计算方法

Ch05 线性方程组的迭代法

程勇

buctcourse@163.com

http://www.buct.edu.cn

Dept. of Computer
Beijing University of Chemical Technology

Sept. 19, 2019

提纲

本章提纲

雅克比 (Jacobi) 迭代法

设有 n 阶方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

写成矩阵形式为：

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

其中

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

雅克比 (Jacobi) 迭代法 (续)

若系数矩阵 A 非奇异，且 $a_{ii} \neq 0$ $(i = 1, 2, \cdots, n)$，则可将上述方程组改写成如下形式：

$$\begin{cases} x_1 = \frac{1}{a_{11}}(b_1 - a_{12}x_2 - a_{13}x_3 - \cdots - a_{1n}x_n) \\ x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2n}x_n) \\ \cdots \\ x_n = \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - a_{n2}x_2 - \cdots - a_{nn-1}x_{n-1}) \end{cases}$$

改写为迭代格式如下：

$$\begin{cases} x_1^{k+1} = \frac{1}{a_{11}}(b_1 - a_{12}x_2^k - a_{13}x_3^k - \cdots - a_{1n}x_n^k) \\ x_2^{k+1} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^k - a_{23}x_3^k - \cdots - a_{2n}x_n^k) \\ \cdots \\ x_n^{k+1} = \frac{1}{a_{nn}}(b_n - a_{n1}x_1^k - a_{n2}x_2^k - \cdots - a_{nn-1}x_{n-1}^k) \end{cases}$$

雅克比 (Jacobi) 迭代法 (续)

将上述迭代格式可简写为：

$$x_i^{k+1} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1,j\neq i}^{n} a_{ij}x_j^k\right) \qquad i = 1, 2, 3, \cdots, n$$

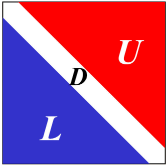
终止迭代条件为：

$$\max_{1\leq i\leq n} |x_i^{k+1} - x_i^k| < \varepsilon$$

其中 ε 为事前给定的阈值。

雅克比 (Jacobi) 迭代法的矩阵形式

矩阵按区域分为三部分 L, D, U，如下：



$$
\begin{aligned}
Ax = b &\Leftrightarrow (D + L + U)\,x = b \\
&\Leftrightarrow Dx = -(L + U)x + b \\
&\Leftrightarrow x = \underbrace{-D^{-1}(L + U)}_{B}x + \underbrace{D^{-1}b}_{f}
\end{aligned}
$$

因此，雅克比 (Jacobi) 迭代法的矩阵形式为：

$$
\begin{aligned}
x^{k+1} &= Bx^k + f \\
&= -D^{-1}(L + U)x^k + D^{-1}b
\end{aligned}
$$

Jacobi 迭代法例子

例：用 Jacobi 迭代法解方程组

$$\begin{cases} 10x_1 + 3x_2 + x_3 = 14 \\ 2x_1 - 10x_2 + 3x_3 = -5 \\ x_1 + 3x_2 + 10x_3 = 14 \end{cases}$$

初值 $x_1^0 = x_2^0 = x_3^0 = 0$，要求精度 $\varepsilon = 0.02$。
解：迭代格式：

$$\begin{cases} x_1^{k+1} = -0.3x_2^k - 0.1x_3^k + 1.4 \\ x_2^{k+1} = 0.2x_1^k + 0.3x_3^k + 0.5 \\ x_3^{k+1} = -0.1x_1^k - 0.3x_2^k + 1.4 \end{cases}$$

反复迭代，计算结果见下表。

Jacobi 迭代法例子 (续)

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	1.4	0.5	1.4
2	1.11	1.20	1.11
3	0.929	1.055	0.929
4	0.9906	0.9645	0.9906
5	1.01159	0.9953	1.01159
6	1.000251	1.005795	1.000251

