

# 第六章 解线性方程组的迭代解法

## 6.1 Picard 迭代回顾

**内容** 有  $f(x) = 0$ ，将其化为等价方程  $x = \varphi(x)$ ，选取初值  $x_0$ ，令  $x_k = \varphi(x_{k-1})$ ，则有  $\lim_{k \rightarrow \infty} x_k = x^*$

$x^*$  即为原方程的解。  $|e_k| = |x_k - x^*| \leq \frac{L^k}{1-L} |x_1 - x_0|$

**向量的迭代公式** 对于 **线性方程组**  $Ax = b$ ，可以写为 **等价方程组**  $x = Bx + f$   
为此构造序列  $x^{(k+1)} = Bx^{(k)} + f$   
要考虑 **向量列的收敛性、收敛速度、误差** 等内容，和  $B$  有关。  
**收敛性判断**：迭代过程中前后两次的距离越来越小。

## 6.2 基本概念

### 6.2.1 向量与矩阵的范数

**考虑问题** 1. 能否定义 **一个函数（映射）** 将 **向量或者矩阵** 投影成 **一个实数**？  
2. 如果存在这个映射，这个映射应该满足什么性质？  
3. 作为 **距离函数** 应该满足什么性质？ 正定性、齐次性、三角不等式

#### 6.2.1.1 向量范数

**定义 1** 设  $V$  为线性空间， $V$  上的实值函数  $N(x) = \|x\|$  满足：

- ① **正定性**：  $\forall x \in V: \|x\| \geq 0$  且  $\|x\| = 0 \Leftrightarrow x = 0$  **当且仅当零向量时为零**（不存在负距离）
  - ② **齐次性**：  $\forall k \in R, \forall x \in V: \|kx\| = |k| \|x\|$  空间中任意向量模长伸缩，不存在负距离，加绝对值
  - ③ **三角不等式**：  $\forall x, y \in V: \|x + y\| \leq \|x\| + \|y\|$  **三角形中边长关系**
- 则称  $N(x) = \|x\|$  为  $V$  上向量  $x$  的范数

**性质** ① 若  $x \neq 0$ ，  $\left\| \frac{x}{\|x\|} \right\| = 1$  ②  $\|-x\| = \|x\|$  ③  $\left| \|x\| - \|y\| \right| \leq \|x - y\|$

**常用范数** 基本条件  $\forall x = (x_1, x_2, \dots, x_n) \in R^n (C^n)$

- ① **欧式范数（2-范数）**  $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} = \sqrt{\sum_{i=1}^n \bar{x}_i x_i}$  模长
- ② **最大模范数（ $\infty$ -范数）**  $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$  绝对值最大的一个
- ③ **绝对值范数（1-范数）**  $\|x\|_1 = \sum_{i=1}^n |x_i|$  绝对值之和（曼哈顿距离）
- ④  **$p$  范数**  $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$

**注意** ① 可以证明上述定义均满足向量的范数定义  
② 2-范数和1-范数都是  $p$ -范数的特例  
③  $\infty$ -范数也是  $p$ -范数的特例（ $p \rightarrow \infty, \|x\|_p \rightarrow \|x\|_\infty$ ）  
④ 范数可以认为是函数的函数。

**A 的 F 范数** 如果认为  $m \times n$  矩阵是一个  $m \times n$  维的向量，  $\forall A_{m \times n}: \|A\|_2 = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \triangleq \|A\|_F$   
一般 **不能简单的** 认为矩阵就是向量

### 6.2.1.2 矩阵范数

**条件** 简化起见, 先考虑方阵情况

**定义 2** 若 $\forall A_{n \times n}$ 对应一个实数 $\|A\|$ , 满足有:

- ①  $\|A\| \geq 0$ , 且  $\|A\| = 0 \Leftrightarrow A = 0$
- ②  $\|kA\| = |k|\|A\|$
- ③  $\|A + B\| \leq \|A\| + \|B\|$
- ④  $\|AB\| \leq \|A\|\|B\|$

则称 $\|A\|$ 为方阵的范数: **矩阵范数** 例如: **F范数**为矩阵范数, 行列式不是范数

**注意** 矩阵理化及运算常要考虑**矩阵与向量的乘积**, 我们希望范数 $\|Ax\|_{\text{向量范数}} \leq \|A\|_{\text{矩阵范数}} \|x\|_{\text{向量范数}}$  由向量范数出发, 寻找矩阵范数, 使得满足条件。则称矩阵范数为由向量范数诱导出的范数。

