第六章 解线性方程组的迭代法

几个数学概念

- 泛函：以函数为自变量的函数（以数量为因变量）。
- 希尔伯特空间：完备的内积空间；
- 自伴算子：希尔伯特空间的线性算子，T*=T
- 变分：研究求泛函极大值（极小值）的方法。即求泛函极值的问题。
  - 泛函与函数的变分问题，类似于函数和函数的极值问题。

内积：

内积 $<f,g>$ 是三维空间向量点积 $a\cdot b=\alpha_1\beta_1+\alpha_2\beta_2+\alpha_3\beta_3$ 的推广。不同的希尔伯特空间有不同的内积形式，在矩量法中应用的内积一般是：

一维：$<f,g>=\int f(x)g(x)dx$

二维：$<f,g>=\int f(x,y)g(x,y)dxdy$

三维：$<f,g>=\int f(x,y,z)g(x,y,z)dxdydz$

向量与矩阵的范数

一、．向量范数：

对n维实空间R^n中任一向量X，按一定规则有一确定的实数与其相对应，该实数记为$\|X\|$，若$\|X\|$满足下面三个性质：

(1)(非负性)$\|X\|\geq 0$，$\|X\|=0$当且仅当X=0。

(2)(齐次性)对任意实数λ，$\|\lambda X\|=|\lambda|\,\|X\|$。

(3)(三角不等式)对任意向量$Y\in R^n$，$\|X+Y\|\leq\|X\|+\|Y\|$

则称该实数$\|X\|$为向量X的范数

几种常用的向量范数

$X=\{x_1,x_2,\ldots\ldots,x_n\}$

（1）向量的1—范数：

$$\|X\|_1=\sum_{i=1}^{n}|x_i|=|x_1|+|x_2|+\ldots+|x_n|$$

（2）向量的2—范数：

$$\|X\|_2=\sqrt{\sum_{i=1}^{n}x_i^2}=\sqrt{x_1^2+x_2^2+\ldots+x_n^2}$$

（3）向量的∞—范数：$\|X\|_\infty=\max\limits_{1\leq i\leq n}|x_i|$

（4）向量p范数

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

例：设 $x=(1,-4,0,2)^T$ 求它的向量范数

$$\|X\|_1=\sum_{k=1}^{n}|x_k|=7\qquad \|X\|_2=\sqrt{\sum_{i=1}^{n}x_i^2}=\sqrt{21}$$

$$\|X\|_\infty=\max_{1\leq i\leq n}|x_i|=4$$

$$\|X\|_p=\sqrt[p]{\sum_{i=1}^{n}|x_i|^p}=\sqrt[p]{1+4^p+2^p}$$

注：前三种范数都是p—范数的特殊情况。其中

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

向量范数的连续性：

定理 6.1 设$f(X)=\|X\|$为R^n上的任一向量范数,则$f(X)$为X的分量$x_1,x_2,\dots,x_n$的连续函数.

向量范数的等价性：

定理 6.2 若$\|X\|_p$与$\|X\|_q$为R^n上任意两种范数，则存在C_1，$C_2>0$，使得对任意$X\in R^n$，都有：

$$C_1\|X\|_p\le\|X\|_q\le C_2\|X\|_p \qquad \text{（证明略）}$$

注：同样有下列结论：存在$C_3,C_4>0$ 使得：

$$C_3\|X\|_q\le\|X\|_p\le C_4\|X\|_q$$

向量序列的收敛性

定义：假定给了R^n空间中的向量序列$X^{(1)},X^{(2)},\dots,X^{(k)},\dots$，简记为$\{X^{(k)}\}$，其中$X^{(k)}=(x_1^{(k)},x_2^{(k)},\dots,x_n^{(k)})^T$，若$X^{(k)}$的每一个分量$x_i^{(k)}$都存在极限$x_i$，即

$$\lim_{k\to\infty}x_i^{(k)}=x_i \quad (i=1,2,\dots,n)$$

则称向量$X=(x_1, x_2, \dots, x_n)^T$为向量序列$\{X^{(k)}\}$的极限，或者说向量序列$\{X^{(k)}\}$收敛于向量X，记为

$$\lim_{k\to\infty}X^{(k)}=X \quad 或 \quad X^{(k)}\to X\,(k\to\infty)$$

$$X^{(k)}=\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\\\vdots\\x_n^{(k)}\end{pmatrix}\begin{matrix}\longrightarrow x_1\\\longrightarrow x_2\\\cdots\cdots\cdots\cdots\\\longrightarrow x_n\end{matrix}\qquad(k\to\infty)$$

$$X^{(k)}=\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\\\vdots\\x_n^{(k)}\end{pmatrix}\to\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=X\qquad(k\to\infty)$$