Figure: 迭代计算结果

可以看到，当迭代次数 k 增大时，迭代值 x_1^k, x_2^k, x_3^k 会越来越接近解 $x_1 = x_2 = x_3 = 1$。

高斯-赛得尔迭代式

$$\begin{cases} x_1^{k+1} = \frac{1}{a_{11}}(-a_{12}x_2^k - a_{13}x_3^k - a_{14}x_4^k - \cdots - a_{1n}x_n^k + b_1) \\ x_2^{k+1} = \frac{1}{a_{22}}(-a_{21}x_1^{k+1} - a_{23}x_3^k - a_{24}x_4^k - \cdots - a_{2n}x_n^k + b_2) \\ x_3^{k+1} = \frac{1}{a_{33}}(-a_{31}x_1^{k+1} - a_{32}x_2^{k+1} - a_{34}x_4^k - \cdots - a_{3n}x_n^k + b_3) \\ \cdots \\ x_n^{k+1} = \frac{1}{a_{nn}}(-a_{n1}x_1^{k+1} - a_{n2}x_2^{k+1} - a_{n3}x_3^{k+1} - \cdots - a_{nn-1}x_{n-1}^{k+1} + b_n) \end{cases}$$

将上述迭代格式可简写为如下通式：

$$x_i^{k+1} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{k+1} - \sum_{j=i+1}^{n} a_{ij}x_j^k\right) \qquad i = 1, 2, 3, \cdots, n$$

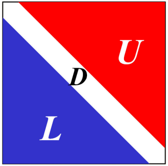
终止迭代条件为：

$$\max_{1\le i\le n} |x_i^{k+1} - x_i^k| < \varepsilon$$

其中 ε 为事前给定的阈值。

高斯-赛得尔迭代法的矩阵形式

矩阵按区域分为三部分 $\boldsymbol{L}, \boldsymbol{D}, \boldsymbol{U}$，如下：



$$
\begin{aligned}
&\boldsymbol{x}^{k+1} = -\boldsymbol{D}^{-1}(\boldsymbol{L}\boldsymbol{x}^{k+1} + \boldsymbol{U}\boldsymbol{x}^{k}) + \boldsymbol{D}^{-1}\boldsymbol{b} \\
&\Leftrightarrow (\boldsymbol{D}+\boldsymbol{L})\boldsymbol{x}^{k+1} = -\boldsymbol{U}\boldsymbol{x}^{k} + \boldsymbol{b} \\
&\Leftrightarrow \boldsymbol{x}^{k+1} = \underbrace{-(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}}_{\boldsymbol{B}}\boldsymbol{x}^{k} + \underbrace{(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{b}}_{\boldsymbol{f}}
\end{aligned}
$$

因此，高斯-赛得尔迭代法的矩阵形式为：

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

高斯-赛得尔迭代法例子

例：用 Gauss-Seidel 迭代法解方程组

$$\begin{cases} 10x_1 + 3x_2 + x_3 = 14 \\ 2x_1 - 10x_2 + 3x_3 = -5 \\ x_1 + 3x_2 + 10x_3 = 14 \end{cases}$$

初值 $x_1^0 = x_2^0 = x_3^0 = 0$，要求精度 $\varepsilon = 0.05$。
解：迭代格式：

$$\begin{cases} x_1^{k+1} = -0.3x_2^k - 0.1x_3^k + 1.4 \\ x_2^{k+1} = 0.2x_1^{k+1} + 0.3x_3^k + 0.5 \\ x_3^{k+1} = -0.1x_1^{k+1} - 0.3x_2^{k+1} + 1.4 \end{cases}$$

反复迭代，计算结果见下表。

高斯-赛得尔迭代法例子 (续)

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	1.4	0.78	1.026
2	1.06340	1.02048	0.98752
3	0.99510	0.99528	1.00191
4	1.00122	1.00082	0.99963

Figure: 迭代计算结果

可以看到，当迭代次数增大时，迭代值 x_1^k, x_2^k, x_3^k 会越来越接近解 $x_1 = x_2 = x_3 = 1$。一般来说，Gauss-Seidel 要比 Jacobi 迭代好，收敛速度快，但也有特例，甚至有时 Jacobi 迭代收敛而 Gauss-Seidel 迭代发散。

迭代法的收敛条件

对角占优矩阵定义

如果矩阵的每一行中，不在主对角线上的所有元素绝对值之和小于主对角线上元素的绝对值，即：

$$\sum_{j=1,j\neq i}^{n} |a_{ij}| < |a_{ii}| \quad i = 1, 2, 3, \cdots, n$$

则称矩阵 $\boldsymbol{A}$ 按行严格对角占优。类似地，也有按列严格对角占优。

定理

若线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 的系数矩阵 $\boldsymbol{A}$ 按行严格对角占优，则雅克比迭代法和高斯-赛得尔迭代法对任意给定初值均收敛。