**定义 3** 设 $\|\cdot\|$ 为向量范数,  $\|\cdot\|_M$ 为矩阵范数。若 $\forall A \in R^{n \times n}, \forall x \in R^n \Rightarrow \|Ax\| \leq \|A\|_M \|x\|$  则称 $\|A\|_M$ 为与向量范数 $\|\cdot\|$ **相容的矩阵范数**

例如,  $\|A\|_F$ 与 $\|\cdot\|_1$ 不相容, 如果 $A = \begin{pmatrix} 1 & 4 \\ 1 & 4 \end{pmatrix} \quad x = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

**定义 4** 设 $\forall A \in R^{n \times m}$ ,  $\|\cdot\|$ 为向量范数, 称 $\|A\| = \max_{x \in R^n, x \neq 0} \frac{\|Ax\|}{\|x\|}$  为**矩阵A的算子范数 (诱导范数)**

**注意** ①  $\|A\| = \max_{\|x\|=1} \|Ax\|$

② 可以证明算子范数满足矩阵范数的 4 个条件. 故为矩阵范数.

③ 矩阵范数**不一定**都是算子范数, 如 F-范数.

④ 算子范数与向量范数相容。( $\|A\| = \max_{\|x\|=1} \frac{\|Ax\|}{\|x\|}$ ,  $\|Ax\| = \frac{\|Ax\|}{\|x\|} \|x\| \leq \|A\| \|x\|$ )

#### 常用范数

① **行范数**  $\|A\|_\infty = \max_{\|x\|=1} \|Ax\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$  行范数(每行相加, 取最大)

② **列范数**  $\|A\|_1 = \max_{\|x\|=1} \|Ax\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$  列范数(每列相加, 取最大)

③ **2 范数**  $\|A\|_2 = \max_{\|x\|=1} \|Ax\|_2 = \sqrt{\lambda_{\max}(A^T A)}$  难以计算, 但很有用

其中 $\lambda_{\max}(A^T A)$ 为矩阵 $A^T A$ 的绝对值最大的特征值

**例题** 设 $A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix}$ , 求它的三个范数

$$\|A\|_1 = \max\{2, 5, 2\} = 5 \quad \|A\|_\infty = \max\{3, 4, 2\} = 4$$

$$A^T A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 9 & -1 \\ 1 & -1 & 2 \end{pmatrix} \Rightarrow |\lambda I - A^T A| = \begin{vmatrix} \lambda - 2 & 0 & -1 \\ 0 & \lambda - 9 & 1 \\ -1 & 1 & \lambda - 2 \end{vmatrix} = 0 \quad \text{则 } \lambda^3 - 13\lambda^2 + 38\lambda - 25 = 0$$

可得:  $\lambda_1 = 9.14, \lambda_2 = 2.92, \lambda_3 = 0.9331$  即 $\|A\|_2 = \sqrt{9.14} = 3.023$

**定义 5** 设 $\lambda_i (i = 1, 2, \dots, n)$ 为 $A \in R^{n \times n}$ 的 $n$ 个特征值, 称 $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$ 为**A的谱半径**

$\rho(A) \leq \|A\|_p$  **谱半径小于任意的范数**, 可用于估计特征值的上界

**定理** 若 $\|A\| < 1$ , 则 $I + A$ 可逆, 且 $\|(I + A)^{-1}\| \leq \frac{1}{1 - \|A\|}$

① 因为 $\|A\| < 1$ , 所以 $\rho(A) < 1 \Rightarrow \pm 1$ 不是A的特征值, 所以 $I + A$ 可逆

② 令 $D = (I + A)^{-1}$ , 则 $1 = \|I\| = \|(I + A)D\| = \|D + AD\| \geq \|D\| - \|AD\| \geq \|D\| - \|A\|\|D\| = \|D\|(1 - \|A\|)$  因为

## 6.2.2 误差分析介绍

### 6.2.2.1 问题

**问题** 设有方程组  $\begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ , 解得  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

若有小误差  $\begin{pmatrix} 1 & 1 \\ 1 & 0.9999 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1.9999 \end{pmatrix}$ , 解得  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

注: 初始数据  $A, b$  的微小变化引起解的巨大变化。——病态方程组

### 6.2.2.2 分析

**扰动** 设  $Ax = b + \delta b$  加入扰动 有解  $\tilde{x} = x + \delta x$  误差. ( $A$  可逆,  $b \neq 0$ )

代入解则  $A(x + \delta x) = b + \delta b \Rightarrow Ax + A\delta x = b + \delta b \Rightarrow A\delta x = \delta b$

$\Rightarrow \delta x = A^{-1}\delta b \Rightarrow \|\delta x\| \leq \|A^{-1}\| \|\delta b\|$  误差扰动的倍数取决于  $A^{-1}$ , 改为相对误差:

