 1 & -4 & 1 & 0 \\ 0 & 1 & -4 & 1 \\ 0 & 0 & 1 & -4 \end{bmatrix}$ 为严格对角占优阵

■ $B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$ 为弱对角占优阵



可约与不可约矩阵

□ 定义 **8.5** （可约与不可约矩阵） 设 $\mathbf{A} = (a_{ij})_{n \times n} \in \mathbf{R}^{n \times n}$ ，当 $n \geq 2$ 时，如果存在 n 阶置换阵 $\mathbf{P}$ 使

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix} \quad (8.3.8)$$

成立，其中 $\mathbf{A}_{11}$ 为 r 阶子矩阵， $\mathbf{A}_{22}$ 为 $n - r$ 阶子矩阵（ $1 \leq r \leq n$ ），则称 $\mathbf{A}$ 是 **可约矩阵**。如果不存在置换阵 $\mathbf{P}$ 使式 (8.3.8) 成立，则称 $\mathbf{A}$ 是 **不可约矩阵**。



可约与不可约矩阵（续）

- A 是可约矩阵，意味着 $Ax = b$ 可经过若干行列重排，化为两个低阶方程组求解
 - 若 A 经过两行交换的同时进行相应的两列的交换，称对 A 进行一次行列重排
- 由 $Ax = b$ 可化为 $P^T A P (P^T x) = P^T b$ ，记

$$y = P^T x = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad P^T b = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

其中 y_1, d_1 为 r 维向量

- $Ax = b$ 可化为求解 $\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$



可约与不可约矩阵（续）

- A 是可约矩阵，意味着 $Ax = b$ 可经过若干行列重排，化为两个低阶方程组求解
 - 若 A 经过两行交换的同时进行相应的两列的交换，称对 A 进行一次行列重排
- 由 $Ax = b$ 可化为 $P^T AP(P^T x) = P^T b$ ，记

$$y = P^T x = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad P^T b = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

其中 y_1, d_1 为 r 维向量

- $Ax = b$ 可化为求解
$$\begin{cases} A_{11}y_1 + A_{12}y_2 = d_1 \\ A_{22}y_2 = d_2 \end{cases}$$



例8.10

□ 在例8.9中矩阵***B***是可约矩阵

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

■ 即存在置换阵 $\mathbf{P} = \mathbf{I}_{13}$, 使

$$\mathbf{P}^T \mathbf{B} \mathbf{P} = \left[\begin{array}{c|cc} 2 & 1 & 0 \\ \hline 0 & 1 & 1 \\ 0 & 1 & 1 \end{array} \right]$$



对角占优定理

□ 定理8.6 如果 $A = (a_{ij})_{n \times n} \in \mathbf{R}^{n \times n}$ 为严格对角占优阵或不可约弱对角占优阵，则 A 是非奇异矩阵

■ 设 A 为严格对角占优阵，下面只就这种情况证明此定理

■ 若 $\det A = 0$ ，则 $Ax = 0$ 有非零解，记作

$$x = (x_1, x_2, \dots, x_n)^T$$

■ 又记 $x_k = \max_{1 \leq i \leq n} |x_i| \neq 0$



对角占优定理（续）

- 由齐次方程组 $\mathbf{Ax} = \mathbf{0}$ 的第 k 个方程

$$\sum_{j=1}^n a_{kj}x_j = 0$$

可得

$$|a_{kk}x_k| = \left| \sum_{\substack{j=1 \\ j \neq k}}^n a_{kj}x_j \right| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}| |x_j| \leq |x_k| \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}|$$

- 因此

$$|a_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}|$$

与假设矛盾，故 $\det \mathbf{A} \neq 0$ ，即 $\mathbf{A}$ 非奇异阵



面向对角占优矩阵的收敛条件

□ **定理8.7** 如果 $A \in \mathbf{R}^{n \times n}$ 为严格对角占优阵或为不可约弱对角占优阵，则对于任意的 $x^{(0)}$ ，解方程组 $Ax = b$ 的Jacobi迭代法，Gauss-Seidel迭代法均收敛

