6 解线性方程组的迭代法

/* 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？

定义1 对于给定的方程组x=Bx+f,用公式

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

一阶定常迭代

逐步代入求近似解的方法，称为迭代法

如果 $\lim_{k\to\infty} x^{(k)} = x^*$ 存在，称此迭代法收敛，

x*就是方程组的解，否则称此迭代法发散。

如何建立迭代方程?

M为分裂矩阵，非奇异

A=M-N

Ax=b $\Longleftrightarrow$ **Mx=Nx+b** $\Longleftrightarrow$ **X=M^{-1}Nx+M^{-1}b**

B=M^{-1}N=M^{-1}(M-A)=I-M^{-1}A

§ 6.1 迭代法的收敛性 /* Convergence of Iterative methods */

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

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

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

充分条件: … $\|B\|^k \to 0$ as k …

… $\| \to 0$ …

必要条 … ∞ $\Rightarrow$ …

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

等价于对任何
向量x,有
$\lim_{k\to\infty} A_k x = Ax$

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

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

定理1 $\lim\limits_{k\to\infty} A_k = A \Longleftrightarrow \lim\limits_{k\to\infty}\|A_k - A\| = 0$

定理2 $\lim\limits_{k\to\infty} A_k = 0 \Longleftrightarrow \lim\limits_{k\to\infty} A_k x = 0$

证明: $\|A_k x\| \le \|A_k\| \|x\|$

$\lim\limits_{k\to\infty} A_k = 0 \Longleftrightarrow \lim\limits_{k\to\infty}\|A_k\| = 0$

$\Longleftrightarrow \lim\limits_{k\to\infty} A_k x = 0$

反之,

$$\lim_{k\to\infty} A_k x = 0 \Longleftrightarrow \lim_{k\to\infty} A_k e_j = 0 \Longleftrightarrow \lim_{k\to\infty} A_k = 0$$

定理3 $\lim\limits_{k\to\infty} B^k = 0 \Longleftrightarrow \rho(B) < 1$

证明：1）$\Longrightarrow$ 2)反证

若B存在一个大于1的特征值λ， $|\lambda| \geq 1$

$B^k x = \lambda^k x \longrightarrow \|B^k x\| = |\lambda|^k \|x\| \longrightarrow \{B^k x\}$ 不收敛

2)$\Longrightarrow$1) $Bv_i = \lambda_i v_i$

$\forall x \in R^n \qquad x = \sum\limits_{i=1}^{n} c_i v_i$

$\Longrightarrow B^k x = \sum\limits_{i=1}^{n} c_i \lambda_i^k v_i = \lambda_1^k \sum\limits_{i=1}^{n} c_i (\frac{\lambda_i}{\lambda_1})^k v_i \to 0$

$\Longrightarrow B^k \to 0$

定理3 $\lim\limits_{k\to\infty} B^k = 0 \Longrightarrow \rho(B) < 1$

证明：对任何矩阵，存在非奇异

Jordan标准型

$$P^{-1}BP = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{bmatrix} = J$$

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

$$B = PJP^{-1} \Longrightarrow B^k = PJ^kP^{-1}$$

$$J^k = \begin{bmatrix} J_1^k & & & \\ & J_2^k & & \\ & & \ddots & \\ & & & J_r^k \end{bmatrix}$$

$$\lim_{k\to\infty} B^k = 0 \Longleftrightarrow \lim_{k\to\infty} J^k = 0 \Longleftrightarrow \lim_{k\to\infty} J_i^k = 0$$

$$J_i^k = (\lambda_i I + E_{t,1})^k = \sum_{j=0}^{k} C_k^j \lambda_i^{k-j} (E_{t,1})^j = \sum_{j=0}^{k} C_k^j \lambda_i^{k-j} E_{t,j}$$

$E_{t,0}$=I,

$$E_{t,k} = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \text{ t-k}$$

$$= \begin{bmatrix} \lambda_i^k & C_k^1 \lambda_i^{k-1} & C_k^2 \lambda_i^{k-2} & \cdots & C_k^{t-1} \lambda_i^{k-(t-1)} \\ & \lambda_i^k & C_k^1 \lambda_i^{k-1} & \ddots & C_k^{t-1} \lambda_i^{k-(t-2)} \\ & & \ddots & \ddots & C_k^2 \lambda_i^{k-2} \\ & & & \ddots & C_k^1 \lambda_i^{k-1} \\ & & & & \lambda_i^k \end{bmatrix}$$

利用极限 $\lim_{k\to\infty} k^r c^k = 0$ **(0<c<1,r ⩾ 0)** $\Longleftrightarrow$ $\lim_{k\to\infty} C_k^j \lambda^{k-j} = 0$