另一方面:  $\|b\| = \|Ax\| \leq \|A\| \|x\|$  仅零矩阵时可能为零, 此时不可逆, 所以能除过来  $\Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$

得到  $\frac{\|\delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta b\|}{\|b\|}$  同理, 若  $(A + \delta A)(x + \delta x) = b$ , 可得  $\frac{\|\delta x\|}{\|x + \delta x\|} \leq \frac{\|A\| \|A^{-1}\| \frac{\|\delta A\|}{\|A\|}}{1 - \|A\| \|A^{-1}\| \frac{\|\delta A\|}{\|A\|}}$

注: 解的相对误差不超过初始数据相对误差的  $\|A\| \|A^{-1}\|$  倍, 即  $\|A\| \|A^{-1}\|$  刻画了方程组的形态

### 6.2.2.3 条件数

**定义 3** 设  $A$  非奇异,  $\|\cdot\|$  为算子范数, 则称  $\text{cond}(A) = \|A\| \|A^{-1}\|$  为矩阵  $A$  的条件数

$\|A\| \|A^{-1}\|$  刻画了解的稳定性

**注意** ①  $\text{Cond}(A) = \|A\| \|A^{-1}\| \geq \|AA^{-1}\| = \|I\| = 1 > 0$  严格大于 1

② 条件数的值与范数的类型有关  $\text{Cond}(A)_{\infty} = \|A\|_{\infty} \|A^{-1}\|_{\infty}$  (行模)

$$\text{cond}(A)_2 = \|A\|_2 \|A^{-1}\|_2 = \frac{\sqrt{\lambda_{\max}}}{\sqrt{\lambda_{\min}}} \quad (\text{谱模})$$

其中  $\lambda_{\max}, \lambda_{\min}$  分别是  $A^T A$  的最大, 最小模特征值

**定义 7** 设  $A$  非奇异, 若  $\text{Cond}(A) \gg 1$ , 则称  $Ax = b$  为病态方程组

已知条件数远大于未知条件数, 且相互矛盾 若  $\text{Cond}(A)$  相对较小, 则称  $Ax = b$  为良态方程组

**例题** 本节开始时的方程组中, 系数矩阵  $A = \begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix}$   $A^{-1} = \begin{pmatrix} 10001 & -10000 \\ -10000 & 10000 \end{pmatrix}$

$\therefore \text{Cond}(A)_{\infty} = \|A\|_{\infty} \|A^{-1}\|_{\infty} = 2.0001 \times 20001 \approx 40004$  所以  $Ax = b$  为病态

### 6.2.2.4 病态方程组

**现状** 理论上有条件数, 但标准很难定, 且  $\|A^{-1}\|$  极其难求

**判断方法** **直观判断:** ① 用主元消去法求解时出现小主元

② 某些行、列几乎线性相关 如  $A = \begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix}$

③  $A$  的元素间数量级很大, 且无规律

若方程组出现上述情况之一, 方程组有可能“病态”

**处置措施** 对于“病态”方程组的求解, 需要采用特殊处理或专用方法:

- ① 提高原始数据和运算的精度, 如原始数据和运算采用双精度等 (算力目前不值钱, 默认最高)
- ② 用适当方法改善原始模型的性态, 如对矩阵进行“预处理”以降低其条件数

**例题** 1. 设线性方程组  $\begin{pmatrix} 1 & 10^7 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10^7 \\ 2 \end{pmatrix}$ , 试计算  $\text{Cond}(A)_{\infty}$ , 并取3位有效数字求解

① 直接计算  $\text{Cond}(A)_{\infty}$ , 并求解.  $A^{-1} = \frac{-1}{10^7-1} \begin{pmatrix} 1 & -10^7 \\ -1 & 1 \end{pmatrix} = \frac{1}{10^7-1} \begin{pmatrix} -1 & 10^7 \\ 1 & 1 \end{pmatrix}$

$\text{Cond}(A)_{\infty} = \|A\|_{\infty} \|A^{-1}\|_{\infty} = \frac{(10^7+1)^2}{10^7-1} \approx 10^7$  使用列主元素消去法求解:

$\begin{pmatrix} 1 & 10^7 & 10^7 \\ 1 & 1 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 10^7 & 10^7 \\ 0 & -10^7 & -10^7 \end{pmatrix}$  回代:  $x_2 = 1$ , 代入  $x_1 + 10^7 x_2 = 10^7$ , 解得  $x_1 = 0$

② 对  $A$  做预处理。在方程组两边左乘矩阵  $\begin{pmatrix} 10^{-7} & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 10^{-7} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 10^7 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10^{-7} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 10^7 \\ 2 \end{pmatrix}$