■ 设 A 为不可约弱对角占优阵，现证明Gauss-Seidel迭代法收敛

■ 对于方程组 $Ax = b$ ，Gauss-Seidel迭代法为

$$x^{(k+1)} = Gx^{(k)} + f$$

■ 迭代矩阵为

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



定理8.7证明

$$\mathbf{G} = (\mathbf{D} - \mathbf{L})^{-1} \mathbf{U}$$

- 根据定理8.3，如果 $\mathbf{G}$ 的特征值小于1，则Gauss-Seidel迭代法收敛
- 令 λ 为 $\mathbf{G}$ 的特征值，则

$$\begin{aligned}\det(\lambda \mathbf{I} - \mathbf{G}) &= \det(\lambda \mathbf{I} - (\mathbf{D} - \mathbf{L})^{-1} \mathbf{U}) \\ &= \det((\mathbf{D} - \mathbf{L})^{-1} (\lambda (\mathbf{D} - \mathbf{L}) - \mathbf{U})) \\ &= \det((\mathbf{D} - \mathbf{L})^{-1}) \cdot \det(\lambda (\mathbf{D} - \mathbf{L}) - \mathbf{U}) = 0\end{aligned}$$

- 可得

$$\det(\lambda (\mathbf{D} - \mathbf{L}) - \mathbf{U}) = 0$$

- 接下来，将证明当 $|\lambda| \geq 1$ 时 $\det(\lambda (\mathbf{D} - \mathbf{L}) - \mathbf{U}) \neq 0$



定理8.7证明（续）

■ 根据***D***、***L***和***U***的定义

$$\mathbf{C} = \lambda(\mathbf{D} - \mathbf{L}) - \mathbf{U}$$

$$= \begin{bmatrix} \lambda a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \lambda a_{31} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{n-1,n} \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{n,n-1} & \lambda a_{nn} \end{bmatrix} = (c_{ij})_{n \times n}$$

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

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

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



定理8.7证明（续）

- 由 $\mathbf{A}$ 为不可约矩阵，则 $\mathbf{C}$ 亦为不可约矩阵
- 由 $\mathbf{A}$ 为弱对角占优阵得到（当 $|\lambda| \geq 1$ 时）

$$\begin{aligned} |c_{ii}| = |\lambda| |a_{ii}| &\geq \sum_{j=1}^{i-1} \lambda |a_{ij}| + \sum_{j=i+1}^n \lambda |a_{ij}| \\ &\geq \sum_{j=1}^{i-1} \lambda |a_{ij}| + \sum_{j=i+1}^n |a_{ij}| = \sum_{j \neq i} |c_{ij}| \end{aligned}$$

且至少有一不等式严格成立

- 因此，当 $|\lambda| \geq 1$ 时， $\mathbf{C}$ 为不可约弱对角占优阵
- 根据定理8.6，此时 $\mathbf{C}$ 是非奇异矩阵，即 $\det \mathbf{C} \neq 0$



目录

- 引言
- Jacobi迭代法
- Gauss-Seidel迭代法
- 迭代法的收敛性
- 解线性方程组的超松弛迭代法



解线性方程组的超松弛迭代法

□ 逐次超松弛迭代法（Successive Over Relaxation method, 简称SOR方法）

- 是Gauss-Seidel方法的一种加速方法
- 是解大型稀疏矩阵方程组的有效方法之一
- 具有计算公式简单，程序设计容易，占用计算机内存较少等优点
- 但需要选择好的加速因子（即最佳松弛因子）

□ 设有方程组

$$Ax = b \quad (8.4.1)$$

- $A \in R^{n \times n}$ 为非奇异矩阵，且设 $a_{ii} \neq 0$ ($i = 1, 2, \dots, n$)



逐次超松弛迭代法

□ Gauss-Seidel方法

$$\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T \quad (\text{初始向量})$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) \\ (k = 0, 1, 2, \dots; i = 1, 2, \dots, n), \quad (8.2.5)$$

