科学计算(Scientific Computing)

第八章　线性方程组的解法: 迭代解法

任课老师: Xiao-liang Cheng(程晓良)

Department of Mathematics, Zhejiang University, Hangzhou 310027, P.R. China.

since 2018

线性方程组的解法:迭代解法

1. 范数
   - 向量范数
   - 矩阵范数
   - 线性方程组的敏感性与条件数

线性方程组的解法:迭代解法

1. 范数
    - 向量范数
    - 矩阵范数
    - 线性方程组的敏感性与条件数
2. 迭代法
    - Jacobi迭代法
    - Gauss-Seidel迭代法
    - 超松弛迭代法(SOR)
    - 迭代法收敛条件与误差估计

线性方程组的解法:迭代解法

1. 范数
   - 向量范数
   - 矩阵范数
   - 线性方程组的敏感性与条件数
2. 迭代法
   - Jacobi迭代法
   - Gauss-Seidel迭代法
   - 超松弛迭代法(SOR)
   - 迭代法收敛条件与误差估计
3. 共轭斜量法*

向量范数

定义： 对$x \in R^n$ 是列向量, 若实值函数$N(x) =: ||x||$ 满足下列关系式:

(1) 正定性: 对所有的$x \in R^n$ 有 $||x|| \geq 0$, 而且 $||x|| = 0$ 当且仅当 $x = 0$;

(2) 齐次性: 对所有的$x \in R^n$ 和$\alpha \in R$ 有 $||\alpha x|| = |\alpha| \cdot ||x||$;

(3) 三角不等式: 对所有的$x, y \in R^n$ 有 $||x + y|| \leq ||x|| + ||y||$,

则称$N(x) =: ||x||$为R^n上的向量x的**范数(或模)**.

最常用的向量范数是下列三个:

$$||x||_1 = \sum_{j=1}^{n} |x_j| \quad ;$$

$$||x||_2 = (\sum_{j=1}^{n} |x_j|^2)^{\frac{1}{2}} \quad ;$$

$$||x||_\infty = \max_{1 \le j \le n} |x_j| \quad .$$

它们分别称为1范数,2范数和∞ 范数.

在有限空间里，可以引进各种范数，但由下面的等价定理.

定理: 设$\|\cdot\|_\alpha$和$\|\cdot\|_\beta$是$\mathbf{R}^n$上任意两种范数，则存在正常数C_1和C_2, 使得对一切$x \in \mathbf{R}^n$有

$$C_1\|x\|_\alpha \leq \|x\|_\beta \leq C_2\|x\|_\alpha.$$

可以证明常见范数之间等价的精确常数:

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2;$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty;$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty.$$

定义： 若矩阵$A \in R^{n\times n}$ 的某个实值函数$N(A) =: ||A||$ 满足下列条件:

(1) 正定性: 对所有的$A \in R^{n\times n}$ 有 $||A|| \geq 0$, 而且 $||A|| = 0$ 当且仅当 $A = 0$;

(2) 齐次性: 对所有的$A \in R^{n\times n}$ 和$\alpha \in R$ 有 $||\alpha A|| = |\alpha| \cdot ||A||$;

(3) 三角不等式: 对所有的$A, B \in R^{n\times n}$ 有 $||A + B|| \leq ||A|| + ||B||$;

(4) 相容性: 对所有的$A, B \in R^{n\times n}$ 有 $||AB|| \leq ||A|| \cdot ||B||$,

则称$N(A) =: ||A||$为$R^{n\times n}$上矩阵A的**范数**(**或模**).

注意到： $R^{n\times n}$上的矩阵范数可以看做是R^{n^2}上的向量范数.

矩阵范数和向量范数"协调":

$$\|Ax\| \leq \|A\|\,\|x\|.$$

对于向量1范数,2范数和∞ 范数$\|x\|_r$, 容易验证

$$\|A\|_r = \max_{x\neq 0}\frac{\|Ax\|_r}{\|x\|_r}$$

满足矩阵定义中的(1)(2)(3)(4).
可以计算

$$\|A\|_1 = \max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|;$$

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

$$\|A\|_\infty = \max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|.$$

对于一般的向量范数$\|x\|$，我们定义

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

它是一个矩阵范数. 这样的矩阵范数$\|A\|$称为向量范数$\|x\|$的诱导范数, 或算子范数.

