# 第六章解线性代数方程组的迭代法

内容提要

6.1 引言

6.2 基本迭代法

6.3 迭代法的收敛性

## 6.1 引言

考虑线性方程组

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

即$Ax=b$ 其中$A$为非奇异矩阵，当$A$为低维稠密矩阵时，线性方程组用直接法(如高斯消去法和三角分解法)是有效的，但对于由工程技术中产生的大型稀疏矩阵方程组（$A$的维数$n$很大，但零元素较多），利用迭代法求解是适合的。在计算机内存和运算两方面，迭代通常都可利用$A$中有大量零元素的特点。

本章将介绍迭代法的一般理论及雅可比迭代法、高斯—塞德尔迭代法、超松弛迭代法，研究它们的收敛性。

**例6-1** 求解线性方程组

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

解：方程记为 $Ax = b$ ，即

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

改写原方程为

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

或改写为 $x = Bx + f$，其中

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

方程为 $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & \frac{3}{8} & -\frac{2}{8} \\ -\frac{4}{11} & 0 & \frac{1}{11} \\ -\frac{6}{12} & -\frac{3}{12} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \frac{20}{8} \\ \frac{33}{11} \\ \frac{36}{12} \end{bmatrix}.$$

构造迭代格式：$\begin{cases} x_1^{(k+1)} = (3x_2^{(k)} - 2x_3^{(k)} + 20)/8, \\ x_2^{(k+1)} = (-4x_1^{(k)} + x_3^{(k)} + 33)/11, \\ x_3^{(k+1)} = (-6x_1^{(k)} - 3x_2^{(k)} + 36)/12. \end{cases}$

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

任取初值$x = (0,0,0)^{\mathrm{T}}$带入，得$x^{(1)} = (2.5, 3, 3)^{\mathrm{T}}$

反复迭代得，$x^{(10)}$=$(3.000\,032, 1.999\,838, 0.999\,881\,3)^{\mathrm{T}}$,

此时，精确解 $x^* = (3,2,1)^{\mathrm{T}}$

$\left\|\varepsilon^{(10)}\right\|_\infty = 0.000\,162$，其中 $\varepsilon^{(10)} = x^{(10)} - x^*.$

无穷最大

从此例可以看出，由迭代法产生的向量序列逐步逼近此方程组的精确解。对于任何一个线性方程组$x$=$Bx$+$f$ （由 $Ax = b$ 变形得到的等价线性方程组），由迭代法产生的向量序列$x^{(k)}$ 是否一定逐步逼近此方程组的解$x^*$ 呢？回答是不一定。

例如：线性方程组

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

解：改写为 $x^{(k+1)} = Bx^{(k)} + f$ 形式

$$\begin{cases} x_1^{(k+1)} = 2x_2^{(k)} + 5 \\ x_2^{(k+1)} = 3x_1^{(k)} + 5 \end{cases}$$

任取初值$x = (0, 0)^{\mathrm{T}}$带入，得$x^{(1)} = (5, 5)^{\mathrm{T}}$，$x^{(2)} = (15, 20)^{\mathrm{T}}$

$x^{(3)} = (35, 65)^{\mathrm{T}}$……。迭代不到精确解$x = (-3, -4)^{\mathrm{T}}$。

故迭代法的收敛性研究是非常重要的。

迭代法概念：

一般地，由$Ax = b$变形得到等价的$x = Bx + f$.

设有唯一解$x*$, 则

$$x* = Bx* + f$$

又设$x^{(0)}$为任取的初始向量，按下述公式构造向量序列

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

其中$k$表迭代次数（迭代索引iteration index）。

**定义6-1** (1)对于方程组$x = Bx + f$，用公式(*)逐步代入求近似解的方法称为**迭代法**（一阶线性定常迭代法）.

(2)若$\lim\limits_{k\to\infty} x^{(k)}$存在(记为$x^*$)，则称此**迭代法收敛**，显然$x^*$是方程组的解，否则迭代法发散.