二、矩阵范数：设A是n×n 阶矩阵，$A\in R^{n\times n}$ $X\in R^n$，$\|X\|$为R^n中的某范数，称

$$\|A\|=\max_{\substack{X\in R^n\\\|X\|\ne0}}\frac{\|AX\|}{\|X\|}=\max_{\|X\|=1,X\in R^n}\|AX\|$$

为矩阵A的从属于该向量范数的范数，或称为矩阵A的算子，记为$\|A\|$。

几种常用的矩阵范数

设 $A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$

$\|A\|_1=\max\limits_{1\le j\le n}\sum\limits_{i=1}^{n}|a_{ij}|$ 列范数

$\|A\|_\infty=\max\limits_{1\le i\le n}\sum\limits_{j=1}^{n}|a_{ij}|$ 行范数

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

$\|A\|_F=\sqrt{\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n}a_{ij}^2}$

弗罗贝尼乌斯 (Frobenius) 范数简称F范数

Matlab中计算矩阵的范数的命令(函数):

(1) n = norm(A) 矩阵A的谱范数(2范数), $=A'A$的最大特征值的算术根.

(2) n = norm(A,1) 矩阵A的列范数（1-范数） 等于A的最大列之和.

(3) n = norm(A,inf) 矩阵A的行范数(无穷范数) 等于A的最大行之和.

(4) n = norm(A, 'fro') 矩阵A的Frobenius范数.

$$A=\begin{bmatrix}1&2&3\\4&5&6\\7&8&0\end{bmatrix}$$

•矩阵范数的性质：

（1）对任意A∈$R^{n\times n}$,有$\|A\|\geq 0$，当且仅当A=0时，$\|A\|=0$.

（2）$\|\lambda A\|=|\lambda|\|A\|$（λ为任意实数）

（3）对于任意A、B ∈$R^{n\times n}$，恒有

$\|A+B\|\leq\|A\|+\|B\|$.

（4）对于矩阵A ∈$R^{n\times n}$，X ∈R^{n}，恒有：

$\|AX\| \leq \|A\|\bullet \|X\|$.

（5）对于任意A、B ∈$R^{n\times n}$ 恒有 $\|AB\| \leq \|A\| \bullet \|B\|$

•谱半径： $\boxed{\rho(A^m)=[\rho(A)]^m}$

设 n×n 阶矩阵A的特征值为λ_i(i=1,2,3……n)，则称 $\rho(A)=\max\limits_{1\le i\le n}|\lambda_i|$ 为矩阵A的谱半径。

谱半径=A的特征值中绝对值的最大者

例5.求矩阵 $A=\begin{bmatrix}4 & 6 & 0\\ -3 & -5 & 0\\ -3 & -6 & 1\end{bmatrix}$ 的谱半径

解： $\because |\lambda E-A|=(\lambda-1)^2(\lambda+2)$

$\therefore \rho(A)=2$

函数的线性组合表示

- 未知函数 *f* 近似表示成函数线性组合，即

$$f\approx f^N=\sum_{n=1}^{M}a_n f_n$$

在实际问题中，不可能在无穷维数的函数空间展开未知函数。

原函数、近似函数与误差函数可以用欧几里德空间上的矢量分解来形象地理解；

正交函数的展开过程，即把任意函数向坐标基做投影，对于完备基，误差为零。

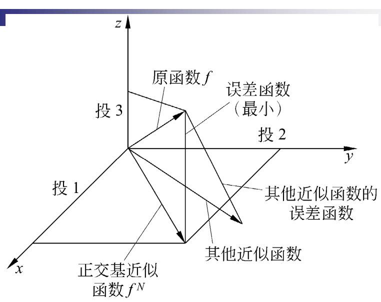
$$f\approx f^N=\sum_{n=1}^{M}a_n f_n$$

- 对于不完备基，误差函数为

$$d(f,f^N)=\left\|f-f^N\right\|=\left\|f-\sum_{n=1}^{M}a_n f_n\right\|$$

通过选择标量系数a_n，使得误差与*N*维基函数都正交，则为正交投影函数。即

$$\langle f_n, f-f^N\rangle=0,\quad n=1,2,\cdots,N$$



- 一个三维相量，在三维基下是精确的；在二维基下，差相量与xoy平面垂直的近似误差最小。

线性方程组求解问题

工程技术中常产生大型稀疏矩阵方程组，例如由某些偏微分方程数值解所产生的线性方程组 *Ax=b*，*A*的阶数很大，$(n\geq 10^4)$ 但零元素较多，迭代法是能够充分利用系数矩阵稀疏性特点的有效算法。

直接法是通过有限步运算后得到线性方程组的解. 包含:高斯消元法(列主元消去法)、三角分解法、追赶法.