□ 逐次超松弛迭代法

- 设已知第 k 次迭代向量 $\mathbf{x}^{(k)}$ ，及第 $k+1$ 次迭代向量 $\mathbf{x}^{(k+1)}$ 的分量 $x_j^{(k+1)}$ ($j = 1, 2, \dots, i-1$)，要求计算分量 $x_i^{(k+1)}$



逐次超松弛迭代法（续）

- 首先用Gauss-Seidel迭代法定义辅助量

$$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right)$$
$$(i = 1, 2, \dots, n) \quad (8.4.3)$$

- 再把 $x_i^{(k+1)}$ 取为 $x_i^{(k)}$ 与 $\tilde{x}_i^{(k+1)}$ 某个平均值（即加权平均）

$$\begin{aligned} x_i^{(k+1)} &= (1 - \omega) x_i^{(k)} + \omega \tilde{x}_i^{(k+1)} \\ &= x_i^{(k)} + \omega \left(\tilde{x}_i^{(k+1)} - x_i^{(k)} \right) \end{aligned} \quad (8.4.4)$$

- ✓ ω 为松弛因子， $\omega = 1$ 为Gauss-Seidel迭代法



逐次超松弛迭代法（续）

- 用式(8.4.3)中的 $\tilde{x}_i^{(k+1)}$ 代入式(8.4.4)即得到解方程组 $\mathbf{Ax} = \mathbf{b}$ 的逐次超松弛迭代公式

$$\begin{cases} x_i^{(k+1)} = x_i^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right) \\ x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})^T \quad (k = 0, 1, \dots; i = 1, 2, \dots, n) \end{cases}$$

- 可改写为

$$\begin{cases} x_i^{(k+1)} = x_i^{(k)} + \Delta x_i \quad (k = 0, 1, \dots; i = 1, 2, \dots, n) \\ \Delta x_i = \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right) \end{cases}$$



逐次超松弛迭代法（续）

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

- 在**SOR**方法中，迭代一次主要的运算量是计算一次矩阵与向量的乘法
- 在计算机上应用**SOR**方法解方程组时只需一组工作单元，以便存放近似解
- 在用计算机计算时，可用 $|p_0| = \max_i |\Delta x_i| = \max_{1 \leq i \leq n} |x_i^{(k+1)} - x_i^{(k)}| < \varepsilon$ 控制迭代终止
- 当 $\omega < 1$ 时，称式(8.4.5)为低松弛法；当 $\omega > 1$ 时，称式(8.4.5)为超松弛法



例8.11

□ 用SOR方法解方程组

$$\begin{bmatrix} -4 & 1 & 1 & 1 \\ 1 & -4 & 1 & 1 \\ 1 & 1 & -4 & 1 \\ 1 & 1 & 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

它的精确解为 $x^* = (-1, -1, -1, -1)^T$

■ 取 $x^{(0)} = \mathbf{0}$, 迭代公式为

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} - \omega (1 + 4x_1^{(k)} - x_2^{(k)} - x_3^{(k)} - x_4^{(k)}) / 4 \\ x_2^{(k+1)} = x_2^{(k)} - \omega (1 - x_1^{(k+1)} + 4x_2^{(k)} - x_3^{(k)} - x_4^{(k)}) / 4 \\ x_3^{(k+1)} = x_3^{(k)} - \omega (1 - x_1^{(k+1)} - x_2^{(k+1)} + 4x_3^{(k)} - x_4^{(k)}) / 4 \\ x_4^{(k+1)} = x_4^{(k)} - \omega (1 - x_1^{(k+1)} - x_2^{(k+1)} - x_3^{(k+1)} + 4x_4^{(k)}) / 4 \end{cases}$$



