第四讲 线性方程组迭代解法

目录

https://math.ecnu.edu.cn/~jypan/Teaching/SC

4-2 收敛性分析

4.2 收敛性分析

https://math.ecnu.edu.cn/~jypan/Teaching/SC

4-2-1 向量序列与矩阵序列的收敛性

定义 4.2 (向量序列的收敛) 设 $\{x^{(k)}\}_{k=0}^{\infty}$ 是 $\mathbb{R}^n$ 中的一个向量序列. 如果存在向量 $x = [x_1, x_2, \ldots, x_n]^{\mathsf{T}} \in \mathbb{R}^n$ 使得

$$\boxed{\lim_{k\to\infty} x_i^{(k)} = x_i}, \quad i = 1, 2, \ldots, n,$$

其中 $x_i^{(k)}$ 表示 $x^{(k)}$ 的第 i 个分量. 则称 $\{x^{(k)}\}$ (**按分量**) **收敛** 到 x, 记为

$$\boxed{\lim_{k\to\infty} x^{(k)} = x}.$$

我们称 x 为序列 $\{x^{(k)}\}$ 的 **极限**.

矩阵序列的收敛性

定义 4.3 (矩阵序列的收敛) 设 $\left\{A^{(k)}=\left[a_{ij}^{(k)}\right]\right\}_{k=0}^{\infty}$ 是 $\mathbb{R}^{n\times n}$ 中的一个矩阵序列. 如果存在矩阵 $A=[a_{ij}]\in\mathbb{R}^{n\times n}$ 使得

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

则称 $A^{(k)}$ 收敛到 A, 记为

$$\boxed{\lim_{k\to\infty} A^{(k)}=A}.$$

我们称 A 为 $A^{(k)}$ 的 **极限**.

矩阵序列收敛举例

例 4.1 设 $0 < |a| < 1$, 考虑矩阵序列 $\{A^{(k)}\}$, 其中

$$A^{(k)} = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}^k, \quad k = 1, 2, \ldots.$$

易知当 $k \to \infty$ 时, 有

$$A^{(k)} = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}^k = \begin{bmatrix} a^k & ka^{k-1} \\ 0 & a^k \end{bmatrix} \to \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

收敛性定理

定理 4.1 设向量序列 $\{x^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$, 矩阵序列 $\left\{A^{(k)} = \left[a_{ij}^{(k)}\right]\right\}_{k=0}^{\infty} \subset \mathbb{R}^{n\times n}$, 则

(1) $\boxed{\lim\limits_{k\to\infty} x^{(k)} = x \iff \lim\limits_{k\to\infty} \|x^{(k)} - x\| = 0}$, 其中 $\|\cdot\|$ 为任一向量范数;

(2) $\boxed{\lim\limits_{k\to\infty} A^{(k)} = A \iff \lim\limits_{k\to\infty} \|A^{(k)} - A\| = 0}$, 其中 $\|\cdot\|$ 为任一矩阵范数;

该结论将向量 (矩阵) 序列的收敛性转化为数列的收敛性.

两类特殊的矩阵序列

收敛到 0 的矩阵序列

$$\lim_{k\to\infty} A^{(k)} = 0 \iff \lim_{k\to\infty} \left\| A^{(k)} \right\| = 0$$

矩阵的幂组成的序列

$$\lim_{k\to\infty} A^{k} = 0 \iff \lim_{k\to\infty} \left\| A^{k} \right\| = 0$$

更多判别方法

定理 4.2 设矩阵序列 $\left\{A^{(k)}=\left[a_{ij}^{(k)}\right]\right\}_{k=0}^{\infty}\subset\mathbb{R}^{n\times n}$, 则

$$\lim_{k\to\infty}A^{(k)}=0\Longleftrightarrow\lim_{k\to\infty}A^{(k)}x=0,\quad\forall\,x\in\mathbb{R}^n.$$

(板书)

更多判别方法

定理 4.2 设矩阵序列 $\left\{A^{(k)}=\left[a_{ij}^{(k)}\right]\right\}_{k=0}^{\infty}\subset\mathbb{R}^{n\times n}$, 则

$$\lim_{k\to\infty}A^{(k)}=0\Longleftrightarrow\lim_{k\to\infty}A^{(k)}x=0,\quad\forall\, x\in\mathbb{R}^n.$$

