《数值计算方法》课程

解方程组（高斯赛德尔迭代法）

胡建芳

http://sdcs.sysu.edu.cn/content/5143

计算机学院

课程回顾

- **直接法：** 直接解方程组

高斯法

消元，回代，

复杂度O(n^3)；

LU分解法

LU分解

分解的充要条件：前n-1个顺序主子式不为0

误差

范数

条件数

部分选主元LU分解法

消元过程中的乘子，绝对值小于1

迭代法

- **Jacobi法：**

  不动点迭代的形式：改写方程组，然后不停的迭代

应用 Jacobi 方法解方程组 $3u+v=5, u+2v=5$

$$\begin{cases} u = \frac{5-v}{3}, \\ v = \frac{5-u}{2}. \end{cases} \Rightarrow$$

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \frac{5-v_0}{3} \\ \frac{5-u_0}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-0}{3} \\ \frac{5-0}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{5}{2} \end{bmatrix},$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{5-v_1}{3} \\ \frac{5-u_1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-5/2}{3} \\ \frac{5-5/3}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ \frac{5}{3} \end{bmatrix},$$

$$\begin{bmatrix} u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} \frac{5-5/3}{3} \\ \frac{5-5/6}{2} \end{bmatrix} = \begin{bmatrix} \frac{10}{9} \\ \frac{25}{12} \end{bmatrix}.$$

继续迭代，收敛到[1,2]

迭代法

- **Jacobi法：**

  不动点迭代的形式：改写方程组，然后不停的迭代

应用 Jacobi 方法解方程组 $3u+v=5, u+2v=5$

$$\begin{cases} u = 5-2v, \\ v = 5-3u \end{cases} \Rightarrow$$

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 5-2v_0 \\ 5-3u_0 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix},$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} 5-2v_1 \\ 5-3u_1 \end{bmatrix} = \begin{bmatrix} -5 \\ -10 \end{bmatrix},$$

$$\begin{bmatrix} u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} 5-2(-10) \\ 5-3(-5) \end{bmatrix} = \begin{bmatrix} 25 \\ 20 \end{bmatrix}.$$

继续迭代发散，为什么？

迭代法

- **Jacobi法：**

  严格对角占优矩阵

定义 2.9 若对每个 $1 \leqslant i \leqslant n$, $|a_{ii}| > \sum\limits_{j \neq i} |a_{ij}|$, 则称 $n \times n$ 矩阵 $A = (a_{ij})$ 是**严格对角占优的**. 也就是说, 每个主对角元在绝对值上要比所在行的其他所有元素的绝对值和更大.

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right].$$

迭代法

- **Jacobi法：**

严格对角占优矩阵

定义 2.9 若对每个 $1 \leqslant i \leqslant n$, $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, 则称 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 是**严格对角占优**的. 也就是说, 每个主对角元在绝对值上要比所在行的其他所有元素的绝对值和更大.

确定矩阵

$$\boldsymbol{A} = \begin{bmatrix} 3 & 1 & -1 \\ 2 & -5 & 2 \\ 1 & 6 & 8 \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} 3 & 2 & 6 \\ 1 & 8 & 1 \\ 9 & 2 & -2 \end{bmatrix}$$

是否严格对角占优?

迭代法

- **Jacobi法:**

  严格对角占优矩阵

定义 2.9 若对每个 $1 \leqslant i \leqslant n, |a_{ii}| > \sum_{j \neq i} |a_{ij}|$, 则称 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 是**严格对角占优**的. 也就是说, 每个主对角元在绝对值上要比所在行的其他所有元素的绝对值和更大.

