第六章 线性方程组的迭代解法

§1 基本迭代法

§2 共轭梯度法

§3 广义极小残差法

一个例子

求解 $\begin{cases} 8x_1 - 3x_2 + 2x_3 = 20 \\ 4x_1 + 11x_2 - x_3 = 33 \\ 6x_1 + 3x_2 + 12x_3 = 36 \end{cases}$ 即$Ax = b$, 精确解：$x^* = (3,2,1)^T$

方案1 Jacobi迭代法

$$Ax = b \Leftrightarrow \begin{cases} x_1 = \frac{1}{8}(20 + 3x_2 - 2x_3) \\ x_2 = \frac{1}{11}(-4x_1 + x_3 + 33) \\ x_3 = \frac{1}{12}(-6x_1 - 3x_2 + 36) \end{cases}$$

建立迭代：

$$\begin{cases} x_1^{(k+1)} = \frac{1}{8}(20 + 3x_2^{(k)} - 2x_3^{(k)}) \\ x_2^{(k+1)} = \frac{1}{11}(33 - 4x_1^{(k)} + x_3^{(k)}) \\ x_3^{(k+1)} = \frac{1}{12}(36 - 6x_1^{(k)} - 3x_2^{(k)}) \end{cases}$$

$$k = 0,1,2,\ldots$$

取初值：$x^{(0)} = (0,0,0)^T$

$$x^{(1)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)})^T = (5/2, 3, 3)^T$$

$$x^{(2)} = (x_1^{(2)}, x_2^{(2)}, x_3^{(2)})^T = (3/8, 20/11, 5/4)^T$$

••••••

$$x^{(10)} = (3.000032, 1.999838, 0.9998813)^T$$

矩阵形式：$Ax = b$

$$\Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & \frac{3}{8} & \frac{-2}{8} \\ \frac{-4}{11} & 0 & \frac{1}{11} \\ \frac{-6}{12} & \frac{-3}{12} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} \frac{20}{8} \\ \frac{33}{11} \\ \frac{36}{12} \end{pmatrix}$$

$$\Leftrightarrow x = \mathbf{M}x + g$$

$$\Leftrightarrow x^{(k+1)} = \mathbf{M}x^{(k)} + g$$

✍方案2　Gauss-Seidel迭代法

$$Ax=b \Leftrightarrow \begin{cases} x_1 = \frac{1}{8}(20+3x_2-2x_3) \\ x_2 = \frac{1}{11}(33-4x_1+x_3) \\ x_3 = \frac{1}{12}(36-6x_1-3x_2) \end{cases}$$

$$\begin{cases} x_1^{(k+1)} = \frac{1}{8}(20+3x_2^{(k)}-2x_3^{(k)}) \\ x_2^{(k+1)} = \frac{1}{11}(33-4x_1^{(k)}+x_3^{(k)}) \\ x_3^{(k+1)} = \frac{1}{12}(36-6x_1^{(k)}-3x_2^{(k)}) \end{cases}$$

$$\Rightarrow x^{(0)}=(0,0,0)^T, x^{(1)}=(2.5,3,3)^T$$

…… ……

$$x^{(10)}=(3.00003,1.99983,0.99988)^T$$

$$\Rightarrow \begin{cases} x_1^{(k+1)} = \frac{1}{8}(20+3x_2^{(k)}-2x_3^{(k)}) \\ x_2^{(k+1)} = \frac{1}{11}(33-4x_1^{(k)}+x_3^{(k)}) \\ x_3^{(k+1)} = \frac{1}{12}(36-6x_1^{(k)}-3x_2^{(k)}) \end{cases}$$

把已有的结果用上，

期待能有更好结果！

建立迭代：
$$\begin{cases} x_1^{(k+1)} = \frac{1}{8}(20+3x_2^{(k)}-2x_3^{(k)}) \\ x_2^{(k+1)} = \frac{1}{11}(33-4x_1^{(k+1)}+x_3^{(k)}) \\ x_3^{(k+1)} = \frac{1}{12}(36-6x_1^{(k+1)}-3x_2^{(k+1)}) \end{cases}$$

取初值：$x^{(0)}=(0,0,0)^T$

✍方案3　SOR迭代法

将G-S迭代等价变形：

$$\begin{cases} x_1^{(k+1)} = \frac{1}{8}(20+3x_2^{(k)}-2x_3^{(k)}) \\ x_2^{(k+1)} = \frac{1}{11}(33-4x_1^{(k+1)}+x_3^{(k)}) \\ x_3^{(k+1)} = \frac{1}{12}(36-6x_1^{(k+1)}-3x_2^{(k+1)}) \end{cases}$$

$$\Leftrightarrow \begin{cases} x_1^{(k+1)} = x_1^{(k)} + \frac{1}{8}(20-8x_1^{(k)}+3x_2^{(k)}-2x_3^{(k)}) \\ x_2^{(k+1)} = x_2^{(k)} + \frac{1}{11}(33-4x_1^{(k+1)}-11x_2^{(k)}+x_3^{(k)}) \\ x_3^{(k+1)} = x_3^{(k)} + \frac{1}{12}(36-6x_1^{(k+1)}-3x_2^{(k+1)}-12x_3^{(k)}) \end{cases}$$