$\Longleftrightarrow$ **| λ |<1**

定理4 $$\lim_{k\to\infty}\left\|B^k\right\|^{\frac{1}{k}}=\rho(B)$$

证明: $$\rho(B)=[\rho(B^k)]^{\frac{1}{k}}\le\left\|B^k\right\|^{\frac{1}{k}}$$

$$B_\varepsilon=[\rho(B)+\varepsilon]^{-1}B \Longrightarrow \rho(B_\varepsilon)<1$$

$$\exists N, k>N, \left\|B_\varepsilon^k\right\|=\frac{\left\|B^k\right\|}{[\rho(B)+\varepsilon]^k}<1$$

$$\lim_{k\to\infty}B_\varepsilon^k=0$$

$$\rho(B)\le\left\|B^k\right\|^{\frac{1}{k}}\le\rho(B)+\varepsilon$$

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

$\Leftrightarrow \| B^k \vec{x} \| \to 0$ 对任意非零向量 $\vec{x}$ 成立

$\Leftrightarrow B^k \vec{x} \to \vec{0}$ 对任意非零向量 $\vec{x}$ 成立

$\Leftrightarrow$ 从任意 $\vec{x}^{(0)}$ 出发，记 $\vec{e}^{(0)} = \vec{x}^{(0)} - \vec{x}^*$，则

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

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

定理5（迭代法基本定理）方程组x=Bx+f

及一阶定常迭代法$x^{(k+1)}=Bx^{(k)}+f$，对任意初始向量x(0)，迭代收敛的充分必要条件是矩阵的谱半径 $\rho(B)<1$

A=I-B

证明: $\Longrightarrow$ $\rho(B)<1 \Longrightarrow Ax=f$有唯一解 $\Longrightarrow x^*=Bx^*+f$

$\Longrightarrow \varepsilon^{(k)}=x^{(k)}-x^*$

$\rho(B)<1 \Longrightarrow \lim_{k\to\infty} B^k = 0$

$\Longrightarrow \varepsilon(k) \to 0$

$\Longleftarrow \lim_{k\to\infty} x^{(k)} = x^*$ $\Longrightarrow$ x^*是$x=Bx+f$的根

$\Longrightarrow \varepsilon^{(k)}=x^{(k)}-x^*=B^k\varepsilon^{(0)}\to 0$

$\Longrightarrow \lim_{k\to\infty} B^k = 0 \Longrightarrow \rho(B)<1$

例3 考察迭代法解线性方程组 $x^{(k+1)} = Bx^{(k)} + f$ 的收敛性

情况一

$$B = \begin{pmatrix} 0 & \frac{3}{8} & \frac{-2}{8} \\ \frac{-4}{11} & 0 & \frac{1}{11} \\ \frac{-6}{12} & \frac{-3}{12} & 0 \end{pmatrix} \qquad f = \begin{pmatrix} \frac{20}{8} \\ \frac{33}{8} \\ 3 \end{pmatrix}$$

$$\det(\lambda I - B) = \lambda^2 + 0.03409\lambda + 0.03977 = 0$$

$$\lambda_1 = -0.3082, \lambda_{2,3} = 0.1541 \pm i0.3245$$

$$\rho(B) = \left|\lambda_{2,3}\right| = 0.3592 < 1$$

收敛

情况二

$$B = \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix} \qquad f = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

$$\det(\lambda I - B) = \lambda^2 - 6 = 0$$

$$\lambda_{1,2} = \pm\sqrt{6}$$

$$\rho(B) > 1$$

不收敛

定理5　设有方程组x=Bx+f,B∈Rn×n

一阶定常迭代法　$x^{(}$

若有B的某种算子范数$\|$

由定理4,
显然成立。

1)迭代法收敛。

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

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

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

证明: 2) $x^*-x^{(k+1)}=B(x^*-x^{(k)})=\ldots=B^k(x^*-x^{(0)})$

$$\|x^*-x^{(k+1)}\| \le \|B^k\|\|x^*-x^{(0)}\| \le \|B\|^k\|x^*-x^{(0)}\| = q^k\|x^*-x^{(0)}\|$$

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

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

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

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

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

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

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

例5 考察迭代法 $x^{(k+1)} = Bx^{(k)} + f$ **的收敛性，**

其中 $B = \begin{pmatrix} 0.9 & 0 \\ 0.3 & 0.8 \end{pmatrix}$ $f = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

解:

$$\|B\|_1 = 1.2, \|B\|_2 = 1.043, \|B\|_F = \sqrt{1.54}$$

$$\rho(B) = 0.9 < 1$$

收敛.