确定矩阵

$$\boldsymbol{A} = \begin{bmatrix} 3 & 1 & -1 \\ 2 & -5 & 2 \\ 1 & 6 & 8 \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} 3 & 2 & 6 \\ 1 & 8 & 1 \\ 9 & 2 & -2 \end{bmatrix}$$

对角占优，第一个例子 **非对角占优，第二个例子**

迭代法

- **Jacobi法：**

  迭代形式：A=D+L+U

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b},\\
(\boldsymbol{D}+\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x} &= \boldsymbol{b},\\
\boldsymbol{D}\boldsymbol{x} &= \boldsymbol{b}-(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x},\\
\boldsymbol{x} &= \boldsymbol{D}^{-1}(\boldsymbol{b}-(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x}).
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{x}_0 &= \text{初始向量},\\
\boldsymbol{x}_{k+1} &= \boldsymbol{D}^{-1}(\boldsymbol{b}-(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x}_k), \quad k=0,1,2,\cdots.
\end{aligned}
$$

迭代公式

$$
\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix},
$$

带有 $\boldsymbol{x}_k = \begin{bmatrix} u_k \\ v_k \end{bmatrix}$ 的不动点迭代

$$
\begin{bmatrix} u_{k+1} \\ v_{k+1} \end{bmatrix} = \boldsymbol{D}^{-1}(\boldsymbol{b}-(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x}_k) = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\left(\begin{bmatrix} 5 \\ 5 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}_k\right) = \begin{bmatrix} \frac{5-v_k}{3} \\ \frac{5-u_k}{2} \end{bmatrix},
$$

这与我们原先的格式是一致的.

迭代法

- **Gauss-Seidel法和SOR：**

迭代形式：A=D+L+U

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \frac{5-v_0}{3} \\ \frac{5-u_0}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-0}{3} \\ \frac{5-0}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{5}{2} \end{bmatrix},$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{5-v_1}{3} \\ \frac{5-u_1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-5/2}{3} \\ \frac{5-5/3}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ \frac{5}{3} \end{bmatrix},$$

$$\begin{bmatrix} u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} \frac{5-5/3}{3} \\ \frac{5-5/6}{2} \end{bmatrix} = \begin{bmatrix} \frac{10}{9} \\ \frac{25}{12} \end{bmatrix}.$$

Jacobi迭代

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \frac{5-v_0}{3} \\ \frac{5-u_1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-0}{3} \\ \frac{5-5/3}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{5}{3} \end{bmatrix},$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{5-v_1}{3} \\ \frac{5-u_2}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-5/3}{3} \\ \frac{5-10/9}{2} \end{bmatrix} = \begin{bmatrix} \frac{10}{9} \\ \frac{35}{18} \end{bmatrix},$$

$$\begin{bmatrix} u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} \frac{5-v_2}{3} \\ \frac{5-u_3}{2} \end{bmatrix} = \begin{bmatrix} \frac{5-35/18}{3} \\ \frac{5-55/54}{2} \end{bmatrix} = \begin{bmatrix} \frac{55}{54} \\ \frac{215}{108} \end{bmatrix}$$

Gauss-Seidel迭代

区别?

迭代法

- **Gauss-Seidel法和SOR：**

  迭代形式：A=D+L+U

$$(\boldsymbol{L}+\boldsymbol{D})\boldsymbol{x}_{k+1} = -\boldsymbol{U}\boldsymbol{x}_k + \boldsymbol{b}.$$

$$\boldsymbol{x}_0 = \text{初始向量},$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{D}^{-1}(\boldsymbol{b} - \boldsymbol{U}\boldsymbol{x}_k - \boldsymbol{L}\boldsymbol{x}_{k+1}), \quad k = 0, 1, 2, \cdots.$$

迭代法

- **Gauss-Seidel法和SOR：**

应用 Gauss-Seidel 方法, 求解方程组

$$\begin{bmatrix} 3 & 1 & -1 \\ 2 & 4 & 1 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$$

$$u_{k+1} = \frac{4 - v_k + w_k}{3}, \qquad v_{k+1} = \frac{1 - 2u_{k+1} - w_k}{4}, \qquad w_{k+1} = \frac{1 + u_{k+1} - 2v_{k+1}}{5}$$

从 $\boldsymbol{x}_0 = [u_0, v_0, w_0] = [0, 0, 0]$ 开始, 我们计算

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} \frac{4-0-0}{3} = \frac{4}{3} \\ \frac{1-8/3-0}{4} = -\frac{5}{12} \\ \frac{1+4/3+5/6}{5} = \frac{19}{30} \end{bmatrix} \approx \begin{bmatrix} 1.333\,3 \\ -0.416\,7 \\ 0.633\,3 \end{bmatrix},$$

$$\begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{101}{60} \\ -\frac{3}{4} \\ \frac{251}{300} \end{bmatrix} \approx \begin{bmatrix} 1.683\,3 \\ -0.750\,0 \\ 0.836\,7 \end{bmatrix}.$$

方程组是严格对角占优的, 因此迭代将收敛到解 $[2, -1, 1]$.