例8.11（续）

■ 取 $\omega = 1.3$ ，第11次迭代结果为

$$x^{(11)} = (-0.99999646, -1.00000310, -0.99999953, -0.99999912)^T$$

$$\|\epsilon^{(11)}\|_2 \leq 0.46 \times 10^{-5}$$

■ 对 ω 取其他值，迭代次数如右表所示，从此例可以看到，松弛因子选择得好，会使SOR方法的收敛大大加速

■ 本例中， $\omega = 1.3$ 是最佳松弛因子

ω	误差 $< 10^{-5}$ 的迭代次数
1.0	22
1.1	17
1.2	12
<u>1.3</u>	11（最少迭代次数）
1.4	14
1.5	17
1.6	23
1.7	33
1.8	53
1.9	109



SOR迭代公式的矩阵形式

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

□ 上述SOR迭代公式亦可写为

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

$(i = 1, 2, \dots, n) \quad (8.4.7)$

□ 用***A***的分解***A = D - L - U***，得

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



SOR迭代公式的矩阵形式（续）

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

■ 等价于

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

■ 显然对于任何一个 ω 值， $\mathbf{D} - \omega\mathbf{L}$ 非奇异

✓ 由假设 $a_{ii} \neq 0; i = 1, 2, \dots, n$ ，可知 $\det(\mathbf{D} - \omega\mathbf{L}) \neq 0$

■ 因此

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

□ 换言之，SOR方法的迭代公式为

$$\mathbf{x}^{(k+1)} = \mathbf{L}_\omega \mathbf{x}^{(k)} + \mathbf{f} \quad (8.4.8)$$

$$\mathbf{L}_\omega = (\mathbf{D} - \omega\mathbf{L})^{-1}[(1 - \omega)\mathbf{D} + \omega\mathbf{U}], \quad \mathbf{f} = \omega(\mathbf{D} - \omega\mathbf{L})^{-1}\mathbf{b}$$



SOR方法的收敛性

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

- L_{ω} 称为SOR方法的迭代矩阵，这说明SOR方法相当于对方程组 $\mathbf{x} = L_{\omega}\mathbf{x} + \mathbf{f}$ 应用一般迭代法，于是关于一般迭代法的理论可应用到式(8.4.8)，到下述定理
- 定理8.8 设有线性方程组 $A\mathbf{x} = \mathbf{b}$ ，且 $a_{ii} \neq 0$ ($i = 1, 2, \dots, n$)，SOR方法收敛的充要条件是
$$\rho(L_{\omega}) < 1$$



SOR方法的收敛性（续）

- 容易验证， $Ax = b$ 的最优解 x^* 满足

$$x^* = L_{\omega} x^* + f$$

- 因此，SOR方法一定收敛到 x^*

- 引进超松弛迭代法的想法是希望能选择松弛因子 ω 使得SOR方法收敛较快，也就是应选择因子 ω 使 $\rho(L_{\omega}) = \min_{\omega} \rho(L_{\omega})$

- 下面研究，对于一般方程组 $Ax = b$ （ $a_{ii} \neq 0; i = 1, 2, \dots, n$ ），松弛因子 ω 在什么范围内取值，SOR方法才可能收敛



SOR方法收敛的必要条件

□ 定理**8.9** 设解式 $Ax = b$ ($a_{ii} \neq 0; i = 1, 2, \dots, n$) 的SOR方法收敛, 则

$$0 < \omega < 2$$

■ 由设SOR方法收敛, 根据定理8.8

$$\rho(L_\omega) < 1$$

■ 设 L_ω 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则

$$|\det L_\omega| = |\lambda_1 \lambda_2 \cdots \lambda_n| \leq (\rho(L_\omega))^n$$

$$\Rightarrow |\det L_\omega|^{1/n} \leq \rho(L_\omega) < 1$$



定理8.9的证明

■ 根据 $\mathbf{L}_\omega = (\mathbf{D} - \omega \mathbf{L})^{-1}[(1 - \omega)\mathbf{D} + \omega \mathbf{U}]$