记作： $x_i^{(k+1)} = x_i^{(k)} + \Delta x_i^{(k)} \ (i=1,2,3,4)$

建立迭代： $x_i^{(k+1)} = x_i^{(k)} + \omega \cdot \Delta x_i^{(k)} \ (i=1,2,3,4)$

即：

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \dfrac{\omega}{8}(20-8x_1^{(k)}+3x_2^{(k)}-2x_3^{(k)}) \\ x_2^{(k+1)} = x_2^{(k)} + \dfrac{\omega}{11}(33-4x_1^{(k+1)}-11x_2^{(k)}+x_3^{(k)}) \\ x_3^{(k+1)} = x_3^{(k)} + \dfrac{\omega}{12}(36-6x_1^{(k+1)}-3x_2^{(k+1)}-12x_3^{(k)}) \end{cases}$$

一　常用迭代法

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

1　Jacobi迭代法

假设$a_{ii} \neq 0$，则从第i个方程中解出x_i：$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ii}x_i + \cdots + a_{in}x_n = b_i$

$$\begin{cases} x_1 = \dfrac{1}{a_{11}}\left(b_1 - (a_{12}x_2 + \cdots + a_{1n}x_n)\right) \\ x_2 = \dfrac{1}{a_{22}}\left(b_2 - (a_{21}x_1 + \cdots + a_{2n}x_n)\right) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = \dfrac{1}{a_{nn}}\left(b_n - (a_{n1}x_1 + \cdots + a_{n,n-1}x_{n-1})\right) \end{cases} \Rightarrow \begin{cases} x_1^{(k+1)} = \dfrac{1}{a_{11}}\left(b_1 - (a_{12}x_2^{(k)} + \cdots + a_{1n}x_n^{(k)})\right) \\ x_2^{(k+1)} = \dfrac{1}{a_{22}}\left(b_2 - (a_{21}x_1^{(k)} + \cdots + a_{2n}x_1^{(k)})\right) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n^{(k+1)} = \dfrac{1}{a_{nn}}\left(b_n - (a_{n1}x_1^{(k)} + \cdots + a_{n,n-1}x_{n-1}^{(k)})\right) \end{cases}$$

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

2　G-S迭代法

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

3 SOR迭代法

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

$$\Updownarrow$$

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

令：

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

其中，ω为松弛因子；$\omega=1$，G-S迭代法；

$\omega>1$，超松弛迭代法；

$\omega<1$，低松弛迭代法。

二 迭代法的矩阵形式

对$Ax=b$，将A进行分解：

$$A=\begin{pmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & a_{33} & & \\ & & & \ddots & \\ & & & & a_{nn} \end{pmatrix}-\begin{pmatrix} 0 & & & & \\ -a_{21} & 0 & & & \\ -a_{31} & -a_{32} & 0 & & \\ \dots & \dots & & \ddots & \\ -a_{n1} & -a_{n2} & -a_{n3} & \dots & 0 \end{pmatrix}-\begin{pmatrix} 0 & -a_{12} & -a_{13} & \dots & -a_{1n} \\ & 0 & -a_{23} & \dots & -a_{2n} \\ & & 0 & & -a_{3n} \\ & & & \ddots & \vdots \\ & & & & 0 \end{pmatrix} \triangleq D-L-U$$

则 $Ax=b \Leftrightarrow (D-L-U)x=b \Leftrightarrow Dx=b+(L+U)x \Leftrightarrow x=D^{-1}\big(b+(L+U)x\big)$

1 Jacobi迭代法

$$\begin{aligned} x^{(k+1)} &= D^{-1}\big(b+(L+U)x^{(k)}\big) \\ &= D^{-1}b+D^{-1}(L+U)x^{(k)} \\ &= D^{-1}b+D^{-1}(D-A)x^{(k)} \\ &= (E-D^{-1}A)x^{(k)}+D^{-1}b \end{aligned}$$

2 G-S迭代法

$$x^{(k+1)} = D^{-1}\big(b+Lx^{(k+1)}+Ux^{(k)}\big)$$

$$\Rightarrow (D-L)x^{(k+1)} = \big(b+Ux^{(k)}\big)$$

$$\Rightarrow x^{(k+1)} = (D-L)^{-1}\big(b+Ux^{(k)}\big)$$

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

3 SOR迭代法

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

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

综上可见，三种迭代法可统一为：

$$x^{k+1}=\mathbf{M}x^{k}+g$$

其中，M为迭代矩阵：$\mathbf{M}_{\mathrm{J}}=(E-D^{-1}A);$

$$\mathbf{M}_{\text{G-S}}=(D-L)^{-1}U;$$

$$\mathbf{M}_{\mathrm{SOR}}=(D-\omega L)^{-1}\left[(1-\omega)D+\omega U\right];$$

-----迭代法的矩阵形式

三 迭代法的收敛性

1 基本概念

1）向量序列极限

设$x^{(k)}=\left(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)}\right)^T\in R^n\ (k=0,1,\cdots),x^*=\left(x_1^*,x_2^*,\cdots,x_n^*\right)^T\in R^n$,

若$\lim\limits_{k\to+\infty}x_i^{(k)}=x_i^*(i=1,\cdots,n)$，则称向量序列$\left\{x^k\right\}$收敛于$x^*$，记作：$\lim\limits_{k\to+\infty}x^k=x^*$.