极限的概念: 分量意义下

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x};$$

$$\lim_{k\to\infty} A^{(k)} = A.$$

谱半径

设$n \times n$阶矩阵A的特征值为$\lambda_i \ (i = 1, 2, \cdots, n)$, 则称

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

为矩阵A的谱半径.

矩阵范数和谱半径有如下关系:

$$\rho(A) \leq \|A\|.$$

利用谱半径，矩阵2范数可写成

$$\|A\|_2 = \sqrt{\rho(A^T A)}.$$

定理:

$$\lim_{k \to \infty} A^k = 0 \Longleftrightarrow \rho(A) < 1.$$

数据扰动的影响

方程组

$$\begin{cases} x_1 + x_2 = 2, \\ x_1 + 1.0001x_2 = 2 \end{cases}$$

的解为$x_1 = 2, x_2 = 0.$

若系数和右端有扰动，方程组

$$\begin{cases} x_1 + x_2 = 2, \\ x_1 + 0.9999x_2 = 1.9999 \end{cases}$$

的解为$x_1 = 1, x_2 = 1.$

解相差很大!

估计式(I):

方程$Ax=b$的右端扰动δb, 解为$x+\delta x$, 则

$$\frac{\|\delta x\|}{\|x\|} \le \|A^{-1}\|\,\|A\|\,\frac{\|\delta b\|}{\|b\|}.$$

证明: 由于

$$Ax=b, \quad A(x+\delta x)=b+\delta b.$$

因此

$$A\delta x=\delta b, \implies \|\delta x\| \le \|A^{-1}\|\,\|\delta b\|.$$

再利用

$$\|b\|=\|Ax\| \le \|A\|\,\|x\|, \implies \|x\| \ge \frac{\|b\|}{\|A\|}.$$

估计式(II):

方程$Ax=b$的系数矩阵扰动δA, 解为$x+\delta x$, 则

$$\frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\|\,\|A\|\,\frac{\|\delta A\|}{\|A\|}.$$

说明: 若$\|A^{-1}\|\,\|\delta A\| < 1$, 则$A+\delta A$非奇异. 由

$$Ax=b,\ \ (A+\delta A)(x+\delta x)=b,$$

推出

$$\delta x = -A^{-1}\,\delta A(x+\delta x)(\approx -A^{-1}\,\delta A\,x).$$

可推出

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\|A^{-1}\|\,\|A\|\,\frac{\|\delta A\|}{\|A\|}}{1-\|A^{-1}\|\,\|A\|\,\frac{\|\delta A\|}{\|A\|}}.$$

条件数

数$\|A^{-1}\|\,\|A\|$反映了方程组$Ax=b$的解对初始数据A,b扰动的灵敏度，可用来刻画方程组的病态程度.
我们称$\|A^{-1}\|\,\|A\|$为矩阵A的条件数，记作$cond(A)$, 即

$$cond(A) = \|A^{-1}\|\,\|A\|.$$

取不同的范数

$$cond(A)_\infty = \|A^{-1}\|_\infty\,\|A\|_\infty.$$

$$cond(A)_2 = \|A^{-1}\|_2\,\|A\|_2 = \sqrt{\frac{\lambda_{max}(A^TA)}{\lambda_{min}(A^TA)}}.$$

近似解可靠性判别法

什么是方程组$Ax = b$的解?

假设$\hat{x}$是一个近似解，计算残量$r = b - A\hat{x}$. 残量越小，$\hat{x}$越近似x?

例如. 方程组

$$\begin{cases} 0.780x_1 + 0.563x_2 = 0.217 \\ 0.913x_1 + 0.659x_2 = 0.254 \end{cases}$$

的准确解为$x_1 = 1, x_2 = -1$.