得到新的矩阵  $\begin{pmatrix} 10^{-7} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  设  $\bar{A} = \begin{pmatrix} 10^{-7} & 1 \\ 1 & 1 \end{pmatrix}$ ,  $\bar{A}^{-1} = \frac{1}{1-10^{-7}} \begin{pmatrix} -1 & 1 \\ 1 & -10^{-7} \end{pmatrix}$

从而  $\text{Cond}(\bar{A})_\infty = \|\bar{A}\|_\infty \|\bar{A}^{-1}\|_\infty = \frac{2 \times 2}{1-10^{-7}} \approx 4$  是个很小的值

使用列主元素消去法求解:  $\begin{pmatrix} 10^{-7} & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 2 \\ 10^{-7} & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$

回代:  $x_2 = 1$ , 代入  $x_1 + x_2 = 2$ , 解得  $x_1 = 1$  得到准确结果

### 6.2.2.5 迭代改善法

**情况** 首先指出, 用近似解  $x'$  代入方程组  $Ax = b$  的左端, 观察  $Ax'$  与  $b$  是否近似, 即用观察残差向量  $r = b - Ax'$  的“大小”来判断  $x'$  是否可以接受的方法是不可靠的

例如用  $x' = (1, 1)^T$  计算方程组  $\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.0001x_2 = 2 \end{cases}$

的残差向量  $r = b - Ax' = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.0001 \end{bmatrix}$  非常“小”, 但  $x'$  与其精确解  $x = (2, 0)^T$

相差很大, 显然不能接受。下面介绍一种事后估计近似解相对误差的方法。

**定理 2** 设  $|A| \neq 0$  且  $b \neq 0$ ,  $x^*$  和  $x'$  分别为方程组  $Ax = b$  的精确解和近似解, 残差向量  $r = b - Ax'$ , 则有

$$\frac{\|x^* - x'\|}{\|x^*\|} \leq \text{Cond}(A) \frac{\|r\|}{\|b\|} \quad \text{相对误差不仅取决于残差误差, 还取决于条件数。所以残差判定不能用。}$$

### 改善方法

采用高精度的计算方法(如列主元素消去法)与高精度的算术运算(如双精度), 虽然可以提高方程组近似解的精确程度, 但是对于病态方程组也不一定能获得高精度的近似解, 而且方程组的病态越严重, 求解越困难。

当用某种方法求得方程组  $Ax = b$  ( $|A| \neq 0$  且  $b \neq 0$ ) 的某个近似解  $x^{(1)}$  后, 若尚未达到精度要求, 可采用下述改善迭代过程获得较精确的近似解。

① 用双精度计算残差向量  $r^{(1)} = b - Ax^{(1)}$

② 用列主元素消去法(或选主元素三角分解法)解方程组  $Ax = r^{(1)}$  得近似解  $d^{(1)}$

③ 用  $d^{(1)}$  修正  $x^{(1)}$  得  $Ax = b$  的新近似解  $x^{(2)} = x^{(1)} + d^{(1)}$

如果在计算  $r^{(1)}$  与  $d^{(1)}$  中没有误差, 则  $x^{(2)} = x^{(1)} + A^{-1}r^{(1)} = x^{(1)} + A^{-1}(b - Ax^{(1)}) = A^{-1}b = x^*$  即  $x^{(2)}$  为  $Ax = b$  的精确解。(但在实际计算中, 由于舍入影响,  $x^{(2)}$  仍为  $Ax = b$  的近似)

④ 计算  $e = \frac{\|d^{(1)}\|_\infty}{\|x^{(1)}\|_\infty}$ , 若  $e < \varepsilon$  (精度常数), 则取  $x^* \approx x^{(2)}$ ; 否则视  $x^{(2)}$  为  $x^{(1)}$ , 重新进行上述过程, 直到

满足条件  $e < \varepsilon$  为止。修正和原来的值比起来很小, 则修正无意义, 停止。

上述过程称为方程组  $Ax = b$  近似解的迭代改善。当方程组的病态不十分严重时, 通过迭代改善方法获得的近似解序列可以很快地收敛到方程组的精确解。

## 6.3 解线性方程组的迭代法

条件

设线性方程组  $Ax = b$ , 其中  $A \in R^{n \times n}$   $b, x \in R^n$

随着计算问题的日益复杂, 20 世纪 30 年代起, 上述系数矩阵  $A$  具有两个明显特点: **大型化与稀疏性**。  
大型化指其阶数可达上万阶, 稀疏性指  $A$  的零元素占绝大部分。