2）迭代法的收敛性

设$x^{(k)}$是由迭代公式$x^{k+1}=\mathrm{M}x^k+g$生成的向量序列，若$\lim\limits_{k\to+\infty}x_i^{(k)}$存在，则称迭代法收敛，否则发散，且收敛时有$x^*=\lim\limits_{k\to+\infty}x^{(k)}$。

3）定理(补充)

$\lim\limits_{k\to+\infty}x^k=x^*\Leftrightarrow\lim\limits_{k\to+\infty}\|x^k-x^*\|=0$，其中$\|\cdot\|$为向量任一范数。

$$\lim_{k\to+\infty}x^k=x^*\Leftrightarrow\lim_{k\to+\infty}|x_i^{(k)}-x_i^{(*)}|=0,(i=1,\cdots,n)$$

$$\Leftrightarrow\lim_{k\to+\infty}\max_i|x_i^{(k)}-x_i^{(*)}|=0$$

范数的等价性

$$\Leftrightarrow\lim_{k\to+\infty}\|x^k-x^*\|_\infty=0\Leftrightarrow\lim_{k\to+\infty}\|x^k-x^*\|=0$$

2　判别条件

1）收敛基本定理【充要条件, P131−Th1.1】

$$x^{(k+1)}=\mathbf{M}x^{(k)}+g\text{对}\boxed{\text{任意初值}}\text{收敛}\Leftrightarrow \rho(M)=\max\left|\lambda_i\right|<1\text{。}$$

例1　判断迭代格式$x^{(k+1)}=\begin{bmatrix}0 & 2\\ 3 & 0\end{bmatrix}x^{(k)}+\begin{bmatrix}5\\ 5\end{bmatrix}$的收敛性。

解　$\because$ 迭代矩阵$\mathbf{M}=\begin{pmatrix}0 & 2\\ 3 & 0\end{pmatrix}$

$$\therefore\ f(\lambda)=|\lambda E-\mathbf{M}|=\begin{pmatrix}\lambda & -2\\ -3 & \lambda\end{pmatrix}=\lambda^2-6=0$$

$$\therefore \lambda_1=\sqrt{6},\lambda_2=-\sqrt{6}$$

$$\therefore \rho(\mathbf{M})=\max(|\lambda_1|,|\lambda_2|)=\sqrt{6}>1,$$ 迭代发散！

例2　设$Ax=b$,其中$A=\begin{pmatrix}1 & 0 & a\\ 0 & 1 & -2a\\ a & 2a & 1\end{pmatrix}$，$a\in R$。

求使Jacobi迭代过程收敛和发散的a的取值范围。

解

$$\because \mathrm{M_J}=E-D^{-1}A=\begin{pmatrix}0 & 0 & -a\\ 0 & 0 & 2a\\ -a & -2a & 0\end{pmatrix}$$

$$\therefore f(\lambda)=|\lambda E-\mathrm{M_J}|=\begin{pmatrix}\lambda & 0 & a\\ 0 & \lambda & -2a\\ a & 2a & \lambda\end{pmatrix}=\lambda(\lambda^2+3a^2)=0$$

$\therefore \mathrm{M_J}$的特征值为：$\lambda_1=0$, $\lambda_{2,3}=\pm\sqrt{3a^2}\,i=\pm\sqrt{3}\,|a|\,i$

$\therefore$迭代收敛$\Leftrightarrow \rho(\mathrm{M_J})<1 \Leftrightarrow \sqrt{3}\,|a|<1 \Leftrightarrow |a|<\dfrac{1}{\sqrt{3}}$

$\therefore$当$|a|<\dfrac{1}{\sqrt{3}}$时，Jacobi迭代收敛；当$|a|\geq\dfrac{1}{\sqrt{3}}$时，Jacobi迭代发散。

例3 对 $\begin{cases} x_1 - 2x_2 + x_3 = 5 \\ ax_1 + x_2 = 4, \quad a \in R, \\ 0.5x_2 + x_3 = 6 \end{cases}$ 1) 建立G-S迭代公式；2) 确定使G-S迭代收敛和发散的a的范围。

解 1) G-S迭代公式: $\begin{cases} x_1^{(k+1)} = 5 + x_2^{(k)} - x_3^{(k)} \\ x_2^{(k+1)} = 4 - ax_1^{(k+1)} \\ x_3^{(k+1)} = 6 - 0.5x_2^{(k+1)} \end{cases}$