引进误差向量$\varepsilon^{(k+1)} = x^{(k+1)} - x^*$，则由(*)式减去$x^* = Bx^* + f$得

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

收敛： $\lim\limits_{k\to\infty}\varepsilon^{(k)} = 0 \Leftrightarrow \lim\limits_{k\to\infty} B^k = \mathbf{0}$.

要研究$B$满足什么条件下$B^k \to \mathbf{0}$ (当$k \to \infty$).

课堂推导

>1发散。=1呢？

# 定理

对任意初始向量$x^{(0)}$, 上述迭代法收敛$\Leftrightarrow \rho(B)<1$.

证：若$\|B^k\|\to 0$, 则$\rho^k(B)=\rho(B^k)\le\|B^k\|\to 0$, 所以$\rho(B)<1$.

若$\rho(B)<1$, 则存在$\varepsilon>0$, 使得$\rho(B)+\varepsilon<1$.

则$\|B^k\|\le\|B\|^k\le(\rho(B)+\varepsilon)^k\to 0$.

此证明不做要求

设$n$维矩阵$\mathbf{A}$的$n$个特征值为$\lambda_1, \lambda_2, \ldots, \lambda_n$, 则称

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

为矩阵A的谱半径. 对矩阵的任何一种相容范数都有

$$\rho(A)\le\|A\|$$

另外, $\forall\varepsilon>0$, $\exists$一种相容范数, 使 $\|A\|\le\rho(A)+\varepsilon$

## 6.2 基本迭代

考虑线性方程组

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

就是 $Ax = b$ 进行矩阵分裂 $A = M - N$

且$M$为可选择的非奇异矩阵，且使得$Mx = d$ 容易求解，于是

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

推导

可得定常迭代法

$$\begin{cases} \text{取初始向量} x^{(0)}, \\ x^{(k+1)} = Bx^{(k)} + f,\ k = 0,1,\cdots, \end{cases}$$

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

取$M$为对角阵，其逆是倒数而已

# 一、雅可比迭代法

$$
\begin{cases}
x_1 = \dfrac{1}{a_{11}}(b_1 - a_{12}x_2 - \cdots - a_{1n}x_n) \\
\cdots \quad \cdots \quad \cdots \\
\boxed{x_i = \dfrac{1}{a_{ii}}\left(b_i - \sum\limits_{j=1}^{i-1} a_{ij}x_j - \sum\limits_{j=i+1}^{n} a_{ij}x_j\right)} \\
\cdots \quad \cdots \quad \cdots \\
x_n = \dfrac{1}{a_{nn}}(b_n - a_{n1}x_1 - \cdots - a_{n,n-1}x_{n-1})
\end{cases}
$$

可以得到计算公式（雅可比迭代法），对$k = 0,1,\cdots$

$$
\begin{cases}
x^{(0)} = (x_1^{(0)}, \cdots, x_n^{(0)})^{\mathrm{T}}, \\
x_i^{(k+1)} = \left(b_i - \sum\limits_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j^{(k)}\right) / a_{ii}, \quad (i = 1, \cdots, n)
\end{cases}
$$

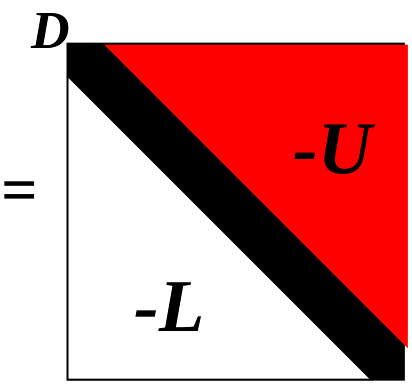
采用矩阵分裂记号

$$A = D - L - U,$$

其中

$$D = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix},\ L = \begin{bmatrix} 0 & & & \\ -a_{21} & 0 & & \\ \vdots & \ddots & \ddots & \\ -a_{n1} & \cdots & -a_{n,n-1} & 0 \end{bmatrix},$$

$$U = \begin{bmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ & 0 & \ddots & \vdots \\ & & \ddots & -a_{n-1,n} \\ & & & 0 \end{bmatrix}.$$