18

列主元素法的精度虽稍低些,但计算简单,且具有良好的数值稳定性。

三角分解法 A=LU

线性方程组 AX=b ⟺ $\begin{cases} LY=b \\ UX=Y \end{cases}$

解线性方程组的所有直接的方法比较适用于中小型方程组.对高阶方程组,即使系数矩阵是稀疏的,但在计算中很难保持稀疏性,因而有存储量大,程序复杂等不足，这些不足之处可用迭代法来弥补解决.

迭代法的构造

迭代法的基本思想是用逐次逼近的方法求线性方程组的解。

设有方程组 $Ax=b$，将其转化为等价的便于迭代的形式 $x=Bx+f$

（这种转化总能实现，如令 $B=I-A,\ f=b$）并由此构造迭代公式

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

其中,B 称为迭代矩阵,f 称为迭代向量。对任意的初始向量 $x^{(0)}$，由迭代式可求得向量序列 $\{x^{(k)}\}_0^{\infty}$

若 $\lim\limits_{k\to\infty}x^{(k)}=x^*$，则 x^* 就是方程组 $Ax=b$ 的解.

§6.1 解线性方程组的三种迭代法

6.1.1. 雅克比（Jacobi）迭代法（以三阶方程组为例）

设有方程组：

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2 \\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3 \end{cases}$$

假设 $a_{ii}\neq 0$ i=1,2,3 则方程组可写为:

$$\begin{aligned} x_1&=\frac{1}{a_{11}}(b_1 \quad -a_{12}x_2-a_{13}x_3) \\ x_2&=\frac{1}{a_{22}}(b_2-a_{21}x_1 \quad -a_{23}x_3) \qquad (6.1) \\ x_3&=\frac{1}{a_{33}}(b_3-a_{31}x_1-a_{32}x_2 \quad) \end{aligned}$$

任选一向量$X^{(0)}$作为解的初值. $X^{(0)}=(x_1^{(0)},x_2^{(0)},x_3^{(0)})$

代入式(6.1)中得方程组的一次近似. $X^{(1)}=(x_1^{(1)},x_2^{(1)},x_3^{(1)})$

把$X^{(1)}$再代入到(6.1)中得方程组的二次近似.

$$X^{(2)}=(x_1^{(2)},x_2^{(2)},x_3^{(2)})$$

重复这一过程,假设得到了m次近似$X^{(m)}$。代入到(6.1)中可得m+1次近似$X^{(m+1)}$。

$$\begin{aligned} x_1^{(m+1)}&=\frac{1}{a_{11}}(b_1 \quad -a_{12}x_2^{(m)}-a_{13}x_3^{(m)}) \\ x_2^{(m+1)}&=\frac{1}{a_{22}}(b_2-a_{21}x_1^{(m)} \quad -a_{23}x_3^{(m)}) \\ x_3^{(m+1)}&=\frac{1}{a_{33}}(b_3-a_{31}x_1^{(m)}-a_{32}x_2^{(m)} \quad) \end{aligned}$$

称此迭代公式为原方程组的雅可比迭代公式.

对于n阶方程组 $Ax=b,\quad A=(a_{ij})_{n\times n}, b=(b_j)_{n\times 1}$

设$a_{ii}\neq 0$ i=1,2,⋯,n

则雅可比迭代公式为:

$$\begin{cases} x^{(0)}=(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)})^T \\ x_i^{(m+1)}=\dfrac{1}{a_{ii}}(b_i-\sum\limits_{j=1}^{i-1}a_{ij}x_j^{(m)}-\sum\limits_{j=i+1}^{n}a_{ij}x_j^{(m)}) \\ i=1,2,\cdots,n. \qquad m=0,1,2,\cdots. \end{cases}$$

