第六章 解线性方程组的迭代法

朱升峰

华东师范大学 数学科学学院

2022 年 12 月

线性方程组的迭代法

求解线性方程组（A 非奇异）

$$Ax = b$$

工程技术中产生的 A：大型、稀疏

- 雅可比迭化法
- 高斯-塞德尔迭代法
- 超松驰迭代法
- 共轭梯度法

例 求解线性方程组

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

记作 $Ax = b$, 其中

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

方程组精确解 $x^* = (3, 2, 1)^T$

例

方程组改写为

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

即 $x = B_0 x + f$

$$\boldsymbol{B}_0 = \begin{bmatrix} 0 & \frac{3}{8} & -\frac{2}{8} \\ -\frac{4}{11} & 0 & \frac{1}{11} \\ -\frac{6}{12} & -\frac{3}{12} & 0 \end{bmatrix}, \quad f = \begin{bmatrix} \frac{20}{8} \\ \frac{33}{11} \\ \frac{36}{12} \end{bmatrix} \tag{1}$$

任取初值如 $x^{(0)} = (0,0,0)^T$，代入一般不满足(1)，通过迭代得到 $x^{(1)} = [2.5, 3, 3]$，利用 $x^{(1)}$ 得到 $x^{(2)}$，反复迭代得

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

矩阵的条件数

简写为

$$x^{(k+1)} = B_0 x^{(k)} + f, \quad k = 0, 1, 2, \cdots \tag{3}$$

其中 k 为迭代次数。迭代 10 步后,

$$\boldsymbol{x}^{(10)} = (3.000032, 1.999874, 0.999881)^{\mathrm{T}}$$

$$\left\|\boldsymbol{\varepsilon}^{(10)}\right\|_\infty = 0.000125 \quad \left(\boldsymbol{\varepsilon}^{(0)} = \boldsymbol{x}^{(10)} - \boldsymbol{x}^*\right)$$

考虑由 $Ax = b$ 变形得到的等价线性方程组

$$x = Bx + f$$

由迭代法产生的向量迭代序列 $x^{(k)}$ 是否收敛于线性方程组的解 x^* 呢?
设 $x^{(0)}$ 为任取的初始向量, 按迭代公式构造向量序列

$$x^{(k+1)} = Bx^{(k)} + f, \quad k = 0, 1, 2, \cdots.$$

定义 1

(1) 用逐步代入求近似解的方法称为迭代法，亦称为一阶定常迭代法（这里 B 与 k 无关）;

(2) 如果 $\lim_{k\to\infty} x^{(k)}$ 存在 (记 x^*), 称此迭代法收敛, 显然 x^* 就是此程组的解, 否则称此迭代法发散。

考虑误差向量

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

易知，$\varepsilon^{(k+1)} = B\varepsilon^{(k)}(k = 0, 1, 2, \cdots)$, 递推得

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

考察 $\{x^{(k)}\}$ 的收敛性，就要研究 B 在什么条件下有

$$\lim_{k\to\infty} \varepsilon^{(k)} = 0$$

亦即当 $k \to \infty$ 时有 $B^k \to 0$ (零矩阵)。

向量序列与矩阵序列的收敛性

定义 2

设向量序列 $\{x^{(k)}\} \subset \mathbb{R}^n$, $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^{\mathrm{T}} \in \mathbb{R}^n$, 如存在 $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$, 使得

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

则称向量序列 $\{x^{(k)}\}$ 收敛于 x, 记作 $\lim_{k\to\infty} x^{(n)} = x$.

$$\lim_{k \to \infty} x^{(k)} = x \Leftrightarrow \lim_{k \to \infty} \|x^{(n)} - x\| = 0,$$

其中 $\|\cdot\|$ 为任一向量范数。

定义 3

设有矩阵序列 $A_k = \left(a_{ij}^{(k)}\right) \in \mathbb{R}^{n\times n}$ 及 $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, 如果 n^2 个数列极限存在且有

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

则称 $\{A_k\}$ 收敛于 A, 记作 $\lim_{k\to\infty} A_k = A$.

例 设有矩阵序列

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

且设 $|\lambda| < 1$, 考察其极限.

解 当 $|\lambda| < 1$ 时, 则有

$$\lim_{k\to\infty} A_k \equiv \lim_{k\to\infty} A^k = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

等价性质

定理 4

$$\lim_{k\to\infty} A_k = A \Leftrightarrow \lim_{k\to\infty} \|A_k - A\| = 0,$$

其中 $\|\cdot\|$ 为矩阵的任意一种算子范数.

定理 5

$\lim_{k\to\infty} A_k = 0$ 的充分必要条件是

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

定理 6

设 $B \in \mathbb{R}^{n\times n}$, 则下面 *3* 个命题等价: *(1)* $\lim_{k\to\infty} B^k = 0$ *(2)* $\rho(B) < 1$; *(3)* 至少存在一种从属的矩阵范数 $\|\cdot\|_\epsilon$, 使 $\|B\|_\epsilon < 1$.