迭代收敛判定例子

例题：解方程组

$$\begin{cases} 20x_1 + 2x_2 + 3x_3 = 24 \\ x_1 + 8x_2 + x_3 = 12 \\ 2x_1 - 3x_2 + 15x_3 = 30 \end{cases}$$

取 $x_1^0 = x_2^0 = x_3^0 = 0$。问：Jacobi 和 Gauss-Seidel 迭代法是否收敛？

解：由于

$$20 > 2 + 3 = 5$$
$$8 > 1 + 1 = 2$$
$$15 > 2 + |-3| = 5$$

因此系数矩阵是按行严格对角占优矩阵，根据定理 Jacobi 和 Gauss-Seidel 迭代法应用于此方程组均收敛。

向量和矩阵的范数

为了研究线性方程组近似解的误差估计和迭代法的收敛性，有必要对向量及矩阵的“大小”引进某种度量，即范数的概念。向量范数是用来度量向量长度的，它可以看成是二、三维解析几何中向量长度概念的推广。用 R^n 表示 n 维实向量空间。

定义

对任一向量 $\boldsymbol{X} \in R^n$，按照一定规则确定一个实数与它对应，该实数记为 $\|\boldsymbol{X}\|$。若 $\|\boldsymbol{X}\|$ 满足下面三个性质：

1. $\|\boldsymbol{X}\| \geq 0$，$\|\boldsymbol{X}\| = 0$ 当且仅当 $\boldsymbol{X} = 0$；
2. 对任意实数 λ，$\|\lambda\boldsymbol{X}\| = |\lambda|\|\boldsymbol{X}\|$；
3. 对任意向量 $\boldsymbol{Y} \in R^n$，$\|\boldsymbol{X} + \boldsymbol{Y}\| \leq \|\boldsymbol{X}\| + \|\boldsymbol{Y}\|$；

则称该实数 $\|\boldsymbol{X}\|$ 为向量 $\boldsymbol{X}$ 的范数。

常见的范数

在 R^n 中，常见的范数有以下几种：

1. $\|\boldsymbol{X}\|_1 = |x_1| + |x_2| + \cdots + |x_n| = \sum_{i=1}^n |x_i|$;
2. $\|\boldsymbol{X}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$;
3. $\|\boldsymbol{X}\|_\infty = \max\{|x_1|, |x_2|, \cdots, |x_n|\} = \max_{1\le i\le n}\{|x_i|\}$;

其中 $x_1, x_2, \cdots, x_n$ 分别是 $\boldsymbol{X}$ 的 n 个分量。以上定义的范数分别称为 1-范数，2-范数和 ∞-范数。可以验证它们都是满足范数性质的，其中 $\|x\|_2$ 是由内积导出的向量范数。

上述范数都是 p 范数

$$\|\boldsymbol{X}\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$$

当不需要指明使用哪一种向量范数时，就用记号 $\|\cdot\|$ 泛指任何一种向量范数。

∞-范数

定理

对任意向量 $\boldsymbol{x} \in R^n$，有 $\lim_{p\to\infty} ||\boldsymbol{x}||_p = ||\boldsymbol{x}||_\infty$。

证明：由于 $||\boldsymbol{x}||_\infty = \max_{1\le i\le n} |x_i|$
所以

$$||\boldsymbol{x}||_\infty = \left(\max_{1\le i\le n} |x_i|^p\right)^{\frac{1}{p}} \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \le \left(n \max_{1\le i\le n} |x_i|^p\right)^{\frac{1}{p}}$$

即：

$$||\boldsymbol{x}||_\infty \le ||\boldsymbol{x}||_p \le n^{\frac{1}{p}} ||\boldsymbol{x}||_\infty$$

当 $p \to \infty$，$n^{\frac{1}{p}} \to 1$。
所以：

$$\lim_{p\to\infty} ||\boldsymbol{x}||_p = ||\boldsymbol{x}||_\infty$$

范数计算的例子

例题：设 $\boldsymbol{x}=(1,0,-1,2)^\top$，试计算 1-范数，2-范数和 ∞-范数。

解:

$$
\begin{aligned}
\|\boldsymbol{x}\|_1 &= 1+0+|-1|+2=4 \\
\|\boldsymbol{x}\|_2 &= \sqrt{1^2+0^2+(-1)^2+2^2} \\
&= \sqrt{6} \\
\|\boldsymbol{x}\|_\infty &= \max\{1,0,|-1|,2\} \\
&= 2
\end{aligned}
$$