若用矩阵来记录雅可比矩阵,可作如下的推导:
令$A=D-L-U$,其中

$$D=\begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix} \quad -L=\begin{pmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \cdots & \cdots & \cdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{pmatrix},$$

$$-U=\begin{pmatrix} 0 & a_{12} & a_{13} & & a_{1n} \\ & 0 & a_{23} & & a_{2n} \\ & & 0 & & \\ & & & \ddots & a_{n-1n} \\ & & & & 0 \end{pmatrix}$$

则有$AX=DX-LX-UX=b$.即$DX=b+(L+U)X$

从而有$DX^{(m+1)}=b+(L+U)X^{(m)}$.若 $a_{ii}\neq 0$ (i=1,2,$\cdots$,n)

则D可逆,于是得

$$x^{(m+1)}=D^{-1}(L+U)x^{(m)}+D^{-1}b$$
$$=B_J x^{(m)}+f_J$$
$$B_J=D^{-1}(L+U),\ f_J=D^{-1}b$$

称B_J为雅可比迭代矩阵.这种迭代格式称为雅可比迭代格式。

定义6.1 如果向量序列$\{X^{(m)}\}=\{(x_1^{(m)}, x_2^{(m)},\dots x_n^{(m)}\}$ 有 $x_i^{(m)}$ $\to x_i^*$ (i=1,2,3,…n)($m\to\infty$)

则称向量$X^*=(x_1^*,x_2^*,\dots x_n^*)$为向量序列$\{X^{(m)}\}$的极限,记为:

$$\lim_{m\to\infty} X^{(m)}=X^*$$

在某种条件下，按雅可比迭代所产生的向量序列的极限会存在，且等于原方程组的解。这种求解方法被称为雅可比迭代法，或简单迭代法。

例 用简单迭代法解下列方程组

$$\begin{cases} 10x_1-x_2-2x_3=7.2 \\ -x_1+10x_2-2x_3=8.3 \\ -x_1-x_2+5x_3=4.2 \end{cases}$$

解将方程组写成等价形式

$$\begin{cases} x_1= \quad 0.1x_2+0.2x_3+0.72 \\ x_2=0.1x_1 \quad +0.2x_3+0.83 \\ x_3=0.2x_1+0.2x_2 \quad +0.84 \end{cases}$$

取初始值$x^{(0)}=0$,按迭代公式

$$\begin{cases} x_1^{(k+1)}=0.1x_2^{(k)}+0.2x_2^{(k)} \quad +0.72 \\ x_2^{(k+1)}=0.1x_1^{(k)} \quad +0.2x_3^{(k)}+0.83 \\ x_3^{(k+1)}=0.2x_1^{(k)}+0.2x_2^{(k)} \quad +0.84 \end{cases}$$

k	0	1	2	3	4	5	6	7
$x_1^{(k)}$	0	0.72	0.971	1.057	1.085 3	1.095 1	1.098 3	…
$x_2^{(k)}$	0	0.83	1.070	1.157 1	1.185 3	1.195 1	1.198 3	…
$x_3^{(k)}$	0	0.84	1.150	1.248 2	1.282 8	1.294 1	1.298 0	…

6.1.2 高斯——赛德尔迭代法

对雅可比迭代法作如下的改进:将初值$X^{(0)}=(x_1^{(0)},x_2^{(0)},x_3^{(0)})$代入6.1的第一个方程可得$x_1^{(1)}$,用$(x_1^{(1)},x_2^{(0)},x_3^{(0)})$代入第二个方程得 $x_2^{(1)}$,用$(x_1^{(1)},x_2^{(1)},x_3^{(0)})$代入第三个方程得$x_3^{(1)}$，这样一直做下去,直到得到满意的解为止.之所以作这样的改进是希望更快的得到近似解.

$$\begin{aligned} x_1&=\frac{1}{a_{11}}(b_1 \quad -a_{12}x_2-a_{13}x_3) \\ x_2&=\frac{1}{a_{22}}(b_2-a_{21}x_1 \quad -a_{23}x_3) \qquad (9.1) \\ x_3&=\frac{1}{a_{33}}(b_3-a_{31}x_1-a_{32}x_2 \quad) \end{aligned}$$

这种迭代的方法用公式写出来就是:

$$x_1^{(m+1)} = \frac{1}{a_{11}}(b_1 \qquad - a_{12}x_2^{(m)} - a_{13}x_3^{(m)})$$

$$x_2^{(m+1)} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^{(m+1)} \qquad - a_{23}x_3^{(m)})$$

$$x_3^{(m+1)} = \frac{1}{a_{33}}(b_3 - a_{31}x_1^{(m+1)} - a_{32}x_2^{(m+1)} \qquad)$$

