线性方程组的数值解法

直接法	• 低阶稠密矩阵
迭代法	• 大型稀疏矩阵
极小化法	• 优化、非线性及病态问题

第六节　极小化方法

一、线性方程组的等价问题

二、最速下降法

三、共轭斜量法(共轭梯度法)

四、预条件共轭斜量法

一、与线性方程组等价的变分问题

设A对称正定，求解的线性方程组为

$$Ax=b \qquad (1)$$

其中$A=(a_{ij})\in R^{n\times n}, x=(x_1,x_2,\ldots,x_n)^T, b=(b_1,b_2,\ldots,b_n)^T$

对应的二次函数φ：$R^{n\times n}\to R$，称为模函数，定义为

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

例： 设 $A=\begin{bmatrix}1 & 2\\ 2 & 6\end{bmatrix}$, $b=\begin{bmatrix}4\\ 10\end{bmatrix}$

$$\varphi(x)=\frac{1}{2}(x_1^2+6x_2^2+4x_1x_2)-(4x_1+10x_2)$$

$$\varphi(x)=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_i x_j-\sum_{i=1}^{n}b_i x_i$$

φ有如下性质：

（1）对一切$x\in R^n$，有$\nabla\varphi(x)=\mathrm{grad}\varphi(x)=Ax-b=-r$

证：
$$\frac{\partial\varphi}{\partial x_i}=\sum_{j=1}^{n}a_{ij}x_j-b_i=-r_i,\qquad i=1,2,\ldots,n$$

$$grad\,\varphi(x)=\begin{bmatrix}\frac{\partial\varphi}{\partial x_1} & \cdots & \frac{\partial\varphi}{\partial x_n}\end{bmatrix}^T=Ax-b=-r$$

例：设 $A=\begin{bmatrix}1 & 2\\ 2 & 6\end{bmatrix}$, $b=\begin{bmatrix}4\\ 10\end{bmatrix}$

$$\varphi(x)=\frac{1}{2}(x_1^2+6x_2^2+4x_1x_2)-(4x_1+10x_2)$$

$$\frac{\partial\varphi}{\partial x_1}=x_1+2x_2-4=-r_1,\qquad \frac{\partial\varphi}{\partial x_2}=2x_1+6x_2-10=-r_2$$

(2) 对一切$x, y \in R^n, \alpha \in R$

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

$$=\frac{1}{2}(Ax, x)-(b, x)+\alpha(Ax, y)-\alpha(b, y)+\frac{\alpha^2}{2}(Ay, y)$$

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

$$\boxed{\varphi(x)=\frac{1}{2}(Ax, x)-(b, x)}$$

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

（3）设$x^* = A^{-1}b$为$Ax = b$的解，则

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

对一切$x \in R^n$，有

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

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

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

定理：设A对称正定，则$x^*=A^{-1}b$为$Ax=b$解的充分必要条件是：x^*是二次函数$\varphi(x)$的极小值点，即

$$x^*=A^{-1}b \Leftrightarrow \varphi(x^*)=\min_{x\in R^n}\varphi(x)$$

证明：

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

设$x^*=A^{-1}b$，由上式及A的正定性，

$$\varphi(x)-\varphi(x^*)\geq 0$$

所以有$\varphi(x^*)\leq\varphi(x)$，$\forall x\in R^n$，即$x^*$使$\varphi(x)$达到最小。

充分性：若x^*使$\varphi(x)$取极小值，则有

$$\operatorname{grad}\varphi(x)\Big|_{x=x^*}=Ax^*-b=0,$$

即x^*是方程组$Ax=b$的解。

$$\nabla\varphi(x)=\operatorname{grad}\varphi(x)=Ax-b$$

求二次函数$\varphi(x)$极小值点的一般方法是：

构造一个向量序列$\left\{x^{(k)}\right\}$，使$\varphi(x^{(k)}) \to \min \varphi(x)$

可以采取以下方法：

(1)任取一个初始向量$x^{(0)}$,

(2)构造迭代格式

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \qquad (k = 0, 1, \ldots)$$

其中p$^{(k)}$是搜索方向，α_k是搜索步长，

可以采取以下方法：

(1)任取一个初始向量$x^{(0)}$，

(2)构造迭代格式

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \qquad (k = 0, 1, \ldots)$$

其中$p^{(k)}$是搜索方向，α_k是搜索步长，

(3)选择$p^{(k)}$和α_k使得

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

则当$k \to \infty$时，有$\varphi(x^{(k)}) \to \varphi(x^*) = \min\limits_{x \in R^n} \varphi(x)$

(4)给出误差限ε，直到

$$\left\| x^{(k)} - x^{(k-1)} \right\| < \varepsilon \text{或} \left\| r^{(k)} \right\| = \left\| b - Ax^{(k)} \right\| < \varepsilon$$

迭代终止。

对迭代格式

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \qquad (k = 0, 1, \dots)$$

关键是要确定搜索方向$p^{(k)}$和搜索步长α_k。

(1) 确定搜索方向$p^{(k)}$

最速下降法：$p^{(k)}$取为模函数$\varphi(x)$减少最快的方向，

即：$\varphi(x)$的负梯度方向$-\mathrm{grad}(\varphi(x))$.