■迭代法的收敛速度

$$\varepsilon^{(k)} = B^k \varepsilon^{(0)} \Longrightarrow \|\varepsilon^{(k)}\| \le \|B\|^k \|\varepsilon^{(0)}\|$$

若**B**为对称矩阵 $\|\varepsilon^{(k)}\|_2 \le \|B\|_2^k \|\varepsilon^{(0)}\|_2 = [\rho(B)]^k \|\varepsilon^{(0)}\|_2$

欲使初始误差缩小10^{-s}所需的迭代次数

$$[\rho(B)]^k \le 10^{-s} \Longrightarrow k \ge \frac{s\ln 10}{-\ln \rho(B)}$$

定义4 $R_k(B) = -\ln \left\|B^k\right\|^{\frac{1}{k}}$ **---平均收敛速度**

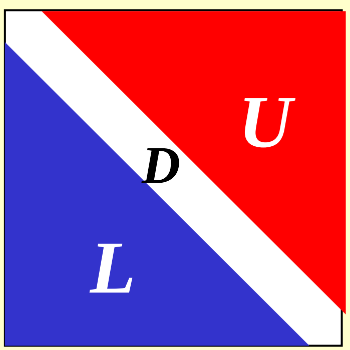
定义5 R(B)=-ln ρ (B)---渐进收敛速度，简称迭代法收敛速度。

§2 Jacobi 法、 Gauss - Seidel 法、SOR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:

$A =$ 

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

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

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

Jacobi 迭代阵

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

Algorithm: Jacobi Iterative Method

Solve $A\vec{x}=\vec{b}$ given an initial approximation $\vec{x}^{(0)}$.

Input: the number of equations and unknowns n; the matrix entries a[ ][ ]; the entries b[ ]; the initial approximation X0[ ]; tolerance TOL; maximum number of iteratio

Output: approximate solution X[ ]

Step 1 Set $k = 1$;

Step 2 Wh … steps 3-6

$X0$

; /* compute x_k */

a_{ii}

Step 4 If $\| X - X0 \|_\infty = \max_{1\le i \le n} | X_i - X0_i | < TOL$ then Output (X[ ]);

STOP; /* successful */

Step 5 For $i = 1, \ldots, n$ Set $X0$[ ] = X[ ]; /* update $X0$ */

Step 6 Set k ++;

Step 7 Output (Maximum number of iterations exceeded);

STOP. /* unsuccessful */

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**。

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。

➢ Gauss - Seidel Iterative Method

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

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

只存一组向量即可。

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

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

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

写成矩阵形式:

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

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

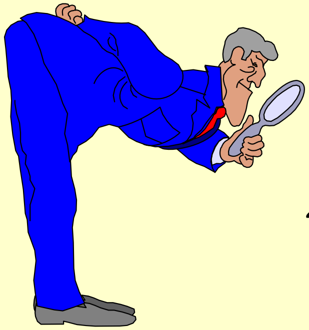
$$\Leftrightarrow \quad \vec{x}^{(k+1)} = \underbrace{-(D+L)^{-1}U}_{B}\vec{x}^{(k)} + \underbrace{(D+L)^{-1}\vec{b}}_{\vec{f}}$$

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。

松弛法 /* Relaxation Methods */



换个角度看Gauss - Seidel 方法:

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

$$= x_i^{(k)} + \frac{r_i^{(k+1)}}{a_{ii}} \quad \text{其中} \; r_i^{(k+1)} = b_i - \sum_{j<i} a_{ij} x_j^{(k+1)} - \sum_{j\ge i} a_{ij} x_j^{(k)}$$

/* residual */

相当于在 $x_i^{(k)}$ 的基础上加个余项生成 $x_i^{(k+1)}$。

下面令 $x_i^{(k+1)} = x_i^{(k)} + \omega \dfrac{r_i^{(k+1)}}{a_{ii}}$，希望通过选取合适的 ω 来加速收敛，这就是松弛法 /* Relaxation Methods */ 。

$0 < \omega < 1$	低松弛法 /* Under- Relaxation methods */
$\omega = 1$	Gauss - Seidel 法
$\omega > 1$	(渐次)超松弛法 /* Successive Over- Relaxation methods */

写成矩阵形式:

$$x_i^{(k+1)} = x_i^{(k)} + \omega \frac{r_i^{(k+1)}}{a_{ii}} = (1-\omega)x_i^{(k)} + \frac{\omega}{a_{ii}}[-\sum_{j<i} a_{ij} x_j^{(k+1)} - \sum_{j>i} a_{ij} x_j^{(k)} + b_i]$$

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

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

松弛迭代阵

定理 设 A 可逆，且 $a_{ii} \neq 0$，松弛法从任意 $\vec{x}^{(0)}$ 出发对某个 ω 收敛 $\Leftrightarrow \rho(H_\omega) < 1$。