对这样的  $A$  作直接三角分解, 稀疏性会受到破坏。迭代法在这样的背景下得到关注和发展。

解法

将  $Ax = b$  改写为等价方程组:  $x = Bx + f$

$$\text{取初始向量: } x^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix} \quad \text{令 } x^{(1)} = Bx^{(0)} + f, x^{(2)} = Bx^{(1)} + f, \dots$$

$$\text{一般地令: } x^{(k+1)} = \begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{pmatrix} = B \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} + \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} = Bx^{(k)} + f \quad (k = 0, 1, 2, \dots, n)$$

称上式为求解线性方程组的迭代法(迭代过程, 迭代格式)

$B$  称为迭代矩阵, 若当  $k \rightarrow \infty$  时,  $x^{(k)} \rightarrow x^*$ , 则称该迭代法收敛。否则称迭代法发散

注意

①  $x^{(k)} \rightarrow x^*, (k \rightarrow \infty)$ , 记作  $\lim_{k \rightarrow \infty} x^{(k)} = x^*$  即  $\|x^{(k)} - x^*\| \rightarrow 0, (k \rightarrow \infty)$

② 两端取极限得  $x^* = Bx^* + f$ , 即  $x^*$  是两个方程的解。

### 6.3.1 简单迭代法

#### 6.3.1.1 雅可比迭代法

条件

$$\text{记 } Ax = b \text{ 为 } \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \quad \text{设 } A = (a_{ij})_{n \times n} \text{ 可逆, 且 } a_{ii} \neq 0 (i = 1, 2, \dots, n)$$

推导

$$\text{则有 } a_{ii}x_i = b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^n a_{ij}x_j \Rightarrow x_i = \frac{1}{a_{ii}}(b_i - \sum_{j=i}^n a_{ij}x_j) \quad (i = 1, 2, \dots, n)$$

$$\text{写成迭代格式: } x_i^{(k+1)} = \frac{1}{a_{ii}}(b_i - \sum_{j \neq i}^n a_{ij}x_j^{(k)}) \quad (i = 1, 2, \dots, n, k = 0, 1, 2, \dots) \quad \text{或}$$

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

称此为分量形式的雅可比迭代法(Jacobi)

矩阵形式

将  $a_{ii}x_i = b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^n a_{ij}x_j$  写成矩阵形式:

$$\begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} - \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & & \\ \dots & & \ddots & \\ a_{n1} & a_{n2} & \dots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & & \dots & a_{2n} \\ & & \dots & \\ & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$A = L_{\text{下方}} + D_{\text{对角}} + U_{\text{上方}} \quad (L + D + U)x = b \Rightarrow Dx = b - Lx - Ux$$

$$\Rightarrow Dx = Lx + Ux + b \quad \text{其中 } D - L - U = A \Rightarrow x = D^{-1}(L + U)x + D^{-1}b$$

$$\text{令 } B_J = D^{-1}(L + U), f_J = D^{-1}b, \text{ 则有 } x = B_Jx + f_J \quad \text{即有 Jacobi 迭代的矩阵形式: } x^{(k+1)} = B_Jx^{(k)} + f_J$$

例题

$$\text{将方程组 } \begin{cases} 8x_1 - 3x_2 + 2x_3 = 20 \\ 4x_1 + 11x_2 - x_3 = 33 \\ 2x_1 + x_2 + 4x_3 = 12 \end{cases} \text{ 写成雅可比迭代格式: } \begin{cases} x_1^{(k+1)} = \frac{1}{8}(20 + 3x_2^{(k)} - 2x_3^{(k)}) \\ x_2^{(k+1)} = \frac{1}{11}(33 - 4x_1^{(k)} + x_3^{(k)}) \\ x_3^{(k+1)} = \frac{1}{4}(12 - 2x_1^{(k)} - x_2^{(k)}) \end{cases}$$

$$\text{其矩阵形式: } \begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} 1/8 & & \\ & 1/11 & \\ & & 1/4 \end{pmatrix} \begin{pmatrix} 0 & 3 & -2 \\ -4 & 0 & 1 \\ -2 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} + \begin{pmatrix} 2.5 \\ 3 \\ 3 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0.375 & -0.25 \\ -0.3636 & 0 & 0.0909 \\ -0.5 & -0.25 & 0 \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} + \begin{pmatrix} 2.5 \\ 3 \\ 3 \end{pmatrix}.$$

取初始向量  $x^{(0)} = (0,0,0)^T$  进行迭代

0	0.375	-0.25	2.5	0	2.5	2.875	3.136364	3.024148
-0.36364	0	0.09091	3	0	3	2.363636	2.045455	1.947831
-0.5	-0.25	0	3	0	3	1	0.971591	0.920455