共轭梯度法：取A–共轭方向$p^{(k)}$。

(2) 确定搜索步长α_k

确定α_k使得从k步到k+1步是最优的，即：

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

这称为沿$p^{(k)}$方向的一维极小搜索。$\varphi(x^{(k+1)})$是局部极小。

对确定的搜索方向$p^{(k)}$，构造一个α的函数

$$F(\alpha)=\varphi(x^{(k+1)})=\varphi(x^{(k)}+\alpha p^{(k)})$$

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

$$=\frac{1}{2}(Ax^{(k)},x^{(k)})-(b,x^{(k)})+\alpha(Ax^{(k)},p^{(k)})-\alpha(b,p^{(k)})$$

$$+\frac{\alpha^2}{2}(Ap^{(k)},p^{(k)})$$

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

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

令$F'(\alpha)=0$，即：$-(r^{(k)},p^{(k)})+(\alpha Ap^{(k)},p^{(k)})=0$

$$F'(\alpha)=-(r^{(k)},p^{(k)})+\alpha(Ap^{(k)},p^{(k)})$$

得 $$\alpha=\frac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})}$$

$$\because F''(\alpha)=(Ap^{(k)},p^{(k)})>0,\qquad (A正定)$$

取 $\boxed{\alpha_k=\dfrac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})}}$， α_k是$\varphi(x^{(k)}+\alpha p^{(k)})$下降的极小值点，

即α_k是$\mathbf{k}\to\mathbf{k+1}$步的最优步长。

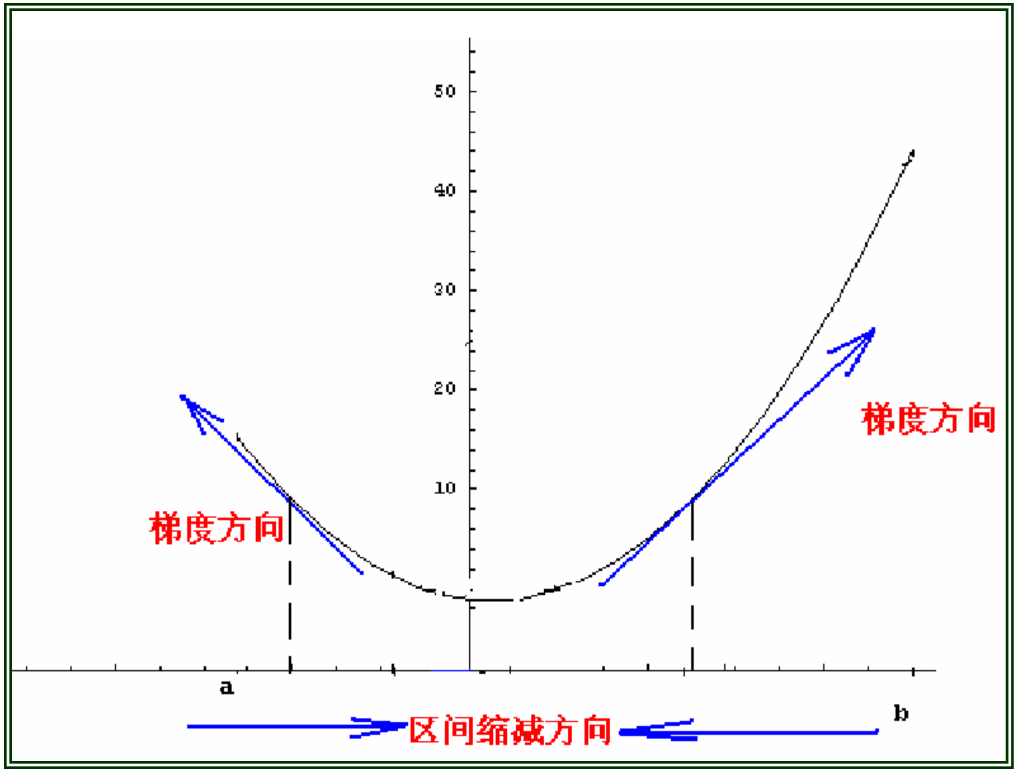
二、最速下降法

最速下降法：

$p^{(k)}$取模函数$\varphi(x)$减少最快的方向，

即：$\varphi(x)$的负梯度方向

$$p^{(k)} = -\mathrm{grad}(\varphi(x^{(k)}))$$
$$= r^{(k)}$$



例： 设 $A=\begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$, $b=\begin{bmatrix} 4 \\ 10 \end{bmatrix}$

$$\varphi(x)=\frac{1}{2}(x_1^2+6x_2^2+4x_1x_2)-(4x_1+10x_2)$$

$$\frac{\partial\varphi}{\partial x_1}=x_1+2x_2-4=-r_1, \qquad \frac{\partial\varphi}{\partial x_2}=2x_1+6x_2-10=-r_2$$

$$\varphi(x^*)=\min\varphi(x)=-9, \qquad x^*=\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