$A =$ 


雅可比迭代法的矩阵表示形式为

$$x^{(k+1)} = \boxed{D^{-1}(L+U)}x^{(k)} + D^{-1}b \stackrel{\Delta}{=} B_{\mathrm{J}}x^{(k)} + f$$

推导

## 二、高斯—塞德尔迭代法（ Gauss-Seidel迭代法）

学用思想

在Jacobi迭代中，使用最新计算出的分量值

$$
\begin{cases}
x_1^{(k+1)} = \dfrac{-1}{a_{11}}(a_{12}x_2^{(k)} + \cdots + a_{1n}x_n^{(k)} - b_1) \\
x_2^{(k+1)} = \dfrac{-1}{a_{22}}(\boxed{a_{21}x_1^{(k+1)}} + a_{23}x_3^{(k)} + \cdots + a_{1n}x_n^{(k)} - b_2) \\
\quad\vdots \\
x_n^{(k+1)} = \dfrac{-1}{a_{nn}}(\boxed{a_{n1}x_1^{(k+1)} + \cdots + a_{n\,n-1}x_{n-1}^{(k+1)}} - b_n)
\end{cases}
$$

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

## 二、高斯—塞德尔迭代法

Gauss 1823

Seidel 1874

$$\begin{cases} x^{(0)} = (x_1^{(0)}, \cdots, x_n^{(0)})^{\mathrm{T}}, \\ x_i^{(k+1)} = \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right) / a_{ii}, \quad (i = 1, \cdots, n) \end{cases}$$

称为高斯-赛德尔迭代法。

采用矩阵$A$ 的分裂记号，迭代法等价于

$$Dx^{(k+1)} = Lx^{(k+1)} + Ux^{(k)} + b$$

高斯赛德尔迭代法的矩阵表示形式为

$$x^{(k+1)} = (D-L)^{-1}Ux^{(k)} + (D-L)^{-1}b \stackrel{\Delta}{=} B_{\mathrm{G}}x^{(k)} + f$$

例6-2　解例6-1中的线性方程组，取初值$x^{(0)} = (0,0,0)^{\mathrm{T}}$

例6-1 求解线性方程组

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

J迭代格式：

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

G-S迭代格式：

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

$$x^{(7)}=(3.000\ 002,\ 1.999\ 998\ 7,\ 0.999\ 993\ 2)^{\mathrm{T}},$$

$$\left\|\varepsilon^{(7)}\right\|_{\infty}=\left\|x^*-x^{(7)}\right\|_{\infty}<6.8\times10^{-6}.$$

无穷最大

验算一下

上例计算结果表明，Gauss-Seidel迭代法比Jacobi迭代法效果好。事实上，对有些问题Gauss-Seidel迭代法确实比Jacobi迭代法收敛得快（有时快1倍），但也有Gauss-Seidel迭代比Jacobi迭代收敛得慢，甚至还有Jacobi迭代收敛，Gauss-Seidel迭代发散的情形。

评价：与Jacobi相比，只需一组工作单元存放近似解。

# 注意的问题

**（1）Jacobi迭代法和Gauss-Seidel迭代法的迭代矩阵不同：**

$$\boldsymbol{B_J = D^{-1}(L+U),\ B_G = (D-L)^{-1}U}$$

**（2）Jacobi迭代法和Gauss-Seidel迭代法的收敛性没有必然的联系：**

**即当Gauss-Seidel法收敛时，Jacobi法可能不收敛；**

**而Jacobi法收敛时， Gauss-Seidel法也可能不收敛。**

例如 $\begin{cases} 2x_1 - x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 1 \\ x_1 + x_2 - 2x_3 = 1 \end{cases}$