例3 用SOR方法解方程组

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

- **用Jacobi求解，其迭代公式为**

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

➢用SOR求解，迭代公式为

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

定理 设Ax=b,其中A=D-L-U为非奇异矩

$J=D^{-1}(L+U)$

1) 解方程组的Jacobi迭代 … $)<1$

$L_w=(D-wL)^{-1}((1-w)D+wU)$

2) 解方程组的Gau … $)<1$

3) 解方程组的SOR迭代法收敛 $\Leftrightarrow$ $\rho(L_w)<1$

定义3（对角占优阵）

1)若A的元素满足 $$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

称A为严格对角占优阵。

2)若A的元素满足 $$|a_{ii}| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

且上式至少有一个不等式严格成立，称A为弱对角占优阵。

定义4（可约与不可约阵）若存在置换阵P使

$$P^T A P = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

称A为可约矩阵，否则，若不存在这样的置换阵P使上式成立，则称A为不可约矩阵。

定理 （充分条件）若A 为严格对角占优阵(SDD) /* 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: $B_J = D^{-1}(L+U)$

$|\lambda I - B$

关于Gauss-Seidel迭代的证明与此类似(p.252) 。

SDD阵

定理9 设Ax=b.

1) A为严格对角占优矩阵，则解Ax=b的Jacobi, Gauss-Seidel迭代收敛。

2) A为弱对角占优不可约矩阵，则解Ax=b的Jacobi迭代法，Gauss-Seidel迭代法收敛。

定理10 设矩阵A对称，且对角元大于零

1) Jacobi迭代收敛 $\Longleftrightarrow$ A, 2D-A正定

2）Gauss-Seidel收敛 $\Longleftrightarrow$ A正定

例8 线性方程组Ax=b, $A=\begin{bmatrix}1 & a & a\\ a & 1 & a\\ a & a & 1\end{bmatrix}$

证明1)当-0.5<a<1时，Gauss-Seidel迭代收敛;

2)当-0.5<a<0.5时，Jacobi迭代收敛。

证明：1)A对称，且对角元素大于0，。

Gauss-Seidel迭代收敛 $\Longleftrightarrow$ **A正定**

A的顺序主子式

$$\left.\begin{aligned}&\Delta_1=1>0 \quad \Delta_2=\begin{vmatrix}1 & a\\ a & 1\end{vmatrix}=1-a^2>0\\ &\Delta_3=\det A=1+2a^3-3a^2=(1-a)^2(1+2a)>0\end{aligned}\right\}\ \textbf{-0.5<a<1}$$

2）Jacobi迭代矩阵的谱半径<1时，收敛

$$\det(\lambda I-J)=\lambda^3-3\lambda^2 a+2a^3=(\lambda-a)^2(\lambda+2a)=0 \quad \rho(J)=|2a|<1$$

定理（SOR收敛必要条件）设 A 可逆，且 $a_{ii} \neq 0$，松弛法从任意 $\vec{x}^{(0)}$ 出发收敛 $\Rightarrow 0 < \omega < 2$。

证明：从 $H_\omega = (D + \omega L)^{-1}[(1-\omega)D - \omega U]$ 出发

利用 $\det(H_\omega) = \prod_{i=1}^{n} \lambda_i$，而且收敛 $\Leftrightarrow |\lambda_i| < 1$ 总成立

可知收敛 $\Rightarrow |\det(H_\omega)| < 1$

$$\det((D + \omega L)^{-1}) = \frac{1}{\det(D + \omega L)} = \prod_{i=1}^{n} \frac{1}{a_{ii}}$$

$$\det((1-\omega)D - \omega U) = (1-\omega)^n \prod_{i=1}^{n} a_{ii}$$

$$\longrightarrow \det(H_\omega) = (1-\omega)^n$$

$$\Rightarrow |\det(H_\omega)| = |1-\omega|^n < 1 \Rightarrow 0 < \omega < 2$$

∎

定理（**Ostrowski-Reich 充分条件**）若 A 对称正定，且有 $0<\omega<2$，则松弛法从任意 $\vec{x}^{(0)}$ 出发收敛。

证明：定理的结论就是要证明 $|\lambda|<1$

$$L_w y=\lambda y \iff (D-wL)^{-1}((1-w)D+wU)y=\lambda y$$

$$\Rightarrow ((1-w)D+wU)y=\lambda(D-wL)y$$

$$\Rightarrow (((1-w)D+wU)y,y)=$$

$$(Dy,y)=\sum_{i=1}^{n}a_{ii}|y_i|^2=\sigma>0$$

$$\Rightarrow \lambda=\frac{(1-w)(Dy,y)+w(Uy,y)}{(Dy,y)-w(Ly,y)}=\frac{(\sigma-w\sigma-\alpha w)+iw\beta}{(\sigma+\alpha w)+iw\beta}$$

$$-(Uy,y)=-(y,U^Ty)=-(y,Ly)=-\overline{(Ly,y)}=\alpha-i\beta$$

$$-(Ly,y)=\alpha+i\beta$$

$$|\lambda|^2=\frac{\boxed{(\sigma-w\sigma-\alpha w)^2}+w^2\beta^2}{\boxed{(\sigma+\alpha w)^2}+w^2\beta^2}\ \mathbf{<1}$$

$$(\sigma-w\sigma-\alpha w)^2-(\sigma+\alpha w)^2=w\sigma\boxed{(\sigma+2\alpha)}(w-2)<0$$

