第 8 章 解线性方程组的迭代法

姚遥

南京大学智能科学与技术学院

本章目录

- 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 | 引言

考虑线性方程组，其中A为非奇异矩阵

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{8.1.1}$$

- 当A为低阶稠密矩阵时，第 7 章所讨论的选主元素消去法是一种有效方法
- 但是，对于由工程技术中产生的大型稀疏矩阵方程组（$\boldsymbol{A}$ 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 $\boldsymbol{A}$ 中有大量零元素的特点

如何操作？

- 改写成 $\boldsymbol{x} = f(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{x})$ 形式

例8.1 求解方程组

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

- 记作 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$，其中

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

- 方程组的精确解是

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

迭代法 | 引言：迭代法例子

- 将方程组 (8.1.2) 改写成

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

- 记为 $\boldsymbol{x} = \boldsymbol{B_0}\boldsymbol{x} + \boldsymbol{f}$，其中

$$
\boldsymbol{B_0} = \begin{bmatrix} 0 & \dfrac{3}{8} & -\dfrac{2}{8} \\ -\dfrac{4}{11} & 0 & \dfrac{1}{11} \\ -\dfrac{6}{12} & -\dfrac{3}{12} & 0 \end{bmatrix}
\qquad
\boldsymbol{f} = \begin{bmatrix} \dfrac{20}{8} \\ \dfrac{33}{11} \\ \dfrac{36}{12} \end{bmatrix}
$$

迭代法 | 引言：迭代法例子

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

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

- 再将 $\boldsymbol{x}^{(1)}$ 分量代入式 (8.1.3) 右端，得到 $\boldsymbol{x}^{(2)}$，反复计算，得到一向量序列

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

- 迭代公式如下

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

迭代法 | 引言：迭代法例子

- 迭代过程可写为，其中 k $(k = 0, 1, 2, \dots)$ 表示迭代次数

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{B}_0 \boldsymbol{x}^{(k)} + \boldsymbol{f}$$

- 迭代到第10次有

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

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

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

- $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f}$ 是由 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 变形得到的等价方程组
- 不一定收敛
- 一个例子：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 | 引言：迭代法收敛性

定义8.1

- 对于给定的方程组 $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f}$，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 $\boldsymbol{B}$ 与 k 无关）

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

- 如果 $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)}$ 存在（记作 $\boldsymbol{x}^*$），称此迭代法收敛，显然 $\boldsymbol{x}^*$ 就是方程组的解，否则称此迭代法发散

为了研究 $\{x^{(k)}\}$ 的收敛性，引进误差向量

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

- 由式(8.1.6)减去(8.1.5)得

$$\left.\begin{aligned} \boldsymbol{x}^* &= \boldsymbol{B}\boldsymbol{x}^* + \boldsymbol{f} \\ \boldsymbol{x}^{(k+1)} &= \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f} \end{aligned}\right\} \Rightarrow \boldsymbol{\varepsilon}^{(k+1)} = \boldsymbol{B}\boldsymbol{\varepsilon}^{(k)} \quad (k = 0, 1, 2 \cdots)$$

- 递推得

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

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

本章目录

- 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 | Jacobi迭代法

设有方程组 $\boldsymbol{Ax = b}$

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

- $\boldsymbol{A}$ 为非奇异矩阵且 $a_{ii} \neq 0\ (i = 1, 2, \dots, n)$
- 将 $\boldsymbol{A}$ 分解为 $\boldsymbol{A} = \boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U}$, 其中

$$\boldsymbol{D} = \begin{bmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & a_{nn} \end{bmatrix} \quad \boldsymbol{L} = -\begin{bmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \quad \boldsymbol{U} = -\begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 0 & a_{23} & \cdots & a_{2n} \\ & & 0 & \ddots & \vdots \\ & & & \ddots & a_{n-1,n} \\ & & & & 0 \end{bmatrix}$$

迭代法 | Jacobi迭代法

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

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

- 简记作

$$\boldsymbol{x} = \boldsymbol{B}_0\boldsymbol{x} + \boldsymbol{f}$$

其中

$$\begin{aligned} \boldsymbol{B}_0 &= \boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{A} \\ &= \boldsymbol{D}^{-1}(\boldsymbol{D} - \boldsymbol{A}) \\ &= \boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U}) \end{aligned}$$

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

迭代法 | Jacobi迭代法

- 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

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

其中 $\boldsymbol{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^{\mathrm{T}}$ 为第 k 次迭代向量

- 迭代公式(8.2.3)的矩阵形式为

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

其中 $\boldsymbol{B}_0$ 称为Jacobi方法迭代矩阵

迭代法 | Jacobi迭代法：讨论

Jacobi迭代法公式

- 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 $\boldsymbol{x}^{(k)}$ 及 $\boldsymbol{x}^{(k+1)}$