**用Jacobi迭代法求解不收敛，但用 G-S法收敛。**

又如 $\begin{cases} x_1 + 2x_2 - 2x_3 = 1 \\ x_1 + x_2 + x_3 = 2 \\ 2x_1 + 2x_2 + x_3 = 3 \end{cases}$

**用Jacobi迭代法求解收敛，但用 G-S法不收敛。**

$$B_J = \begin{pmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{pmatrix}, \quad B_G = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{pmatrix}$$

$B_J$的特征值为 0，0，0;
而$B_G$的特征值为 0，2，2。

Matlab用vrho()求谱半径

再例：用Gauss-Seidel 迭代法解方程组 $Ax$=$b$

$$A=\begin{pmatrix} 9 & -1 & -1 \\ -1 & 8 & 0 \\ -1 & 0 & 9 \end{pmatrix},\ b=\begin{pmatrix} 7 \\ 7 \\ 8 \end{pmatrix}$$

迭代公式: 给初始向量 $x^{(0)}$

$$\begin{cases} x_1^{(k+1)} = 1/9(x_2^{(k)} + x_3^{(k)} + 7) \\ x_1^{(k+1)} = 1/8(x_1^{(k+1)} + 7) \\ x_1^{(k+1)} = 1/9(x_1^{(k+1)} + 8) \end{cases} \qquad k = 0,1,\cdots$$

计算结果: $x^{(1)} = (0.7778, 0.9722, 0.9753)'$

$$x^{(2)} = (0.9942, 0.9993, 0.9994)'$$

$$x^{(3)} = (0.9999, 0.9999, 0.9999)'$$

$$x^{(4)} = (1.0000, 1.0000, 1.0000)'$$

G –S 迭代法每次迭代只需计算一次矩阵与向量的乘法。 G –S 迭代法比 Jacobi 迭代法有一个明显的优点就是计算时仅需一组工作单位用来保存

$x^{(k)}$ 分量(或 $x^{(k+1)}$ 分量）。当计算出 $x_i^{(k+1)}$ 就冲掉旧分量 $x_i^{(k)}$ 。从 G- S 迭代公式可以看出在

$x^{(k)} \rightarrow x^{(k+1)}$ 的一步迭代中，计算分量 $x_i^{(k+1)}$ 时利用了已经计算出的最新分量 。因此，G – S 迭代法可看作是Jacobi 迭代法的一种修正。

反复讲，也可略

反复讲，重要，学用思想

| Jacobi iteration | Gauss-Seidel iteration |
| --- | --- |
| 计算$x^{(k+1)}$时需要$x^{(k)}$的所有分量，因此需开两组存储单元分别存放$x^{(k)}$和$x^{(k+1)}$ | 计算$x_i^{(k+1)}$时只需要$x^{(k)}$的$i$+1~$n$个分量，因此$x^{(k+1)}$的前$i$个分量可存贮在$x^{(k)}$的前$i$个分量所占的存储单元，无需开两组存储单元. |

高斯–塞德尔迭代法又等价于：

$$x_i^{(k+1)} = x_i^{(k)} + \Delta x_i \text{，其中}$$

推导

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

若修正项中加入松弛因子$\omega$ ，则 $x_i^{(k+1)} = x_i^{(k)} + \omega \Delta x_i$ (其中$\omega = 1$) 此时引入松弛迭代法。

## 三、逐次超松驰(Successive Over-Relaxation )迭代法

逐次超松驰迭代（Successive Over-Relaxation, SOR）, 简称SOR方法是G-S迭代法的一种加速方法，是解大型稀疏矩阵方程组的有效方法之一，它有着较为广泛的应用。

