第三章 线性代数方程组的迭代解法

/* Iterative Techniques for Solving Linear Systems */

如何快速求解系数为大型稀疏矩阵的方程组

$$A\vec{x}=\vec{b}$$

思路：与解方程 $f(x)=0$ 的不动点迭代相似。

研究 内容：

- 如何建立迭代格式？
- 向量序列的收敛条件？
- 收敛速度？
- 误差估计？

迭代法思想:

$$\vec{x} = B\vec{x} + \vec{f}$$

$$\vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{f}$$

- **如何建立迭代格式？**

分裂 $A=M-N$, $(M-N)x=b$ ，

构造 $Mx^{k+1}=Nx^{k}+b$

若M**可逆，** $x^{k+1}=M^{-1}Nx^{k}+M^{-1}b$

$B=M^{-1}N$**被称为迭代法的迭代矩阵**

目标：

精度可控，迭代收敛。

§1 三种迭代法

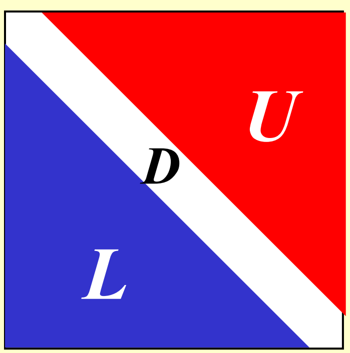
/* Jacobi，Gauss-Seidel and SOR Iterative Methods */

➤ Jacobi Iterative Method

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \quad \dots \quad \dots \quad \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \xrightarrow{a_{ii} \neq 0} \begin{cases} x_1 = \dfrac{1}{a_{11}}(-a_{12}x_2 - \dots - a_{1n}x_n + b_1) \\ x_2 = \dfrac{1}{a_{22}}(-a_{21}x_1 - \dots - a_{2n}x_n + b_2) \\ \dots \quad \dots \quad \dots \quad \dots \\ x_n = \dfrac{1}{a_{nn}}(-a_{n1}x_1 - \dots - a_{nn-1}x_{n-1} + b_n) \end{cases}$$

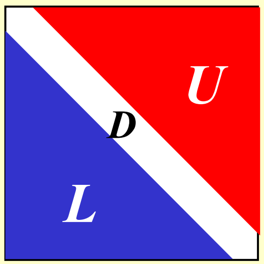
写成矩阵形式：

$A =$



$$A\vec{x} = \vec{b} \Leftrightarrow (D + L + U)\vec{x} = \vec{b}$$
$$\Leftrightarrow D\vec{x} = -(L + U)\vec{x} + \vec{b}$$
$$\Leftrightarrow \vec{x} = -D^{-1}(L + U)\vec{x} + D^{-1}\vec{b}$$

$$\vec{x}^{(k+1)} = \underbrace{-D^{-1}(L + U)}_{B}\vec{x}^{(k)} + \underbrace{D^{-1}\vec{b}}_{\vec{f}}$$

$A =$ 

Jacobi 迭代法：

$$A\vec{x}=\vec{b} \iff (D+L+U)\vec{x}=\vec{b}$$
$$\iff D\vec{x}=-(L+U)\vec{x}+\vec{b}$$
$$\iff \vec{x}=-D^{-1}(L+U)\vec{x}+D^{-1}\vec{b}$$

$$\vec{x}^{(k+1)}=\underbrace{-D^{-1}(L+U)}_{B}\vec{x}^{(k)}+\underbrace{D^{-1}\vec{b}}_{\vec{f}}$$

Jacobi 迭代阵

➢ Gauss - Seidel Iterative Method

$$x_1^{(k+1)} = \frac{1}{a_{11}}(-a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - a_{14}x_4^{(k)} - \cdots - a_{1n}x_n^{(k)} + b_1)$$

只存一组向量即可。

$$x_2^{(k+1)} = \frac{1}{a_{22}}(-a_{21}x_1^{(k+1)} - a_{23} \cdots$$

$$x_3^{(k+1)} = \frac{1}{a_{33}}(-a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} - a_{34}x_4^{(k)} - \cdots - a_{3n}x_n^{(k)} + b_3)$$

… … … …

$$x_n^{(k+1)} = \frac{1}{a_{nn}}(-a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - a_{n3}x_3^{(k+1)} - \cdots - a_{nn-1}x_{n-1}^{(k+1)} + b_n)$$