最速下降算法:

(1) 选取 $x^{(0)} \in R^n$

(2) 对 $k = 0, 1, 2, \ldots.$

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

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

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

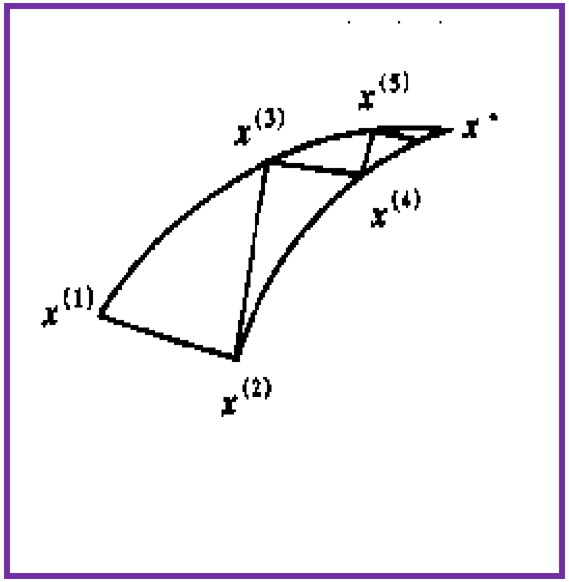
(3)当 $\left\| x^{(k+1)} - x^{(k)} \right\| < \varepsilon$ 时，终止迭代。

$$p^{(k)} = -grad(\varphi(x^k)) = r^{(k)}$$

$$\alpha = \frac{(r^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})}$$

最速下降法的特点

- （1） 初始点可任选，每次迭代计算量小，存储量少，程序简短。即使从一个不好的初始点出发，开始的几步迭代，目标函数值下降很快，然后慢逼近局部极小点
- （2） 任意相邻两点的搜索方向是**正交**的，它的迭代路径为绕道逼近极小点。当迭代点接近极小点时，步长变得很小，越走越慢



三、共轭斜量法(CG)(共轭梯度法)

设按方向$p^{(0)},p^{(1)},\ldots,p^{(k-1)}$已进行k次一维搜索，求得$x^{(k)}$，下一步就是确定$p^{(k)}$，再求解一维极小化问题

$$\min_{\alpha}\varphi(x^{(k)}+\alpha p^{(k)})$$

可得 $$\alpha=\alpha_k=\frac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})} \qquad (7)$$

现在考虑$p^{(0)},p^{(1)},\ldots$取什么方向

定义: A对称正定，若R^n中向量组$\left\{p^{(0)},\ldots,p^{(l)}\right\}$满足

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

则称它为R^n中的一个$A-$共轭向量组。

计算$p^{(k)}$：

取$p^{(0)}=r^{(0)}$，$p^{(k)}$就取为与$p^{(0)},\ldots,p^{(k-1)}$ $A-$共轭的向量，这样的向量不是唯一的，CG法中取$p^{(k)}$为$r^{(k)}$与$p^{(k-1)}$的线性组合，

设
$$p^{(k)}=r^{(k)}+\beta_{k-1}p^{(k-1)} \tag{13}$$

可以证明这样得到的向量序列$\{p^{(k)}\}$是一个$A-$共轭向量组.

经化简得到
$$\beta_{k\text{-}1}=\frac{(r^{(k)},r^{(k)})}{(r^{(k\text{-}1)},r^{(k\text{-}1)})}$$

计算α_k：

由$A-$共轭及$p^{(k)}$的定义得到

$$\alpha_k=\frac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})}=\frac{(r^{(k)},r^{(k)})}{(Ap^{(k)},p^{(k)})} \qquad （16）$$

CG算法：

$$(1)x^{(0)} \in R^{(n)}$$

$$(2)r^{(0)} = b - Ax^{(0)}, p^{(0)} = r^{(0)}$$

$$(3)k = 0, 1, \ldots,$$

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

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

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

$$\beta_k = \frac{(r^{(k+1)}, r^{(k+1)})}{(r^{(k)}, r^{(k)})}$$

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

在计算过程中，若遇 $r^{(k)}$=0，或$(p^{(k)}, Ap^{(k)})$=0时，计算终止, $x^{(k)} = x^*$.