本章目录

迭代法 | Gauss–Seidel迭代法

由Jacobi方法迭代公式 (8.2.4) **可知**

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

- 迭代的每一步计算过程，都是用 $\boldsymbol{x}^{(k)}$ 的全部分量来计算过 $\boldsymbol{x}^{(k+1)}$ 的所有分量
- 在计算第 i 个分量工的时，已经计算出的最新分量 $x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}$ 没有被利用
- 可以认为，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第 $k+1$ 次近似 $\boldsymbol{x}^{(k+1)}$ 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）

迭代法 | Gauss–Seidel迭代法

Gauss-Seidel迭代法（利用 $x_j^{(k+1)}$ ）

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

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

$$(k = 0, 1, 2, \dots; i = 1,2, \dots, n), \qquad (8.2.5)$$

Jacobi迭代公式（只利用 x_j^k ）

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

Gauss-Seidel迭代法（利用 $x_j^{(k+1)}$ ）

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

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

$$(k = 0, 1, 2, \dots; i = 1,2, \dots, n), \qquad (8.2.5)$$

- 可写成

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

迭代法｜Gauss–Seidel迭代法

Gauss-Seidel迭代矩阵形式

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

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

作为对比，Jacobi迭代公式

$$\left.\begin{aligned} \boldsymbol{x}^{(k+1)} &= \boldsymbol{B}_0\boldsymbol{x}^{(k)} + \boldsymbol{f} \\ \boldsymbol{B}_0 &= \boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U}) \\ \boldsymbol{f} &= \boldsymbol{D}^{-1}\boldsymbol{b} \end{aligned}\right\} \Rightarrow \boldsymbol{D}\boldsymbol{x}^{(k+1)} = (\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x}^{(k)} + \boldsymbol{b}$$

迭代法 | Gauss–Seidel迭代法

Gauss-Seidel迭代矩阵形式

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

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

$$\Rightarrow \boldsymbol{x}^{(k+1)} = (\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)} + (\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{b} \qquad (\text{设}(\boldsymbol{D}-\boldsymbol{L})^{-1}\text{存在})$$

- 可改写为

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{G}\boldsymbol{x}^{(k)} + \boldsymbol{f} \qquad (8.2.6)$$

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

作为对比，Jacobi迭代公式

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{B}_0\boldsymbol{x}^{(k)} + \boldsymbol{f}$$

$$\boldsymbol{B}_0 = \boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U}), \qquad \boldsymbol{f} = \boldsymbol{D}^{-1}\boldsymbol{b}$$

迭代法 | Gauss–Seidel迭代法

应用Gauss-Seidel迭代法解式(8.2.1)就是对方程组 $\boldsymbol{x}=\boldsymbol{G}\boldsymbol{x}+\boldsymbol{f}$ 应用迭代法

- $\boldsymbol{G}$ 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

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

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

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

- 迭代到第 5 次时，有

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

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

迭代法 | Gauss–Seidel迭代法

- 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

本章目录

迭代法 | 矩阵收敛

定义8.2 设有矩阵序列 $A_k=(a_{ij}^{(k)})_{n\times n}(k=1,2,\cdots)$ 及 $A=(a_{ij})_{n\times n}$，若

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

成立，则称 $\{\boldsymbol{A}_k\}$ 收敛于 $\boldsymbol{A}$，记作 $\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}$

例8.4 矩阵序列

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

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

定理8.1 $\lim_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}$ 的充要条件是 $\|\boldsymbol{A}_k - \boldsymbol{A}\| \to 0\ (k \to \infty)$

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

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

迭代法 | 矩阵收敛

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

- 由矩阵 $\boldsymbol{B}$ 的Jordan标准形，存在非奇异矩阵 $\boldsymbol{P}$，使

$$\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}=\begin{bmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_r\end{bmatrix}\equiv\boldsymbol{J} \qquad \boldsymbol{J}_i=\begin{bmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i\end{bmatrix}_{n_i\times n_i}$$

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

- 显然 $\boldsymbol{B}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$，故

$$\boldsymbol{B}^k=(\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1})^{\boldsymbol{k}}=\boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}$$

迭代法 | 矩阵收敛

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

$$\boldsymbol{B}^k=(\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1})^k=\boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}$$

其中

$$\boldsymbol{J}^k=\begin{bmatrix}\boldsymbol{J}_1^k & & & \\ & \boldsymbol{J}_2^k & & \\ & & \ddots & \\ & & & \boldsymbol{J}_r^k\end{bmatrix} \qquad (k=1,2,\ldots)$$

- 于是

$$\boldsymbol{B}^k \to \boldsymbol{O} \quad (k\to\infty)$$

$$\Leftrightarrow \boldsymbol{J}^k \to \boldsymbol{O} \quad (k\to\infty)$$

$$\Leftrightarrow \boldsymbol{J}_i^k \to \boldsymbol{O} \quad (k\to\infty) \quad (i=1,2,\cdots,r)$$