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

■ 可得

$$\det \mathbf{L}_\omega = \det((\mathbf{D} - \omega \mathbf{L})^{-1}) \cdot \det((1 - \omega)\mathbf{D} + \omega \mathbf{U})$$

$$= \frac{1}{a_{11}a_{22} \cdots a_{nn}} \cdot (1 - \omega)^n \cdot a_{11}a_{22} \cdots a_{nn}$$

$$= (1 - \omega)^n$$

■ 结合 $|\det \mathbf{L}_\omega|^{1/n} < 1$, 可得

$$|1 - \omega| < 1 \Leftrightarrow 0 < \omega < 2$$



SOR方法收敛的充分条件

- 定理8.9说明对于解一般方程组 $\mathbf{Ax} = \mathbf{b}$ ($a_{ii} \neq 0; i = 1, 2, \dots, n$)，SOR方法只有取松弛因子 ω 在 $(0, 2)$ 范围内才能收敛
- 而当 $\mathbf{A}$ 是对称正定矩阵时，若 ω 满足 $0 < \omega < 2$ ，则SOR方法一定收敛
- 定理8.10 如果 $\mathbf{A}$ 为对称正定矩阵，且 $0 < \omega < 2$ ，则解式 $\mathbf{Ax} = \mathbf{b}$ 的SOR方法收敛



最佳松弛因子理论

- 最佳松弛因子理论是由Young(1950年)针对一类椭圆型微分方程数值解得到的代数方程组 $\mathbf{Ax} = \mathbf{b}$ 所建立的理论, 给出了最佳松弛因子公式

$$\omega_{\text{opt}} = \frac{2}{1 + \sqrt{1 - \rho^2(\mathbf{B}_0)}}$$

- $\rho(\mathbf{B}_0)$ 是Jacobi方法迭代矩阵 $\mathbf{B}_0$ 的谱半径
- 一般来说, 在实际应用中计算 $\rho(\mathbf{B}_0)$ 较困难, 对某些微分方程数值解问题可考虑用第9章求近似值的方法
- 亦可由计算实践摸索出 (近似) 最佳松弛因子



SOR算法流程

□ 用SOR方法解式 $A\mathbf{x} = \mathbf{b}$ ，其中 A 为对称正定矩阵

■ 数组 $\mathbf{x}$ 为一组工作单元，开始存放初始向量，然后存放近似值解 $\mathbf{x}^{(k)}$ ，最后存放结果

■ 用

$$|p_0| = \max_{1 \leq i \leq n} |\Delta x_i| = \max_{1 \leq i \leq n} |x_i^{(k+1)} - x_i^{(k)}| < \varepsilon \quad (\text{精度要求})$$

控制控制迭代终止， k 表示迭代次数

1. $k \leftarrow 0$
2. $x_i \leftarrow 0 \ (i = 1, 2, \dots, n)$
3. $k \leftarrow k + 1$



SOR算法流程（续）

4. $p_0 \leftarrow 0$

5. 对于 $i = 1, 2, \dots, n$, 有

$$(1) \ p \leftarrow \Delta x_i = \omega \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^n a_{ij}x_j \right) / a_{ii}$$

(2) 如果 $|p| > |p_0|$, 则 $p_0 \leftarrow p$

(3) $x_i \leftarrow x_i + p$

6. 输出 p_0

7. 如果 $|p_0| > \varepsilon$, 则转步3

8. 输出结果 $\mathbf{x}, k$



总结

□ 引言

- 定义、迭代法收敛、发散

□ Jacobi迭代法

- Jacobi迭代公式、Jacobi方法迭代矩阵

□ Gauss-Seidel迭代法

- Gauss-Seidel迭代法
- Gauss-Seidel迭代法的迭代矩阵



总结（续）

□ 迭代法的收敛性

- 迭代法基本定理
- 迭代法收敛的充分条件
- 对角占优定理、面向对角占优矩阵的收敛条件

□ 解线性方程组的超松弛迭代法

- 逐次超松弛迭代法
- SOR方法收敛的充要条件
- SOR方法收敛的必要条件
- SOR方法收敛的充分条件