第 11 周： P280 2/1),补充题（见邮箱），P328 3（用最优步长）4，5

P280 2/1)： $\min\varphi(\lambda)=3\lambda^4-4\lambda^3-12\lambda^2$，$\lambda_1=-1.2$

解：$\varphi'(\lambda)=12\lambda^3-12\lambda^2-24\lambda$，$\varphi''(\lambda)=36\lambda^2-24\lambda-24$

$\varphi'(\lambda_1)=-9.216,\ \varphi''(\lambda_1)=56.64,\ \lambda_2=\lambda_1-\dfrac{\varphi'(\lambda_1)}{\varphi''(\lambda_1)}=-1.037$，

$\varphi'(\lambda_2)=-1.398,\ \varphi''(\lambda_2)=39.601,\ \lambda_2=\lambda_1-\dfrac{\varphi'(\lambda_2)}{\varphi''(\lambda_2)}=-1.002$

$\varphi'(\lambda_3)=-0.072,\ \varphi''(\lambda_3)=36.192,\ \lambda_4=\lambda_3-\dfrac{\varphi'(\lambda_3)}{\varphi''(\lambda_3)}=-1.000$

补充题：用两种可接受一维搜索方法求解：

$$\min_{\lambda\geq -4}\varphi(\lambda)=3\lambda^4-4\lambda^3-12\lambda^2$$

取初始点 $\lambda_1=-1.2$，$\sigma_1=0.2,\sigma_2=\sigma_3=0.6,\alpha=1.2$。

解：令 $\mu=\lambda+4$，则原问题变为

$$\min_{\mu\geq 0}\varphi(\mu-4)=3(\mu-4)^4-4(\mu-4)^3-12(\mu-4)^2$$

其中 $\varphi'(\mu-4)\big|_{\mu=0}<0,\mu_1=2.8$。

A-G 准则：$\mu_1=2.8$ 时，$\varphi(\mu_1-4)\leq\varphi(-4)+\sigma_1\varphi'(-4)\mu_1$，$\varphi(\mu_1-4)\geq\varphi(-4)+\sigma_2\varphi'(-4)\mu_1$，因此 $\tilde{\mu}=\mu_1=2.8$，$\tilde{\lambda}=\tilde{\mu}-4=-1.2$。

W-P 准则：$\mu_1=2.8$ 时，$\varphi(\mu_1-4)\leq\varphi(-4)+\sigma_1\varphi'(-4)\mu_1$，$\varphi'(\mu_1-4)\geq\sigma_3\varphi'(-4)$，因此 $\tilde{\mu}=\mu_1=2.8$，$\tilde{\lambda}=\tilde{\mu}-4=-1.2$。

P328/3（用最优步长）：

$\nabla f(\boldsymbol{x})=\begin{pmatrix}2x_1-2x_2+1\\-2x_1+8x_2-3\end{pmatrix}$，$\boldsymbol{H}=\nabla^2 f(\boldsymbol{x})=\begin{pmatrix}2&-2\\-2&8\end{pmatrix}$正定

k=1：$\boldsymbol{x}^1=\begin{pmatrix}1\\1\end{pmatrix}$，$\nabla f(\boldsymbol{x}^1)=\begin{pmatrix}1\\3\end{pmatrix}$，$\boldsymbol{d}^1=\begin{pmatrix}-1\\-3\end{pmatrix}$，$\lambda_1=-\dfrac{\nabla f(\boldsymbol{x}^1)^T\boldsymbol{d}^1}{(\boldsymbol{d}^1)^T\boldsymbol{H}\boldsymbol{d}^1}=\dfrac{5}{31}$，$\boldsymbol{x}^2=\boldsymbol{x}^1+\lambda_1\boldsymbol{d}^1=\begin{pmatrix}26/31\\16/31\end{pmatrix}$。

k=2 ：$\boldsymbol{x}^2=\begin{pmatrix}26/31\\16/31\end{pmatrix}$，$\nabla f(\boldsymbol{x}^2)=\dfrac{17}{31}\begin{pmatrix}3\\-1\end{pmatrix}$，取 $\boldsymbol{d}^2=\begin{pmatrix}-3\\1\end{pmatrix}$，$\lambda_2=-\dfrac{\nabla f(\boldsymbol{x}^2)^T\boldsymbol{d}^2}{(\boldsymbol{d}^2)^T\boldsymbol{H}\boldsymbol{d}^2}=\dfrac{17}{31}\dfrac{5}{19}$，

$\boldsymbol{x}^3=\boldsymbol{x}^2+\lambda_2\boldsymbol{d}^2=\begin{pmatrix}239/589\\389/589\end{pmatrix}$。

P328/4：

(1) 等值线（略），$\nabla f(\boldsymbol{x})=\begin{pmatrix}2x_1-4\\8x_2-8\end{pmatrix}$，$\boldsymbol{H}=\nabla^2 f(\boldsymbol{x})=\begin{pmatrix}2&0\\0&8\end{pmatrix}$正定，$f$严格凸，最优解 $\overline{\boldsymbol{x}}=\begin{pmatrix}2\\1\end{pmatrix}$。

