（一） 最速下降法

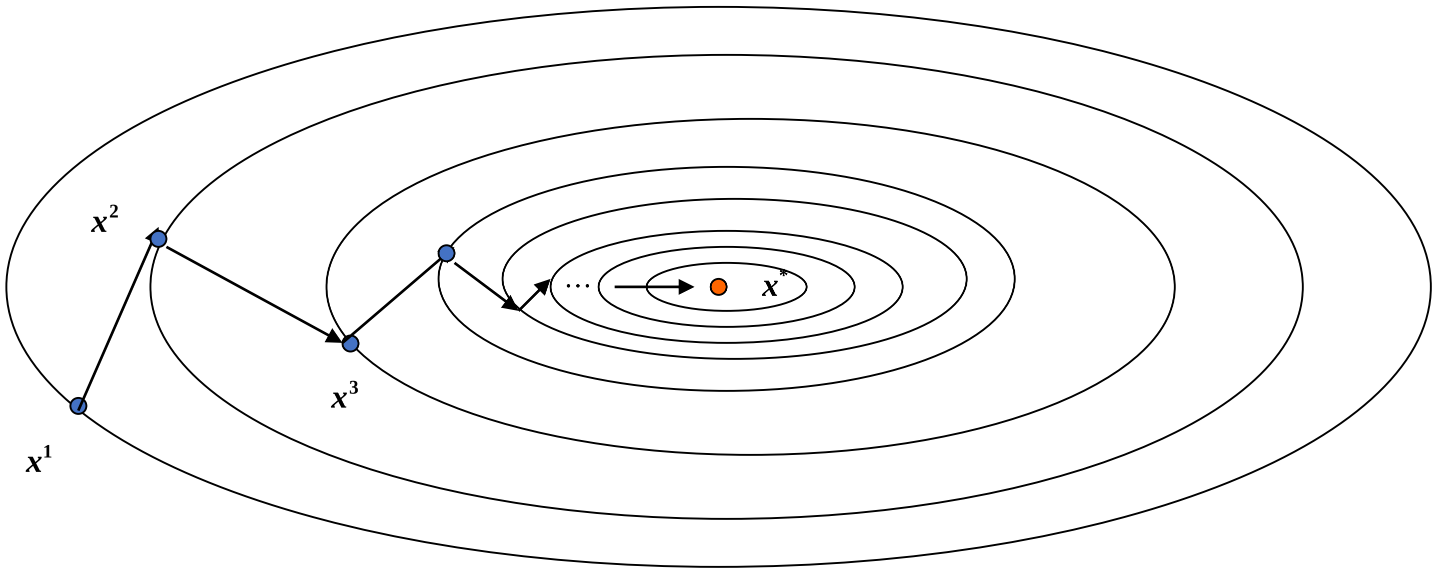
最速下降法：取模函数$\varphi(x)$减少最快的方向，

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

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

从$x^{(0)}$出发寻找$\varphi(x)$的极小值点：$\varphi(x)=\varphi(x^{(0)})$是$\varphi(x)$的等值面，因为$A$正定，它是$n$维空间的一个椭球面，从$x^{(0)}$出发先找一个使$\varphi(x)$减小最快的方向，这就是正交于椭球面的$\varphi(x)$负梯度方向$-\nabla\varphi(x^{(0)})$.

目标函数为二次函数，其等值面为椭球面.



注：最速下降方向反映了目标函数的一种局部性质.
它只是局部目标函数值下降最快的方向.

最速下降法是线性收敛的算法.

最速下降算法:

(1) 选取$x^{(0)} \in R^n$

(2) 对$k = 0,1,2,\ldots.$

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

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

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

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

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

关于最速下降法，有以下两个结论：

（1）最速下降法中相邻两次的搜索方向是正交的，即

$$(r^{(k+1)},\ r^{(k)})=0 \tag{6}$$

证：

$$\begin{aligned}(r^{(k+1)},r^{(k)}) &= (b-Ax^{(k+1)},r^{(k)})\\ &= (b-A(x^{(k)}+\alpha_k r^{(k)}),r^{(k)})\\ &= (b-Ax^{(k)}-\alpha_k Ar^{(k)},r^{(k)})\\ &= (r^{(k)}-\alpha_k Ar^{(k)},r^{(k)})\\ &= (r^{(k)},r^{(k)})-\frac{(r^{(k)},r^{(k)})}{(Ar^{(k)},r^{(k)})}(Ar^{(k)},r^{(k)})\\ &= 0\end{aligned}$$