一般地,对n阶线性方程组的迭代格式改为:

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

$$i = 1, 2, \cdots, n; m = 0, 1, 2, \cdots. \qquad (9.8)$$

对给定的初值,用此迭代公式求线性方程组的方法被称为高斯—塞德尔迭代法。（G—S）

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

$$i = 1, 2, \cdots, n; m = 0, 1, 2, \cdots. \qquad (9.8)$$

A_{ii}元素组成矩阵D，*A=D-L-U*

用矩阵表示此方法为：$DX^{(m+1)} = b + LX^{(m+1)} + UX^{(m)}$

即：$X^{(m+1)} = (D-L)^{-1}UX^{(m)} + (D-L)^{-1}b$

$= B_G X^{(m)} + f_G$

称B_G为高斯——塞德尔迭代矩阵

例　用赛德尔迭代法解方程组

$$\begin{cases} 10x_1 - x_2 - 2x_3 = 7.2 \\ -x_1 + 10x_2 - 2x_3 = 8.3 \\ -x_1 - x_2 + 5x_3 = 4.2 \end{cases}$$

解　将原方程组写成等价形式并构造赛德尔迭代公式

$$\begin{cases} x_1^{(k+1)} = 0.1x_2^{(k)} + 0.2x_3^{(k)} + 0.72 \\ x_2^{(k+1)} = 0.1x_1^{(k+1)} + 0.2x_3^{(k)} + 0.83 \\ x_3^{(k+1)} = 0.2x_1^{(k+1)} + 0.2x_2^{(k+1)} + 0.84 \end{cases}$$

k	0	1	2	3	4	5	6
$x_1^{(k)}$	0	0.72	1.043 1	1.093 1	1.099 1	1.099 9	…
$x_2^{(k)}$	0	0.902	1.167 2	1.195 7	1.199 5	1.199 9	…
$x_3^{(k)}$	0	1.164 4	1.282 0	1.297 8	1.299 7	1.300 0	…

在多数情况下用高斯—赛德尔迭代法比雅克比迭代法收敛快。

但也有相反的情况，即高斯—赛德尔迭代法比雅克比迭代法收敛慢，甚至还有雅克比迭代法收敛，高斯—赛德尔迭代法发散的情形。

6.1.3 超松弛迭代法

超松弛迭代法是高斯—赛德尔迭代法的一种改进,是解大型稀疏矩阵方程组的有效方法之一.

设已知第k次迭代向量$X^{(k)}$，及第$k+1$次迭代向量$X^{(k+1)}$的前$i-1$个分量 $x_j^{(k+1)}(j=1,2,\dots,i-1)$,

现在研究如何求向量$X^{(k+1)}$的第i个分量$x_i^{(k+1)}$。

首先，由高斯—赛德尔迭代法求出一个值，记

$$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}) \quad (i=1,2,\dots,n)$$

39

首先，由高斯—赛德尔迭代法求出一个值，记

$$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}) \quad (i=1,2,\dots,n)$$

再将第k次迭代向量的第i个分量$x_i^{(k)}$与$\tilde{x}_i^{(k+1)}$进行加权平均，得$x_i^{(k+1)}$，即

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \omega\tilde{x}_i^{(k+1)} = x_i^{(k)} + \omega(\tilde{x}_i^{(k+1)} - x_i^{(k)})$$

于是得SOR迭代公式.

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

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

40

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

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

用此公式求解线性方程组的方法称为带有松弛因子ω的松弛迭代法.

当ω>1时称为超松弛迭代法;(SOR法)

当ω<1时称为低松弛迭代法;

当ω=1时就是G—S迭代法.

当某些方程组用高斯—赛德尔迭代法不收敛时，可以用低松弛方法获得收敛.