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

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

即 $\boldsymbol{H}\boldsymbol{d}^k=\dfrac{1}{\lambda_k}\boldsymbol{d}^k$，即 $\boldsymbol{d}^k$ 是 $\boldsymbol{H}$ 的特征向量，即 $\boldsymbol{d}^k\ //\left(1\ 0\right)^T$ 或 $\boldsymbol{d}^k\ //\left(0\ 1\right)^T$。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 $\bar{\boldsymbol{x}}$。

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

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

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

第 12 周： P328 2（用牛顿法求解，迭代 1 次），P330 12，13，14/2)4)

P328/2（用牛顿法求解，迭代 1 次）

$f(\boldsymbol{x})=(6+x_1+x_2)^2+(2-3x_1-3x_2-x_1x_2)^2$ 凸函数

$$\nabla f(\boldsymbol{x})=\begin{pmatrix}2(6+x_1+x_2)+2(2-3x_1-3x_2-x_1x_2)(-3-x_2)\\ 2(6+x_1+x_2)+2(2-3x_1-3x_2-x_1x_2)(-3-x_1)\end{pmatrix}$$

$$\nabla^2 f(\boldsymbol{x})=\begin{pmatrix}2(x_2^2+6x_2+10) & 2(2x_1x_2+6x_1+6x_2+8)\\ 2(2x_1x_2+6x_1+6x_2+8) & 2(x_1^2+6x_1+10)\end{pmatrix}$$

$\boldsymbol{x}^1=\hat{\boldsymbol{x}}=\begin{pmatrix}-4\\6\end{pmatrix}$： $\nabla f(\boldsymbol{x}^1)=\begin{pmatrix}-356\\24\end{pmatrix}$，$\nabla^2 f(\boldsymbol{x}^1)=\begin{pmatrix}164 & -56\\ -56 & 4\end{pmatrix}$，$\nabla^2 f(\boldsymbol{x}^1)^{-1}=-\dfrac{1}{2480}\begin{pmatrix}4 & 56\\ 56 & 164\end{pmatrix}$,

牛顿方向 $\boldsymbol{d}^1=-[\nabla^2 f(\boldsymbol{x}^1)]^{-1}\nabla f(\boldsymbol{x}^1)=\begin{pmatrix}22/31\\ -126/31\end{pmatrix}$，$\boldsymbol{x}^2=\boldsymbol{x}^1+\boldsymbol{d}^1=\begin{pmatrix}-102/31\\ 60/31\end{pmatrix}$

P330/12:

设 $\boldsymbol{p}^i$ 是 $\boldsymbol{A}$ 对应于特征值 λ_i 的特征向量，即 $\boldsymbol{A}\boldsymbol{p}^i=\lambda_i\boldsymbol{p}^i$，则 $j\neq i$ 时，

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

P330/13：用归纳法。

k=2 时，$\boldsymbol{d}^2=\boldsymbol{p}^2-\dfrac{(\boldsymbol{d}^1)^T\boldsymbol{A}\boldsymbol{p}^2}{(\boldsymbol{d}^1)^T\boldsymbol{A}\boldsymbol{d}^1}\boldsymbol{d}^1$，$(\boldsymbol{d}^1)^T\boldsymbol{A}\boldsymbol{d}^2=(\boldsymbol{d}^1)^T\boldsymbol{A}\boldsymbol{p}^2-\dfrac{(\boldsymbol{d}^1)^T\boldsymbol{A}\boldsymbol{p}^2}{(\boldsymbol{d}^1)^T\boldsymbol{A}\boldsymbol{d}^1}(\boldsymbol{d}^1)^T\boldsymbol{A}\boldsymbol{d}^1=0$，即 $\boldsymbol{d}^1,\boldsymbol{d}^2$ 是 $\boldsymbol{A}$-共轭的。

设 $\boldsymbol{d}^1,\cdots,\boldsymbol{d}^k$ 是 A-共轭的，则

$\boldsymbol{d}^{k+1}=\boldsymbol{p}^{k+1}-\displaystyle\sum_{i=1}^{k}\frac{(\boldsymbol{d}^i)^T\boldsymbol{A}\boldsymbol{p}^{k+1}}{(\boldsymbol{d}^i)^T\boldsymbol{A}\boldsymbol{d}^i}\boldsymbol{d}^i$，$j=1,\cdots,k$ 时，

$$(\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{d}^{k+1} = (\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{p}^{k+1} - \sum_{i=1}^{k} \frac{(\boldsymbol{d}^i)^T \boldsymbol{A}\boldsymbol{p}^{k+1}}{(\boldsymbol{d}^i)^T \boldsymbol{A}\boldsymbol{d}^i} (\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{d}^i$$

$$\overset{(\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{d}^i=0, i\neq j\text{时}}{=} (\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{p}^{k+1} - \frac{(\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{p}^{k+1}}{(\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{d}^j} (\boldsymbol{d}^j)^T \boldsymbol{A}\boldsymbol{d}^j = 0$$

即 $\boldsymbol{d}^1, \cdots, \boldsymbol{d}^{k+1}$ 是 $\boldsymbol{A}$-共轭的。由归纳法知，$\boldsymbol{d}^1, \cdots, \boldsymbol{d}^n$ 是 $\boldsymbol{A}$-共轭的。