### 6.3.1.2 Gauss - Seidel 迭代

#### 内容

考虑 Jacobi 迭代分量式:

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

通常计算值  $x_j^{(k+1)}$  比前一步计算值  $x_j^{(k)}$  更精确, 但是雅可比迭代法后部没有更新。可以边算边更。

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

称上式为高斯—赛德尔迭代(分量式)。

考虑 Jacobi 迭代矩阵式:  $\Rightarrow x^{(k+1)} = D^{-1}(L + U)x^{(k)} + D^{-1}b \Rightarrow Dx^{(k+1)} = Lx^{(k)} + Ux^{(k)} + b$

取  $Dx^{(k+1)} = Lx^{(k+1)} + Ux^{(k)} + b \Rightarrow (D - L)x^{(k+1)} = Ux^{(k)} + b$

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

令  $B_G = (D - L)^{-1}U, f_G = (D - L)^{-1}b$ , 则有  $x^{(k+1)} = B_G x^{(k)} + f_G$ , 称(2.10)为 G—S 迭代的矩阵式

#### 例题

在上例  $\begin{cases} 8x_1 - 3x_2 + 2x_3 = 20 \\ 4x_1 + 11x_2 - x_3 = 33 \\ 2x_1 + x_2 + 4x_3 = 12 \end{cases}$  中使用 Gauss - Seidel 迭代

$$\text{迭代公式} \begin{cases} x_1^{(k+1)} = \frac{1}{8}(20 + 3x_2^{(k)} - 2x_3^{(k)}) \\ x_2^{(k+1)} = \frac{1}{11}(33 - 4x_1^{(k+1)} + x_3^{(k)}) \\ x_3^{(k+1)} = \frac{1}{4}(12 - 2x_1^{(k+1)} - x_2^{(k+1)}) \end{cases}$$

0	0	0
2.5	2.090909	1.227273
2.977273	2.028926	1.004132
3.009814	1.996807	0.995891
2.99983	1.999688	1.000163
2.999842	2.000072	1.000061
3.000012	2.000001	0.999994
3.000002	1.999999	0.999999
3	2	1
3	2	1

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{pmatrix} = \begin{pmatrix} a_{11}^{-1} & & & \\ & a_{22}^{-1} & & \\ & & \ddots & \\ & & & a_{nn}^{-1} \end{pmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} - \begin{pmatrix} 0 \\ a_{21}x_1^{(k)} \\ \vdots \\ \sum_{j=1}^{n-1} a_{nj}x_j^{(k)} \end{pmatrix} - \begin{pmatrix} \sum_{j=2}^n a_{1j}x_j^{(k)} \\ \sum_{j=3}^n a_{2j}x_j^{(k)} \\ \vdots \\ 0 \end{pmatrix}$$

### 6.3.2 迭代的收敛性

**准则** 迭代的求解过程为求极限过程, 这个极限过程的收敛性就是迭代法的收敛与发散问题.

**定理 2** 迭代法:  $x^{(k+1)} = Bx^{(k)} + f$  收敛  $\Leftrightarrow \rho(B) < 1$   
但是要检验一个矩阵的谱半径小于 1 比较困难, 所以我们希望用别的办法判断收敛性

**定理 3** 若  $\|B\| < 1$ , 则  $x^{(k+1)} = Bx^{(k)} + f$  收敛 且  $\|x^{(k+1)} - x^*\| < \frac{\|B\|}{1-\|B\|} \|x^{(k+1)} - x^{(k)}\|$

$$\|x^{(k+1)} - x^*\| \leq \frac{\|B\|^{k+1}}{1-\|B\|} \|x^{(1)} - x^{(0)}\| \quad \text{其中 } \|B\| \text{ 为 } B \text{ 的任一算子范数}$$

**证明** 因为  $\rho(B) \leq \|B\| < 1$ , 所以  $x^{(k+1)} = Bx^{(k)} + f$  收敛  
又因为  $x^* - x^{(k+1)} = B(x^* - x^{(k)}) = B(x^* - x^{(k+1)} + x^{(k+1)} - x^{(k)})$   
 $\Rightarrow \|x^* - x^{(k+1)}\| \leq \|B\| (\|x^* - x^{(k+1)}\| + \|x^{(k+1)} - x^{(k)}\|)$   
 $\Rightarrow \|x^{(k+1)} - x^*\| \leq \frac{\|B\|}{1-\|B\|} \|x^{(k+1)} - x^{(k)}\|$  又  $x^{(k+1)} - x^{(k)} = B(x^{(k)} - x^{(k-1)}) = \dots = B^k(x^{(1)} - x^{(0)})$   
 $\Rightarrow \|x^{(k+1)} - x^{(k)}\| \leq \|B\|^k \|x^{(1)} - x^{(0)}\|$   
 $\|x^{(k+1)} - x^*\| \leq \frac{\|B\|}{1-\|B\|} \|x^{(k+1)} - x^{(k)}\| \leq \frac{\|B\|^{k+1}}{1-\|B\|} \|x^{(1)} - x^{(0)}\|.$

