第十一次作业答案

1． 给定函数

$$f(x)=(6+x_1+x_2)^2+(2-3x_1-3x_2-x_1x_2)^2$$

求在点 $\hat{x}=\begin{bmatrix}-4\\6\end{bmatrix}$

处的牛顿方向和最速下降方向。

解：牛顿方向为：$d=\begin{bmatrix}\dfrac{22}{31}\\-\dfrac{126}{31}\end{bmatrix}$

最速下降方向为：$d=\begin{bmatrix}344\\-56\end{bmatrix}$。

2． 设有函数

$$f(x)=\frac{1}{2}x^TAx+b^Tx+c$$

其中 A 为对称正定矩阵。又设 $x^{(1)}(\neq\bar{x})$ 可表示为

$$x^{(1)}=\bar{x}+\mu p$$

其中 $\bar{x}$是$f(x)$的极小点，p是A的属于特征值λ的特征向量。证明：

（1）$\nabla f(x^{(1)})=\mu\lambda p$

（2）如果从 $x^{(1)}$ 出发，沿最速下降方向作一维搜索，则一步达到极小点 $\bar{x}$ 。

证明：（1）因为 $\nabla f(x^{(1)})=Ax^{(1)}+b$，而 $x^{(1)}=\bar{x}+\mu p$，

所以 $\nabla f(x^{(1)})=A\left(\bar{x}+\mu p\right)+b=A\bar{x}+\mu Ap+b=\nabla f(\bar{x})+\mu Ap$

由于 $\bar{x}$是$f(x)$的极小点，p是A的属于特征值λ的特征向量

所以，$\nabla f\left(\bar{x}\right)=0,\ Ap=\lambda p$

即 $\nabla f(x^{(1)})=\mu\lambda p$。

（2）搜索方向 $d=-\nabla f\left(x^{(1)}\right)=-\mu\lambda p$，设从 $x^{(1)}$ 出发，沿最速下降方向作一维搜索得到

$x^{(2)}$，则 $x^{(2)}=x^{(1)}+\alpha d$，其中 $\alpha=-\dfrac{\nabla f\left(x^{(1)}\right)^T d}{d^TAd}$，由(1)，得 $\alpha=\dfrac{1}{\lambda}$，

所以，$x^{(2)}=x^{(1)}+\alpha d=\bar{x}+\mu p+\frac{1}{\lambda}\left(-\lambda\mu p\right)=\bar{x}$，

即从 $x^{(1)}$ 出发，沿最速下降方向作一维搜索，则一步达到极小点 $\bar{x}$。

3. 设 A 为 n 阶实对称正定矩阵，证明 A 的 n 个互相正交的特征向量 $p^{(1)},p^{(2)},\cdots,p^{(n)}$ 关于 A 共轭。

证明：设 $Ap^{(i)}=\lambda_i p^{(i)},i=1,2,\cdots,n$，则对任意的 $i\neq j,i,j=1,2,\cdots,n$，有

$$\left(p^{(i)}\right)^T Ap^{(j)}=\left(p^{(i)}\right)^T \lambda_j p^{(j)}=0$$

所以 $p^{(1)},p^{(2)},\cdots,p^{(n)}$ 关于 A 共轭。

4. 设 A 为 n 阶对称正定矩阵，非零向量 $p^{(1)},p^{(2)},\cdots,p^{(n)}\in E^n$ 关于矩阵 A 共轭。证明：

（1）$x=\sum_{i=1}^{n}\frac{p^{(i)^T}Ax}{p^{(i)^T}Ap^{(i)}}p^{(i)},\quad \forall x\in E^n$

（2）$A^{-1}=\sum\frac{p^{(i)}p^{(i)^T}}{p^{(i)^T}Ap^{(i)}}$

证明：(1) 因为非零向量 $p^{(1)},p^{(2)},\cdots,p^{(n)}\in E^n$ 关于矩阵 A 共轭，所以 $p^{(1)},p^{(2)},\cdots,p^{(n)}$ 线性无关，因而 $p^{(1)},p^{(2)},\cdots,p^{(n)}$ 构成一组基。对 $\forall x\in E^n$，设

$$x=k_1p^{(1)}+k_2p^{(2)}+\cdots+k_np^{(n)}$$

上式两边左乘 $p^{(i)^T}A$，得 $p^{(i)^T}Ax=k_ip^{(i)^T}Ap^{(i)}\Rightarrow k_i=\frac{p^{(i)^T}Ax}{p^{(i)^T}Ap^{(i)}}$

设$A^{-1}=\left(\beta_1,\beta_2,\cdots,\beta_n\right)$

则由（1），得：

$$A^{-1}=\left(\sum_{i=1}^{n}\frac{p^{(i)^T}A\beta_1}{p^{(i)^T}Ap^{(i)}}p^{(i)},\sum_{i=1}^{n}\frac{p^{(i)^T}A\beta_2}{p^{(i)^T}Ap^{(i)}}p^{(i)},\cdots,\sum_{i=1}^{n}\frac{p^{(i)^T}A\beta_n}{p^{(i)^T}Ap^{(i)}}p^{(i)}\right)$$

$$=\sum_{i=1}^{n}\frac{1}{p^{(i)^T}Ap^{(i)}}\left(p^{(i)^T}A\beta_1p^{(i)},p^{(i)^T}A\beta_2p^{(i)},\cdots,p^{(i)^T}A\beta_np^{(i)}\right)$$

$$=\sum_{i=1}^{n}\frac{1}{p^{(i)^T}Ap^{(i)}}p^{(i)}\left(p^{(i)^T}A\beta_1,p^{(i)^T}A\beta_2,\cdots,p^{(i)^T}A\beta_n\right)$$

$$=\sum_{i=1}^{n}\frac{1}{p^{(i)^T}Ap^{(i)}}p^{(i)}p^{(i)^T}A\left(\beta_1,\beta_2,\cdots,\beta_n\right)$$

$$=\sum_{i=1}^{n}\frac{1}{p^{(i)^T}Ap^{(i)}}p^{(i)}p^{(i)^T}AA^{-1}=\sum_{i=1}^{n}\frac{p^{(i)}p^{(i)^T}}{p^{(i)^T}Ap^{(i)}}$$