迭代法 | 矩阵收敛

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

- 引进记号

$$\boldsymbol{E}_{tk} = \begin{array}{c} \overbrace{\qquad\qquad\qquad}^{k\ 个} \\ \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & \ddots & \ddots & 1 \\ & & & & \ddots & \ddots & 0 \\ & & & & & \ddots & \vdots \\ & & & & & & 0 \end{bmatrix}_{t\times t} \end{array} \quad (其中\ t = n_i)$$

可以证明

- 当 $k < t$ 时，$(\boldsymbol{E}_{t1})^k = \boldsymbol{E}_{tk}$
- 当 $k \geq t$ 时 $(\boldsymbol{E}_{t1})^k = \boldsymbol{E}_{tk} = \boldsymbol{O}$

迭代法 | 矩阵收敛

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

- 可得

$$\boldsymbol{J}_i=\begin{bmatrix}\lambda_i & & & & \\ & \lambda_i & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_i\end{bmatrix}+\begin{bmatrix}0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0\end{bmatrix}=\lambda_i\boldsymbol{I}+\boldsymbol{E}_{t1}$$

进一步可得

$$\boldsymbol{J}_i^k=(\lambda_i\boldsymbol{I}+\boldsymbol{E}_{t1})^k=\sum_{j=0}^{k}\binom{k}{j}\lambda_i^{k-j}(\boldsymbol{E}_{t1})^j=\sum_{j=0}^{k}\binom{k}{j}\lambda_i^{k-j}\boldsymbol{E}_{tj}$$

$$（\text{当 } k\geq t \text{ 时 } \boldsymbol{E}_{tk}=\boldsymbol{O}）\quad=\sum_{j=0}^{t-1}\binom{k}{j}\lambda_i^{k-j}\boldsymbol{E}_{tj}$$

迭代法 | 矩阵收敛

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

- 进而

$$\boldsymbol{J}_i^k=(\lambda_i\boldsymbol{I}+\boldsymbol{E}_{t1})^k=\sum_{j=0}^{t-1}\binom{k}{j}\lambda_i^{k-j}\boldsymbol{E}_{tj}=\begin{bmatrix}\lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \cdots & \cdots & \binom{k}{t-1}\lambda_i^{k-(t-1)}\\ & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \cdots & \binom{k}{t-2}\lambda_i^{k-(t-2)}\\ & & \ddots & \ddots & \vdots\\ & & & \ddots & \binom{k}{1}\lambda_i^{k-1}\\ & & & & \lambda_i^k\end{bmatrix}$$

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

迭代法 | 矩阵收敛

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

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

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

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

定理8.3 设有方程组

$$\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f} \tag{8.3.1}$$

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知 $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{B}$ 为非奇异矩阵，故 $\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{f}$ 有唯一解，记作

$$\boldsymbol{x}^* = \boldsymbol{B}\boldsymbol{x}^* + \boldsymbol{f} \tag{8.3.2}$$

迭代法 | 迭代法基本定理

- 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 $\boldsymbol{B}^k \to \boldsymbol{O}\ (k \to \infty)$.
- 于是对于任意 $\boldsymbol{x}^{(0)}$ 及 $\boldsymbol{f}$，有 $\boldsymbol{\varepsilon}^{(k)} \to \boldsymbol{0}\ (k \to \infty)$，即 $\boldsymbol{x}^{(k)} \to \boldsymbol{x}^* (k \to \infty)$

必要性

- 设对于任意 $\boldsymbol{x}^{(0)}$ 及任意 $\boldsymbol{f}$，皆有 $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$，其中 $\boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f}$

迭代法 | 迭代法基本定理

- 显然
  - 极限 $\boldsymbol{x}^*$ 是方程组 $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f}$ 的解
  - 对于任意 $\boldsymbol{x}^{(0)}$ 及任意 $\boldsymbol{f}$，有

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

- 由第2点可知， $\boldsymbol{B}^k \to \boldsymbol{O}\ (k \to \infty)$
- 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

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

- Jacobi方法

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

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

迭代法 | 迭代法收敛

- 迭代矩阵 $\boldsymbol{B}_0$ 的特征方程为

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

- 解得

$$\lambda_1=-0.3082$$

$$\lambda_2=0.1541+\mathrm{i}0.3245$$

$$\lambda_3=0.1541-\mathrm{i}0.3245,$$

- 可知

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

即 $\rho(\boldsymbol{B}_0)<1$，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

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

其中

$$\boldsymbol{B} = \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix}, \qquad \boldsymbol{f} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

