第六章 解线性方程组的迭代法

/*Chapter6 Iterative Techniques for Solving Linear Systems */

求解 $A\vec{x}=\vec{b}$

思路　与解 $f(x)=0$ 的不动点迭代相似，将 $A\vec{x}=\vec{b}$ 等价改写为 $\vec{x}=B\vec{x}+\vec{f}$ 形式，建立迭代 $\vec{x}^{(k+1)}=B\vec{x}^{(k)}+\vec{f}$。从初值 $\vec{x}^{(0)}$ 出发，得到序列 $\{\vec{x}^{(k)}\}$ 。

计算精度可控，特别适用于求解系数为大型稀疏矩阵 /* sparse matrices */ 的方程组。

研究内容:

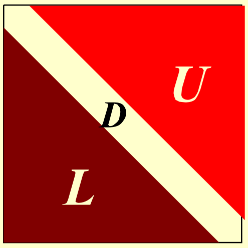
- 如何建立迭代格式？
- 收敛速度？
- 向量序列的收敛条件？
- 误差估计？

➢1. Jacobi 迭代法

➢ Jacobi Iterative Method

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

写成矩阵形式:

$A =$ 

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

Jacobi 迭代阵

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


必须等$X^{(k)}$完全计算好了才能计算$X^{(k+1)}$，因此需要两组向量存储。
迭代过程中，A的元素不改变，故可以事先调整好A使得$a_{ii} \neq 0$，否则A不可逆。
What if a_{ii}
A bit wasteful, isn’t it?

➢2. Gauss - Seidel 迭代法

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

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

只存一组向量即可。

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

$$\cdots\quad\cdots\quad\cdots\quad\cdots$$

$$x_n^{(k+1)}=\frac{1}{a_{nn}}(\qquad +b_n)$$

写成矩阵形式: $\vec{x}^{(k+1)}=-D^{-1}(L\vec{x}^{(k+1)}+U\vec{x}^{(k)})+D^{-1}\vec{b}$

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

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

Gauss-Seidel
迭代阵

注：二种方法都存在收敛性问题。

有例子表明：**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。

➤3.迭代法的收敛性

$\vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{f}$ 的收敛条件

$$\underline{\vec{e}^{(k+1)}} = \vec{x}^{(k+1)} - \vec{x}^* = (B\vec{x}^{(k)} + \vec{f}) - (B\vec{x}^* + \vec{f}) = B(\vec{x}^{(k)} - \vec{x}^*) = \underline{B\vec{e}^{(k)}}$$

$$\Rightarrow \boxed{\vec{e}^{(k)} = B^k \vec{e}^{(0)}} \Rightarrow \|\vec{e}^{(k)}\| \le \|B\| \cdot \|\vec{e}^{(k-1)}\| \le \ldots \le \|B\|^k \cdot \|\vec{e}^{(0)}\|$$

充分条件：$\|B\| < 1$ $\quad \Rightarrow \|B\|^k \to 0$ a

$\Rightarrow \|\vec{e}^{(k)}\| -$

等价于对
任何算子范数有
$\|A_k - A\| \to 0$ as $k \to \infty$

必要条件：$\vec{e}^{(k)} \to \vec{0}$ as $k \to \infty$ $\Rightarrow$

定义 设：$A = (a_{ij})_{n\times n}$，$A_k = (a_{ij}^{(k)})_{n\times n} \in R^{n\times n}$.

$\lim\limits_{k\to\infty} A_k = A$ 是指 $\lim\limits_{k\to\infty} a_{ij}^{(k)} = a_{ij}$ 对所有 $1 \le i, j \le n$ 成立。

定理 设 $\vec{x} = B\vec{x} + \vec{f}$ 存在唯一解，则从任意 $\vec{x}^{(0)}$ 出发，迭代 $\vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{f}$ 收敛 $\Leftrightarrow$ $B^k \to 0$

证明： $B^k \to 0 \Leftrightarrow \| B^k \| \to 0 \Leftrightarrow \max \dfrac{\| B^k \vec{x} \|}{\| \vec{x} \|} \to 0$

But hey, you don't
seriously expect me to compute B^k
whenever I want to check
the convergence, do you?