$0<(Ay,y)=((D-L-U)y,y)=\sigma+2\alpha$

定理10 A为严格对角占优矩阵（或弱对角占优不可约矩阵），0<w≤1,则SOR迭代法收敛。

定理 若 A 为对称正定三对角阵，则 $\rho(B_{G-S}) = [\rho(B_J)]^2 < 1$

且SOR的最佳松弛因子 /* optimal choice of ω for SOR method */

为 $\omega = \dfrac{2}{1+\sqrt{1-[\rho(B_J)]^2}}$，此时 $\rho(H_\omega) = \omega - 1$。

例：$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$，$\vec{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$，考虑迭代格式 $\vec{x}^{(k+1)} = \vec{x}^{(k)} + \omega(A\vec{x}^{(k)} - \vec{b})$

问：① ω 取何值可使迭代收敛？

② ω 取何值时迭代收敛最快？

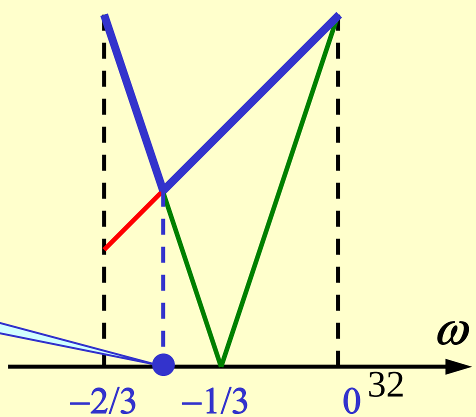
解：考察 $B = I + \omega A$ 的特征根 ➡ $\lambda_1 = 1+\omega$，$\lambda_2 = 1+3\omega$

① 收敛要求 $\rho(B)<1$ ➡ $-2/3 < \omega < 0$

② $\rho(B) = \max\{\,|1+\omega|, |1+3\omega|\,\}$

当 ω 取何值时最小？

$\omega = -1/2$



分块迭代法

A=D-L-U

非奇异矩阵

$$A=\begin{bmatrix} \boxed{A_{11}} & A_{12} & \cdots & A_{1q} \\ A_{21} & \boxed{A_{22}} & \cdots & A_{2q} \\ \vdots & \vdots & & \vdots \\ A_{q1} & A_{q2} & \cdots & \boxed{A_{qq}} \end{bmatrix} \qquad D=\begin{bmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{qq} \end{bmatrix}$$

$$L=\begin{bmatrix} 0 & & & \\ -A_{12} & 0 & & \\ \vdots & & \ddots & \\ -A_{q1} & -A_{q2} & \cdots & 0 \end{bmatrix} \qquad U=\begin{bmatrix} 0 & -A_{12} & \cdots & -A_{1q} \\ & 0 & & -A_{2q} \\ & & \ddots & \vdots \\ & & & 0 \end{bmatrix}$$

■块Jacobi迭代法

$$Dx^{(k+1)} = (L+U)x^{(k)} + b$$

$$\Downarrow$$

$$A_{ii}x^{(k+1)} = b_i - \sum_{\substack{j=1 \\ j\neq i}}^{q} A_{ij}x_j^{(k)} \quad = \mathbf{g_i}$$

■块SOR迭代法

$$(D - wL)x = [(1-w)D + wU]x + wb$$

$$\Downarrow$$

$$A_{ii}x^{(k+1)} = A_{ii}x_i^{(k)} + w(b_i - \sum_{j=1}^{i-1} A_{ij}x_j^{(k+1)} - \sum_{j=i}^{q} A_{ij}x_j^{(k)})$$

定理14 设Ax=b，其中A=D-L-U（分块形式）

1）若A为对称正定矩阵

2）0<w<2

则解Ax=b的块SOR迭代法收敛。

定理15 设A为非奇异T-矩阵，

$$A=\begin{bmatrix} D_1 & F_1 & & \\ E_2 & D_2 & F^2 & \\ & \ddots & \ddots & \ddots \\ & E_{q-1} & D_{q-1} & F_{q-1} \\ & & E_q & D_q \end{bmatrix}$$

D非奇， $J=I-D^{-1}A$ 则当 $\rho(J)<1$ 对0<w<2

$\Longrightarrow$ $\rho(L_w)<1$ $\quad w_{opt}=\dfrac{2}{1+\sqrt{1-[\rho(J)]^2}}$ $\quad \rho(L_{w_{opt}})=w_{opt}-1$