2) $\mathrm{M_G} = (D-L)^{-1}U = \begin{pmatrix} 1 & & \\ a & 1 & \\ 0 & 0.5 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 2 & -1 \\ & 0 & 0 \\ & & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & -1 \\ 0 & -2a & a \\ 0 & a & -0.5a \end{pmatrix}$

$\therefore f(\lambda) = |\lambda E - M_J| = \lambda^2(\lambda + 2.5a) = 0$

$\therefore \mathrm{M_{G\text{-}S}}$的特征值为：$\lambda_{1,2} = 0$，$\lambda_3 = -2.5a$

$\therefore \rho(\mathrm{M_{G\text{-}S}}) = |-2.5a| = 2.5|a|$

$\therefore$迭代收敛 $\Leftrightarrow \rho(\mathrm{M_{G\text{-}S}}) < 1 \Leftrightarrow 2.5|a| < 1 \Leftrightarrow |a| < \dfrac{2}{5}$

$\therefore$当$|a| < \dfrac{2}{5}$时，迭代收敛；当$|a| \geq \dfrac{2}{5}$时，迭代发散。

2) 判定条件II【充分条件, P131-推论1.1】

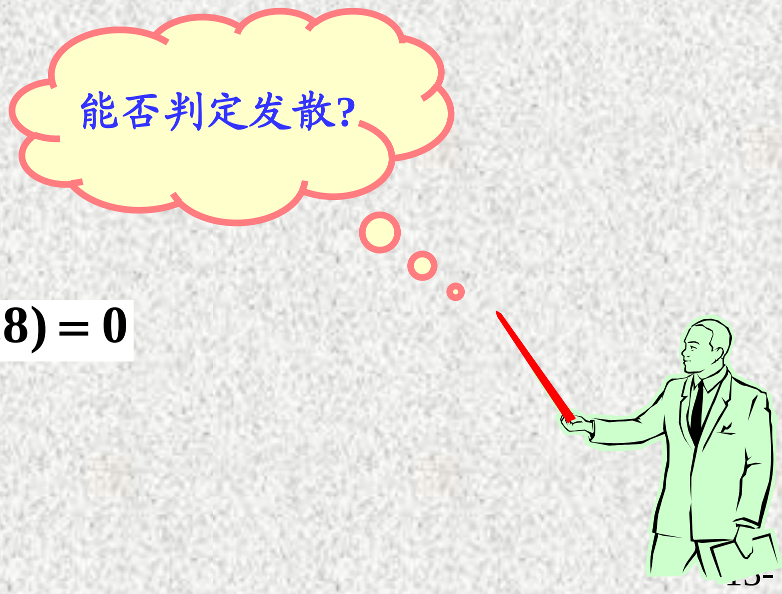
设迭代矩阵的某种范数满足$\|\mathbf{M}\| = q < 1$，则迭代$x^{(k+1)} = \mathbf{M}x^{(k)} + g$收敛，且有

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

例4 设迭代矩阵$\mathbf{M} = \begin{bmatrix} 0.9 & 0 \\ 0.3 & 0.8 \end{bmatrix}$,判断迭代$x^{(k+1)} = \mathbf{M}x^{(k)} + g$是否收敛?

易见: $\|\mathbf{M}\|_1 = \max\limits_{1\le j\le n}\sum\limits_{i=1}^{n}|\mathbf{M}_{ij}| = 1.2 > 1$

$$\|\mathbf{M}\|_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|\mathbf{M}_{ij}| = 1.1 > 1$$



但: $|\lambda\mathbf{E}\text{-}\mathbf{M}| = 0 \Rightarrow (\lambda - 0.9)(\lambda - 0.8) = 0$

$$\Rightarrow \rho(\mathbf{M}) = 0.9 < 1$$

$\Rightarrow$ 收敛

3) 判定条件III【充分条件, P135−Th1.4,1.5】

对$Ax=b$,①若A为严格对角占优方阵，则Jacobi和G-S迭代收敛；
②若A为对称正定矩阵，则G-S迭代收敛。

例5 判定下面方程组用Jacobi迭代是否收敛，如不收敛，能否将方程组变形使之收敛。

$$\begin{cases} x_1-10x_2+20x_3=11 \\ -10x_1+x_2-5x_3=-14 \\ x_1-5x_2-x_3=3 \end{cases}$$