# 三、逐次超松驰(SOR)迭代法

SOR迭代法的计算公式：对$k$=0, 1, ⋯,

$$\begin{cases} x^{(0)} = (x_1^{(0)}, \cdots, x_n^{(0)})^{\mathrm{T}}, \\ x_i^{(k+1)} = x_i^{(k)} + \omega\left( b_i - \sum\limits_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum\limits_{j=i}^{n} a_{ij} x_j^{(k)} \right) / a_{ii}, \\ (i = 1, 2, \cdots, n), \\ \text{松弛因子}\ \omega > 0. \end{cases}$$

采用矩阵$A$ 的分裂记号，化为

$$Dx^{(k+1)} = Dx^{(k)} + \omega(b + Lx^{(k+1)} + Ux^{(k)} - Dx^{(k)})$$

推导

SOR迭代法的矩阵表示形式为

$$x^{(k+1)} = \underline{(D - \omega L)^{-1}\{(1-\omega)D + \omega U\}} x^{(k)} + \omega(D - \omega L)^{-1} b.$$

**说明:** 1) $\omega$=1, 即为G-S（高斯-赛德尔）迭代法;

2) 2>$\omega$>1，称为超松驰法;

$\omega$<1，称为低松驰法;

3) SOR方法每迭代一次主要运算量相当于计算一次矩阵与向量的乘法。

注：迭代法求解线性方程组在计算机实现时可用

无穷最大

解间差 $\left\| x^{(k+1)} - x^{(k)} \right\|_\infty = \max\limits_{1 \le i \le n} \left| x_i^{(k+1)} - x_i^{(k)} \right| < \varepsilon$

控制迭代终止，或用 $\left\| r^{(k+1)} \right\|_\infty = \left\| b - Ax^{(k+1)} \right\|_\infty < \varepsilon$ 控制迭代终止。

残差

推导

例6-3　用SOR迭代法解线性代数方程组

$$\begin{bmatrix} -4 & 1 & 1 & 1 \\ 1 & -4 & 1 & 1 \\ 1 & 1 & -4 & 1 \\ 1 & 1 & 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix}$$

解: 取$x^{(0)}=0$，迭代公式为

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} - \omega(1+4x_1^{(k)} - x_2^{(k)} - x_3^{(k)} - x_4^{(k)})/4 \\ x_2^{(k+1)} = x_2^{(k)} - \omega(1-x_1^{(k+1)} + 4x_2^{(k)} - x_3^{(k)} - x_4^{(k)})/4 \\ x_3^{(k+1)} = x_3^{(k)} - \omega(1-x_1^{(k+1)} - x_2^{(k+1)} + 4x_3^{(k)} - x_4^{(k)})/4 \\ x_4^{(k+1)} = x_4^{(k)} - \omega(1-x_1^{(k+1)} - x_2^{(k+1)} - x_3^{(k+1)} + 4x_4^{(k)})/4 \end{cases}$$

取$\omega=1.3$, 第11次迭代结果为

$$x^{(11)} = (-0.999\ 996\ 46,\ -1.000\ 003\ 10,\ -0.999\ 999\ 53,\ -0.999\ 999\ 12)^{\mathrm{T}}$$

此时，$\left\|\varepsilon^{(11)}\right\|_2 = 0.46\times10^{-5}$

向量2范数：平方和开根号

(1) 当取松弛因子 $\omega = 1.3$ 时，迭代次数$k$=11.

(2) 当取$\omega = 1.0$ 时，初始向量相同，达到同样精度所需要迭代次数$k$=22.

(3) 当取 $\omega = 1.7$ 时，初始向量相同，达到同样精度，则需要迭代$k$=33.