41

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

A_{ii}元素组成矩阵D， *A=D-L-U*

将上式写成矩阵的形式，得:

$$DX^{(k+1)} = (1-\omega)DX^{(k)} + \omega(b + LX^{(k+1)} + UX^{(k)})$$

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

于是得SOR迭代的矩阵表示

$$X^{(k+1)} = B_\omega X^{(k)} + f_\omega$$

其中$B_\omega = (D-\omega L)^{-1}[(1-\omega)D + \omega U]$

$$f_\omega = \omega(D-\omega L)^{-1}b$$

42

6.2 迭代法的收敛条件

前面介绍了三种迭代法.从例子看到这三种迭代法都有成功的时候.但我们也可以预计,某种迭代法可能会失效.下面我们试图从理论上来探讨这一问题.三种迭代法的统一写法为:

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

48

6.2.1 迭代法收敛的概念

定义　设给定R^n中的向量序列$\{x^{(m)}\}$，即 $x^{(0)}, x^{(1)}, \cdots, x^{(m)}, \cdots$

其中 $x^{(m)} = \left(x_1^{(m)}, x_2^{(m)}, \cdots, x_n^{(m)}\right)^T$

若对任何i ($i=1,2,\dots,n$)都有 $\lim_{m\to\infty} x_i^{(m)} = x_i^*$

则向量$x^* = (x_1^*, \cdots, x_n^*)^T$ 称为向量序列$\{x^{(m)}\}$的极限,

或者说向量序列$\{x^{(m)}\}$依坐标收敛于向量，记为

$$\lim_{m\to\infty} x^{(m)} = x^*$$

49

引理　向量序列$\{x^{(m)}\}$依坐标收敛于x^*的充要条件是

$$\lim_{m\to\infty}\left\|x^{(m)}-x^*\right\|=0$$

如果满足此式，称$x^{(m)}$依范数收敛于x^*

向量序列依范数收敛与依坐标收敛是等价的。

证：由 $\lim\limits_{m\to\infty} x^{(m)}=x^* \Leftrightarrow \lim\limits_{m\to\infty}(x_i^{(m)}-x_i^*)=0\ (i=1,2,\cdots,n)$

$$\Leftrightarrow \lim_{m\to\infty}\max_{1\le i\le n}\left|x_i^{(m)}-x_i^*\right|=0 \Leftrightarrow \lim_{m\to\infty}\left\|x^{(m)}-x^*\right\|_\infty=0$$

再由范数的等价性有

$$\lim_{m\to\infty}x^{(m)}=x^* \Leftrightarrow \lim_{m\to\infty}\left\|x^{(m)}-x^*\right\|=0$$

定义6.2　设x^*是方程组$Ax=b$的解,对于给定的初始向量$x^{(0)}$,若由某种迭代法产生的向量序列$\{x^{(m)}\}$有

$$\lim_{m\to\infty}x^{(m)}=x^*$$

则称该方法收敛,否则称该方法发散.

6.2.2 迭代法收敛的判定定理

定理6.1　设 $x^{(m+1)}=Bx^{(m)}+f$，若 $\|B\|_p=q<1$

则对任意的初始向量 $x^{(0)}$,该迭代过程收敛于 $x=Bx+f$

的唯一解 x^* ,且有估计式

$$(1)\quad \left\|x^*-x^{(m)}\right\|_p\le\frac{1}{1-q}\left\|x^{(m+1)}-x^{(m)}\right\|_p$$

$$(2)\quad \left\|x^*-x^{(m)}\right\|_p\le\frac{q^m}{1-q}\left\|x^{(1)}-x^{(0)}\right\|_p$$

证:先证　若 $\|B\|_p<1$, 则**E-B**非奇异.

用反证法:

设**E-B**是奇异的,则存在非零向量x,使$(E-B)x=0$.即有$x=Bx$.

两边取范数,再由范数的性质得

$$\|x\|_p=\|Bx\|_p\le\|B\|_p\|x\|_p$$

由于$\|x\|_p>0$, 得 $\|B\|_p\ge1$

与$\|B\|_p<1$　矛盾

由于$E-B$是非奇异的，所以方程组$(E-B)x=f$的解存在且唯一.

设为x^*,即　$x^*=Bx^*+f$,

进而有

$$(E-B)x=f \qquad x^{(m+1)}=Bx^{(m)}+f$$

两式相减：　$x^*-x^{(m+1)}=B(x^*-x^{(m)})$

取范数得：

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

$$=q\left\|x^*-x^{(m)}\right\|_p\le q^2\left\|x^*-x^{(m-1)}\right\|_p$$

$$\le\cdots\le q^m\left\|x^*-x^{(0)}\right\|_p$$

由于$0<q<1$,所以

$$\lim_{m\to\infty}\left\|x^*-x^{(m)}\right\|_p=0$$

所以迭代过程收敛.又

$$\left\|x^*-x^{(m)}\right\|_p=\left\|x^*-x^{(m+1)}+x^{(m+1)}-x^{(m)}\right\|_p$$

$$\le\left\|x^*-x^{(m+1)}\right\|_p+\left\|x^{(m+1)}-x^{(m)}\right\|_p$$

$$\le q\left\|x^*-x^{(m)}\right\|_p+\left\|x^{(m+1)}-x^{(m)}\right\|_p$$

即：　$\left\|x^*-x^{(m)}\right\|_p\le q\left\|x^*-x^{(m)}\right\|_p+\left\|x^{(m+1)}-x^{(m)}\right\|_p$