解

$$A=\begin{pmatrix} 1 & -10 & 20 \\ -10 & 1 & -5 \\ 1 & -5 & -1 \end{pmatrix}$$

1）若直接建立Jacobi迭代，不满足判定条件III，故无法判定。

2）利用收敛基本判定定理判定

$$A=\begin{pmatrix} 1 & -10 & 20 \\ -10 & 1 & -5 \\ 1 & -5 & -1 \end{pmatrix}$$

$$\because \mathrm{M_J}=E-D^{-1}A=\begin{pmatrix} 0 & 10 & -20 \\ 10 & 0 & 5 \\ 1 & -5 & 0 \end{pmatrix}$$

$$\therefore f(\lambda)=|\lambda E-\mathrm{M_J}|=\begin{vmatrix} \lambda & -10 & 20 \\ -10 & \lambda & -5 \\ -1 & 5 & \lambda \end{vmatrix}=\lambda^3-55\lambda-1050$$

$$\because f(1)=1-55-1050<0$$

$$f(100)=1000000-1050-5500>0$$

$\therefore f(\lambda)=0$在$(1,100)$之间有根

$\therefore \rho(\mathrm{M_J})>1$，Jacobi迭代发散。

注：实际上，$\lambda_1=11.9520, \lambda_{2,3}=-5.9760\pm 7.22071i$

$\therefore \rho(\mathrm{M_J})=11.9520>1$，发散！

将方程等价变形：

$$\begin{cases} x_1 - 10x_2 + 20x_3 = 11 \\ -10x_1 + x_2 - 5x_3 = -14 \\ x_1 - 5x_2 - x_3 = 3 \end{cases} \Leftrightarrow \begin{cases} -10x_1 + x_2 - 5x_3 = -14 \\ x_1 - 5x_2 - x_3 = 3 \\ x_1 - 10x_2 + 20x_3 = 11 \end{cases}$$

则$A' = \begin{pmatrix} -10 & 1 & -5 \\ 1 & -5 & -1 \\ 1 & -10 & 20 \end{pmatrix}$，严格对角占优！

∴由判定条件III知，Jacobi迭代收敛

建立的Jacobi迭代格式为：

$$\begin{cases} x_1^{(k+1)} = (-14 - x_2^{(k)} + 5x_3^{(k)}) \big/ (-10) \\ x_2^{(k+1)} = (3 - x_1^{(k)} + x_3^{(k)}) \big/ (-5) \\ x_3^{(k+1)} = (11 - x_1^{(k)} + 10x_2^{(k)}) \big/ (20) \end{cases}$$

注：A对称正定时，不能保证Jacobi迭代收敛。

如，对称阵$A=\begin{bmatrix}1&a&a\\a&1&a\\a&a&1\end{bmatrix}$，$A$正定 $\Leftrightarrow\begin{cases}|1|>0\\ \begin{vmatrix}1&a\\a&1\end{vmatrix}=1-a^2>0\\ |A|=(2a+1)(1-a)>0\end{cases}$ $\Leftrightarrow -\dfrac{1}{2}<a<1$

$\therefore$ 当$-\dfrac{1}{2}<a<1$时，G-S迭代收敛；

但是，$\mathrm{M_J}=E-D^{-1}A=-\begin{bmatrix}0&a&a\\a&0&a\\a&a&0\end{bmatrix}$

$$\therefore |\lambda E-\mathrm{M_J}|=(\lambda+2a)(\lambda-a)^2$$

$\therefore$ J迭代收敛 $\Leftrightarrow \rho(\mathrm{M_J})<1 \Leftrightarrow \max(|-2a|,|a|)=2|a|<1 \Leftrightarrow -\dfrac{1}{2}<a<\dfrac{1}{2}$

这表明，当$-\dfrac{1}{2}<a<1$时，A对称正定，但Jacobi不收敛。

4) 判定条件IV【SOR迭代法, P136 - Th6,7】

- SOR迭代法收敛的必要条件是 $0<\omega<2$。

- 对 $Ax=b$，若 A 对称正定，则SOR迭代收敛的充要条件是 $0<\omega<2$。

第六章 线性方程组的迭代解法

§1 基本迭代法

§2 共轭梯度法

§3 广义极小残差法

共轭梯度法(CG法)是一种不需要任何参数求解对称正定线性方程组的方法，它是由50年代提出来的，近20年来有关研究得到了前所未有的发展，目前相关方法和理论已相当成熟，并成为求解大型稀疏线性方程组最受欢迎的一类方法。

一 Ritz变分原理

1 设$A \in R^{n\times n}$为对称正定矩阵，定义：

A内积： $(x,y)_A = x^T Ay, \forall x,y \in R^n$

A共轭(正交)： $(x,y)_A = 0$

A范数： $\|x\|_A = \sqrt{x^T Ax}$

2 Ritz变分原理(P138-Th2.1)

设$A \in R^{n\times n}$是对称正定矩阵，$b \in R^n$是常向量，定义

$$f(x) = \frac{1}{2}x^T Ax - b^T x = \frac{1}{2}(x,x)_A - (b,x), \quad x \in R^n$$

则

$$Ax^* = b \Leftrightarrow f(x^*) = \min_{x\in R^n} f(x)$$

证明　记$x^* = A^{-1}b$，可得

$$2f(x) + (x^*)^T Ax^* = (x - x^*)^T A(x - x^*)$$

即

$$2f(x) + \left\|x^*\right\|_A^2 = \left\|x - x^*\right\|_A^2$$

所以

$$f(x) = \min \Leftrightarrow \left\|x - x^*\right\|_A = \min \Leftrightarrow x = x^*$$

意义 把正定方程组$Ax = b$的求解问题转化为极小化函数$f(x)$的问题.