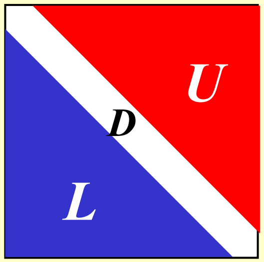
写成矩阵形式：

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

$$\Leftrightarrow (D+L)\vec{x}^{(k+1)} = -U\vec{x}^{(k)} + \vec{b}$$

$$\Leftrightarrow \vec{x}^{(k+1)} = \underbrace{-(D+L)^{-1}U}_{B}\vec{x}^{(k)} + \underbrace{(D+L)^{-1}\vec{b}}_{\vec{f}}$$

Gauss-Seidel 迭代阵

$A =$ 

$$Ax = b \Leftrightarrow (D+L)x^{k+1} = -U\,x^{k} + b$$

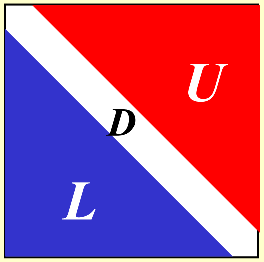
Gauss-Seidel 迭代法：

$$\Leftrightarrow (D+L)\vec{x}^{(k+1)} = -U\vec{x}^{(k)} + \vec{b}$$

$$\Leftrightarrow \boxed{\vec{x}^{(k+1)} = \underbrace{-(D+L)^{-1}U}_{B}\,\vec{x}^{(k)} + \underbrace{(D+L)^{-1}\vec{b}}_{\vec{f}}}$$

Gauss-Seidel
迭代阵

➢ Successive Over-relaxation (SOR) Iterative Method

$A =$ 

$$\omega Ax = \omega b \overset{\omega\neq 0}{\Longleftrightarrow} M = D + \omega L,$$

$$N = (1-\omega)D - \omega U$$

$$\omega A = M - N$$

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

超松弛(SOR)迭代法：

$$x^{k+1} = \underbrace{(D+\omega L)^{-1}[(1-\omega)D - \omega U]}_{B}x^{k} + \underbrace{\omega(D+\omega L)^{-1}b}_{\vec{f}}$$

SOR
迭代阵

问题：ω <1时称为低松弛, ω >1时称为超松弛，那么ω =1时，SOR迭代法是什么迭代法？

§2 迭代法收敛性

/* Convergence of the three Iterative Methods */

- 迭代序列的收敛条件

回顾：向量范数 /* vector norms */

常用向量范数：

$$\|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|\vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

$$\|\vec{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

$$\|\vec{x}\|_\infty = \max_{1\le i\le n} |x_i|$$

注： $\lim_{p\to\infty} \|\vec{x}\|_p = \|\vec{x}\|_\infty$

定义 向量序列 $\{\vec{x}^{(k)}\}$ **收敛**于向量 $\vec{x}^*$ 是指对每一个 $1\le i\le n$ 都有 $\lim_{k\to\infty} x_i^{(k)} = x_i^*$。可以理解为 $\|\vec{x}^{(k)} - \vec{x}^*\|_\infty \to 0$

矩阵范数：$\| A \|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} | a_{ij} |$（行和范数）

$\| A \|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} | a_{ij} |$（列和范数）

$\| A \|_2 = \sqrt{\lambda_{\max}(A^T A)}$（谱范数 /* spectral norm */ ）

Frobenius 范数 $\| A \|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} | a_{ij} |^2}$ —— 向量$\| \cdot \|_2$的直接推广

引理： 设$||x||$是在R^n 上一个范数，则$||x||$是x的分量$x_1, x_2, \cdots, x_n$的连续函数.

定理： R^n 上任意两种范数都等价.

定理： 设$A \in R^{n\times n}$,

（1）对任意从属范数 $||\cdot||$，有

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

（2）任给$\varepsilon > 0$，存在一种A的从属范数$\|\cdot\|_\varepsilon$，使得

$$\|A\|_\varepsilon \le \rho(A) + \varepsilon.$$

定义 矩阵序列 $A^{(k)}=(a_{ij}^{(k)})\in R^{n\times n}$ **收敛**于矩阵 A 是指对每一个 $1\le i, j\le n$ 都有 $\lim\limits_{k\to\infty} a_{ij}^{(k)}=a_{ij}$。此时称 A 为矩阵序列 $A^{(k)}$ 的极限，记为 $A=\lim\limits_{k\to\infty} A^{(k)}$

