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$\lim\limits_{k\to\infty}x_k=x^*$

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

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**$\boldsymbol{x}=\boldsymbol{Bx}+\boldsymbol{f}$ 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**

概述　范数是$\boldsymbol{R^n}$中**向量长度**概念的**直接推广**，用于对线性空间中元素大小进行度量。

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**：$\forall x\in V:\|x\|\ge 0$ 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**：$\forall k\in R,\forall x\in V:\|kx\|=|k|\cdot\|x\|$　　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**：$\forall x,y\in V:\|x+y\|\le\|x\|+\|y\|$　　**三角形中边长关系**

则称$\boldsymbol{N(x)=\|x\|}$ **为$\boldsymbol{V}$上向量$\boldsymbol{x}$的范数**

性质　① 若$x\ne 0$，$\left\|\frac{x}{\|x\|}\right\|=1$ 归一化　　② $\|-x\|=\|x\|$　　③ $\big|\|x\|-\|y\|\big|\le\|x-y\|$

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）**　$\|x\|_2=\sqrt{\sum_{i=1}^n|x_i|^2}=\sqrt{\sum_{i=1}^n\bar{x}_ix_i}$　　**模长**

② **最大模范数（∞-范数）**　$\|x\|_\infty=\max\limits_{1\le i\le n}|x_i|$　　**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**　$\boldsymbol{\|x\|_p=\sqrt[p]{\sum_{i=1}^n|x_i|^p}}$

注意
① 可以证明上述定义均满足向量的范数定义
② 显而易见地，**2-范数**和 **1-范数**都是**p范数**的特例，**∞-范数**也是**p范数**的特例 （$p\to\infty$，$\|x\|_p\to\|x\|_\infty$）
③ 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，$\forall A_{m\times n}$：$\|A\|_2=\sqrt{\sum_{i=1}^m\sum_{j=1}^n\left|a_{ij}\right|^2}\triangleq\|A\|_F$

但是，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 为简化起见，先考虑**方阵**$A_{n\times n}$的情况

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

① 正定 $\|A\| \ge 0$, 且 $\|A\| = 0 \Leftrightarrow A = 0$ ② 齐次 $\|kA\| = |k|\|A\|$

③ 三角 $\|A+B\| \le \|A\| + \|B\|$ **④ 相容 $\|AB\| \le \|A\|\|B\|$** 保证不超过单独作用之乘积

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(负值)

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$\|\boldsymbol{Ax}\|_{\text{向量范数}} \le \|\boldsymbol{A}\|_{\text{矩阵范数}}\|\boldsymbol{x}\|_{\text{向量范数}}$

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<u>为由向量范数诱导出的范数</u>。

定义 3 设$\|\cdot\|$为向量范数，$\|\cdot\|_M$为矩阵范数。 若$\forall A \in R^{n\times n}, \forall x \in R^n \Rightarrow$ $\|\boldsymbol{Ax}\| \le \|\boldsymbol{A}\|_{\boldsymbol{M}}\|\boldsymbol{x}\|$

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，有反例$A = \begin{pmatrix} 1 & 4 \\ 1 & 4 \end{pmatrix}$ $x = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

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

性质 ① $\|A\| = \max\limits_{\|x\|=1} \|Ax\|$

② 可以证明算子范数满足矩阵范数的 4 个条件.

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。($\|A\| = \max \frac{\|Ax\|}{\|x\|}$，$\|Ax\| = \frac{\|Ax\|}{\|x\|}\|x\| \le \|A\|\|x\|$)

常用范数

① 行范数 $\|\boldsymbol{A}\|_\infty = \max\limits_{\|x\|=1} \|Ax\|_\infty = \max\limits_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$ 行范数(**每行绝对相加，取最大**)

② 列范数 $\|\boldsymbol{A}\|_1 = \max\limits_{\|x\|=1} \|Ax\|_1 = \max\limits_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$ 列范数(**每列绝对相加，取最大**)

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

其中$\lambda_{max}(A^TA)$为矩阵A^TA的绝对值最大的特征值。**特征值计算：$|\boldsymbol{\lambda I} - \boldsymbol{A^TA}| = 0$**

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

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

$A^TA = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 9 & -1 \\ 1 & -1 & 2 \end{pmatrix} \Rightarrow |\lambda I - A^TA| = \begin{vmatrix} \lambda-2 & 0 & -1 \\ 0 & \lambda-9 & 1 \\ -1 & 1 & \lambda-2 \end{vmatrix} = 0$ 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$\boldsymbol{\rho(A)} = \max\limits_{1\le i\le n}|\boldsymbol{\lambda_i}|$为$A$的谱半径