而$\hat{x}_1 = 0.341$, $\hat{x}_2 = -0.087$, 则残量$r = (-10^{-6}, 0)^T$. 换句话说

$$\begin{cases} 0.780x_1 + 0.563x_2 = 0.217 + 10^{-6} \\ 0.913x_1 + 0.659x_2 = 0.254 \end{cases}$$

的解就是$\hat{x}_1 = 0.341$, $\hat{x}_2 = -0.087$.

定理: 设x^*是方程组$Ax=b$ (A非奇异，$b \neq 0$)的精确解. 若$\hat{x}$是该方程组的近似解，其残量$r=b-A\hat{x}$, 则有

$$\frac{\|x^* - \hat{x}\|}{\|x^*\|} \leq Cond(A)\frac{\|r\|}{\|b\|}.$$

证明: 因为

$$\|b\| = \|Ax^*\| \leq \|A\|\,\|x^*\|.$$

$$\|\hat{x} - x^*\| = \|A^{-1}(A\hat{x} - b)\| \leq \|A^{-1}\|\,\|r\|.$$

易证.

病态方程组的一个测试例子

Hilbert矩阵：$H=(h_{ij})_{n\times n}$，

$$h_{i,j}=\frac{1}{i+j-1}.$$

例如:

- $n=10,\ Cond(H)=1.6025\times 10^{13};$
- $n=15,\ Cond(H)=2.6388\times 10^{17};$
- $n=20,\ Cond(H)=2.5072\times 10^{18}.$

研究: 预条件处理方法

求解$Ax=b$转化成求解$P^{-1}Ax=P^{-1}b$要求

(1). $Cond(P^{-1}A) \ll Cond(A)$;

(2). 求解$Py=d$比求解$Ax=b$容易的多!

解线性方程组的迭代法

基本思想: 构造一个向量序列$x^{(n)}$, 使其收敛至某个极限向量x^*, 则x^*就是要求的方程组

$$Ax = b$$

的准确解.

对迭代法来说一般有几个问题:
(1) 如何构造迭代序列?
(2) 构造的序列在什么情况下收敛?
(3) 如果收敛，收敛的速率如何?
(4) 近似解的误差估计.

Jacobi迭代法

解线性方程组$Ax=b$. 对于第i个方程解出x_i(把其它变量看作已知):

$$x_i = \frac{1}{a_{ii}}\Big(b_i - \sum_{j=1,j\neq i}^{n} a_{ij}x_j\Big), \quad i=1,2,\cdots,n.$$

将右边用第k步的值$x^{(k)}$代入，得到第$k+1$步的值$x^{(k+1)}$:

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

把线性方程组中的矩阵A 记为

$$A = D - L - U$$

其中

$$D = diag(a_{11}, a_{22}, \cdots, a_{nn})$$

$$L = \begin{bmatrix} 0 & & & & 0 \\ -a_{21} & 0 & & & \\ -a_{31} & -a_{32} & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ -a_{n1} & -a_{n2} & \cdots & -a_{n,n-1} & 0 \end{bmatrix}$$

$$U = \begin{bmatrix} 0 & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ & 0 & -a_{23} & \cdots & -a_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & 0 & -a_{n-1,n} \\ 0 & & & & 0 \end{bmatrix}$$

则Jacobi迭代法的矩阵表示:

$$Ax = b \Longleftrightarrow (D - L - U)x = b \Longleftrightarrow Dx = b + (L + U)x$$

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

即

$$x^{(k+1)} = B_J x^{(k)} + f_J\ , \quad (k = 0, 1, 2, \cdots)$$

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

Gauss-Seidel迭代法

解线性方程组$Ax = b$. 对于第i个方程解出x_i(把其它变量看作已知):

$$x_i = \frac{1}{a_{ii}}\Big(b_i - \sum_{j=1, j\neq i}^{n} a_{ij}x_j\Big), \quad i = 1, 2, \cdots, n.$$

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

将右边用第k步的值$x^{(k)}$代入，得到第$k+1$步的值$x^{(k+1)}$:

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

则Guass-Seidel迭代法的矩阵表示:

$$Ax = b \Longleftrightarrow (D - L - U)x = b \Longleftrightarrow Dx = b + (L + U)x$$