且 $$\rho(L_w)=\begin{cases} \dfrac{1}{4}[w\boxed{\mu}+\sqrt{w^2\mu^2-4(w-1)}]^2 & 0<w<w_{opt} \\ w-1 & w_{opt}\le w<2 \end{cases}$$

$$\mu=\rho(J)$$

共轭梯度法(CG) ---共轭斜量法

- 与方程组等价的变分问题

Ax=b, A对称正定 $A=(a_{ij})\in R^{n\times n}$ $b=(b_1,b_2,\cdots,b_n)^T$

$$\varphi(x)=\frac{1}{2}(Ax,x)-(b,x)=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j-\sum_{j=1}^{n}b_jx_j$$

性质： **1**） $\nabla\varphi(x)=Ax-b$

2） $\varphi(x+ay)=\frac{1}{2}(A(x+ay),x+ay)-(b,x+ay)$

$$=\varphi(x)+a(Ax-b,y)+\frac{a^2}{2}(Ay,y)$$

3） $x^*=A^{-1}b$ $\quad\varphi(x^*)=-\frac{1}{2}(b,A^{-1}b)=-\frac{1}{2}(Ax^*,x^*)$

$$\varphi(x)-\varphi(x^*)=\frac{1}{2}(Ax,x)-(Ax^*,x)+\frac{1}{2}(Ax^*,x^*)\quad=\frac{1}{2}(A(x-x^*),x-x^*)$$

定理16 **设A对称正定，则x*为线性方程组的解的充要条件为** $\varphi(x^*)=\min\limits_{x\in R^n}\varphi(x)$

证明: $\Rightarrow$ 设 $x^*=A^{-1}b$ **由A的正定性，得**

$$\varphi(x)-\varphi(x^*)=\frac{1}{2}(A(x-x^*),x-x^*)\geq 0 \Rightarrow \varphi(x)\geq\varphi(x^*)$$

$\Leftarrow$ 若 $\forall x\in R^n,\varphi(\overline{x})\leq\varphi(x) \Rightarrow \varphi(\overline{x})-\varphi(x^*)=0$

$$\frac{1}{2}(A(\overline{x}-x^*),\overline{x}-x^*)=0$$

由A正定，推出 $\overline{x}=x^*$

最速下降法

求 $\varphi(x)$ 极小点x*可转化为求一维问题的极小，

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$$

已知

s.t. $$\varphi(x^{(k+1)}) = \min_{\alpha \in R} \varphi(x^{(k)} + \alpha_k \boxed{p^{(k)}})$$

因为 $$\varphi(x^{(k)} + \alpha p^{(k)}) = \varphi(x^{(k)}) + \alpha(Ax^{(k)} - b, p^{(k)}) + \frac{\alpha^2}{2}(Ap^{(k)}, p^{(k)})$$

$$\frac{\partial \varphi(x^{(k)} + \alpha p^{(k)})}{\partial \alpha} = (Ax^{(k)} - b, p^{(k)}) + \alpha(Ap^{(k)}, p^{(k)}) = 0$$

$$\alpha_k = -\frac{(Ax^{(k)} - b, p^{(k)})}{(Ap^{(k)}, p^{(k)})}$$

求 $\varphi(x)$ 的下降法

若选一方向 $p^{(k)}$ 使得 $\varphi(x)$ 在点 $x^{(k)}$ 沿 $p^{(k)}$ 下降最快

$\Leftrightarrow$ $\varphi(x)$ 的负梯度方向 $-\nabla\varphi(x) = -(\frac{\partial\varphi(x)}{\partial x_1}, \frac{\partial\varphi(x)}{\partial x_2}, \cdots, \frac{\partial\varphi(x)}{\partial x_n})^T$

$$p^{(k)} = -\nabla\varphi(x) = -(Ax-b) = r^{(k)}$$

$\Rightarrow$

$$\left.\begin{aligned} \alpha_k &= -\frac{(r^{(k)}, r^{(k)})}{(Ar^{(k)}, r^{(k)})} \\ x^{(k+1)} &= x^{(k)} + \alpha_k r^{(k)} \end{aligned}\right\}$$

解线性方程组的最速下降法

$$(r^{(k+1)}, r^{(k)}) = (b - A(x^{(k)} + \alpha_k r^{(k)}), r^{(k)}) = (r^{(k)}, r^{(k)}) - \alpha_k (Ar^{(k)}, r^{(k)}) = 0$$

两个相邻的搜索方向是正交的

$$\lim_{k\to\infty} x^{(k)} = x^* = A^{-1}b \qquad \left\|x^{(k)} - x^*\right\|_A \le \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k \left\|x^{(0)} - x^*\right\|_A$$

共轭梯度法（CG）

若按方向 $\boxed{p^{(0)}, p^{(1)}, \cdots, p^{(k-1)}}$ 已进行k次的一维搜索，求得 $x^{(k)}$

不是具有正交性的 $r^{(0)}, r^{(1)}, \cdots, r^{(k-1)}$

求 $p^{(k)}$ 使 $x^{(k+1)}$ 更快的求得**x***

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$$

$x^{(0)} = 0$

$$\Rightarrow \quad x^{(k)} = \alpha_0 p^{(0)} + \alpha_1 p^{(1)} + \cdots + \alpha_{k-1} p^{(k-1)}$$