对于此例，最佳松弛因子是 $\omega^* = 1.3$ ，即达到同样精度所需迭代次数最少。由此可知，用SOR 方法解线性方程时， 松弛因子选择得较好，常常会使SOR迭代收敛大大加速。

也即，SOR方法收敛的快慢与松弛因子ω的选择有密切关系。

但是如何选取最佳松弛因子, 即选取$\omega=\omega^*$, 使$\rho(B_\omega)$或残差达到最小，是一个尚未很好解决的问题。实际上可采用试算的方法来确定较好的松弛因子。经验上可取$1.4<\omega<1.6$。

Report

解 $Ax=b$ 的SOR 方法收敛的必要条件是：

$$\boxed{0<\omega<2}$$

因此， 用SOR 方法解 $Ax=b$ 时松弛因子应在

$$\boxed{0<\omega<2}$$

内选择（这时SOR方法可能收敛）。

## 6.3 迭代法的收敛性

### 一、一阶线性定常迭代法的基本定理

设 $Ax=b$

其中$A\in R^{n\times n}$非奇异，有精确解$x^*$，并有

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

则 $$x^*=Bx^*+f.$$

并有一阶线性定常迭代法 $x^{(k+1)}=Bx^{(k)}+f$

向量序列与矩阵序列的极限

定义6-2 设向量序列$\left\{x^{(k)}\in R^n\right\}$，$x^{(k)}=\left(x_1^{(k)},\quad x_2^{(k)},\quad \cdots,\quad x_n^{(k)}\right)^{\mathrm{T}}\in R^n$，如果存在$x^*=(x_1,\quad x_2,\quad \cdots,\quad x_n)^{\mathrm{T}}\in R^n$，使

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

则称向量序列$\left\{x^{(k)}\right\}$收敛于$x^*$，记作$\lim\limits_{k\to\infty} x^{(k)}=x^*$。

定义6-3 设矩阵序列$A_k=(a_{ij}^{(k)})\in R^{n\times n}$及，$A^*=(a_{ij}^{*})\in R^{n\times n}$，如果$n^2$个数列极限存在且有

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

则称矩阵序列$\{A_k\}$收敛于$A^*$，记作$\lim\limits_{k\to\infty} A_k=A^*$。

例如 设矩阵序列

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

且$|\lambda|<1$，考查其极限。

$$\lim_{k\to\infty} k\lambda^k=\lim_{k\to\infty}\frac{k}{\lambda^{-k}}=\lim_{k\to\infty}\frac{k}{(1/\lambda)^k}\overset{\frac{\infty}{\infty}}{=}\lim_{k\to\infty}\frac{1}{(\ln(1/\lambda))(1/\lambda)^k}=\frac{1}{\infty}=0.$$

引进误差向量$\varepsilon^{(k)}=x^{(k)}-x^*$,则

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

要研究何时收敛，即$B$满足什么条件有$B^k\to 0$（当$k\to\infty$）。

**定理6-1** $\lim\limits_{k\to\infty} A_k = A \Leftrightarrow \lim\limits_{k\to\infty} \|A_k - A\| = 0$，其中$\|\cdot\|$为矩阵的任一算子范数。

**定理6-2** $\lim\limits_{k\to\infty} A_k = 0 \Leftrightarrow \lim\limits_{k\to\infty} A_k x = 0$，$\forall x \in R^n$, 其中 两个极限右端分别指零矩阵与零向量。

**定理6-3** 设矩阵$B = (b_{ij}^{(k)}) \in R^{n\times n}$，则下列三个条件等价：

(1) $\lim\limits_{k\to\infty} B^k = 0$;

(2) $B$ 的谱半径$\rho(B) < 1$;

(3) 至少存在一种从属矩阵范数$\|\cdot\|_\varepsilon$，使$\|B\|_\varepsilon < 1$.

**定理6-4** 设线性方程组$Ax = b$对应有

$$x = Bx + f$$

及一阶线性定常迭代法

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

则，对任意初始向量$x^{(0)}$，该迭代法收敛的充要条件是迭代矩阵$B$ 的谱半径$\rho(B) < 1$.