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

即

$$x^{(k+1)} = B_G x^{(k)} + f_G \;, \quad (k = 0, 1, 2, \cdots)$$

$$B_G = (D - L)^{-1}U \;, \quad f_G = (D - L)^{-1}b.$$

例子

方程组

$$\begin{cases} 10x_1 - 2x_2 - x_3 = 3, \\ -2x_1 + 10x_2 - x_3 = 15, \\ -x_1 - 2x_2 + 5x_3 = 10. \end{cases}$$

改写成

$$\begin{cases} x_1 = 0.3 + 0.2x_2 + 0.1x_3, \\ x_2 = 1.5 + 0.2x_1 + 0.1x_3, \\ x_3 = 2.0 + 0.2x_1 + 0.4x_2. \end{cases}$$

则Jacobi迭代法: 给定$x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^T$, 对$k = 0, 1, 2, \cdots$,

$$\begin{cases} x_1^{(k+1)} = 0.3 + 0.2x_2^{(k)} + 0.1x_3^{(k)}, \\ x_2^{(k+1)} = 1.5 + 0.2x_1^{(k)} + 0.1x_3^{(k)}, \\ x_3^{(k+1)} = 2.0 + 0.2x_1^{(k)} + 0.4x_2^{(k)}. \end{cases}$$

则Gauss-Seidel迭代法: 给定$x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^T$, 对$k = 0, 1, 2, \cdots$,

$$\begin{cases} x_1^{(k+1)} = 0.3 + 0.2x_2^{(k)} + 0.1x_3^{(k)}, \\ x_2^{(k+1)} = 1.5 + 0.2x_1^{(k+1)} + 0.1x_3^{(k)}, \\ x_3^{(k+1)} = 2.0 + 0.2x_1^{(k+1)} + 0.4x_2^{(k+1)}. \end{cases}$$

或

$$\begin{cases} x_1^{(k+1)} = 0.3 + 0.2x_2^{(k)} + 0.1x_3^{(k)}, \\ x_2^{(k+1)} = 1.56 + 0.04x_2^{(k)} + 0.12x_3^{(k)}, \\ x_3^{(k+1)} = 2.684 + 0.056x_2^{(k)} + 0.068x_3^{(k)}. \end{cases}$$

注意: 写算法程序无需这一步!下面收敛分析需要这一步!!

超松弛迭代法(SOR)

在Gauss-Seidel迭代中加入松弛因子$\omega > 0$

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

当$\omega = 1$时即为Gauss-Seidel迭代法. 也可看作加权平均:

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

矩阵形式:

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

即

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

即

$$x^{(k+1)} = B_S x^{(k)} + f_S,$$

$$B_S = (D - \omega L)^{-1}((1-\omega)D + \omega U);$$

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

$\omega = 1$即Gauss-Seidel迭代公式. 通常$0 < \omega < 2$.

当$1 < \omega < 2$称为超松弛迭代;

当$0 < \omega < 1$称为低松弛迭代.

一般迭代法

▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽▽

一般迭代法的形式: $k = 0, 1, 2, \cdots,$

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

矩阵B称为该迭代法的迭代矩阵, $x^{(0)}$为初始值.

△△△△△△△△△△△△△△△△△△△△△△△△△△△△

收敛? 收敛速度? 误差估计?

定理: 对任意初始向量$x^{(0)}$和常数项f, 由迭代格式

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

产生的向量序列$x^{(k)}$收敛且极限与初值无关的充分必要条件是

$$\rho(B) < 1$$

其中$\rho(B)$是矩阵B的谱半径.

证明: 必要性，收敛到x^*, 则取极限得$x^* = Bx^* + f$. 于是

$$x^{(k+1)} - x^* = B(x^{(k)} - x^*) = \cdots = B^{k+1}(x^{(0)} - x^*).$$

对任意初始向量均收敛到零，则B^k的极限是0.

充分性: 若$\rho(B) < 1$, 则$x = Bx + f$有唯一解，记x^*. 同理可证.