$\rho(A) \le \|A\|_p$ **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$\|(I+A)^{-1}\| \le \frac{1}{1-\|A\|}$

证明 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$1 = \|I\| = \|(I+A)D\| = \|D + AD\| \ge \|D\| - \|AD\| \ge \|D\| - \|A\|\|D\|$

$= \|D\|(1 - \|A\|)$ $\because 1 - \|A\| > 0$ $\therefore \|D\| \le \frac{1}{1-\|A\|}$ 即：$\|(I+A)^{-1}\| \le \frac{1}{1-\|A\|}$

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$\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}1\\ 1\end{pmatrix}$

由该例可见，初始数据A,b的微小变化引起解的巨大变化，我们称这类方程组为**病态方程组**

6.2.2.2 分析

扰动 设$Ax=b+\delta b_{\text{扰动}}$有解 $\tilde{x}=x+\delta x_{\text{误差}}$ （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:$\|b\|=\|Ax\|\leq\|A\|\|x\|_{\text{仅零矩阵时可能为零，此时不可逆，所以能除过来}} \Rightarrow \frac{1}{\|x\|}\leq\frac{\|A\|}{\|b\|}$

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$\mathbf{cond}(A)=\|A\|\|A^{-1}\|$ **为矩阵A的条件数**

$\|A\|\|A^{-1}\|$ 能描述解的稳定性

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

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

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

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

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

若$Cond(A)$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 $\because A=\begin{pmatrix}1 & 1\\ 1 & 1.0001\end{pmatrix}$ $A^{-1}=\begin{pmatrix}10001 & -10000\\ -10000 & 10000\end{pmatrix}$

$\therefore \mathrm{Cond}(A)_\infty=\|A\|_\infty\|A^{-1}\|_\infty=2.0001\times 20001\approx 40004$ 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**数量级差异很大**，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$\mathrm{Cond}(A)_\infty$，并取3位有效数字求解

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

$Cond(A)_\infty=\|A\|_\infty\|A^{-1}\|_\infty=\frac{(10^7+1)^2}{10^7-1}\approx 10^7$ 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$x_1+10^7x_2=10^7$,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 $Cond(\bar{A})_\infty=\|\bar{A}\|_\infty\|\bar{A}^{-1}\|_\infty=\frac{2\times2}{1-10^{-7}}\approx4$ 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$x'=\begin{pmatrix}1\\1\end{pmatrix}$ 代入方程组$\begin{cases}x_1+x_2=2\\x_1+1.0001x_2=2\end{cases}$。有残差向量：$r=b-Ax'=\begin{bmatrix}2\\2\end{bmatrix}-\begin{bmatrix}1 & 1\\1 & 1.0001\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}$

$=\begin{bmatrix}0\\-0.0001\end{bmatrix}$非常小，但$x'$与其精确解$x=\begin{pmatrix}2\\0\end{pmatrix}$相差很大，显然不能接受。

有必要引入一种事后估计近似解相对误差的可靠方法。

定理 2 设$|A|\neq0$ 且$b\neq0$，$x^*_{\text{精确解}}$和$x'_{\text{近似解}}$分别为方程组$Ax=b$的解，残差向量$r=b-Ax'$,则有

$\frac{\|x^*-x'\|}{\|x^*\|}\le\mathrm{Cond}(A)\frac{\|r\|}{\|b\|}$ 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$Ax=b$ ($|A|\neq0$ 且$b\neq0$)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$x^{(2)}=x^{(1)}+A^{-1}r^{(1)}=x^{(1)}+A^{-1}\left(b-Ax^{(1)}\right)=A^{-1}b=x^*$

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：$\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f}$

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

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

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$k \to \infty$时，$x^{(k)} \to x^*$，则称该迭代法收敛，否则称迭代法发散

注意 ① $x^{(k)} \to x^*, (k \to \infty)$,记作 $\lim\limits_{k\to\infty} x^{(k)} = x^*$ 即$\left\|x^{(k)} - x^*\right\| \to 0$ $(k \to \infty)$

② 两端取极限可得$x^* = Bx^* + f$，即x^*同时是两个方程的解

6.3.1 简单迭代法

6.3.1.1 雅可比迭代法

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

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

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

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

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

两边同时左乘$\boldsymbol{D^{-1}}$ $\Rightarrow x = -\boldsymbol{D^{-1}}(\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x} + \boldsymbol{D^{-1}b}$