注：定理6-4中的矩阵是迭代矩阵，常用格式的迭代矩阵如下：

1) 雅可比迭代法: $B_J=D^{-1}(L+U)$，$f_J=D^{-1}b$;

2) 高斯-赛德尔迭代法: $B_G=(D-L)^{-1}U$，$f_G=(D-L)^{-1}b$;

3) SOR迭代法: $B_S=(D-\omega L)^{-1}\{(1-\omega)D+\omega U\}$，

$f_S=\omega(D-\omega L)^{-1}b$。

**推论** 设$Ax=b$, $A\in R^{n\times n}$非奇异, $A=D-L-U$, $D$非奇异, 则

雅可比迭代法收敛$\Leftrightarrow \rho(B_J)<1$，其中$B_J=D^{-1}(L+U)$;

高斯$-$塞德尔迭代法收敛$\Leftrightarrow \rho(B_G)<1$, 其中$B_G=(D-L)^{-1}U$;

SOR迭代法收敛$\Leftrightarrow \rho(B_S)<1$，其中$B_S=(D-\omega L)^{-1}\{(1-\omega)D+\omega U\}$.

例如 考察用雅可比迭代法求解线性方程组

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

解: 迭代矩阵的特征方程为

$$\det(\lambda I - B_{\mathrm{J}}) = \lambda^3 + 0.034090909\lambda + 0.039772727 = 0$$

解得$\lambda_1 = \text{-}0.3082$

$$\lambda_2 = 0.1541 + 0.3245\mathrm{i}$$

$$\lambda_3 = 0.1541 - 0.3245\mathrm{i}$$

$$|\lambda_2| = |\lambda_3| = 0.35921 < 1, \quad |\lambda_1| < 1$$

即 $\rho(B_{\mathrm{J}}) < 1.$

所以用雅可比迭代法解方程组是收敛的。

例6-4 考察迭代法解方程组 $x^{(k+1)} = Bx^{(k)} + f$ 的收敛性其中 $B = \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix}$ 和 $f = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$

解: 特征方程为 $\det(\lambda I\text{-}B) = \lambda^2\text{-}6 = 0$,

特征根 $\lambda_{1,2} = \pm\sqrt{6}$

即 $\rho(B) > 1$, 这说明所用迭代法解此方程组不收敛。

**定理6-5**(迭代法收敛的充分条件) 设有方程组及一阶线性定常迭代法如定理6-4。如果有$B$ 的某种算子范数$\|B\| = q < 1$，则

(1) 迭代法全局收敛，即对任意$x^{(0)}$， $\lim\limits_{k\to\infty} x^{(k)} = x^*$ ， 且$x^* = Bx^* + f$ ；

(2) $\|x^* - x^{(k)}\| \le q^k \|x^* - x^{(0)}\|$;

(3) $\|x^* - x^{(k)}\| \le \dfrac{q}{1-q}\|x^{(k)} - x^{(k-1)}\|$;

(4) $\|x^* - x^{(k)}\| \le \dfrac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|$.

非线性方程求解也如此

解误差vs解间差vs残差

实际计算中, 通常利用$\|x^{(k)} - x^{(k-1)}\| < \varepsilon$ 作为控制迭代的终止条件。但要注意, 当 $q \approx 1$时, $q/(1-q)$较大, 尽管 $\|x^{(k)} - x^{(k-1)}\|$ 已非常小, 但误差向量的范数 $\|x^* - x^{(k)}\|$ 可能很大, 迭代法收敛将是缓慢的。

由定理6−5还可以得到雅可比迭代法收敛的充分条件是$\|B_J\|<1$。高斯-塞德尔迭代法收敛的充分条件是$\|B_G\|<1$。

注意：一般而言，解间差$\|x^{(k)}-x^{(k-1)}\|$很小时，解误差$\|x^{(k)}-x^*\|$就很小，实际上用$\|x^{(k)}-x^{(k-1)}\|<\varepsilon$作为迭代终止的条件；但$\|B\|$接近1时，非也，用之未必合适。

重要多讲

该定理只是收敛的充分条件，并不必要，如 $B=\begin{pmatrix}0.8 & 0.5\\ 0 & 0.7\end{pmatrix}$

则$\|B\|_1=1.2$, $\|B\|_\infty=1.3$, $\|B\|_2=1.09$, $\|B\|_F=1.17$

1列∞行

但$\rho(B)=0.8<1$, 所以对应的迭代法是收敛的.