定理: 若迭代过程$x^{(k+1)} = Bx^{(k)} + f$中迭代矩阵B的某种算子范数$\|B\|_r = q < 1$, 则

(1) 对任意初始向量$x^{(0)}$, 该迭代过程均收敛于方程$x = Bx + f$的唯一解x^*;

(2)

$$\|x^* - x^{(k)}\|_r \leq \frac{1}{1-q}\|x^{(k+1)} - x^{(k)}\|_r;$$

(3)

$$\|x^* - x^{(k)}\|_r \leq \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|_r.$$

证明: (1)因为$\|B\|_r = q < 1$，则$\rho(B) < 1$, 因此方程$x = Bx + f$存在唯一解x^*.
又

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

$$\|x^{(k+1)} - x^*\|_r \leq q\|x^{(k)} - x^*\|_r \leq q^{k+1}\|x^{(0)} - x^*\|_r.$$

(2) 因为

$$\|x^{(k)} - x^*\|_r = \|x^{(k)} - x^{(k+1)} + x^{(k+1)} - x^*\|_r \leq \|x^{(k+1)} - x^{(k)}\|_r + \|x^{(k+1)} - x^*\|_r.$$

$$\|x^{(k)} - x^*\|_r \leq \|x^{(k+1)} - x^{(k)}\|_r + q\|x^{(k)} - x^*\|_r.$$

(3)

$$x^{(k+1)} - x^{(k)} = B(x^{(k)} - x^{(k-1)}),$$

$$\|x^{(k+1)} - x^{(k)}\|_r \leq q\|x^{(k)} - x^{(k-1)}\|_r.$$

注意到，迭代矩阵范数小于1是收敛的充分条件，不是必要的.

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad k = 0, 1, 2, \cdots$$

迭代矩阵B的$\|B\|_1 = 2$, $\|B\|_\infty = 2$, 但迭代法收敛.

判别收敛的几个常用条件

不可约: 如果矩阵A不能通过行的次序的调换和相应列的次序的调换成为

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

其中A_{11}, A_{22}为方阵，则称A不可约.

可约: 对于矩阵A, 若存在非空指标集S, T

$$S \cap T = \phi, \quad S \cup T = \{1, 2, \cdots, n\}$$

满足

$$a_{ij} = 0, \quad i \in S, j \in T.$$

对角优势(占优): 若矩阵$A=(a_{ij})_{n\times n}$满足

$$|a_{ii}| \geq \sum_{j\neq i, 1\leq j\leq n} |a_{ij}|, \quad i=1,2,\cdots,n$$

且至少有一个i使得上式中严格的不等号成立，则称矩阵A具有对角优势. 特别地，若所有i值上式中严格的不等号成立，则称矩阵A具有**严格对角优势(占优)**.

定理: 若方程组$Ax=b$的系数矩阵A具有严格对角优势，即满足

$$|a_{ii}| > \sum_{j\neq i, 1\leq j\leq n} |a_{ij}|, \quad i=1,2,\cdots,n$$

则方程组$Ax=b$有唯一解, 且对任意初始向量$x^{(0)}$, Jacobi和Gauss-Seidel迭代法都收敛.

用$\|\cdot\|_\infty$来证. 记

$$q = \max_{1\leq i\leq n} \frac{\sum\limits_{j\neq i, 1\leq j\leq n} |a_{ij}|}{|a_{ii}|} < 1.$$

设$\|x^{(k)}-x^*\|_\infty = t$, 按分量, 比如Jacobi迭代法,

$$|x_1^{(k+1)} - x_1^*| = |\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}}(x_j^{(k)} - x_j^*)| \leq qt.$$

$$|x_i^{(k+1)} - x_i^*| = |\sum_{j=2}^{n} \frac{a_{ij}}{a_{ii}} (x_j^{(k)} - x_j^*)| \leq qt.$$

从而可证明: $\|x^{(k+1)} - x^*\|_\infty \leq qt$
即

$$\|x^{(k+1)} - x^*\|_\infty \leq q\|x^{(k)} - x^*\|_\infty.$$

递推可证序列$x^{(k)}$收敛.

