1

令$f(x_1,x_2) = {x_1}^2 - 2\,x_1\,x_2 + x_1 + 4\,{x_2}^2 - 3\,x_2$。

$$\nabla f = \begin{pmatrix} 2\,x_1 - 2\,x_2 + 1 \\ 8\,x_2 - 2\,x_1 - 3 \end{pmatrix}$$

在$(1,1)$处，$\nabla f\,\big|_{(1,1)} = (1,3)^T$，所以最速下降方向为$d^{(1)} = (-1,-3)^T$，

$$\alpha^{(1)} = \min_{\alpha} f(x^{(1)} + \alpha d^{(1)}) = \frac{5}{31}$$

于是$x^{(2)} = x^{(1)} + \alpha^{(1)} d^{(1)} = (\frac{26}{31}, \frac{16}{31})^T$。

$(\frac{26}{31}, \frac{16}{31})$处，$\nabla f\,\big|_{(\frac{26}{31}, \frac{16}{31})} = (\frac{51}{31}, -\frac{17}{31})^T$，$d^{(2)} = (-\frac{51}{31}, \frac{17}{31})^T$，

$$\alpha^{(2)} = \min_{\alpha} f(x^{(2)} + \alpha d^{(2)}) = \frac{5}{19}$$

$x^{(3)} = x^{(2)} + \alpha^{(2)} d^{(2)} = (\frac{289}{589}, \frac{867}{589})^T$。

2

$$\nabla f = Ax + b$$

最速下降法将选取$d^{(1)} = -Ax^{(1)} - b$作为下降方向。

$$\begin{aligned} d^{(1)} &= A\bar{x} + \sum_{i=1}^{m} \mu_i A p^{(i)} + b \\ &= \sum_{i=1}^{m} \mu_i \lambda_i p^{(i)} + A\bar{x} + b \\ &= \sum_{i=1}^{m} \mu_i \lambda_i p^{(i)} \end{aligned}$$

如果要在一次迭代终止，需要$d^{(1)}$与$(x^{(1)} - \bar{x})$共线，即要求

$$\begin{gathered} d^{(1)} = \alpha(x^{(1)} - \bar{x}) \\ \sum_{i=1}^{m} \mu_i \lambda_i p^{(i)} = \alpha \sum_{i=1}^{m} \mu_i p^{(i)} \end{gathered}$$

若要上式成立，需要$\lambda_i = \alpha, \forall i$。从题目条件中知道$A$有多个特征值，显然上式不成立。所以不可能在一次迭代后终止。

3

原方程的梯度为

$$g(x) = Ax - b$$

$x^{(k)}$处的下降方向为$d^{(k)} = -g^{(k)} = -g(x^{(k)})$。

$$f(x^{(k)} - \alpha g^{(k)}) = \frac{1}{2}\left(x^{(k)} - \alpha g^{(k)}\right)^T A \left(x^{(k)} - \alpha g^{(k)}\right) - b^T \left(x^{(k)} - \alpha g^{(k)}\right)$$

对于$\forall x^{(k)} \neq \bar{x}$

$$\alpha^{(k)} = \min_{\alpha} f(x^{(k)} - \alpha g^{(k)}) = \frac{{g^{(k)}}^T g^{(k)}}{{g^{(k)}}^T A g^{(k)}}$$

记$y^{(k)}=x^{(k)}-\bar{x}$，

$$\frac{E(x^{(k)})-E(x^{(k+1)})}{E(x^{(k)})}=\frac{2\alpha^{(k)}{g^{(k)}}^T A y^{(k)}-{\alpha^{(k)}}^2 {g^{(k)}}^T A g^{(k)}}{{y^{(k)}}^T A y^{(k)}}$$

由于$g(\bar{x})=0$，所以$A\bar{x}=b$，$Ay^{(k)}=A(x^{(k)}-\bar{x})=Ax^{(k)}-b=g^{(k)}$。

$$\begin{aligned}\frac{E(x^{(k)})-E(x^{(k+1)})}{E(x^{(k)})}&=\frac{\frac{2({g^{(k)}}^T g^{(k)})^2}{{g^{(k)}}^T A g^{(k)}}-\frac{({g^{(k)}}^T g^{(k)})^2}{{g^{(k)}}^T A g^{(k)}}}{{g^{(k)}}^T A^{-1} g^{(k)}}\\&=\frac{\left({g^{(k)}}^T g^{(k)}\right)^2}{\left({g^{(k)}}^T A g^{(k)}\right)\left({g^{(k)}}^T A^{-1} g^{(k)}\right)}\end{aligned}$$

由Kantorovich 不等式，

$$\begin{aligned}\frac{E(x^{(k)})-E(x^{(k+1)})}{E(x^{(k)})}&=\frac{\left({g^{(k)}}^T g^{(k)}\right)^2}{\left({g^{(k)}}^T A g^{(k)}\right)\left({g^{(k)}}^T A^{-1} g^{(k)}\right)}\\&\geq\frac{4mM}{(m+M)^2}\end{aligned}$$