迭代法

- **Gauss-Seidel法和SOR：**

称为**逐次超松弛**(Successive Over-Relaxation, SOR) 的方法采取 Gauss-Seidel 趋向于解的方向并试图加速收敛. 设 ω 是一实数, 定义新估计量 x_{k+1} 的每个分量为 ω 乘上 Gauss-Seidel 公式与 $1-\omega$ 乘上当前估计量 x_k 的加权平均. 数 ω 叫做**松弛参数**(relaxation parameter), $\omega > 1$ 时被认为是**超松弛的**(over-relaxation).

$$u_{k+1} = \frac{4 - v_k + w_k}{3},$$
$$v_{k+1} = \frac{1 - 2u_{k+1} - w_k}{4}.$$
$$w_{k+1} = \frac{1 + u_{k+1} - 2v_{k+1}}{5}$$

Gauss-Seidel迭代

$$u_{k+1} = (1-\omega)u_k + \omega\frac{4 - v_k + w_k}{3},$$
$$v_{k+1} = (1-\omega)v_k + \omega\frac{1 - 2u_{k+1} - w_k}{4},$$
$$w_{k+1} = (1-\omega)w_k + \omega\frac{1 + u_{k+1} - 2v_{k+1}}{5}.$$

SOR 迭代

对应于物理中的动量（momentum），保持运动惯性；深度学习中重要的优化技巧

迭代法

- **Gauss-Seidel法和SOR：**

$$u_{k+1} = \frac{4 - v_k + w_k}{3},$$
$$v_{k+1} = \frac{1 - 2u_{k+1} - w_k}{4}$$
$$w_{k+1} = \frac{1 + u_{k+1} - 2v_{k+1}}{5}$$

$$u_{k+1} = (1-\omega)u_k + \omega\frac{4 - v_k + w_k}{3},$$
$$v_{k+1} = (1-\omega)v_k + \omega\frac{1 - 2u_{k+1} - w_k}{4},$$
$$w_{k+1} = (1-\omega)w_k + \omega\frac{1 + u_{k+1} - 2v_{k+1}}{5}.$$

Gauss-Seidel迭代

SOR 迭代

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} \frac{4-0-0}{3} = \frac{4}{3} \\ \frac{1-8/3-0}{4} = -\frac{5}{12} \\ \frac{1+4/3+5/6}{5} = \frac{19}{30} \end{bmatrix} \approx \begin{bmatrix} 1.333\,3 \\ -0.416\,7 \\ 0.633\,3 \end{bmatrix},$$

$$\begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{101}{60} \\ -\frac{3}{4} \\ \frac{251}{300} \end{bmatrix} \approx \begin{bmatrix} 1.683\,3 \\ -0.750\,0 \\ 0.836\,7 \end{bmatrix}.$$

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} \approx \begin{bmatrix} 1.666\,7 \\ -0.729\,2 \\ 1.031\,2 \end{bmatrix}, \quad \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} \approx \begin{bmatrix} 1.983\,5 \\ -1.067\,2 \\ 1.021\,6 \end{bmatrix}$$

迭代法

- **Gauss-Seidel法和SOR：**

$$(\omega \boldsymbol{L}+\omega \boldsymbol{D}+\omega \boldsymbol{U})\boldsymbol{x}=\omega \boldsymbol{b},$$

$$(\omega \boldsymbol{L}+\boldsymbol{D})\boldsymbol{x}=\omega \boldsymbol{b}-\omega \boldsymbol{U}\boldsymbol{x}+(1-\omega)\boldsymbol{D}\boldsymbol{x},$$