在Gauss-Seidel迭代法中，同理通过逐个分量证明，最后得到上面的式子.

定理: 若方程组$Ax=b$的系数矩阵A具有对角优势且不可约，则Jacobi迭代法必定收敛.

证明: 设迭代矩阵为B, 则只要证明$\rho(B)<1$即可. 设B的特征值为μ, 对应的的特征向量为$x=(x_1,x_2,\cdots,x_n)^T$, $|x_p|=\max_i|x_i|>0$，则

$$\mu a_{ii}x_i=-\sum_{j\neq i}a_{ij}x_j,\ \ i=1,2,\cdots,n, \text{let } i=p\Longrightarrow|\mu|\leq 1.$$

若$\mu=1$,则必有

$$\sum_{j\neq p}\frac{|a_{pj}|}{|a_{pp}|}=1,\quad \frac{|x_j|}{|x_p|}=1, \text{when } a_{pj}\neq 0.$$

故A不可能具有严格对角优势. 下面说明: 若$|\mu|=1$, A也不可能不可约且具有对角优势.

若A不可约且具有对角优势，则那些与p相连$(a_{pj}\neq 0)$的j必有$|x_j|=|x_p|$. 将所有下标分成两个集合: 所有模为$|x_p|$的指标记为S, 其余指标集为T, 在T的j, $|x_j|<|x_p|$.

显然T非空，因为具有对角优势的定义，至少有一个q, 有

$$|a_{qq}|>\sum_{j\neq q}|a_{qj}|$$

此时必有$|x_q|<|x_p|$, 否则$|x_q|=|x_p|$就矛盾.

于是，存在非空集合S,T, 有$a_{ij}=0, i\in S, j\in T$, 和不可约矛盾. 因此$|\mu|<1$.

定理: 若方程组$Ax=b$的系数矩阵为对称正定矩阵，则对任意初始向量$x^{(0)}$, Gauss-Seidel迭代法收敛.

证明. 设$A=D-L-L^T$, Gauss-Seidel迭代矩阵为$(D-L)^{-1}L^T$.

$$(D-L)^{-1}L^Tx=\lambda x.$$

则

$$L^Tx=\lambda(D-L)x.$$

记$(x,L\bar{x})=\alpha+\beta i$, $c_0=(Dx,\bar{x})>0$.

$$(Ax,\bar{x})=(Dx,\bar{x})-(Lx,\bar{x})-(L^Tx,\bar{x})=c_0-2\alpha>0.$$

由$L^Tx = \lambda(D-L)x$知，

$$|\lambda|^2 = |\frac{(L^Tx, \bar{x})}{((D-L)x, \bar{x})}|^2 = \frac{\alpha^2+\beta^2}{(c_0-\alpha)^2+\beta^2}.$$

$$|\lambda|^2 = \frac{\alpha^2+\beta^2}{c_0(c_0-2\alpha)+\alpha^2+\beta^2} < 1.$$

即

$$|\rho(B_G)| < 1.$$

Gauss-Seidel迭代法收敛.

关于SOR方法:

定理: 若方程组$Ax=b$的系数矩阵A的主对角线上元素$a_{ii}\neq 0$, 则SOR法求解方程组$Ax=b$收敛的必有条件是$0<\omega<2$.

证明. 计算迭代矩阵

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

的行列式

$$|det(B_\Omega)|=(1-\omega)^n=\lambda_1\lambda_2\cdots\lambda_n.$$

于是收敛

$$|1-\omega|\leq\rho(B_\omega)<1.$$

得到$0<\omega<2$, 即收敛的必要条件.

SOR方法收敛的充分条件:

定理: 设方程组$Ax=b$的系数矩阵是对称正定矩阵, 且$0<\omega<2$, 则对任意初始值向量$x^{(0)}$, SOR方法收敛.

定理: 设方程组$Ax=b$的系数矩阵为严格对角占优矩阵, 且$0<\omega<1$, 则对任意初始值向量$x^{(0)}$, SOR方法收敛.