二 最速下降法

1 梯度 $\mathbf{grad}\vec{f} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)^T$

方向导数

$$\frac{\partial f}{\partial l} = \frac{\partial f}{\partial x}\cos\alpha + \frac{\partial f}{\partial y}\cos\beta = \left(\mathbf{grad}\vec{f}, \vec{n}\right) = \left|\mathbf{grad}\vec{f}\right| \cdot \left|\vec{n}\right| \cdot \cos\left\langle \vec{f}, \vec{n}\right\rangle$$

所以，梯度方向是函数变化(增加)最快的方向；

负梯度方向：$-\mathbf{grad}\vec{f} = \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}\right)^T$ 为f下降最快的方向.

2 最速下降法

任给初始点$x^{(0)}$,假设$x^{(k)}$已求得,在过点$x^{(k)}$方向为负梯度方向的直线

$$\pi_1 = \left\{x \in R^n \,\middle|\, x = x^{(k)} + \alpha r^{(k)}, \alpha \in R\right\}$$

上搜索一点$x^{(k+1)} = x^{(k)} + \alpha_k r^{(k)}$,使得

$$f(x^{(k+1)}) = f(x^{(k)} + \alpha_k r^{(k)}) = \min_{\alpha \in R} f(x^{(k)} + \alpha r^{(k)})$$

则显然序列$\left\{x^{(k)}\right\}$是$f(x)$的下降序列,即满足:

$$f(x^{(0)}) \geq f(x^{(1)}) \geq f(x^{(2)}) \geq \cdots$$

易知: $r^{(k)} = -\mathrm{grad}\, f(x^{(k)}) = b - Ax^{(k)}$

下面求α_k,使$f(x)$在直线π_1上的x^{k+1}处极小:

考察: $\varphi(\alpha) = f(x^{(k)} + \alpha r^{(k)})$

$$= \frac{1}{2}(x^{(k)} + \alpha r^{(k)})^T A(x^{(k)} + \alpha r^{(k)}) - b^T (x^{(k)} + \alpha r^{(k)})$$

$$= \frac{1}{2}\left((x^{(k)}\right)^T Ax^{(k)} + \frac{1}{2}\left((x^{(k)}\right)^T A\alpha r^{(k)} + \frac{1}{2}\alpha\left((r^{(k)}\right)^T A x^{(k)}$$

$$+\frac{1}{2}\alpha^2\left((r^{(k)}\right)^T A r^{(k)} - b^T x^{(k)} - \alpha b^T r^{(k)}$$

$$= \frac{1}{2}(r^{(k)}, r^{(k)})_A \alpha^2 + \boxed{\left(r^{(k)^T} Ax^{(k)} - b^T r^{(k)}\right)}\alpha + f(x^{(k)})$$

$$\boxed{\begin{aligned} & r^{(k)^T} Ax^{(k)} - r^{(k)^T} b \\ = {} & r^{(k)^T}\left(Ax^{(k)} - b\right) \\ = {} & -r^{(k)^T} \cdot r^{(k)} \end{aligned}}$$

$$= \frac{1}{2}(r^{(k)}, r^{(k)})_A \alpha^2 - (r^{(k)}, r^{(k)})\alpha + f(x^{(k)})$$

$\therefore$ 当 $\alpha_k = \dfrac{(r^{(k)}, r^{(k)})}{(r^{(k)}, r^{(k)})_A}$ 时, $\varphi(\alpha)$ 取得最小, 即 $x^{(k+1)} = x^{(k)} + \dfrac{(r^{(k)}, r^{(k)})}{(r^{(k)}, r^{(k)})_A} r^{(k)}$

3 算法2.1(最速下降法)

任取初值 $x^{(0)} \in R^n$，计算 $r^{(0)} = b - Ax^{(0)}$

for $k = 0, 1, 2, \cdots$

$$\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(r^{(k)}, r^{(k)})_A}$$

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

$$r^{(k+1)} = b - Ax^{(k+1)}$$

end

4 收敛性

- **引理2.0**

设 (λ, x) 为方阵 A 的特征对, $P(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n$ 为矩阵多项式,

则 $(P(\lambda), x)$ 是 $P(A)$ 的特征对.