（2） $\{\varphi(x^{(k)})\}$是单调下降有界序列，它存在极限，

可以证明 $\lim\limits_{k\to\infty} x^{(k)} = x^* = A^{-1}b$

而且 $\|x^{(k)} - x^*\|_A \le \left(\dfrac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k \|x^{(0)} - x^*\|_A$

其中λ_1, λ_n分别是对称正定矩阵A的最大、最小特征值，

范数$\|\bullet\|_A$定义为$\|u\|_A = (Au, u)^{1/2}$

当$\lambda_1 >> \lambda_n$时，收敛是很慢的，

当$\|r^{(k)}\|$很小时，因舍入误差的影响，计算将出现不稳定现象.

例 用最速下降法解方程组

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

方程组的精确解为$x=(4,-1)^{\mathrm{T}}$.

解 系数矩阵A对称正定，取$x^{(0)}=(0,0)^{\mathrm{T}}$，有

$$r^{(0)}=b-Ax^{(0)}=(6,3)^T$$

$$\alpha_0=\frac{(r^{(0)},r^{(0)})}{(Ar^{(0)},r^{(0)})}=\frac{5}{21}$$

第一步的迭代结果为

$$x^{(1)}=x^{(0)}+\alpha_0 r^{(0)}=\begin{pmatrix} \frac{10}{7} \\ \frac{5}{7} \end{pmatrix}$$

$$r^{(1)} = b - Ax^{(1)} = \begin{pmatrix} \frac{12}{7} \\ -\frac{24}{7} \end{pmatrix}$$

$$\alpha_1 = \frac{(r^{(1)}, r^{(1)})}{(Ar^{(1)}, r^{(1)})} = \frac{45}{126}$$

第二步的迭代结果为

$$x^{(2)} = x^{(1)} + \alpha_1 p^{(1)} = \begin{pmatrix} 2.0408 \\ -0.5102 \end{pmatrix}$$

继续迭代下去，即可求得方程组的解.

（二） 共轭梯度法 (CG) (共轭斜量法)

设按方向$p^{(0)}, p^{(1)}, ..., p^{(k-1)}$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
$$\alpha = \alpha_k = \frac{(r^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})} \qquad (7)$$

下一个近似解和对应的剩余向量是

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \qquad (8)$$

$$r^{(k+1)} = b - Ax^{(k+1)} = r^{(k)} - \alpha_k Ap^{(k)} \qquad (9)$$

不失一般性地设$x^{(0)}$=0，反复利用(8)有

$$x^{(k+1)} = \alpha_0 p^{(0)} + \alpha_1 p^{(1)} + ... + \alpha_k p^{(k)}$$

现在考虑$p^{(0)}, p^{(1)}, \ldots$取什么方向.

设$p^{(0)}=r^{(0)}$，一般$k \geq 1$时$p^{(k)}$的确定，我们不但希望使

$$\varphi(x^{(k+1)}) = \min_{\alpha} \varphi(x^{(k)} + \alpha p^{(k)}) \qquad (10)$$

而且希望$\left\{ p^{(k)} \right\}$的选择使

$$\varphi(x^{(k+1)}) = \min_{x \in span\left\{ p^{(0)}, \ldots, p^{(k)} \right\}} \varphi(x) \qquad (11)$$

若$x \in span\left\{ p^{(0)}, \ldots, p^{(k)} \right\}$，可记成

$$x = y + \alpha p^{(k)}, \quad y \in span\left\{ p^{(0)}, \ldots, p^{(k-1)} \right\}, \quad \alpha \in R$$

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

所以有

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

$$= \varphi(y) + \alpha(Ay, p^{(k)}) - \alpha(b, p^{(k)}) + \frac{\alpha^2}{2}(Ap^{(k)}, p^{(k)}) \qquad (12)$$

令 $(Ay, p^{(k)}) = 0, \quad y \in span\left\{p^{(0)}, \ldots, p^{(k-1)}\right\}$

即 $(Ap^{(j)}, p^{(k)}) = 0, \qquad j = 0,1,\ldots,k-1$

如果$k=1,2\cdots$,每步都如此选择$p^{(k)}$，则它们符合下面定义.