类似向量序列收敛性结论：$A^{(k)}\to A \Leftrightarrow \| A^{(k)}-A\| \to 0$

定理 设 $A\in R^{n\times n}$，当 $k\to\infty$ 时，$A^k\to 0$ 的充要条件是 $\rho(A)<1$。

证明：若 $\rho(A)<1$，则由上述定理可以选择一种范数，使得 $\|A\|<1$，因此由矩阵的相容性得

$$\|A^k\|\le\|A\|^k\to 0$$

于是**充分性**得证。

定理 设 $A \in R^{n\times n}$，当 $k \to \infty$ 时，$A^k \to 0$ 的充要条件是 $\rho(A) < 1$。

必要性证明：

设 $\rho(A) \geq 1$，令 λ 为某一使 $|\lambda| \geq 1$ 的特征值，设 x 为对应特征向量，则

$$\|A^k\|\|x\| \geq \|A^k x\| = \|\lambda^k x\| \geq \|x\|.$$

意味着对所有的k，$\|A^k\| \geq 1$，矛盾。
于是必要性得证。

➢ 迭代法的收敛性判别

对任何初始向量x^0，上述迭代法

$$\vec{x}^{(k+1)} = B\,\vec{x}^{(k)} + \vec{f}$$

收敛的充要条件是 $\rho(B) < 1.$

Matlab主程序：谱半径判断敛散性

```
function H=testspect(B)
        H=eig(B);
        Sp=norm(H, inf); %按模取最大特征值
if Sp>=1
        disp(’迭代序列发散，谱半径Sp和特征值H如下：’)
else
        disp(’迭代序列收敛，谱半径Sp和特征值H如下：’)
end
Sp
```

Neumann引理 矩阵幂级数$\sum_{k=0}^{\infty} A^k$收敛的充要条件是 $\rho(A)<1$，且当 $\rho(A)<1$ 时有

$$I+A+\cdots+A^k\cdots=(I-A)^{-1}$$

证明：若$\sum_{k=0}^{\infty} A^k$收敛，则 $A^k \to 0$ ，从而 $\rho(A)<1$

反之，若 $\rho(A)<1$ ，则I - A的特征值为 $1-\lambda_i, i=1,2,\cdots,n$

由于 $(I-A)(I+A+\cdots+A^k)=(I-A^{k+1})$

得 $(I+A+\cdots+A^k)=(I-A)^{-1}(I-A^{k+1})$

当 $k\to\infty$ 时，有 $A^k\to 0$ 故

$$I+A+\cdots+A^k+\cdots=(I-A)^{-1}$$

Banach引理 若矩阵 $A\in R^{n\times n}$ 非奇异，对任何 $E\in R^{n\times n}$ 且当 $\|A^{-1}\|\cdot\|E\|<1$ 时，则有 $A+E$ 非奇异，并且

$$\|(A+E)^{-1}\|\le\frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|E\|}$$

证明：设 $B=-A^{-1}E$，则 $\|B\|<1$，由上一定理可知 I-B非奇异，且

$$I+B+\cdots+B^k\cdots=(I-B)^{-1}$$

因此

$$\|(I-B)^{-1}\|\le\frac{1}{1-\|B\|}\le\frac{1}{1-\|A^{-1}\|\cdot\|E\|}$$

但是

$$A+E=A(I+A^{-1}E)=A(I-B)$$

所以$A+E$是非奇异的，并

$$\|(A+E)^{-1}\|=\|(I-B)^{-1}A^{-1}\|\le\frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|E\|}$$

∎

定理 设$A \in R^{n\times n}$，当$k\to\infty$时，$A^k\to 0$的充要条件是$\rho(A)<1$.

定理(Neumann引理) 矩阵幂级数 $\sum_{k=0}^{\infty} A^k$ 收敛的充要条件是 $\rho(A)<1$ 且当$\rho(A)<1$时有

$$I+A+\cdots+A^k\cdots=(I-A)^{-1}$$

定理 若矩阵 B 对某个从属范数满足 $\|B\|<1$，则必有

(1) $I\pm B$可逆； (2) $\left\|(I\pm B)^{-1}\right\|\le\dfrac{1}{1-\|B\|}$

定理 若A对称，则有 $\|A\|_2=\rho(A)$

$\rho(B)$一般难以计算，故而考虑迭代法收敛的充分条件$||B|| < 1$.