引理2.1

设A是n阶对称正定矩阵, 特征值为$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$, $P(t)$为t的一个多项式,

则

$$\|P(A)x\|_A \le \max_{1\le i\le n} |p(\lambda_i)| \cdot \|x\|_A \text{ , } \forall x \in R^n$$

证明 设$\alpha_1, \alpha_2, \cdots, \alpha_n$为$A$的$n$个单位正交特征向量系，则对$\forall x \in R^n$，设

$$x = \sum_{i=1}^{n} k_i \alpha_i$$

则

$$\|x\|_A^2 = x^T A x = \left(\sum_{i=1}^{n} k_i \alpha_i\right)^T A \left(\sum_{i=1}^{n} k_i \alpha_i\right) = \sum_{i=1}^{n} k_i^2 \lambda_i$$

$$P(A)x = P(A)\sum_{i=1}^{n} k_i \alpha_i = \sum_{i=1}^{n} k_i P(A)\alpha_i = \sum_{i=1}^{n} k_i P(\lambda_i)\alpha_i$$

所以

$$\|p(A)x\|_A^2 = (p(A)x, Ap(A)x) = (\sum_{i=1}^{n} k_i p(\lambda_i)\alpha_i, \sum_{i=1}^{n} k_i \lambda_i p(\lambda_i)\alpha_i)$$

$$= \sum_{i=1}^{n} k_i^2 \lambda_i p^2(\lambda_i) \le \max_{1\le i\le n} |p(\lambda_i)|^2 \sum_{i=1}^{n} k_i^2 \lambda_i = \left(\max_{1\le i\le n} |p(\lambda_i)|\right)^2 \|x\|_A^2$$

定理2.2

设A是n阶对称正定矩阵，特征值为$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$，则由最速下降法产生的迭代序列$x^{(k)}$满足

$$\left\|x^{(k)} - x^*\right\|_A \le \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k \left\|x^{(0)} - x^*\right\|_A = \left(\frac{\kappa_2 - 1}{\kappa_2 + 1}\right)^k \left\|x^{(0)} - x^*\right\|_A$$

其中x^*是$f(x)$的极小点，$\kappa_2 = \mathrm{cond}_2(A) = \lambda_1/\lambda_n$.

证明　$\because\ f(x^{(k)}) \le f(x^{(k-1)} + \alpha r^{(k-1)}), \forall \alpha \in R$

$$\therefore (x^{(k)} - x^*)^T A(x^{(k)} - x^*) \le (x^{(k-1)} + \alpha r^{(k-1)} - x^*)^T A(x^{(k-1)} + \alpha r^{(k-1)} - x^*)$$

$$= [(I - \alpha A)(x^{(k-1)} - x^*)]^T A[(I - \alpha A)(x^{(k-1)} - x^*)]$$

记$p_\alpha(t) = 1 - \alpha t$，则上式为

$$\left\|x^{(k)} - x^*\right\|_A^2 \le \left\|P_\alpha(A)(x^{(k-1)} - x^*)\right\|_A^2$$

$$f(x) = \frac{1}{2} x^T A x - b^T x$$

由引理1，得

$$\left\|P(A)x\right\|_A \le \max_{1 \le i \le n} \left|p(\lambda_i)\right| \cdot \left\|x\right\|_A, \forall x \in R^n$$

$$\left\|x^{(k)} - x^*\right\|_A \le \max_{1 \le i \le n} \left|P_\alpha(\lambda_i)\right| \cdot \left\|x^{(k-1)} - x^*\right\|_A, \forall \alpha \in R$$

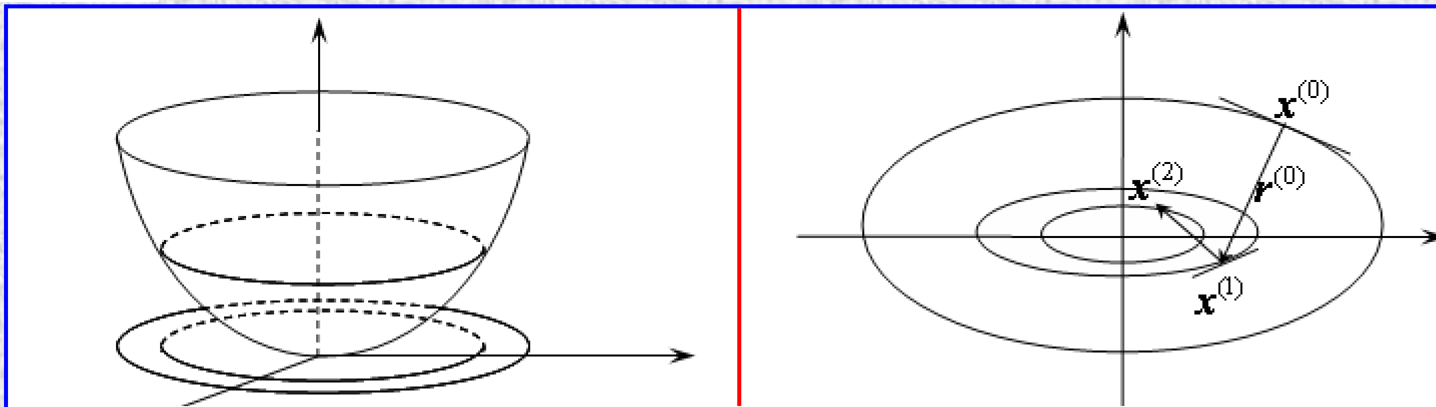
而
$$\min_{\alpha\in R}\max_{\lambda_n\le t\le\lambda_1}|p_\alpha(t)|=\min_{\alpha\in R}\max_{\lambda_n\le t\le\lambda_1}|1-\alpha t|=\frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}$$