- 特征方程为

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

- 特征根$\lambda_{1,2} = \pm\sqrt{6}$，即$\rho(\boldsymbol{B}) > 1$
- 说明此迭代过程不收敛

迭代法 | 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 $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$，相应的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$
- 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，$\lambda_i^k \to 0\ (i = 1,2, \ldots, n; k \to 0)$ 愈快，即 $\boldsymbol{\varepsilon}^{(k)} \to \boldsymbol{0}$ 愈快，故可用量 $\rho(\boldsymbol{B})$ 来刻画迭代法的收敛快慢

迭代法 | 迭代法的收敛速度

- 可以依据给定精度要求来确定迭代次数 k，即使

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

- 取对数得

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

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

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

迭代法 | 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 $\rho(\boldsymbol{B}) < 1$
- 收敛速度由 $R(\boldsymbol{B}) = -\ln \rho(\boldsymbol{B})$ 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 $\rho(\boldsymbol{B}) \leq \|\boldsymbol{B}\|_v (v = 1,2,\infty$ 或 $F)$，所以可用 $\|\boldsymbol{B}\|_v$ 来作为 $\rho(\boldsymbol{B})$ 上界的一种估计

迭代法 | 迭代矩阵的特征值

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

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

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

- 迭代法收敛
- $\left\|\boldsymbol{x}^{*}-\boldsymbol{x}^{(k)}\right\|_v \le \frac{q}{1-q}\left\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)}\right\|_v$
- $\left\|\boldsymbol{x}^{*}-\boldsymbol{x}^{(k)}\right\|_v \le \frac{q^k}{1-q}\left\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\right\|_v$

证明

- 根据定理8.3和 $\rho(\boldsymbol{B}) \le \|\boldsymbol{B}\|_v$，迭代法收敛显然成立
- 由

$$\boldsymbol{x}^* = \boldsymbol{B}\boldsymbol{x}^* + \boldsymbol{f}$$

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

可得

$$\boldsymbol{x}^* - \boldsymbol{x}^{(k+1)} = \boldsymbol{B}\left(\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\right) \tag{8.3.4}$$

- 根据范数性质，可得

$$\left\|\boldsymbol{x}^* - \boldsymbol{x}^{(k+1)}\right\|_v = \left\|\boldsymbol{B}\left(\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\right)\right\|_v \tag{8.3.6}$$

$$\le \|\boldsymbol{B}\|_v \left\|\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\right\|_v = q\left\|\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\right\|_v \qquad (k = 1,2,\cdots)$$

证明

- 同理可得

$$\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)} = \boldsymbol{B}\left(\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right)$$

- 因此

$$\left\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\right\|_v \leq q\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right\|_v \tag{8.3.5}$$

- 另一方面

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

迭代法 | 迭代矩阵的特征值

证明

- 结合

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

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

  可证明第 2 点成立

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

- 对上式重复利用(8.3.5)，可得第 3 点成立：

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

例8.7 仍然考察例8.1，Jacobi方法

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

- 迭代矩阵 $\boldsymbol{B}_0$ 的 ∞-范数为

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

- 所以对例8.1应用Jacobi迭代方法是收敛的

迭代法丨迭代矩阵的特征值

例8.8 设有迭代过程 $\boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f}\ (k = 0,1,2,\ \dots)$，其中

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

- 可得

$$\|\boldsymbol{B}\|_\infty = 1.1, \qquad \|\boldsymbol{B}\|_1 = 1.2$$

$$\|\boldsymbol{B}\|_2 = 1.021, \qquad \|\boldsymbol{B}\|_F = \sqrt{1.54}$$

- 虽然 $\boldsymbol{B}$ 的这些范数都大于 1，但 $\boldsymbol{B}$ 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

迭代法 | 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 $\boldsymbol{B}$ 的某一种范数 $\|\boldsymbol{B}\| < 1$ 时，如果相邻两次迭代

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

ε_0 为给定的精度要求，则

$$\left\|\boldsymbol{x}^{*} - \boldsymbol{x}^{(k)}\right\| \le \frac{q}{1-q}\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right\|_v \le \frac{q}{1-q}\varepsilon_0$$

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

迭代法 | 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

- 不过要注意，当 $q \approx 1$ 时，$\frac{q}{1-q}$ 大，尽管 $\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\|$ 已非常小，但误差向量的模 $\|\boldsymbol{\varepsilon}^{(k)}\| = \|\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\|$ 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 $\|\boldsymbol{\varepsilon}^{(k)}\| < \varepsilon_0$

$$\left\|\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\right\|_v \le \frac{q^k}{1-q}\left\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\right\|_v$$

迭代法 | Gauss–Seidel迭代法的收敛性

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

- 对Gauss-Seidel迭代法应用定理8.3即可证

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{G}\boldsymbol{x}^{(k)}+\boldsymbol{f} \tag{8.2.6}$$