移项并合并，于是有

$$\left\|x^*-x^{(m)}\right\|_p\le\frac{1}{1-q}\left\|x^{(m+1)}-x^{(m)}\right\|_p$$

即(1)式成立.

再由于 $\left\|x^{(m+1)}-x^{(m)}\right\|=\left\|Bx^{(m)}+f-Bx^{(m-1)}-f\right\|_p$

$\le\|B\|_p\left\|x^{(m)}-x^{(m-1)}\right\|_p=q\left\|x^{(m)}-x^{(m-1)}\right\|_p$

$\le\cdots\le q^m\left\|x^{(1)}-x^{(0)}\right\|_p$

根据前面 $x^{(m+1)}=Bx^{(m)}+f$

$$\left\|x^*-x^{(m)}\right\|_p\le\frac{1}{1-q}\left\|x^{(m+1)}-x^{(m)}\right\|_p$$

所以有

$$\left\|x^*-x^{(m)}\right\|_p\le\frac{q^m}{1-q}\left\|x^{(1)}-x^{(0)}\right\|_p$$

即(2)式成立.

(2)式提示我们可以利用 $\left\|x^{(m+1)}-x^{(m)}\right\|$ 来控制误差

G-S： $X^{(m+1)}=(D-L)^{-1}UX^{(m)}+(D-L)^{-1}b$

$=B_G X^{(m)}+f_G$

例　写出用雅可比迭代法和G-S迭代法解线性方程组收敛的迭代格式。$\begin{cases}3x_1+x_2=5\\x_1+2x_2=5\end{cases}$

解：对A分解

$$A=\begin{pmatrix}3&1\\1&2\end{pmatrix},D=\begin{pmatrix}3&0\\0&2\end{pmatrix}\quad L=\begin{pmatrix}0&0\\-1&0\end{pmatrix},U=\begin{pmatrix}0&-1\\0&0\end{pmatrix}$$

于是

$$B_J=D^{-1}(L+U)=\begin{pmatrix}0&-\frac{1}{3}\\-\frac{1}{2}&0\end{pmatrix}\quad B_G=(D-L)^{-1}U=\begin{pmatrix}0&-\frac{1}{3}\\0&\frac{1}{6}\end{pmatrix}$$

由此得 $\|B_J\|_\infty=\frac{1}{2}<1,\|B_G\|_\infty=\frac{1}{3}<1$

所以用雅可比迭代法和G-S迭代法求解方程组都收敛。分别为：

雅可比迭代格式　　G-S迭代格式

$x_1^{(0)}=0,x_2^{(0)}=0,$　　$x_1^{(0)}=0,x_2^{(0)}=0,$

$$\begin{cases}x_1^{(m+1)}=\frac{5}{3}-\frac{1}{3}x_2^{(m)}\\x_2^{(m+1)}=\frac{5}{2}-\frac{1}{2}x_1^{(m)}\end{cases}\qquad\begin{cases}x_1^{(m+1)}=\frac{5}{3}-\frac{1}{3}x_2^{(m)}\\x_2^{(m+1)}=\frac{5}{2}-\frac{1}{2}x_1^{(m+1)}\end{cases}$$

定义6.3　如果方阵A满足

$$\sum_{j=1}^{i-1}|a_{ij}|+\sum_{j=i+1}^{n}|a_{ij}|<\sum_{i=1}^{n}|a_{ii}|(i=1,2,\cdots,n)$$

则称A按行严格对角占优.

(类似地可定义按列严格对角占优)

注意：$\begin{pmatrix}1&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&1&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&1\end{pmatrix}$ 是对角占优,不是严格对角占优.

迭代法 $x^{(m+1)}=D^{-1}(L+U)x^{(m)}+D^{-1}b$

$=B_J x^{(m)}+f_J$

$B_J=D^{-1}(L+U),f_J=D^{-1}b$

定理6.2　若方程组Ax=b的系数矩阵按行(列)严格对角占优,则雅可比迭代法收敛,G-S迭代法也收敛.