**注意** 因为矩阵范数  $\|B\|_1, \|B\|_\infty$  都可以直接用矩阵的元素计算, 因此用定理容易判别迭代法的收敛性

**例题** 1. 设  $\begin{cases} x_1 + 2x_2 - 2x_3 = 1 \\ x_1 + x_2 + x_3 = 1 \\ 2x_1 + 2x_2 + x_3 = 1 \end{cases}$  问使用雅可比迭代和 G—S 迭代求解是否收敛.

$$\textcircled{1} B_J = D^{-1}(L + U) = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{pmatrix}.$$

行范数  $\|B_J\|_\infty = \max\{4, 2, 4\} = 4 > 1$ ,  $\|B_J\|_1 = 4 > 1$ , 考虑谱半径

$$|\lambda I - B_J| = \begin{vmatrix} \lambda & 2 & -2 \\ 1 & \lambda & 1 \\ 2 & 2 & \lambda \end{vmatrix} = \begin{vmatrix} \lambda & 2 & 0 \\ 1 & \lambda & 1 + \lambda \\ 2 & 2 & \lambda + 2 \end{vmatrix} = \lambda \begin{vmatrix} \lambda & 1 + \lambda \\ 2 & 2 + \lambda \end{vmatrix} - 2 \begin{vmatrix} 1 & 1 + \lambda \\ 2 & 2 + \lambda \end{vmatrix} = \lambda^3 = 0.$$

$\Rightarrow \rho(B_J) = 0 < 1$ , 所以雅可比迭代收敛。

$$\textcircled{2} B_G = (D - L)^{-1}U = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -2 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 0 & -2 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\text{考虑谱半径: } |\lambda I - B_G| = \begin{vmatrix} \lambda & 2 & -2 \\ 0 & \lambda - 2 & 3 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = \lambda(\lambda - 2)^2 \Rightarrow \rho(B_G) = 2 > 1, \text{ 发散.}$$

可能变化由收敛变为发散。

2. 设方程组  $Ax = b$ , 其中  $A = \begin{pmatrix} 10 & a & 0 \\ b & 10 & b \\ 0 & a & 5 \end{pmatrix}$ , 求 J—迭代法与 G—S 迭代法收敛的充要条件

$$\text{雅克比迭代矩阵: } B_J = \begin{pmatrix} 10 & & \\ & 10 & \\ & & 5 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -a & 0 \\ -b & 0 & -b \\ 0 & -a & 0 \end{pmatrix} = \begin{pmatrix} 0 & -a/10 & 0 \\ -b/10 & 0 & -b/10 \\ 0 & -a/5 & 0 \end{pmatrix}$$

$$|\lambda I - B_J| = \begin{vmatrix} \lambda & -a/10 & 0 \\ -b/10 & \lambda & -b/10 \\ 0 & -a/5 & \lambda \end{vmatrix} = \lambda(\lambda^2 - \frac{3ab}{100}) \quad \text{所以, 雅克比迭代收敛的充要条件是 } |ab| < 100/3$$



$$\text{GS 迭代矩阵: } B_G = \begin{pmatrix} 10 & & \\ b & 10 & \\ & a & 5 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -a & 0 \\ 0 & 0 & -b \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{10} & & \\ -\frac{b}{100} & \frac{1}{10} & \\ \frac{ab}{100} & -\frac{a}{50} & \frac{1}{5} \end{pmatrix} \begin{pmatrix} 0 & -a & 0 \\ 0 & 0 & -b \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{a}{10} & 0 \\ 0 & \frac{ab}{100} & -\frac{b}{10} \\ 0 & -\frac{a^2b}{100} & \frac{ab}{50} \end{pmatrix}$$

$$|\lambda I - B_G| = \begin{vmatrix} \lambda & \frac{a}{10} & \\ \lambda - \frac{ab}{100} & \lambda - \frac{1}{10} & \frac{b}{10} \\ \frac{a^2b}{100} & \lambda - \frac{ab}{50} & \lambda \end{vmatrix} = \lambda \left( \lambda^2 - \lambda \frac{3ab}{100} - \frac{a^2b^2}{1250} \right) \quad \lambda = 0, \lambda = \frac{3 \pm \sqrt{41}}{200} ab \quad \rho(B_G) = \frac{3 + \sqrt{41}}{200} |ab| < 1$$

$$\Rightarrow |ab| < \frac{200}{3 + \sqrt{41}}$$

### 6.3.3 一类特殊方程组

**定义** 设  $A = (a_{ij})_{n \times n}$  满足:  $|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}|$  ( $i = 1, 2, \dots, n$ ), 则称  $A$  为(行)对角占优阵