取 $p^{(0)} = r^{(0)}$

在第**k** 步选择一方向，同时满足以下两个式子

$$\varphi(x^{(k+1)}) = \min_{\alpha \in R} \varphi(x^{(k)} + \alpha p^{(k)})$$

$$\varphi(x^{(k+1)}) = \min_{x \in span\{p^{(0)}, \cdots, p^{(n)}\}} \varphi(x)$$

设 $x = y + \alpha p^{(k)} \quad y \in span\{p^{(0)}, p^{(1)}, \cdots, p^{(k-1)}\}$

$$\varphi(x) = \varphi(\boxed{y} + \boxed{\alpha} p^{(k)}) = \varphi(y) + \boxed{\alpha(Ay, p^{(k)})} - \alpha(b, p^{(k)}) + \frac{\alpha^2}{2}(Ap^{(k)}, p^{(k)})$$

分别求极小

交叉项

$$(Ay, p^{(k)}) = 0 \Rightarrow (Ap^{(j)}, p^{(k)}) = 0 \quad \mathbf{j=0,1,\ldots,k-1}$$

定义8 设A对称正定，若向量组$\{p^{(0)}, p^{(1)}, \cdots, p^{(k-1)}\}$满足

$$(Ap^{(j)}, p^{(k)}) = 0 \qquad i \neq j$$

称为一个A-共轭向量组或A-正交向量组

若取 $\{p^{(0)}, p^{(1)}, \cdots, p^{(k-1)}\}$ 是A共轭的，

$$\Rightarrow \min_{x\in span\{p^{(0)},p^{(1)},\cdots,p^{(k)}\}} \varphi(x^{(k+1)}) = \min_{\alpha,y} \varphi(y+\alpha p^{(k)})$$

$$= \boxed{\min_{y} \varphi(y)} + \boxed{\min_{\alpha}[\frac{\alpha^2}{2}(Ap^{(k)}, p^{(k)}) + \cancel{\alpha(Ay, p^{(k)})} - \alpha(b, p^{(k)})]}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$y = x^{(k)} \qquad\qquad \alpha_k = -\frac{(r^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})}$$

CG法中向量组 $\{p^{(0)}, p^{(1)}, \cdots, p^{(k-1)}\}$**的选择**

令 $p^{(0)} = r^{(0)}$ $\qquad p^{(k)} = r^{(k)} + \beta_{k-1} p^{(k-1)}$

由 $(p^{(k)}, Ap^{(k-1)}) = 0$ $\Rightarrow$ $\beta_{k-1} = -\frac{(r^{(k)}, Ap^{(k-1)})}{(p^{(k-1)}, Ap^{(k-1)})}$

A共轭

进一步简化 $r^{(k+1)} = b - Ax^{(k+1)} = r^{(k)} - \alpha_k Ap^{(k)}$

$$(r^{(k+1)}, p^{(k)}) = (r^{(k)}, p^{(k)}) - \alpha_k (Ap^{(k)}, p^{(k)}) = 0$$

$$\Rightarrow (r^{(k)}, p^{(k)}) = (r^{(k)}, r^{(k)} + \beta_{k-1} p^{(k-1)}) = (r^{(k)}, r^{(k)})$$

$$\Rightarrow \alpha_k = -\frac{(r^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})}$$

定理17 由CG算法得到的序列 $\{r^{(k)}\}$ $\{p^{(k)}\}$ 有以下性质

$$(r^{(i)}, r^{(j)}) = 0 (i \neq j)$$

$$(Ap^{(i)}, p^{(j)}) = (p^{(i)}, Ap^{(j)}) = 0 (i \neq j)$$

证明（用数学归纳法） $r^{(k+1)} = b - Ax^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)}$

$$\alpha_k = -\frac{(r^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})} \qquad \beta_{k-1} = -\frac{(r^{(k)}, Ap^{(k-1)})}{(p^{(k-1)}, Ap^{(k-1)})}$$