也就是

$$E(x^{(k+1)})\leq\left(\frac{M-m}{M+m}\right)^2 E(x^{(k)})$$

所以有

$$\lim_{k\to\infty}E(x^{(k)})=0$$

又因为A正定，所以$\lim_{k\to\infty}(x^{(k)}-\bar{x})=0$，$\lim_{k\to\infty}x^{(k)}=\bar{x}$。

4

任选$i,j\in\{1,2,3,\cdots,n\},i\neq j$，

$${p^{(i)}}^T A p^{(j)}={p^{(i)}}^T\lambda^{(j)}p^{(j)}=\lambda^{(j)}{p^{(i)}}^T p^{(j)}=0$$

其中$\lambda^{(j)}$是$p^{(j)}$对应的特征值，所以A的特征向量关于A共轭。

5

记$f(x)=x^TAx$，其中

$$A=\begin{bmatrix}\frac{1}{2}&0\\0&1\end{bmatrix}$$

该函数的梯度为

$$\nabla f=Ax$$

对于$x^{(1)}=(4,4)^T$，$g^{(1)}=\nabla f(x^{(k)})=(2,4)^T$，$d^{(1)}=-g^{(1)}=(-2,-4)^T$

$$\alpha^{(1)}=-\frac{{g^{(1)}}^T d^{(1)}}{{d^{(1)}}^T A d^{(1)}}=\frac{10}{9}$$

所以$x^{(2)}=x^{(1)}+g^{(1)}\alpha^{(1)}=(\frac{16}{9},-\frac{4}{9})^T$。有

$$g^{(2)}=(\frac{8}{9},-\frac{4}{9})^T$$

$$\beta^{(2)}=\frac{{g^{(2)}}^T(g^{(2)}-g^{(1)})}{{d^{(1)}}^T(g^{(2)}-g^{(1)})}=\frac{4}{81}$$

$$d^{(2)}=-g^{(2)}+\beta^{(2)}d^{(1)}=\left(-\frac{80}{81},\frac{20}{81}\right)^T$$

$$\alpha^{(2)}=-\frac{{g^{(2)}}^Td^{(2)}}{{d^{(2)}}^TAd^{(2)}}=\frac{9}{5}$$

于是$x^{(3)}=x^{(2)}+g^{(2)}\alpha^{(2)}=(0,0)^T$，$g^{(3)}=(0,0)^T$，至此，我们找到了$f$的极小值。

6

（1）由于A是实对称正定矩阵，所以存在可逆矩阵R，使得$A=R^TR$。

由于$p^{(1)},p^{(2)},\cdots p^{(n)}$关于$A$共轭，对任意$i\neq j$，${p^{(i)}}^TAp^{(j)}={p^{(i)}}^TR^TRp^{(j)}=0$。即$Rp^{(1)},Rp^{(2)},\cdots Rp^{(n)}$正交。那么$\frac{Rp^{(1)}}{\|Rp^{(1)}\|},\frac{Rp^{(2)}}{\|Rp^{(2)}\|},\cdots,\frac{Rp^{(n)}}{\|Rp^{(n)}\|}$是n维空间上的一组标准正交基，其中$\|Rp^{(i)}\|=\sqrt{{p^{(i)}}^TR^TRp^{(i)}}=\sqrt{{p^{(i)}}^TAp^{(i)}}$。

n维向量Rx可以表示为标准正交基的线性组合

$$\begin{aligned}Rx&=\sum_{i=1}^n{\frac{Rp^{(i)}}{\|Rp^{(i)}\|}}^TRx\frac{Rp^{(i)}}{\|Rp^{(i)}\|}\\&=\sum_{i=1}^n\frac{{p^{(i)}}^TR^TRx}{\|Rp^{(i)}\|^2}Rp^{(i)}\\&=R\sum_{i=1}^n\frac{{p^{(i)}}^TAx}{{p^{(i)}}^TAp^{(i)}}p^{(i)}\end{aligned}$$

由于R可逆，那么

$$x=\sum_{i=1}^n\frac{{p^{(i)}}^TAx}{{p^{(i)}}^TAp^{(i)}}p^{(i)}$$

（2）

$$\begin{aligned}R\sum_{i=1}^n\frac{p^{(i)}{p^{(i)}}^T}{{p^{(i)}}^TAp^{(i)}}R^T&=\sum_{i=1}^n\frac{Rp^{(i)}{p^{(i)}}^TR^T}{{p^{(i)}}^TAp^{(i)}}\\&=\sum_{i=1}^n\frac{Rp^{(i)}}{\|Rp^{(i)}\|}{\frac{Rp^{(i)}}{\|Rp^{(i)}\|}}^T\\&=I\end{aligned}$$

由于R可逆，$\sum_{i=1}^n\frac{p^{(i)}{p^{(i)}}^T}{{p^{(i)}}^TAp^{(i)}}=R^{-1}{(R^T)}^{-1}=(R^TR)^{-1}=A^{-1}$。