立

成立

$\Leftrightarrow$ 从任意 $\vec{x}^{(0)}$ 出发”：取 $\vec{x}^{(0)} - \vec{x}^{*} = (0...1...0)^T$，则

$\vec{e}^{(k)} = B^k \vec{e}^{(0)}$ 则 $\vec{e}^{(k)} \to 0$ as $k \to \infty$ 第 i 位

$\Leftrightarrow \{ \vec{x}^{(k)} \}$ 收敛

对任意 $\varepsilon > 0$, 存在算子范数 $\|\cdot\|$ 使得 $\| A \| \leq \rho(A) + \varepsilon$。

证明：对 A 做 Jordan 分解，有 $P^{-1}AP = \begin{bmatrix} J_1 & & 0 \\ & \cdots & \\ 0 & & J_r \end{bmatrix}$，其中

$J_i = \begin{bmatrix} \lambda_i & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda_i \end{bmatrix}_{n_i \times n_i}$ ， $\sum_{i=1}^{r} n_i = n$ ， λ_i 为 A 的 eigen value。

令 $D = \begin{pmatrix} 1 & & & \\ & \delta & & \\ & & \delta^2 & \\ & & & \ddots \\ & & & & \delta^{n-1} \end{pmatrix}$，则有 $D^{-1}P^{-1}APD = \begin{bmatrix} \lambda_1 & \delta & & & & \\ & \ddots & \delta & & & \\ & & \lambda_1 & & & \\ & & & \ddots & & \\ & & & & \lambda_r & \delta & \\ & & & & & \ddots & \delta \\ & & & & & & \lambda_r \end{bmatrix}$

易证：$\| A \|_\delta = \| D^{-1}P^{-1}APD \|_1 = \max\limits_{1 \leq i \leq r}(|\lambda_i| + \delta) = \rho(A) + \delta$

是由 $\| \vec{x} \|_v = \| (PD)^{-1} \vec{x} \|_1$ 导出的算子范数。

所以只要取 $\delta < \varepsilon$，就有 $\| A \|_\delta < \rho(A) + \varepsilon$。 ∎

定理 $B^k \to 0 \iff \rho(B) < 1$

证明: "$\Rightarrow$" 若 λ 是 B 的eigenvalue, 则λ^k 是 B^k 的eigenvalue 。

则 $[\rho(B)]^k = [\max|\lambda|]^k = |\lambda_m{}^k| \le \rho(B^k) \le \|B^k\| \to 0$

$\Rightarrow \rho(B) < 1$ ✓

"$\Leftarrow$" 首先需要一个引理 /* Lemma */

由 $\rho(B) < 1$ 可知存在算子范数$\|\cdot\|$使得 $\|B\| < 1$。

$\|B^k\| \le \|B\|^k \to 0$ as $k \to \infty \iff B^k \to 0$ ∎

迭代从任意向量出发收敛 $\Longleftrightarrow$ $B^k \to 0$ $\Longleftrightarrow$ $\rho(B) < 1$

定理（充分条件）若存在一个矩阵范数使得 $\| B \| = q < 1$，则迭代收敛，且有下列误差估计：

① $$\| \vec{x}^* - \vec{x}^{(k)} \| \le \frac{q}{1-q} \| \vec{x}^{(k)} - \vec{x}^{(k-1)} \|$$

② $$\| \vec{x}^* - \vec{x}^{(k)} \| \le \frac{q^k}{1-q} \| \vec{x}^{(1)} - \vec{x}^{(0)} \|$$

证明：

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

$$\Rightarrow \| \vec{x}^* - \vec{x}^{(k)} \| \le q(\| \vec{x}^* - \vec{x}^{(k)} \| + \| \vec{x}^{(k)} - \vec{x}^{(k-1)} \|) \quad \checkmark$$

② $$\vec{x}^{(k)} - \vec{x}^{(k-1)} = B(\vec{x}^{(k-1)} - \vec{x}^{(k-2)}) = \ldots = B^{k-1}(\vec{x}^{(1)} - \vec{x}^{(0)})$$

$$\Rightarrow \| \vec{x}^{(k)} - \vec{x}^{(k-1)} \| \le q^{k-1} \| \vec{x}^{(1)} - \vec{x}^{(0)} \|$$

∎

定理 （充分条件）若A 为严格对角占优阵 /* strictly diagonally dominant matrix */ 则解 $A\vec{x} = \vec{b}$ 的Jacobi 和 Gauss - Seidel 迭代均收敛。

证明：首先需要一个引理 /* Lemma */

显然

若A 为SDD阵，则 $\det(A) \neq 0$，且所有的 $a_{ii} \neq 0$。

我们需要对 Jacobi 迭代和 Gauss-Seidel迭代分别证明：任何一个$|\lambda| \geq 1$ 都不可能是对应迭代阵的特征根，即 $|\lambda I - B| \neq 0$ 。

$$\begin{pmatrix} \lambda a_{11} & & a_{ij} \\ & \ddots & \\ a_{ij} & & \lambda a_{nn} \end{pmatrix}$$

Jacobi

关于Gauss-Seidel迭代的证明与此类似 。

另一种证明引理的方法利用 Geršgorin Disc Theorem 。

$U|$

是SDD阵