1) k=0

$p_0=r_0$

$$(r^{(0)}, r^{(1)}) = (r^{(0)}, r^{(0)}) - \alpha_0 (\boxed{r^{(0)}}, Ar^{(0)}) = 0$$

$$(p^{(1)}, Ap^{(0)}) = (r^{(1)}, Ar^{(0)}) + \beta_0 (r^{(0)}, Ar^{(0)}) = 0$$

假设至到k步，结论成立 **j=k** $(r^{(k+1)}, r^{(k)}) = 0$

j<k $(r^{(k+1)}, r^{(j)}) = (r^{(k)} - \alpha_k Ap^{(k)}, r^{(j)}) = \cancel{(r^{(k)}, r^{(j)})} - \alpha_k (Ap^{(k)}, r^{(j)})$

$$= -\alpha_k (Ap^{(k)}, \boxed{r^{(j)}}) = -\alpha_k (Ap^{(k)}, p^{(j)} - \beta_{j-1} p^{(j-1)}) = 0$$

$$p^{(k)} = r^{(k)} + \beta_{k-1} p^{(k-1)}$$

$$\beta_{k-1} = -\frac{(r^{(k)}, Ap^{(k-1)})}{(p^{(k-1)}, Ap^{(k-1)})}$$

证明A共轭

$$(p^{(k+1)}, Ap^{(k)}) = (r^{(k+1)}, Ap^{(k)}) + \boxed{\beta_k}(p^{(k)}, Ap^{(k)}) = 0$$

j<k $$(p^{(k+1)}, Ap^{(j)}) = (r^{(k+1)}, \boxed{Ap^{(j)}}) + \beta_k \cancel{(p^{(k)}, Ap^{(j)})} = 0$$

$$\uparrow$$

$$\boxed{r^{(k+1)} = r^{(k)} - \alpha_k Ap^{(k)} \Rightarrow \frac{1}{\alpha_j}(r^{(j)} - r^{(j+1)}) = Ap^{(j)}}$$

进一步简化β

$$\beta_k = -\frac{(r^{(k+1)}, Ap^{(k)})}{(p^{(k)}, Ap^{(k)})} = -\frac{(r^{(k+1)}, \alpha_j^{-1}(r^{(k)} - r^{(k+1)}))}{(r^{(k)} + \beta_{k-1}p^{(k-1)}, Ap^{(k)})} = \frac{(r^{(k+1)}, r^{(k+1)})}{\alpha_j(r^{(k)}, Ap^{(k)})} = \frac{(r^{(k+1)}, r^{(k+1)})}{(r^{(k)}, r^{(k)})}$$

CG算法

1) 任取x⁽⁰⁾,计算 $r^{(0)}=b-Ax^{(0)}$ **,取** $p^{(0)}=r^{(0)}$

2)对k=0,1,…,计算 $\alpha_k=-\dfrac{(r^{(k)},r^{(k)})}{(Ap^{(k)},p^{(k)})}$

$$x^{(k+1)}=x^{(k)}+\alpha_k p^{(k)}$$

$$r^{(k+1)}=r^{(k)}-\alpha_k Ap^{(k)} \qquad \beta_k=\frac{(r^{(k+1)},r^{(k+1)})}{(r^{(k)},r^{(k)})}$$

$$p^{(k+1)}=r^{(k+1)}+\beta_{k-1}p^{(k)}$$

3)若 $r^{(k)}=0$ **或** $\boxed{(p^{(k)},Ap^{(k)})=0}$ **停止计算** $x^{(k)}=x^*$

$$\Downarrow$$

$\Rightarrow p^{(k)}=0 \Rightarrow (r^{(k)},r^{(k)})=(r^{(k)},p^{(k)})=0 \Rightarrow r^{(k)}=0$

用CG算法求解线性方程组，理论上最多n步便可求得精确解。

CG算法有如下估计式 $\|x^{(k)}-x^*\|_A \le 2[\frac{\sqrt{K}-1}{\sqrt{K}+1}]^k \|x^{(0)}-x^*\|_A$

其中， $\|x\|_A=(x,Ax)^{\frac{1}{2}}$ $\quad K=cond(A)_2$

例11 用CG法解线性方程组 $\begin{cases} 3x_1+x_2=5 \\ x_1+2x_2=5 \end{cases}$

解：A对称正定，取 $x^{(0)}=0 \Rightarrow p^{(0)}=r^{(0)}=b-Ax^{(0)}=(5,5)^T$

$\Rightarrow \alpha_0=-\frac{(r^{(0)},p^{(0)})}{(Ap^{(0)},p^{(0)})}=\frac{2}{7} \Rightarrow x^{(1)}=x^{(0)}+\alpha_0 p^{(0)}=(\frac{10}{7},\frac{10}{7})^T \Rightarrow r^{(1)}=r^{(0)}-\alpha_0 Ap^{(0)}=(-\frac{5}{7},\frac{5}{7})^T$

$\Rightarrow \beta_0=\frac{(r^{(1)},r^{(1)})}{(r^{(0)},r^{(0)})}=\frac{1}{49} \Rightarrow p^{(1)}=r^{(1)}+\beta_0 p^{(0)}=(-\frac{30}{49},\frac{40}{49})^T \Rightarrow \alpha_2=-\frac{(r^{(1)},r^{(1)})}{(Ap^{(1)},p^{(1)})}=\frac{7}{10}$

$\Rightarrow x^{(2)}=(1,2)^T$