(板书)

证明概要. **必要性** (“$\Longrightarrow$”). 由条件 $\lim_{k\to\infty}A^{(k)}=0$ 可知, 对任意算子范数都有 $\lim_{k\to\infty}\|A^{(k)}\|=0$. 因此, 对任意 $x\in\mathbb{R}^n$ 有

$$\|A^{(k)}x\|\le\|A^{(k)}\|\cdot\|x\|\to 0\ (k\to\infty)\quad 即\quad \lim_{k\to\infty}A^{(k)}x=0.$$

充分性 (“$\Longleftarrow$”). 取 $x=e_i$, 即单位矩阵的第 i 列, 则由 $\lim_{k\to\infty}A^{(k)}e_i=0$ 可知, $A^{(k)}$ 的第 i 列的极限为 0. 令 $i=1,2,\ldots,n$, 则可得 $\lim_{k\to\infty}A^{(k)}=0$.

更多判别方法 (续)

定理 4.3 设 $A \in \mathbb{R}^{n\times n}$, 若存在矩阵范数使得 $\|A\| < 1$, 则 $\lim\limits_{k\to\infty} A^k = 0$.

(板书)

更多判别方法 (续)

定理 4.3 设 $A \in \mathbb{R}^{n\times n}$, 若存在矩阵范数使得 $\|A\| < 1$, 则 $\lim\limits_{k\to\infty} A^k = 0$.

(板书)

证明概要. 由条件 $\|A\| < 1$ 可知,

$$\|A^k\| \le \|A\|^k \to 0 \ (k \to \infty).$$

所以 $\lim\limits_{k\to\infty} A^k = 0$.

更多判别方法 (续)

定理 4.4 设 $A \in \mathbb{R}^{n\times n}$, 则 $\lim\limits_{k\to\infty} A^k = 0$ 当且仅当 $\rho(A) < 1$.

(板书)

更多判别方法 (续)

定理 4.4 设 $A \in \mathbb{R}^{n\times n}$, 则 $\lim\limits_{k\to\infty} A^k = 0$ 当且仅当 $\rho(A) < 1$.

(板书)

证明概要. **必要性** ("$\Longrightarrow$"). 反证法. 假设 $\rho(A) \geq 1$, 则 A 存在特征值 λ 和特征向量为 $x \neq 0$, 使得 $|\lambda| \geq 1$, $A^k x = \lambda^k x$. 由于 $|\lambda| \geq 1$, 当 $k \to \infty$ 时 $\lambda^k x$ 不可能收敛到 0, 矛盾. 所以结论成立, 即 $\rho(A) < 1$.

充分性 ("$\Longleftarrow$"). 令 $\varepsilon = \frac{1}{2}\big(1 - \rho(B)\big)$, 则 $\varepsilon > 0$. 因此, 由矩阵范数与谱半径之间的关系可知, 存在某个矩阵范数 $\|\cdot\|_\varepsilon$, 使得

$$\|B\|_\varepsilon \leq \rho(B) + \varepsilon = \frac{1}{2}(1 + \rho(B)) < 1.$$

所以 $\lim\limits_{k\to\infty} B^k = 0$.

更多判别方法 (续)

定理 4.4 设 $A \in \mathbb{R}^{n\times n}$, 则 $\lim\limits_{k\to\infty} A^k = 0$ 当且仅当 $\rho(A) < 1$.

(板书)

证明概要. **必要性** ("$\Longrightarrow$"). 反证法. 假设 $\rho(A) \geq 1$, 则 A 存在特征值 λ 和特征向量为 $x \neq 0$, 使得 $|\lambda| \geq 1$, $A^k x = \lambda^k x$. 由于 $|\lambda| \geq 1$, 当 $k \to \infty$ 时 $\lambda^k x$ 不可能收敛到 0, 矛盾. 所以结论成立, 即 $\rho(A) < 1$.