令 $\boldsymbol{B_J = -D^{-1}(L + U)}$ $\boldsymbol{f_J = D^{-1}b}$，则有$x = B_Jx + f_J$

即有雅可比迭代的矩阵形式：$x^{(k+1)} = B_Jx^{(k)} + f_J$

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

其矩阵形式：$\begin{pmatrix}x_1^{(k+1)}\\x_2^{(k+1)}\\x_3^{(k+1)}\end{pmatrix}=\begin{pmatrix}1/8&&\\&1/11&\\&&1/4\end{pmatrix}\begin{pmatrix}0&3&-2\\-4&0&1\\-2&-1&0\end{pmatrix}\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\\x_3^{(k)}\end{pmatrix}+\begin{pmatrix}2.5\\3\\3\end{pmatrix}=\begin{pmatrix}0&0.375&-0.25\\-0.3636&0&0.0909\\-0.5&-0.25&0\end{pmatrix}\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\\x_3^{(k)}\end{pmatrix}+\begin{pmatrix}2.5\\3\\3\end{pmatrix}$.

若取初始向量$x^{(0)}=(0,0,0)^{\mathrm{T}}$进行迭代，迭代结果如下表：

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 高斯-赛德尔迭代（G-S 迭代）

内容 考虑雅可比迭代分量式：

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

通常**计算值$x_j^{(k+1)}$**比前一步计算值$x_j^{(k)}$更精确，但是雅可比迭代法算后面时没有更新前面新算的结果。

已知在算$x_i^{(k+1)}$时，前序的**从1到$i-1$的$x_j^{(k+1)}$已知，**考虑一种可以边算边更的新方法：

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

称上式为**高斯-赛德尔迭代(Gauss - Seidel) (分量式)**

矩阵形式 考虑雅可比迭代矩阵式：$Dx^{(k+1)}=-Lx^{(k)}-Ux^{(k)}+b$

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

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

令$\boldsymbol{B_G=-(D+L)^{-1}U \quad f_G=(D+L)^{-1}b}$，则有$x^{(k+1)}=B_Gx^{(k)}+f_G$，称为 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

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

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：$\boldsymbol{x^{(k+1)}=Bx^{(k)}+f}$ **收敛** $\Leftrightarrow \boldsymbol{\rho(B)<1}$

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 $\|x^{(k+1)}-x^*\|<\frac{\|B\|}{1-\|B\|}\|x^{(k+1)}-x^{(k)}\|$

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

证明 因为$\rho(B)\le\|B\|<1$，所以$x^{(k+1)}=Bx^{(k)}+f$ 收敛

又因为$x^*-x^{(k+1)}=B\left(x^*-x^{(k)}\right)=B\left(x^*-x^{(k+1)}+x^{(k+1)}-x^{(k)}\right)$

$\Rightarrow \|x^*-x^{(k+1)}\|\le\|B\|\left(\|x^*-x^{(k+1)}\|+\|x^{(k+1)}-x^{(k)}\|\right)$

$\Rightarrow \left\|x^{(k+1)} - x^*\right\| \le \frac{\|B\|}{1-\|B\|}\left\|x^{(k+1)} - x^{(k)}\right\|$ 又 $x^{(k+1)} - x^{(k)} = B\left(x^{(k)} - x^{(k-1)}\right) = \ldots = B^k\left(x^{(1)} - x^{(0)}\right)$

$\Rightarrow \left\|x^{(k+1)} - x^{(k)}\right\| \le \|B\|^k\left\|x^{(1)} - x^{(0)}\right\|$

$\left\|x^{(k+1)} - x^*\right\| \le \frac{\|B\|}{1-\|B\|}\left\|x^{(k+1)} - x^{(k)}\right\| \le \frac{\|B\|^{k+1}}{1-\|B\|}\left\|x^{(1)} - x^{(0)}\right\|$

注意 因为**矩阵范数**$\|\boldsymbol{B}\|_1, \|\boldsymbol{B}\|_\infty$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

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$\left\|B_J\right\|_\infty = max\{4,2,4\} = 4 > 1$，$\left\|B_J\right\|_1 = 4 > 1$无法简单判断。故考虑谱半径