使松弛法收敛最快的松弛因子叫最优松弛因子，记为ω_{opt}. 可以讨论最优松弛因子的选取问题. 比如A对称正定矩阵，特征值在$[\alpha,\beta]$之间，则迭代法

$$x^{(k+1)}=x^{(k)}-\omega(Ax^{(k)}-b),\ k=0,1,2,\cdots$$

的最优松弛因子是$\omega_{opt}=\frac{2}{\alpha+\beta}$.

分块矩阵的迭代格式

例. 设A, B是n阶矩阵，A非奇异. 考虑如下方程组

$$\begin{cases} Ax + By = b_1 \\ Bx + Ay = b_2 \end{cases}$$

求证: 若$\rho(A^{-1}B) < 1$, 则下列迭代格式(块Jacobi迭代)

$$\begin{cases} Ax^{(k+1)} = b_1 - By^{(k)} \\ Ay^{(k+1)} = b_2 - Bx^{(k)} \end{cases}$$

显然迭代矩阵

$$C = \begin{pmatrix} 0 & -A^{-1}B \\ -A^{-1}B & 0 \end{pmatrix}, \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = C \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} + \begin{pmatrix} A^{-1}b_1 \\ A^{-1}b_2 \end{pmatrix}.$$

设矩阵C的特征值为λ, 对应特征向量$(z_1, z_2)^T$, 则

$$A^{-1}Bz_2 = (-\lambda)z_1, \quad A^{-1}Bz_1 = (-\lambda)z_2.$$

如果$z_1 = z_2$, 则$(-\lambda)$就是$A^{-1}B$的特征值(注意$z_1 = z_2 \neq 0$);
如果$z_1 \neq z_2$, 则

$$A^{-1}B(z_1 - z_2) = \lambda(z_1 - z_2)$$

这说明λ一定是$A^{-1}B$的特征值. 因此当$\rho(A^{-1}B) < 1$时，推出$|\lambda(C)| < 1$. 因此块Jacobi迭代格式收敛.

收敛速度

单步法(Jacobi迭代法，Gauss-Seidel迭代法等)收敛，其极限x^*比满足

$$x^* = Mx^* + g.$$

误差$y_k = x^{(k)} - x^*$就满足

$$y_k = My_{k-1} = \cdots = M^k y_0,$$

从而

$$\|y_k\| \le \|M^k\| \cdot \|y_0\|.$$

因为初始误差未知，用$\|M^k\|$来刻画.

设M是单步线性迭代的迭代矩阵，则称

$$R(M^k) = R_k(M) = \frac{-ln\|M^k\|}{k} (= -ln\|M^k\|^{\frac{1}{k}})$$

为k次迭代的平均收敛速度.

记

$$R_\infty(M) = \lim_{k\to\infty} R_k(M)$$

称为渐进收敛速率.

定理: 设单步线性定常迭代矩阵为M, 则有

$$\lim_{k\to\infty} R_k(M) = -\ln \rho(M).$$

只要证明: $\lim_{k\to\infty} \|M^k\|^{\frac{1}{k}} = \rho(M)$即可.

共轭斜量法*

共轭斜量法(CG)是一种非线性迭代法，适用于方程组系数矩阵是实对称正定矩阵的情况. 理论上，它是直接法，至多n(矩阵阶数)步能得到准确解,但实际上只是用做迭代法解大型稀疏方程组.

命题: 二次函数

$$H(x) = \frac{1}{2}x^T Ax - b^T x$$

极小点x^*当且仅当x^*是线性方程组

$$Ax = b$$

的解.

证明: "$\Longrightarrow$", 只要$H(x)$求极值，求导数，二阶导数矩阵正定，易得.

"$\Longleftarrow$" 计算

$$H(x) = H(x^*+(x-x^*)) = H(x^*)+\frac{1}{2}(x-x^*)^T A(x-x^*) \geq H(x^*).$$

线搜索算法: 设向量p是给定的方向，已知$\hat{x}$, 在$\hat{x}+tp$上寻找一个新的点，$\tilde{x}$, 使得

$$H(\tilde{x}) \leq H(\hat{x}+tp), \ \forall t \in R.$$

容易得到

$$\tilde{x} = \hat{x} + \frac{(b - A\hat{x}, p)}{(p, p)} p.$$

最速下降法: 选取方向向量p为负梯度方向，即

$$p = b - A\hat{x}.$$

这是非线性一般问题求最值问题的算法，我们在第十章讨论.