证：　先证雅可比迭代法收敛.

因为：$\|B_J\|=\left\|D^{-1}(L+U)\right\|_\infty=\max\limits_{1\le i\le n}\left(\sum_{j=1}^{i-1}\frac{|a_{ij}|}{|a_{ii}|}+\sum_{j=i+1}^{n}\frac{|a_{ij}|}{|a_{ii}|}\right)<1$

所以由定理6.1,Ax=b存在唯一解x*.且用雅可比迭代法求解收敛.

可证G-S迭代法收敛(略)

以上两个定理都是收敛的充分条件.

下面给出一个充分必要条件：

定理6.3　对于任意的初始向量$x^{(0)}$，由 $x^{(m+1)}=Bx^{(m)}+f$

产生的向量序列$\{x^{(m)}\}$收敛的充分必要条件是

矩阵B的谱半径

$$\rho(B)<1$$

矩阵A的谱半径等于矩阵A的特征值的模的最大值；若特征值为虚数，则谱半径为实部与虚部的平方和的开方。

例：写出用雅可比迭代法求解方程组一定收敛的迭代格式。 $\begin{cases} \frac{1}{2}x_1 + x_2 = -\frac{1}{2} \\ x_1 + \frac{1}{2}x_2 = \frac{1}{2} \end{cases}$

解：对A进行分解

$$A=\begin{pmatrix} \frac{1}{2} & 1 \\ 1 & \frac{1}{2} \end{pmatrix}, D=\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, L=\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, U=\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$$

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

由B_J的特征多项式

$$|\lambda E - B_J| = \begin{vmatrix} \lambda & 2 \\ 2 & \lambda \end{vmatrix} = \lambda^2 - 4$$

得特征值为：λ_1=2，λ_2=-2。所以$\rho(B_J)$=2>1

由此例可以看到:对原方程组直接写出雅可比迭代公式是不收敛的.

原方程组： $\begin{cases} \frac{1}{2}x_1 + x_2 = -\frac{1}{2} \\ x_1 + \frac{1}{2}x_2 = \frac{1}{2} \end{cases}$

如果写出与原方程组等价的方程组 $\begin{cases} x_1 + \frac{1}{2}x_2 = \frac{1}{2} \\ \frac{1}{2}x_1 + x_2 = -\frac{1}{2} \end{cases}$

其系数矩阵是严格对角占优的所以用雅可比迭代法求解收敛.

其迭代格式为:

$$x_1^{(0)} = 0, x_2^{(0)} = 0,$$

$$\begin{cases} x_1^{(m+1)} = \frac{1}{2} - \frac{1}{2}x_2^{(m)} \\ x_2^{(m+1)} = -\frac{1}{2} - \frac{1}{2}x_1^{(m)} \end{cases}$$

定理6.4 设方程组Ax=b的系数矩阵A为实对称正定矩阵,且0<ω<2,则松弛迭代法收敛.

说明：定理给出当0<ω<2时，松弛迭代法收敛。但是常用的是1<ω<2的情形，所以本定理条件称为SOR方法的收敛条件，仅为充分条件。

对称正定矩阵：各阶主子式都大于零；或特征值大于零。

例5:讨论例2中的方程组用SOR方法求解的收敛性.

解:例2中方程组的系数矩阵为:方程组

$$\begin{pmatrix} 4 & -2 & -4 \\ -2 & 17 & 10 \\ -4 & 10 & 9 \end{pmatrix}$$

首先A是对称矩阵,再由

$$\det(A_1) = 4; \det(A_2) = 64; \det(A) = 2$$

知A是对称正定矩阵.由定理9.4知,当0<ω<2时,用SOR法求解收敛.

如何选取ω的值，使SOR迭代法的收敛速度最快，目前，只有少数特殊类型的矩阵，才有确定的最佳松弛因子的理论公式。例如，当A为对称正定的三对角矩阵时，若$\mu = \rho(B_J) < 1$，则SOR法的最佳松弛因子为 $\omega_b = \frac{2}{1+\sqrt{1-\mu^2}}$，但这在实际应用时也有一定困难

常用的方法是，选不同的ω进行计算，以确定最佳ω的近似值，或者，先选取一个ω $(0<\omega<2)$，然后根据迭代过程收敛的快慢不断修改ω，以此逐步寻找最佳松弛因子ω_b。

The end！

68