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，或称$A-$正交向量组.

注：

1、当$n>1$时，不含零向量的$A-$共轭向量组线性无关；

2、当$A=I$时，$A-$共轭性质就是一般的正交性；

3、给了一组线性无关的向量，可以按Schmidt正交化的方法得到对应的$A-$共轭向量组.

由（12） $\varphi(x)=\varphi(y+\alpha p^{(k)})=\varphi(y)-\alpha(b,p^{(k)})+\frac{\alpha^2}{2}(Ap^{(k)},p^{(k)})$

将极小问题(11)分离为两个极小问题:

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

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

$$=\min_{y\in span\{p^{(0)},\dots,p^{(k-1)}\}}\varphi(y)+\min_{\alpha}\left[\frac{\alpha^2}{2}(Ap^{(k)},p^{(k)})-\alpha(b,p^{(k)})\right]$$

第一个问题的解为$x^{(k)}$,

第二个问题的解为$\alpha=\alpha_k=\dfrac{(b,p^{(k)})}{(Ap^{(k)},p^{(k)})}$.

$\because x^{(k)}\in span\left\{p^{(0)},...,p^{(k-1)}\right\}$, 故$(Ax^{(k)},p^{(k)})=0$,

$\because (b,p^{(k)})=(b-Ax^{(k)},p^{(k)})=(r^{(k)},p^{(k)})$

$\therefore \alpha=\alpha_k=\dfrac{(b,p^{(k)})}{(Ap^{(k)},p^{(k)})}=\dfrac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})}$ 与（7）相同

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

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

利用
$$\begin{aligned}(p^{(k)},Ap^{(k-1)})&=(r^{(k)}+\beta_{k-1}p^{(k-1)},Ap^{(k-1)})\\&=(r^{(k)},Ap^{(k-1)})+\beta_{k-1}(p^{(k-1)},Ap^{(k-1)})=0,\end{aligned}$$

可得
$$\beta_{k-1}=-\frac{(r^{(k)},Ap^{(k-1)})}{(p^{(k-1)},Ap^{(k-1)})} \tag{14}$$

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

公式化简

由(9)式$r^{(k+1)}=r^{(k)}-\alpha_k Ap^{(k)}$和(7)式$\alpha_k=\dfrac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})}$

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

$$(r^{(k)},\ p^{(k)})=(r^{(k)},\ r^{(k)}+\beta_{k-1}p^{(k-1)})=(r^{(k)},\ r^{(k)}) \qquad (15)$$

代回(7)有$\alpha_k=\dfrac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})}=\dfrac{(r^{(k)},r^{(k)})}{(Ap^{(k)},p^{(k)})} \qquad (16)$

当$r^{(k)}\neq 0$时，$\alpha_k>0.$

定理　由(7)——(16)定义的算法有如下性质：

(1) $(r^{(i)},r^{(j)})=0,i\neq j$,即剩余向量构成一个正交向量组.

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

即$\left\{p^{(k)}\right\}$为一个$A-$共轭向量组.

还可简化β_k的计算：

由前 (14) 式 $$\beta_{k-1}=-\frac{(r^{(k)},Ap^{(k-1)})}{(p^{(k-1)},Ap^{(k-1)})} \qquad (14)$$

前 (9) 式 $$r^{(k+1)}=r^{(k)}-\alpha_k Ap^{(k)} \qquad (9)$$

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

得

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

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

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，$x^{(k)} = x^*$.