充分性 ("$\Longleftarrow$"). 令 $\varepsilon = \frac{1}{2}\big(1 - \rho(B)\big)$, 则 $\varepsilon > 0$. 因此, 由矩阵范数与谱半径之间的关系可知, 存在某个矩阵范数 $\|\cdot\|_\varepsilon$, 使得

$$\|B\|_\varepsilon \leq \rho(B) + \varepsilon = \frac{1}{2}(1 + \rho(B)) < 1.$$

所以 $\lim\limits_{k\to\infty} B^k = 0$.

推论 4.5 设 $A \in \mathbb{R}^{n\times n}$, 则 $\lim\limits_{k\to\infty} A^k = 0$ 当且仅当存在矩阵范数 $\|\cdot\|$ 使得 $\|A\| < 1$.

4-2-2 迭代法的收敛性

定义 4.4 对 任意初始向量 $x^{(0)}$, 设 $\{x^{(k)}\}$ 是由迭代法

$$x^{(k+1)} = Gx^{(k)} + g, \quad k = 0, 1, 2, \ldots$$

生成的向量序列, 如果 $\boxed{\lim_{k\to\infty} x^{(k)}}$ 存在, 则称迭代法 收敛, 否则就称为 发散.

收敛到真解

引理 4.6 设迭代序列 $\{x^{(k)}\}$ 收敛, 且 $\lim\limits_{k\to\infty} x^{(k)} = x_*$, 则 x_* 是原方程组的真解.

(板书)

收敛到真解

引理 4.6 设迭代序列 $\{x^{(k)}\}$ 收敛, 且 $\lim\limits_{k\to\infty} x^{(k)} = x_*$, 则 x_* 是原方程组的真解.

(板书)

证明概要. 对迭代格式 $x^{(k+1)} = Gx^{(k)} + g = M^{-1}Nx^{(k)} + M^{-1}b$ 两边取极限可得

$$x_* = \lim_{k\to\infty} x^{(k+1)} = \lim_{k\to\infty} \left(M^{-1}Nx^{(k)} + M^{-1}b\right) = M^{-1}Nx_* + M^{-1}b.$$

整理后可得 $(M-N)x_* = b$, 即 $Ax_* = b$, 结论成立.

定常迭代法基本收敛性定理

定理 4.7 对任意初始向量 $x^{(0)}$, 迭代法 $x^{(k+1)} = Gx^{(k)} + g$ 收敛的充要条件是 $\rho(G) < 1$.

(板书)

定常迭代法基本收敛性定理

定理 4.7 对任意初始向量 $x^{(0)}$, 迭代法 $x^{(k+1)} = Gx^{(k)} + g$ 收敛的充要条件是 $\rho(G) < 1$.

(板书)

证明概要. **必要性**. 对任意向量 $\tilde{x} \in \mathbb{R}^n$, 令 $x^{(0)} = \tilde{x} - x_*$, 则

$$x^{(k)} - x_* = (Gx^{(k-1)} + g) - (Gx_* + g) = G(x^{(k-1)} - x_*) = \cdots = G^k(x^{(0)} - x_*) = G^k\tilde{x}.$$

由迭代法收敛可知 $G^k\tilde{x} = x^{(k)} - x_* \to 0$ $(k \to \infty)$. 根据定理 4.2, 有 $\lim\limits_{k\to\infty} G^k = 0$. 所以 $\rho(G) < 1$.

充分性. 由条件 $\rho(G) < 1$ 可得 $\lim\limits_{k\to\infty} G^k = 0$. 所以对任意 $x^{(0)} \in \mathbb{R}^n$, 当 $k \to \infty$ 时, 有

$$x^{(k)} - x_* = G^k(x^{(0)} - x_*) \to 0.$$

故 $\lim\limits_{k\to\infty} x^{(k)} = x_*$, 即迭代法收敛.

充分条件

定理 4.8 若存在矩阵范数使得 $\|G\| < 1$, 则迭代法 $x^{(k+1)} = Gx^{(k)} + g$ 收敛

注记

- 由于计算 $\rho(G)$ 通常比较复杂, 而 $\|G\|_1$, $\|G\|_\infty$ 相对比较容易计算, 因此在判别迭代方法收敛性时, 可以先验算一下迭代矩阵的 1-范数或 ∞-范数是否小于 1.
- 该定理中的条件是充分条件, 但不是必要条件, 因此判断一个迭代方法不收敛仍然需要使用基本收敛定理.