**注意** ① 若  $|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}|$  则称  $A$  为(列)对角占优阵

② 若(2.14)中严格不等式成立, 则称  $A$  为严格(行, 列)对角占优阵. 统称为严格占优阵

③ 可以证明严格占优阵必可逆, 即  $|A| \neq 0$

**定理** 设  $Ax = b$  若  $A$  为严格(行)对角占优, 则 **J—迭代和 G—S 迭代均收敛.**

**证明** ①  $B_J = D^{-1}(L + U) = - \begin{pmatrix} a_{11}^{-1} & & & \\ & a_{22}^{-1} & & \\ & & \ddots & \\ & & & a_{nn}^{-1} \end{pmatrix} \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{pmatrix} = B_J = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \cdots & -\frac{a_{2n}}{a_{22}} \\ \cdots & \cdots & \ddots & \cdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{pmatrix}$

$$\|B_J\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n \frac{|a_{ij}|}{|a_{ii}|} = \max_{1 \leq i \leq n} \frac{\sum_{j=1}^n |a_{ij}|}{|a_{ii}|} < 1$$

②  $B_G = (D - L)^{-1}U$  (欲证  $\rho(B_G) < 1$ ) 因为  $\lambda I - B_G = \lambda I - (D - L)^{-1}U = (D - L)^{-1}[\lambda(D - L) - U]$   
 $0 = |\lambda I - B_G| = |(D - L)^{-1}[\lambda(D - L) - U]| \Rightarrow |\lambda(D - L) - U| = 0$

$$\text{令 } G = \lambda(D - L) - U = \begin{pmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ \lambda a_{12} & \lambda a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn} \end{pmatrix} \quad \text{则 } |G| = 0.$$

若  $|\lambda| \geq 1$ , 则  $|\lambda a_{ii}| > \sum_{j \neq i} |\lambda a_{ij}| \geq \sum_{j=1}^{i-1} |\lambda a_{ij}| + \sum_{j=i+1}^n |a_{ij}| \Rightarrow G$  严格对角占优  $\Rightarrow G$  可逆, 矛盾

**注意** ① 若将方程组经初等行变换化为同解方程组, **其系数矩阵严格对角占优, 则可使迭代法收敛.**

$$\text{如方程组: } \begin{cases} 2x_1 - 1.8x_2 + 0.4x_3 = 1 \\ 3x_1 + 2x_2 - 1.1x_3 = 0.2 \\ x_1 - x_2 + 7.3x_3 = 1.6 \end{cases} \quad \begin{pmatrix} 2 & -1.8 & 0.4 \\ 3 & 2 & -1.1 \\ 1 & -1 & 7.3 \end{pmatrix} \xrightarrow[r_2 \times 5]{r_1 + r_2} \begin{pmatrix} 5 & 0.2 & -0.7 \\ 15 & 10 & -5.5 \\ 1 & -1 & 7.3 \end{pmatrix} \xrightarrow{r_2 - 3r_1} \begin{pmatrix} 5 & 0.2 & -0.7 \\ 0 & 9.4 & -3.4 \\ 1 & -1 & 7.3 \end{pmatrix}$$

此时, 矩阵化为严格对角占优阵

② 若  $\|B_J\|_{\infty} < 1$ , 则 G—S 迭代收敛

$$\text{因为: } B_J = \begin{pmatrix} 0 & \frac{a_{12}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 0 & \cdots & \frac{a_{2n}}{a_{22}} \\ \cdots & \cdots & \ddots & \cdots \\ \frac{a_{n1}}{a_{nn}} & \frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{pmatrix} \quad \text{所以 } \|B_J\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j \neq i} \frac{|a_{ij}|}{|a_{ii}|} \quad \text{由 } \|B_J\|_{\infty} < 1 \text{ 可知}$$

$\max_{1 \leq i \leq n} \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| < 1 \Rightarrow \sum_{j \neq i} |a_{ij}| < |a_{ii}|$  ( $i = 1, 2, \dots, n$ ). 即  $A$  严格对角占优, 两种迭代都收敛