例 用CG方法解方程组

$$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

解 系数矩阵A对称正定，取$x^{(0)}=(0,0)^{\mathrm{T}}$，有

$$r^{(0)} = p^{(0)} = b - Ax^{(0)} = (5,5)^T$$

$$\alpha_0 = \frac{(r^{(0)}, r^{(0)})}{(Ap^{(0)}, p^{(0)})} = \frac{2}{7}$$

$$x^{(1)} = x^{(0)} + \alpha_0 p^{(0)} = (\frac{10}{7}, \frac{10}{7})^T$$

$$r^{(1)} = r^{(0)} - \alpha_0 Ap^{(0)} = (-\frac{5}{7}, \frac{5}{7})^T$$

$$\beta_0 = \frac{(r^{(1)}, r^{(1)})}{(r^{(0)}, r^{(0)})} \quad = \frac{1}{49}$$

$$p^{(1)} = r^{(1)} + \beta_0 p^{(0)} \quad = (-\frac{30}{49}, \frac{40}{49})^T$$

$$\alpha_1 = \frac{(r^{(1)}, r^{(1)})}{(Ap^{(1)}, p^{(1)})} \quad = \frac{7}{10}$$

$$x^{(2)} = x^{(1)} + \alpha_1 p^{(1)} \quad = (1, 2)^T$$

$$r^{(2)} = r^{(1)} - \alpha_1 Ap^{(1)} \quad = (0, 0)^T$$

故方程组的解为$x=(1,2)^{\mathrm{T}}$.

注：

(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$x^{(0)}=x^{(n)}$.

(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，共轭梯度法收敛很快，但当A是病态严重的矩阵时，共轭梯度法收敛速度很慢. 可采用预处理技术，降低A的条件数.

（三）预条件共轭梯度法(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$\overline{A}\overline{x}=\overline{b}$用共轭梯度法求解，可达到加速的目的.

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

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

预条件共轭梯度法

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

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

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

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

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

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

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

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

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

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

3、$x^{(k)}=C^{-T}\overline{x}^{(k)}$.

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$Mz^{(k+1)} = r^{(k+1)}$

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

(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.

下面介绍几种选取预优矩阵的方案：

（1）A是大型稀疏对称正定的矩阵，取

$$M=diag(a_{11},a_{22},\dots,a_{nn})$$

（2）取A的三条对角线构成的三对角阵

$$M=\begin{bmatrix} a_{11} & a_{12} & & \\ a_{21} & a_{22} & \ddots & \\ & \ddots & \ddots & a_{n-1,n} \\ & & a_{n,n-1} & a_{nn} \end{bmatrix}$$

（3）M取为$SSOR$迭代法的预处理阵，$M = CC^T$

$$C = \left[\omega(2-\omega)\right]^{-1/2}(D-\omega L)D^{-1/2}$$

$$C^T = \left[\omega(2-\omega)\right]^{-1/2}D^{-1/2}(D-\omega L^T)$$

（4）M取为块$Jacobi$迭代的块对角阵，

$$A = \begin{bmatrix} A_{11} & A_{12} & & \\ A_{21} & A_{22} & \ddots & \\ & \ddots & \ddots & A_{n-1,n} \\ & & A_{n,n-1} & A_{nn} \end{bmatrix},$$

$$M = \begin{bmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{nn} \end{bmatrix}$$

(5) 不完全Cholesky分解的预优矩阵

不完全Cholesky分解就是将A 分解成

$$A = LL^T + R$$

其中L是下三角阵，R 称为剩余矩阵.

取 $M = LL^T$

例：$A = \begin{bmatrix} 3 & -1 & 0 & 2 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 2 & 0 & -1 & 3 \end{bmatrix}$,

$A = LL^T + R$, 调节R使 $L = \begin{bmatrix} * & & & \\ * & * & & \\ 0 & * & * & \\ * & 0 & * & * \end{bmatrix}$

$$A = LL^T + R$$

$$L = \begin{bmatrix} \sqrt{3} & & & \\ -\frac{1}{\sqrt{3}} & \frac{4}{\sqrt{3}} & & \\ 0 & -\frac{\sqrt{3}}{4} & \frac{\sqrt{21}}{4} & \\ \frac{2}{\sqrt{3}} & 0 & -\frac{4}{\sqrt{21}} & \frac{3}{\sqrt{7}} \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & 0 & 0 \end{bmatrix}$$

预条件共轭梯度法MATLAB的三种调用格式:

1.不用预优矩阵的共轭梯度法

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

2.用预优矩阵的共轭梯度法

(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梯度法

r=cholinc(sa,’0’)

x=pcg(a,b,tol,kmax,r’,r,x0)

THE

END