所以
$$\|x^{(k)}-x^*\|_A\le\frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}\|x^{(k-1)}-x^*\|_A$$
$$\le\left(\frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}\right)^k\|x^{(0)}-x^*\|_A=\left(\frac{\kappa_2-1}{\kappa_2+1}\right)^k\|x^{(0)}-x^*\|_A$$

注： 1) $\because 0\le\dfrac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}<1$,所以方法对任意$x^{(0)}$都收敛;

但当$\lambda_1\gg\lambda_n$(即$\kappa_2=\mathrm{cond}_2(A)$很大时),收敛速度慢.

2) 示意图



负梯度方向是$f(x)$局部下降最快方向，而从整体看并非最佳.

三 共轭梯度法

➢ 公式推导

step1: 任取$x_0 \in R^n$,令下山方向$p_0 = r_0 = b - Ax_0$，则

$$\alpha_0 = \frac{(r_0, r_0)}{(r_0, r_0)_A} = \frac{r_0^T r_0}{p_0^T A p_0}, \quad x_1 = x_0 + \alpha_0 p_0, \quad r_1 = b - Ax_1$$

…… ……

stepk+1: 下山方向取为过x_k且由r_k和p_{k-1}张成的二维平面

$$\pi_2 = \{x = x_k + \xi r_k + \eta p_{k-1} : \xi, \eta \in R\}$$

找出使函数$\varphi(\xi, \eta)$下降最快的方向作为新的下山方向p_k：

令$\varphi(\xi, \eta) = f(x_k + \xi r_k + \eta p_{k-1})$

$$= \frac{1}{2}(x_k + \xi r_k + \eta p_{k-1})^T A(x_k + \xi r_k + \eta p_{k-1}) - b^T(x_k + \xi r_k + \eta p_{k-1})$$

直接计算得: $\begin{cases} \dfrac{\partial\varphi}{\partial\xi} = \xi r_k^T A r_k + \eta r_k^T A p_{k-1} - r_k^T r_k \\ \dfrac{\partial\varphi}{\partial\eta} = \xi r_k^T A p_{k-1} + \eta p_{k-1}^T A p_{k-1} \end{cases}$

令 $\dfrac{\partial\varphi}{\partial\xi} = \dfrac{\partial\varphi}{\partial\eta} = 0$,知 $\varphi(\xi, \eta)$ 在 π_2 内存在唯一极小点: $\tilde{x} = x_k + \xi_0 r_k + \eta_0 p_{k-1}$

其中 ξ_0, η_0 满足: $\begin{cases} \xi_0 r_k^T A r_k + \eta_0 r_k^T A p_{k-1} = r_k^T r_k \\ \xi_0 r_k^T A p_{k-1} + \eta_0 p_{k-1}^T A p_{k-1} = 0 \end{cases}$

$\because$ 当 $r_k \neq 0$ 时(即 $x_k \neq x^*$), 有 $\xi_0 \neq 0$

$\therefore$ 可取 $p_k = \dfrac{1}{\xi_0}\left(\tilde{x} - x_k\right) = r_k + \dfrac{\eta_0}{\xi_0} p_{k-1}$ 为新下山方向, 令 $\beta_{k-1} = \dfrac{\eta_0}{\xi_0} = -\dfrac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}$

此时: $x_{k+1} = x_k + \alpha_k p_k$ 其中: $\alpha_k = \dfrac{(r_k, p_k)}{(p_k, p_k)_A}$

$$\beta_k = -\frac{r_{k+1}^T A p_k}{p_k^T A p_k}$$

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

$$r_{k+1} = b - A x_{k+1}$$

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

算法2.3(共轭梯度法)

给定容差tol，任选初值$x^{(0)}$，计算$p^{(0)}=r^{(0)}=b-Ax^{(0)}$

for $k=1,2,\cdots$

$$\alpha_{k-1}=\frac{(r^{(k-1)},p^{(k-1)})}{(p^{(k-1)},p^{(k-1)})_A}=\frac{(r^{(k-1)},r^{(k-1)})}{(p^{(k-1)},p^{(k-1)})_A}$$

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

$$r^{(k)}=r^{(k-1)}-\alpha_{k-1}Ap^{(k-1)}$$

如果$\|r^{(k)}\|_2/\|b\|_2<tol$，则停止

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

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

end

收敛性 (TH2.6)

共轭梯度法计算的近似解$x^{(k)}$满足

$$\|x^{(k)}-x^*\|_A\le 2\left(\frac{\sqrt{\kappa_2}-1}{\sqrt{\kappa_2}+1}\right)^k\|x^{(0)}-x^*\|_A$$ 其中$\kappa_2=\mathrm{cond}_2(A)=\lambda_1/\lambda_n$.

Matlab： [x,Flag,RelRes,Iter] = pcg(A，b，tol，maxit)