(1) 剩余向量相互正交，而R^n中至多有n个相互正交的非零向量，所以$r^{(0)}, r^{(1)}, ..., r^{(n)}$中至少有一个向量为零。若$r^{(k)}=0$，则$x^{(k)} = x^*$.

(2) 实际计算中，由于舍入误差的影响，n步内得不到准确解，故还需继续迭代。一般因$p^{(0)}, p^{(1)}, ..., p^{(n-1)}$是一组$A$−共轭向量组，继续迭代时，要取$x^{(n)} = x^{(0)}$.

(3) 由误差估计式

$$\left\| x^{(k)} - x^* \right\|_A \le 2\left[\frac{\sqrt{cond(A)_2} - 1}{\sqrt{cond(A)_2} + 1}\right]^k \left\| x^{(0)} - x^* \right\|_A$$

当A的条件数很小时，共轭斜量法收敛很快，但当A是病态严重的矩阵时，共轭斜量法收敛速度很慢。可采用预处理技术，降低A的条件数。

四、预条件共轭斜量法(PCG)

寻找一个非奇异矩阵C，使$\overline{A}=C^{-1}AC^{-T}$的条件数比原系数矩阵A的条件数得到改善.

$$Ax=b \Leftrightarrow C^{-1}AC^{-T}C^{T}x=C^{-1}b \Leftrightarrow \overline{A}\overline{x}=\overline{b}$$

其中 $\overline{A}=C^{-1}AC^{-T}, \overline{b}=C^{-1}b, \overline{x}=C^{T}x,$

令$M=CC^T$称为预优矩阵，当M接近A时，$\overline{A}$接近单位阵，$\mathrm{cond}(\overline{A})_2$接近1，对$\overline{A}\overline{x}=\overline{b}$用共轭斜量法求解，可达到加速的目的。

$$M=CC^T \approx A,$$

$$\overline{A}=C^{-1}AC^{-T} \approx C^{-1}MC^{-T}=C^{-1}CC^{T}C^{-T}=I，\ cond(\overline{A})_2 \approx 1$$

预条件共轭斜量法

1、计算$\bar{A}=C^{-1}AC^{-T},\bar{b}=C^{-1}b$

2、解$\bar{A}\bar{x}=\bar{b}$,得$\bar{x}^{(k)}$,

(1)$\bar{x}^{(0)}\in R^{(n)}$

(2)$\bar{r}^{(0)}=b-\bar{A}\bar{x}^{(0)},\bar{p}^{(0)}=\bar{r}^{(0)}$

(3)$k=0,1,\ldots,$

$$\bar{\alpha}_k=\frac{(\bar{r}^{(k)},\bar{r}^{(k)})}{(\bar{A}\bar{p}^{(k)},\bar{p}^{(k)})}$$

$$\bar{x}^{(k+1)}=\bar{x}^{(k)}+\bar{\alpha}_k\bar{p}^{(k)},$$

$$\bar{r}^{(k+1)}=\bar{r}^{(k)}-\bar{\alpha}_k\bar{A}\bar{p}^{(k)},$$

$$\bar{\beta}_k=\frac{(\bar{r}^{(k+1)},\bar{r}^{(k+1)})}{(\bar{r}^{(k)},\bar{r}^{(k)})}$$

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

3、$x^{(k)}=C^{-T}\bar{x}^{(k)}$.

实际计算，可通过变换，转化成用原方程组的量来计算

PCG算法：

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

(2)解方程组$Mz^{(0)} = r^{(0)}$,求出$z^{(0)}$，令$p^{(0)} = z^{(0)}$

(3)$k = 0, 1, \ldots,$

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

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

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

解方程组$\mathrm{M}z^{(k+1)} = r^{(k+1)}$

$$\beta_k = \frac{(z^{(k+1)}, r^{(k+1)})}{(z^{(k)}, r^{(k)})} \qquad p^{(k+1)} = z^{(k+1)} + \beta_k p^{(k)}$$

(4)直到$\left\| r^{(k+1)} \right\| < \varepsilon$,输出$x^{(k+1)}$。

预优矩阵的选取

预优矩阵$M=CC^T$应满足:

(1)M是对称正定的矩阵;(2)方程组$Mz^{(k)}=r^{(k)}$容易求解.

预条件共轭斜量法MATLAB的三种调用格式:

1.不用预优矩阵的共轭斜量法

x=pcg(a,b,tol,kmax)

2.用预优矩阵的共轭斜量法

(1)x=pcg(a,b,tol,kmax,m)

(2) r=chol(m)

x=pcg(a,b,tol,kmax,r',r,x0)

3.未给定预优矩阵的共轭斜量法

r=cholinc(sa,'0')

x=pcg(a,b,tol,kmax,r',r,x0)