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

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

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

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

由此例，可见改变简单迭代为 G-S 迭代，可能使收敛变为发散

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

根据 $\rho(B) < 1$ 求解，其中$\rho(B) = \max\limits_{1\le i\le n}|\lambda_i|$

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

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

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

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

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

6.3.3 一类特殊方程组

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

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

② 若上述定义式中严格不等式成立，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，则雅可比迭代和 G-S 迭代均收敛。**

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

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

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

求特征值: $0=|\lambda I-B_G|=|(D-L)^{-1}||\lambda(D-L)+U|\Rightarrow|\lambda(D-L)+U|=0$

令$G=\lambda(D-L)+U=\begin{pmatrix}\lambda a_{11} & a_{12} & \cdots & a_{1n}\\ \lambda a_{12} & \lambda a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn}\end{pmatrix}$ 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**，**使其系数矩阵严格对角占优，就可使迭代法收敛。**

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

② 若$\|B_J\|_\infty<1$, 则 G-S 迭代收敛

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

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

6.3.4 超松弛迭代法

描述 超松弛迭代法(Successive Over Relaxation Method)是高斯-赛德尔迭代法的改进，是解大型稀疏矩阵线性方程组的有效方法

方法 将 G-S 迭代法改写为$x_i^{(k+1)} = x_i^{(k)} + \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)}\right)$ $(i = 1,2,\cdots,n)$

增加一个因子ω，希望如下迭代过程：

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

收敛的更快。这种迭代法方法称为**超松弛迭代法**，简称为 **SOR 方法**，其中ω称为**松弛因子**。

其实质上是对后一项做一个矫正。

矩阵表达 由上式得

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

用矩阵形式表示为 $Dx^{(k+1)} = (1-\omega)Dx^{(k)} + \omega\left(b - Lx^{(k+1)} - Ux^{(k)}\right)$

从而有：$(D+\omega L)x^{(k+1)} = [(1-\omega)D - \omega U]x^{(k)} + \omega b$

即：$x^{(k+1)} = (D+\omega L)^{-1}[(1-\omega)D - \omega U]x^{(k)} + \omega(D+\omega L)^{-1}b$

于是 SOR 方法的矩阵表达形式为 $x^{(k+1)} = B_\omega x^{(k)} + f_\omega$

其中：$\boldsymbol{B_\omega = (D+\omega L)^{-1}[(1-\omega)D - \omega U]}$ 称为**超松弛迭代矩阵**，$\boldsymbol{f_\omega = \omega(D+\omega L)^{-1}b}$

附：原 G-S 迭代 $B_G = -(D+L)^{-1}U$ $f_G = (D+L)^{-1}b$

例题 用 SOR 方法求下面线性方程组的近似解：$\begin{cases} 4x - 2x - 4x = 10 \\ -2x + 17x + 10x = 3 \\ -4x + 10x + 9x = -7 \end{cases}$

（我们已知其精确解为$x^* = (2,1,-1)^{\mathrm{T}}$）用 SOR 求解，迭代公式为：

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \frac{\omega}{4}\left(10 - 4x_1^{(k)} + 2x_2^{(k)} + 4x_3^{(k)}\right) \\ x_{21}^{(k+1)} = x_2^{(k)} + \frac{\omega}{17}\left(3 + 2x_1^{(k+1)} - 17x_2^{(k)} - 10x_3^{(k)}\right) \qquad (k = 0,1,2,\cdots) \\ x_{13}^{(k+1)} = x_3^{(k)} + \frac{\omega}{9}\left(-7 + 4x_1^{(k+1)} - 10x_2^{(k+1)} - 9x_3^{(k)}\right) \end{cases}$$

取$\boldsymbol{\omega = 1.43}$ $, x^{(0)} = (0,0,0)^{\mathrm{T}}$,则计算结果为$(2,1,-1)^{\mathrm{T}}$，

若取$\omega = 1$，即退化为 G-S 迭代法,仍取$x^{(0)} = (0,0,0)^{\mathrm{T}}$，则计算同样精度的结果，需要迭代 110 次以上

注意 SOR 方法也是一阶定常迭代法，故同样可以用 6.3.2 中相关定理讨论其收敛性。

但由上例引起我们注意的是，**改变松弛因子的值，不仅影响迭代过程的收敛速度，而且还会影响迭代过程的收敛性**。关于这点，引入如下定理：

定理 若线性方程组$Ax = b$的系数矩阵$A = (a_{ij})_{n\times n}$的主对角元素a_{ii}均不为零，则 SOR 法的迭代过程收敛的

必要条件是$\boldsymbol{0 < \omega < 2}$ （通俗讲就是加速度不能太猛）