## 二、某些特殊方程组的迭代收敛性

**定义** 若$n$维矩阵$A=(a_{ij})$满足

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

则称矩阵$A$是严格对角占优矩阵.

**引理** 若$A$是严格对角占优矩阵, 则$\det(A)\neq 0$.

**证** $A=D-L-U=D(I-D^{-1}(L+U))=D(I-B)$ 此证明不做要求

因为$A$是严格对角占优矩阵, 所以$\det(D)\neq 0$, 而且

$$\|B\|_{\infty}=\max_{1\le i\le n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|\frac{a_{ij}}{a_{ii}}\right|<1$$

因此, $\rho(B)\le\|B\|_{\infty}<1$, 故$\lambda=1$不是$B$的特征值, $\det(I-B)\neq 0$。

所以, $\det(A)\neq 0$。

**定理** 设$A$是严格对角占优矩阵, 则解线性方程组$Ax=b$的J迭代法和G-S迭代法均收敛。

此证明不做要求

证 由于$\|B\|_\infty<1$, 所以J迭代法收敛（简记$B=B_J$）。

对于G-S迭代法，设$\lambda$是$G$（简记$G=B_G$）的任一特征值, 则$\lambda$满足特征方程$\det(\lambda I-G)=\det(\lambda I-(D-L)^{-1}U)=\det((D-L)^{-1})(\lambda(D-L)-U)$

$$=\det((D-L)^{-1})\det(\lambda(D-L)-U)=0$$

所以有$\det(\lambda(D-L)-U)=0$

$$\lambda(D-L)-U=\begin{pmatrix}\lambda a_{11} & a_{12} & \cdots & a_{1n}\\ \lambda a_{21} & \lambda a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn}\end{pmatrix}$$

反证法

若$|\lambda|\geq 1$, 则矩阵$\lambda(D-L)-U$是严格对角占优矩阵, 这与$\det(\lambda(D-L)-U)=0$矛盾, 所以$|\lambda|<1$, 于是$\rho(G)<1$。

**定理6-8** 若SOR迭代收敛，则松弛因子$0<\omega<2$.

**定理6-9** **对于线性方程组*Ax*=*b*，若**

***(1)*** ***A*为对称正定矩阵，**

**(2)** $0<\omega<2$，

**则解*Ax*=*b*的SOR迭代收敛。**

**定理6-10（部分）对于线性代数方程组*Ax*=*b*， 若*A*严格对角占优，则当$0<\omega\le 1$时，SOR迭代收敛。**

常用线性方程组迭代法收敛性判别顺序：

1、应用定理6-9和定理6-10（系数矩阵）；

2、应用定理6-5（迭代矩阵）。

知识结构图

迭代法解方程组

- 迭代法基本概念
  - 充要条件：谱半径<1
  - 充分条件：解误差、解间差、残差
  - 严格对角占优：J和G收敛
  - SOR还有其他条件
- 雅可比迭代法
  - 迭代格式
  - 收敛条件（充要条件、充分条件）
- 高斯-赛德尔迭代法
  - 迭代格式
  - 收敛条件（充要条件、充分条件）
- SOR迭代法
  - 迭代格式
  - 收敛条件（充要条件、充分条件、必要条件）
- 迭代法收敛速度

复习与思考题(无需提交)
P208: 1, 2, 3, 4

习题(如愿可做，无需提交)
P209: 8, 9

习题(需提交)
P209: 1, 2