任取初值x_0, 其剩余量$r_0 = b - Ax_0$. 可沿某个方向，比如$p_0 = r_0$, 求下一个近似值$x_1 = x_0 + \alpha_0 x_0$, 可以要求新的剩余向量$r_1 = b - Ax_1 = r_0 - \alpha_0 A p_0$与$p_0$正交，有

$$\alpha_0 = \frac{(r_0, p_0)}{(Ap_0, p_0)}.$$

再从x_1出发，沿着搜索方向p_1找x_2, $x_2 = x_1 + \alpha_1 p_1$. 要求搜索方向p_1与p_0是A正交的，即

$$(Ap_0, p_1) = 0.$$

即

$$p_1 = r_1 - \beta_0 p_0, \quad \beta_0 = \frac{(r_1, Ap_0)}{(p_0, Ap_0)}.$$

求x_2还要求r_2与p_1正交. 如此继续，可得如下算法.

算法：共轭斜量法

任取x_0，计算$r_0 = b - Ax_0$，$p_0 = r_0$.

对于$k = 0, 1, 2, \cdots$，计算

$$\alpha_k = \frac{(r_k, p_k)}{(Ap_k, p_k)};$$

$$x_{k+1} = x_k + \alpha_k p_k;$$

$$r_{k+1} = b - Ax_{k+1} = r_k - \alpha_k Ap_k;$$

若$r_{k+1} = 0$，则停止计算;

$$\beta_k = \frac{(r_{k+1}, Ap_k)}{(p_k, Ap_k)};$$

$$p_{k+1} = r_{k+1} - \beta_k p_k.$$

算法$r_{k+1} = 0$停止时，x_{k+1}为准确解.

定理: 只要$r_k \neq 0$, 上述算法可继续，直到$r_l = 0,\ l \leq n$. 对$k \leq l$, 由$r_0, Ar_0, \cdots, A^{k-1}r_0$张成的线性子空间，所谓Krylov子空间

$$V_k = span\{r_0, Ar_0, \cdots, A^{k-1}r_0\}$$

有正交基

$$r_0, r_1, \cdots, r_{k-1},\ (r_i, r_j) = 0,\ i \neq j$$

和A正交基

$$p_0, p_1, \cdots, p_{k-1},\ (Ap_i, p_j) = 0,\ i \neq j$$

它们之间有关系

$$(r_j, p_i) = 0,\ i < j,\ \ (r_j, p_j) = (r_j, r_j).$$

算法'：共轭斜量法

任取x_0，计算$r_0 = b - Ax_0$，$p_0 = r_0$.

对于$k = 0, 1, 2, \cdots$，计算

$$\alpha_k = \frac{(r_k, r_k)}{(Ap_k, p_k)};$$

$$x_{k+1} = x_k + \alpha_k p_k;$$

$$r_{k+1} = b - Ax_{k+1} = r_k - \alpha_k Ap_k;$$

若$r_{k+1} = 0$，则停止计算;

$$\beta_k = \frac{(r_{k+1}, r_{k+1})}{(r_k, r_k)};$$

$$p_{k+1} = r_{k+1} + \beta_k p_k.$$

算法$r_{k+1} = 0$停止时，x_{k+1}为准确解.

引进范数

$$\|x\|_A = (Ax, x)^{\frac{1}{2}}.$$

定理: 设$Ax^* = b$, A是实对称正定矩阵，则共轭斜量法迭代序列$x_0, x_1, \cdots$有下述极小性质

$$\|x^* - x_k\|_A = \min_{x \in x_0 + V_k} \|x^* - x\|_A.$$

事实上，x_k也是函数$H(x)$在$x_0 + V_k$上的最小点值.