第 10 章作业参考答案（1）

P328/4：

(1) 用最速下降法：初始点 $\boldsymbol{x}^1=\begin{pmatrix}0\\0\end{pmatrix}$。

$\nabla f(\boldsymbol{x})=\begin{pmatrix}2x_1-4\\8x_2-8\end{pmatrix}$，$H=\nabla^2 f(\boldsymbol{x})=\begin{pmatrix}2&0\\0&8\end{pmatrix}$

k=1：$\nabla f(\boldsymbol{x}^1)=\begin{pmatrix}-4\\-8\end{pmatrix}$，$\boldsymbol{d}^1=\begin{pmatrix}1\\2\end{pmatrix}$，$\lambda_1=\dfrac{\nabla f(\boldsymbol{x}^1)^T\boldsymbol{d}^1}{(\boldsymbol{d}^1)^T H\boldsymbol{d}^1}=\dfrac{10}{17}$，$\boldsymbol{x}^2=\boldsymbol{x}^1+\lambda_1\boldsymbol{d}^1=\begin{pmatrix}10/17\\20/17\end{pmatrix}$。

k=2 ： $\nabla f(\boldsymbol{x}^2)=\dfrac{24}{17}\begin{pmatrix}-2\\1\end{pmatrix}$ ， $\boldsymbol{d}^2=\begin{pmatrix}2\\-1\end{pmatrix}$ ， $\lambda_2=\dfrac{\nabla f(\boldsymbol{x}^2)^T\boldsymbol{d}^2}{(\boldsymbol{d}^2)^T H\boldsymbol{d}^2}=\dfrac{15}{34}$ ，

$\boldsymbol{x}^3=\boldsymbol{x}^2+\lambda_2\boldsymbol{d}^2=\begin{pmatrix}25/17\\25/34\end{pmatrix}$。$\nabla f(\boldsymbol{x}^3)\neq 0$，$\boldsymbol{x}^3$ 不是最优解。

(2)若存在 $\boldsymbol{x}^{k+1}=\overline{\boldsymbol{x}}$，则 $\nabla f(\boldsymbol{x}^{k+1})=0$，即

$$0=\nabla f(\boldsymbol{x}^{k+1})=H\boldsymbol{x}^{k+1}+\boldsymbol{c}=H(\boldsymbol{x}^k+\lambda_k\boldsymbol{d}^k)+\boldsymbol{c}=\nabla f(\boldsymbol{x}^k)+\lambda_k H\boldsymbol{d}^k=-\boldsymbol{d}^k+\lambda_k H\boldsymbol{d}^k$$

即 $H\boldsymbol{d}^k=\dfrac{1}{\lambda_k}\boldsymbol{d}^k$，即 $\boldsymbol{d}^k$ 是 H 的特征向量，即 $\boldsymbol{d}^k\,//\begin{pmatrix}1&0\end{pmatrix}^T$ 或 $\boldsymbol{d}^k\,//\begin{pmatrix}0&1\end{pmatrix}^T$。但取初始点为 $\boldsymbol{x}^1=(0,0)^T$，搜索方向 $\boldsymbol{d}^k\,//(1,2)^T$ 或 $\boldsymbol{d}^k\,//(-2,1)^T$。因此不会经过有限步迭代得到 $\overline{\boldsymbol{x}}$。

(3) $\boldsymbol{d}^1\,//\begin{pmatrix}1&0\end{pmatrix}^T$ 或 $\boldsymbol{d}^1\,//\begin{pmatrix}0&1\end{pmatrix}^T$，即 $2x_1^1-4=0$ 或 $8x_2^1-8=0$，即 $x_1^1=2$ 或 $x_2^1=1$，即 $\boldsymbol{x}^1=(2,x_2^1)^T$ 或 $\boldsymbol{x}^1=(x_1^1,1)^T$。

P328/5：

（1）$\nabla f(\boldsymbol{x}^1)=A\boldsymbol{x}^1+\boldsymbol{c}=A(\overline{\boldsymbol{x}}+\mu\boldsymbol{p})+\boldsymbol{c}=\nabla f(\overline{\boldsymbol{x}})+\mu A\boldsymbol{p}=\mu\lambda\boldsymbol{p}$。

（2）$\lambda_1=-\dfrac{\nabla f(\boldsymbol{x}^1)^T\boldsymbol{d}^1}{(\boldsymbol{d}^1)^T A\boldsymbol{d}^1}=\dfrac{\|\nabla f(\boldsymbol{x}^1)\|^2}{\nabla f(\boldsymbol{x}^1)^T A\nabla f(\boldsymbol{x}^1)}=\dfrac{\mu^2\lambda^2\|\boldsymbol{p}\|^2}{\mu^2\lambda^2\boldsymbol{p}^T A\boldsymbol{p}}=\dfrac{\mu^2\lambda^2\|\boldsymbol{p}\|^2}{\mu^2\lambda^3\boldsymbol{p}^T\boldsymbol{p}}=\dfrac{1}{\lambda}$，

$\boldsymbol{x}^2=\boldsymbol{x}^1+\lambda_1\boldsymbol{d}^1=\overline{\boldsymbol{x}}+\mu\boldsymbol{p}-\lambda_1\nabla f(\boldsymbol{x}^1)=\overline{\boldsymbol{x}}+\mu\boldsymbol{p}-\dfrac{1}{\lambda}\mu\lambda\boldsymbol{p}=\overline{\boldsymbol{x}}$。