P330/14(2)：

$\min f(\boldsymbol{x}) = x_1^2 + 2x_2^2 - 2x_1x_2 + 2x_2 + 2$，$\boldsymbol{x}^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$。

$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 + 2 \end{pmatrix}$，$\boldsymbol{H} = \nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$正定，$f$是严格凸函数。

k=1：$\boldsymbol{x}^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$，$\nabla f(\boldsymbol{x}^1) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$，$\boldsymbol{d}^1 = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$，$\lambda_1 = -\dfrac{\nabla f(\boldsymbol{x}^1)^T \boldsymbol{d}^1}{(\boldsymbol{d}^1)^T \boldsymbol{H}\boldsymbol{d}^1} = \dfrac{1}{4}$，$\boldsymbol{x}^2 = \boldsymbol{x}^1 + \lambda_1 \boldsymbol{d}^1 = (0, -1/2)^T$。

k=2：$\boldsymbol{x}^2 = \boldsymbol{x}^1 + \lambda_1 \boldsymbol{d}^1 = \begin{pmatrix} 0 \\ -1/2 \end{pmatrix}$，$\nabla f(\boldsymbol{x}^2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$，$\beta_1 = \dfrac{\|\nabla f(\boldsymbol{x}^2)\|^2}{\|\nabla f(\boldsymbol{x}^1)\|^2} = \dfrac{1}{4}$，$\boldsymbol{d}^2 = -\nabla f(\boldsymbol{x}^2) + \beta_1 \boldsymbol{d}^1 = \begin{pmatrix} -1 \\ -1/2 \end{pmatrix}$，

$\lambda_2 = -\dfrac{\nabla f(\boldsymbol{x}^2)^T \boldsymbol{d}^2}{(\boldsymbol{d}^2)\boldsymbol{H}\boldsymbol{d}^2} = 1$，$\boldsymbol{x}^3 = \boldsymbol{x}^2 + \lambda_2 \boldsymbol{d}^2 = (-1, -1)^T$。

k=3：$\boldsymbol{x}^3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$，$\nabla f(\boldsymbol{x}^3) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$，$\boldsymbol{x}^3$ 是平稳点。又 f 是严格凸函数，因此 $\boldsymbol{x}^* = \boldsymbol{x}^3$ 是最优解。

P330/14(4)：

$\min f(\boldsymbol{x}) = 2x_1^2 + 2x_1x_2 + x_2^2 + 3x_1 - 4x_2$，$\boldsymbol{x}^1 = (3, 4)^T$。

$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 4x_1 + 2x_2 + 3 \\ 2x_1 + 2x_2 - 4 \end{pmatrix}$，$\boldsymbol{H} = \nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$正定，$f$是凸函数。

k=1：$\boldsymbol{x}^1 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$，$\nabla f(\boldsymbol{x}^1) = \begin{pmatrix} 23 \\ 10 \end{pmatrix}$，$\boldsymbol{d}^1 = \begin{pmatrix} -23 \\ -10 \end{pmatrix}$，$\lambda_1 = -\dfrac{\nabla f(\boldsymbol{x}^1)^T \boldsymbol{d}^1}{(\boldsymbol{d}^1)^T \boldsymbol{H}\boldsymbol{d}^1} = \dfrac{629}{3236} \approx 0.194$，

$\boldsymbol{x}^2 = \boldsymbol{x}^1 + \lambda_1 \boldsymbol{d}^1 = (-1.462, 2.06)^T$。

k=2：$\boldsymbol{x}^2 = \begin{pmatrix} -1.462 \\ 2.06 \end{pmatrix}$，$\nabla f(\boldsymbol{x}^2) = \begin{pmatrix} 1.272 \\ -2.804 \end{pmatrix}$，$\beta_1 = \dfrac{\|\nabla f(\boldsymbol{x}^2)\|^2}{\|\nabla f(\boldsymbol{x}^1)\|^2} = 0.015$，

$\boldsymbol{d}^2 = -\nabla f(\boldsymbol{x}^2) + \beta_1 \boldsymbol{d}^1 = \begin{pmatrix} -1.617 \\ 2.654 \end{pmatrix}$，$\lambda_2 = -\dfrac{\nabla f(\boldsymbol{x}^2)\boldsymbol{d}^2}{(\boldsymbol{d}^2)\boldsymbol{H}\boldsymbol{d}^2} = 1.287$，$\boldsymbol{x}^3 = \boldsymbol{x}^2 + \lambda_2 \boldsymbol{d}^2 = (-3.543, 5.476)^T$。

k=3：$\boldsymbol{x}^3 = \begin{pmatrix} -3.543 \\ 5.476 \end{pmatrix}$，$\nabla f(\boldsymbol{x}^3) \approx \begin{pmatrix} 0 \\ 0 \end{pmatrix}$，$\boldsymbol{x}^3$ 是近似平稳点。又 f 是凸函数，因此 $\boldsymbol{x}^* = \boldsymbol{x}^3$ 是近似最优解。