定理 7

设 $B \in \mathbb{R}^{n \times n}$, $\|\cdot\|$ 为任一矩阵范数, 则

$$\lim_{k \to \infty} \|B^k\|^{\frac{1}{k}} = \rho(B).$$

考虑 $Ax = b$, 将 A 分裂为

$$A = M - N,$$

其中 M 为可选的非奇异矩阵，称其为分裂矩阵. 则

$$Ax = b \Leftrightarrow x = M^{-1}Nx + M^{-1}b.$$

求解线性方程组 $x = Bx + f$, 其中

$$B = M^{-1}N = M^{-1}(M - A) = I - M^{1}A, \quad f = M^{-1}b$$

一阶定常迭代法：($x^{(0)}$ 为初始向量)

$$x^{(k+1)} = Bx^{(k)} + f, \quad k = 0, 1, \cdots$$

称 $B = I - M^{-1}A$ 为迭代法的迭代矩阵。选取 M, 得到 $Ax = b$ 的各种迭代法。

定理 8 (定常迭代法的基本定理)

给定线性方程组及一阶定常迭代法, 对任意选取初始向量 $x^{(0)}$, 上述迭代法收敛的充要条件是矩阵 B 谱半径 $\rho(B) < 1$.

考察迭代法

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

的收敛性, 其中

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

解 特征方程为 $\det(\lambda I - B) = \lambda^2 - 6 = 0$, 特征根 $\lambda_{1,2} = \pm\sqrt{6}$, 即 $\rho(B) > 1$. 这说用迭代法解此方程组不收敛.

定理 9 (迭代法收敛的充分条件)

设有线性方程组

$$Ax = b \Leftrightarrow x = Bx + f,$$

及一阶定常迭代法

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

如果有 B 的某种算子范数 $\|B\| = q < 1$, 则
(1) 迭代法收敛, 即对任取 $x^{(0)}$ 有

$$\lim_{k\to\infty} x^{(k)} = x^*, \text{ 且} x^* = Bx^* + f.$$

(2)

$$\|x^* - x^{(k)}\| \leqslant q^k \|x^* - x^{(0)}\|$$

(3)

$$\|x^* - x^{(k)}\| \leqslant \frac{q}{1-q} \|x^{(k)} - x^{(k-1)}\|$$

(4)

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

特例

注意定理只给出迭代法收敛的充分条件, 即条件 $\|B\| < 1$ 对任何常用范数均不成立, 迭代序列仍可能收敛.

例 迭代法 $x^{(k+1)} = Bx^{(k)} + f$, 其中

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

显然 $\|B\|_\infty = 1.1$, $\|B\|_1 = 1.2$, $\|B\|_2 = 1,043$, $\|B\|_F = \sqrt{1.54}$, 表明 B 的各种范数均大于 1 , 但由于 $\rho(B) = 0.9 < 1$, 由此迭代法产生的迭代序列 $\{x^{(k)}\}$ 却是收敛的.

迭代法的收敛性

由

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

得

$$\|\epsilon^{(k)}\| \leqslant \|B^k\| \, \|\epsilon^{(0)}\|, \quad \forall \epsilon^{(0)} \neq 0.$$

于是

$$\frac{\|\epsilon^{(k)}\|}{\|\epsilon^{(0)}\|} \leqslant \|B^k\|.$$

根据矩阵从属范数定义，有

$$\|B^k\| = \max_{\epsilon^{(0)} \neq 0} \frac{\|B^k \epsilon^{(0)}\|}{\|\epsilon^{(0)}\|} = \max_{\epsilon^{(0)} \neq 0} \frac{\|\epsilon^{(k)}\|}{\|\epsilon^{(0)}\|},$$

要使次 k 迭代后，

$$\|\epsilon^{(k)}\| \leqslant \sigma \|\epsilon^{(0)}\|,$$

其中 $\sigma << 1$ 可取 $\sigma = 10^{-s}$, 使

$$\|B^k\| \leqslant \sigma$$

因 $\rho(B)<1$, 当 k 充分大时，$\|B^k\|^{\frac{1}{k}}<1$. 对

$$\|B^k\|^{\frac{1}{k}} \leqslant \sigma^{\frac{1}{k}}$$

两边取对数得

$$\ln\|B^k\|^{\frac{1}{k}} \leqslant \frac{1}{k}\ln\sigma,$$

$$k \geqslant \frac{-\ln\sigma}{-\ln\|B^k\|^{\frac{1}{k}}} = \frac{-s\ln 10}{-\ln\|B^k\|^{\frac{1}{k}}}.$$

表明迭代次数 k 与 $-\ln\|B^k\|^{\frac{1}{k}}$ 成反比.

迭代法

$$x^{(k+1)} = Bx^{(k)} + f, \quad k = 0, 1, 2, \cdots$$

定义 10

迭代法的平均收敛速度定义为

$$R_k(B) = -\ln\|B^k\|^{\frac{1}{k}},$$

平均收敛速度 $R_k(B)$ 依赖于迭代次数及所取范数, 给计算分析带来不便.

$$\lim_{k\to\infty}\|B^k\|^{\frac{1}{k}} = \rho(B) => \lim_{k\to\infty} R_k(B) = -\ln\rho(B)$$

定义 11

迭代法渐近收敛速度定义为

$$R(B) = -\ln \rho(B).$$

$R(B)$ 与迭代次数及 B 取何种范数无关,
反映了迭代次数趋于无穷时迭代法的渐近性,
当 $\rho(B)$ 越小时 $-\ln \rho(B)$ 越大, 迭代法收敛越快, 可用

$$k \geqslant \frac{-\ln \sigma}{R(B)} = \frac{s \ln 10}{R(B)}$$

作为迭代法所需的迭代次数的估计.