例 4.2 讨论迭代法 $x^{(k+1)} = Gx^{(k)} + g$ 的收敛性, 其中 $G = \begin{bmatrix} 0.9 & 0 \\ 0.3 & 0.8 \end{bmatrix}$.

解 由于 G 是下三角矩阵, 因此其特征值分别为 $\lambda_1 = 0.9, \lambda_2 = 0.8$.
所以 $\rho(G) = 0.9 < 1$. 故该迭代方收敛.

误差估计

定理 4.9 若存在算子范数使得 $q \triangleq \|G\| < 1$, 则

(1) $\|x^{(k)} - x_*\| \leq q^k \|x^{(0)} - x_*\|$;

(2) $\|x^{(k)} - x_*\| \leq \dfrac{q}{1-q} \|x^{(k)} - x^{(k-1)}\|$;

(3) $\|x^{(k)} - x_*\| \leq \dfrac{q^k}{1-q} \|x^{(1)} - x^{(0)}\|$.

(板书)

(1) $\|x^{(k)} - x_*\| \le q^k \|x^{(0)} - x_*\|$

证明概要. 由 $x^{(k+1)} = Bx^{(k)} + g$ 和 $x_* = Bx_* + g$ 可得

$$x^{(k+1)} - x^{(k)} = B\left(x^{(k)} - x^{(k-1)}\right) \quad \text{和} \quad x^{(k+1)} - x_* = B\left(x^{(k)} - x_*\right).$$

因此有

$$\|x^{(k+1)} - x^{(k)}\| \le q\|x^{(k)} - x^{(k-1)}\|, \tag{4.1}$$

和

$$\|x^{(k+1)} - x_*\| \le q\|x^{(k)} - x_*\|. \tag{4.2}$$

反复利用结论 (4.2) 即可得 $\|x^{(k)} - x_*\| \le q^k \|x^{(0)} - x_*\|$.

---

(2) $\|x^{(k)} - x_*\| \leq \dfrac{q}{1-q}\|x^{(k)} - x^{(k-1)}\|$

证明概要. 由 (4.2) 可得

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

结合 (4.1) 可得

$$\|x^{(k)} - x_*\| \leq \frac{1}{1-q}\|x^{(k+1)} - x^{(k)}\| \leq \frac{q}{1-q}\|x^{(k)} - x^{(k-1)}\|.$$

(3) $\quad \|x^{(k)} - x_*\| \le \dfrac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|$

证明概要. 反复利用 (4.1) 即可得

$$\|x^{(k+1)} - x^{(k)}\| \le q^k\|x^{(1)} - x^{(0)}\|.$$

所以

$$\|x^{(k)} - x_*\| \le \frac{1}{1-q}\|x^{(k+1)} - x^{(k)}\| \le \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|.$$

两点注记

✍ 定理 4.9 中的结论 (2) 和 (3) 都可以用来估计近似解 $x^{(k)}$ 的误差. 通常结论 (2) 会更准确一些. 因此, 实际应用中有时也以相邻两次迭代解之间的相对误差作为算法终止的条件, 即

$$\frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|} < \mathrm{tol},$$

其中 tol 是一个事前给定的精度要求, 比如 10^{-6} 或 10^{-8} 等.

✍ 定理 4.9 的条件中要求是算子范数, 这是为了有相应的向量范数. 事实上也可以改为任意矩阵范数, 因为任意矩阵范数都存在相容的向量范数.

4-3 经典迭代法的收敛性

4.3 经典迭代法的收敛性

4-3 经典迭代法的收敛性

4.3 经典迭代法的收敛性

https://math.ecnu.edu.cn/~jypan/Teaching/SC

基本准则

充要条件: $\rho(G) < 1$, **充分条件**: $\|G\| < 1$

4-3-1 不可约与对角占优

定义 4.5 (不可约) 设 $A \in \mathbb{R}^{n\times n}$, 若存在置换矩阵 P, 使得 PAP^{T} 为块上三角矩阵, 即

$$PAP^{\mathsf{T}} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

其中 $A_{11} \in \mathbb{R}^{k\times k}$ $(1 \le k < n)$, 则称 A 为**可约矩阵**, 否则称为 **不可约矩阵**.