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

逐次超松弛 (SOR)

$$\boldsymbol{x}_0=\text{初始向量},$$

$$\boldsymbol{x}_{k+1}=(\omega \boldsymbol{L}+\boldsymbol{D})^{-1}[(1-\omega)\boldsymbol{P}\boldsymbol{x}_k-\omega(\boldsymbol{x}_k)]+\omega(\boldsymbol{D}+\omega \boldsymbol{L})^{-1}\boldsymbol{b}, \quad k=0,1,2,\cdots.$$

迭代法

Gauss-Seidel法和SOR：

例 2.24 对含 6 个未知量的 6 个方程的方程组, 比较 Jacobi、Gauss-Seidel 和 SOR：

$$\begin{bmatrix} 3 & -1 & 0 & 0 & 0 & \frac{1}{2} \\ -1 & 3 & -1 & 0 & \frac{1}{2} & 0 \\ 0 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & 0 \\ 0 & \frac{1}{2} & 0 & -1 & 3 & -1 \\ \frac{1}{2} & 0 & 0 & 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ \frac{3}{2} \\ 1 \\ 1 \\ \frac{3}{2} \\ \frac{5}{2} \end{bmatrix}. \tag{2.35}$$

解是 $\boldsymbol{x} = [1, 1, 1, 1, 1, 1]$

Jacobi	Gauss-Seidel	SOR
0.987 9	0.995 0	0.998 9
0.984 6	0.994 6	0.999 3
0.967 4	0.996 9	1.000 4
0.967 4	0.999 6	1.000 9
0.984 6	1.001 6	1.000 9
0.987 9	1.001 3	1.000 4

迭代6步的结果

迭代法

- **Gauss-Seidel法和SOR：**

  收敛性

Jacobi 方法写作

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

谱半径

$$\rho(\boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U})) < 1$$

方程不动点迭代收敛条件：一阶导绝对值小于1

方程组不动点迭代收敛条件：谱半径（一阶导）小于1

迭代法

- **Gauss-Seidel法和SOR：**

定理 2.11 若 $n\times n$ 矩阵 $\boldsymbol{A}$ 是严格对角占优的，则 (1) $\boldsymbol{A}$ 是非奇异矩阵，(2) 对每个向量 $\boldsymbol{b}$ 及每个开始猜测，应用于 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的 Gauss-Seidel 方法都收敛到一个解.

证 设 λ 为 (2.37) 的一个特征值，相应的特征向量为 $\boldsymbol{v}$. 如前一定理中的证明那样，选择特征向量使得 $v_m=1$ 且所有其他的分量在绝对值上都小于 1. 注意到 $\boldsymbol{L}$ 的元素是 $a_{ij}, i>j$; $\boldsymbol{U}$ 的元素是 $a_{ij}, i<j$. 则仔细检查 (2.37) 的特征值方程

$$\lambda(\boldsymbol{D}+\boldsymbol{L})\boldsymbol{v}=\boldsymbol{U}\boldsymbol{v}$$

的 m 行得到类似于前一证明中的一串不等式

$$|\lambda|\left(\sum_{i>m}|a_{mi}|\right)<|\lambda|\left(|a_{mm}|-\sum_{i<m}|a_{mi}|\right)\leqslant|\lambda|\left(|a_{mm}|-\left|\sum_{i<m}a_{mi}v_i\right|\right)$$
$$\leqslant|\lambda|\left|a_{mm}+\sum_{i<m}a_{mi}v_i\right|=\left|\sum_{i>m}a_{mi}v_i\right|\leqslant\sum_{i>m}|a_{mi}|.$$

这就得到 $|\lambda|<1$，从而完成了证明.

作业

- **作业：**

  用 Matlab或 python 实现 Gauss-Seidel 和 SOR 迭代法；随机生成矩阵（三角占优和非三角占优），验证算法的收敛性，并研究SOR中的w对算法结果的影响。写成实验报告，并附上代码。

THE END