矩阵的范数

定义

如果矩阵 $\boldsymbol{A}\in R^{n\times n}$ 的某个非负实值函数 $N(\boldsymbol{A})=\|\boldsymbol{A}\|$，满足：

1. $\|\boldsymbol{A}\|\geq 0$，当且仅当 $\boldsymbol{A}=0$ 时，$\|\boldsymbol{A}\|=0$；
2. 对任意实数 λ，$\|\lambda\boldsymbol{A}\|=|\lambda|\|\boldsymbol{A}\|$；
3. 对任意矩阵 $\boldsymbol{A},\boldsymbol{B}\in R^{n\times n}$，$\|\boldsymbol{A}+\boldsymbol{B}\|\leq\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$；
4. $\|\boldsymbol{A}\boldsymbol{B}\|\leq\|\boldsymbol{A}\|\cdot\|\boldsymbol{B}\|$

则称 $N(\boldsymbol{A})$ 是 $R^{n\times n}$ 的一个矩阵范数 (或模)。

对 n 阶方阵 $\boldsymbol{A}=(a_{ij})_n$，共 1,2,$\infty$ 范数计算如下。其中，$\lambda_{\max}(\boldsymbol{A}^\top\boldsymbol{A})$ 表示 $\boldsymbol{A}^\top\boldsymbol{A}$ 的最大特征值，即满足

$$f(\lambda)=|\lambda\boldsymbol{I}_n-\boldsymbol{A}^\top\boldsymbol{A}|=0$$

矩阵范数的计算

定义

1. $\|\boldsymbol{A}\|_\infty$ 范数 (称为 $\boldsymbol{A}$ 的行范数):

$$\|\boldsymbol{A}\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$$

2. $\|\boldsymbol{A}\|_1$ 范数 (称为 $\boldsymbol{A}$ 的列范数):

$$\|\boldsymbol{A}\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|$$

3. $\|\boldsymbol{A}\|_2$ 范数 (称为 $\boldsymbol{A}$ 的 2 范数或谱范数):

$$\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{\max}(\boldsymbol{A}^\top \boldsymbol{A})}$$

矩阵范数计算例子

例题：计算方阵

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & -2 & 4 \end{pmatrix}$$

的三种常用范数。

解：

$$\|\boldsymbol{A}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| = \max\{1,4,8\} = 8$$

$$\|\boldsymbol{A}\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| = \max\{1,6,6\} = 6$$

$$\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{\max}(\boldsymbol{A}^\top \boldsymbol{A})} = \sqrt{32} = 4\sqrt{2}$$

矩阵的谱半径

定义

设 $\boldsymbol{A} \in R^{n\times n}$ 的特征值为 $\lambda_i(i=1,2,3,\cdots,n)$，称

$$\rho(\boldsymbol{A}) = \max_{1\leq i\leq n} |\lambda_i|$$

为 $\boldsymbol{A}$ 的谱半径。

定理

设 $\boldsymbol{A} \in R^{n\times n}$ 为 n 阶方阵，则对任意矩阵范数 $\|\cdot\|$ 都有 $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|$。

迭代法收敛性定理

迭代法收敛性定理

设有线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$，则对于任意的初始向量 $\boldsymbol{X}^{(0)}$，迭代法

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{B}\boldsymbol{X}^{(k)}+\boldsymbol{b}$$

收敛的充分必要条件是

$$\rho(\boldsymbol{B})<1$$

上述定理表明，线性方程组迭代法收敛与否与 $\boldsymbol{X}^{(0)}$ 和 $\boldsymbol{b}$ 无关，而只与迭代矩阵 $\boldsymbol{B}$ 的性质有关。

本章小结

- 雅克比迭代法;
- 高斯-赛得尔迭代法;

练习题

1. 编程实现 Gauss-Seidel 迭代算法。

谢谢!

AUTHOR: Cheng Yong
ADDRESS: Dept. of Computer
Beijing University of Chemical Technology
Beijing, 100029, China
EMAIL: buctcourse@163.com