不可约矩阵

可约矩阵

- 考虑线性方程组 $Ax=b$, 若 A 可约, 即存在置换矩阵 P, 使得 $PAP^{\mathsf{T}}=\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$, 因此 $Ax=b$ 等价于 $(PAP^{\mathsf{T}})(Px)=Pb$. 记 $y \triangleq Px$, $f \triangleq Pb$, 则

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

  所以原方程组就转化为两个更小规模的子方程组.
- 对于特征值问题, A 的特征值就是 A_{11} 和 A_{22} 的特征值的并.
- 如果 A_{22} 或 A_{11} 仍然是可约的, 则可以转化为更小规模的子问题.

在研究线性方程组或计算矩阵特征值时, 可以假定 A 不可约.

对角占优矩阵

定义 4.6 (对角占优) 设 $A \in \mathbb{R}^{n\times n}$, 若

$$|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|$$

对所有 $i = 1, 2, \ldots, n$ 都成立, 且至少有一个不等式严格成立, 则称 A 为**弱行对角占优**, 简称**对角占优**. 若对所有 $i = 1, 2, \ldots, n$ 不等式都严格成立, 则称 A 是**严格行对角占优**, 简称**严格对角占优**.

类似地, 可以定义**列对角占优** 和 **严格列对角占优**.

对角占优矩阵

定义 4.6 (对角占优) 设 $A \in \mathbb{R}^{n\times n}$, 若

$$|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|$$

对所有 $i = 1, 2, \dots, n$ 都成立, 且至少有一个不等式严格成立, 则称 A 为**弱行对角占优**, 简称**对角占优**. 若对所有 $i = 1, 2, \dots, n$ 不等式都严格成立, 则称 A 是**严格行对角占优**, 简称**严格对角占优**.

✍ 类似地, 可以定义**列对角占优** 和 **严格列对角占优**.

定理 4.10 设 $A \in \mathbb{R}^{n\times n}$ 严格对角占优 (或不可约对角占优), 则 A 非奇异.

不可约对角占优: Jacobi, G-S, SOR

定理 4.11 设 $A \in \mathbb{R}^{n\times n}$, 若 A 严格对角占优 (或不可约对角占优), 则

- Jacobi 迭代法收敛;
- G-S 迭代法收敛;
- 当 $0 < \omega \le 1$ 时, SOR 迭代法收敛.

✍ 事实上, SOR 迭代法收敛的一个必要条件是 $0 < \omega < 2$.

4-3-2 对称正定情形

定理 4.12 设 $A \in \mathbb{R}^{n\times n}$ 对称且对角线元素都为正, 则

- Jacobi 迭代法收敛的充要条件是 A 和 $2D - A$ 都正定;
- G-S 迭代法收敛的充要条件是 A 正定;
- SOR 迭代法收敛的充要条件是 A 正定且 $0 < \omega < 2$.

例 4.3 考虑线性方程组 $Ax=b$, 其中 $A=\begin{bmatrix}1 & a & a\\ a & 1 & a\\ a & a & 1\end{bmatrix}$. 试给出 Jacobi, G-S 和 SOR 收敛的充要条件.

解 易知 A 对称且对角线部分都为正, 因此 Jacobi 收敛当且仅当 A 和 $2D-A$ 都正定. 通过计算 A 的各阶顺序主子式可知, A 正定的充要条件是

$$\det\left(\begin{bmatrix}1 & a\\ a & 1\end{bmatrix}\right)=1-a^2>0,\quad \det(A)=1+2a^3-3a^2=(a-1)^2(2a+1)>0,$$

即 $-\frac{1}{2}<a<1$.

类似地, $2D-A$ 正定的充要条件是 $-1<a<\frac{1}{2}$.

所以 Jacobi 收敛的充要条件是 $-\frac{1}{2}<a<\frac{1}{2}$,

G-S 收敛的充要条件是 A 正定, 即 $-\frac{1}{2}<a<1$,

SOR 收敛的充要条件是 A 正定且 $0<\omega<2$, 即 $-\frac{1}{2}<a<1,\quad 0<\omega<2$.

谢谢

THANK YOU