作业 #5

1. **[5.1]** 已知 f 是凸集 S 上的凸函数，证明水平集 $T=\{x\in S|\, f(x)\le k\}$ 对任意实数 k 是凸集。

证明： 设 $y,z\in T$，令 $x=\lambda y+(1-\lambda)z$，$0\le\lambda\le 1$。因为 S 是凸集，$y,z\in S$，所以 $x\in S$。又因为 f 是凸函数，所以 $f(x)=f(\lambda y+(1-\lambda)z)\le\lambda f(y)+(1-\lambda)f(z)$，而 $f(y)\le k, f(z)\le k$，所以 $f(x)\le\lambda k+(1-\lambda)k=k$，即 $x\in T$。由此知 T 是凸集。

2. **[5.2(1)(4)]** 判断以下函数是否为凸函数：

(1) $f(x)=\ln\left(\frac{1}{x}\right)$

解： $f''(x)=\frac{1}{x^2}>0$，对任意的$x>0$恒成立，所以是凸函数。

(4) $f(x_1,x_2)=x_1e^{-(x_1+x_2)}$

解： $\nabla^2 f(x)=e^{-x_1-x_2}\begin{pmatrix}x_1-2 & x_1-1\\ x_1-1 & x_1\end{pmatrix}$

$\begin{vmatrix}x_1-2 & x_1-1\\ x_1-1 & x_1\end{vmatrix}=x_1(x_1-2)-(x_1-1)(x_1-1)=-1<0$

所以不是凸函数。

3. **[5.4]** 考虑无约束极值问题：

$$\min\ f(\mathbf{x})=8x_1^2+3x_1x_2+7x_2^2-25x_1+31x_2-29$$

(1) 求其所有稳定点；

(2) 稳定点是否是局部极小点？该问题是否有全局极小点？

解： (1) $\nabla f(\mathbf{x})=\begin{pmatrix}16x_1+3x_2-25\\ 3x_1+14x_2+31\end{pmatrix}=0$，求解得稳定点为：

$$\boldsymbol{x}^*=(x_1,x_2)=\left(\frac{443}{215},-\frac{571}{215}\right)。$$

(2) $\nabla^2 f(\mathbf{x})=\begin{pmatrix}16 & 3\\ 3 & 14\end{pmatrix}$，该 Hessian 阵为正定矩阵，所以是凸规划，稳定点 $\mathbf{x}^*$ 既是局部极小点，也是全局极小点。

4. **[5.5]** 考虑无约束极值问题：

$$\min\ f(\mathbf{x})=2x_1^3-3x_1^2-6x_1x_2(x_1-x_2-1)$$

(1) 求其所有局部极小点；

(2) 若迭代求解从 $\mathbf{x}_0=(1,-1)^T$ 开始，请问 $\mathbf{p}_0=(-1,1)^T$ 是否为下降方向？

解： (1) 令目标函数一阶导数为零，得：

$$\begin{cases}\dfrac{\partial f}{\partial x_1}=6x_1^2-6x_1-12x_1x_2+6x_2^2+6x_2=0\\ \dfrac{\partial f}{\partial x_2}=-6x_1^2+12x_1x_2+6x_1=0\end{cases}\quad\Rightarrow\quad\begin{cases}(x_1-x_2)(x_1-x_2-1)=0\\ x_1(x_1-2x_2-1)=0\end{cases}$$

由此，求得四个静止点为：$(0,0)$，$(0,-1)$，$(-1,-1)$ 和 $(1,0)$。进一步检验二阶导数，目标函数的 Hessian 阵为：

$$\begin{pmatrix}12x_1-12x_2-6 & -12x_1+12x_2+6\\ -12x_1+12x_2+6 & 12x_1\end{pmatrix}$$

分别代入四个静止点，得：

$$\begin{pmatrix}-6 & 6\\ 6 & 0\end{pmatrix},\ \begin{pmatrix}6 & -6\\ -6 & 0\end{pmatrix},\ \begin{pmatrix}-6 & 6\\ 6 & -12\end{pmatrix}\ \text{和}\ \begin{pmatrix}6 & -6\\ -6 & 12\end{pmatrix}$$

只有第四个矩阵满足半正定，因此 $(1,0)$ 为局部极小点。

(2) 验证 $\nabla f(x_0)^T\ p_0<0$ 是否成立。代入初始点 $x_0=(1,-1)^T$，得：

$$\nabla f(1,-1)=\begin{pmatrix}6-6+12+6-6\\ -6-12+6\end{pmatrix}=\begin{pmatrix}12\\ -12\end{pmatrix}$$

所以 $\nabla f(x_0)^T\ p_0=(12,-12)\begin{pmatrix}-1\\ 1\end{pmatrix}=-24<0$ 成立，即 $p_0=(-1,1)^T$ 是下降方向。

5. 假设 $\mathbf{G}\in\Re^{n\times n}$ 是正定矩阵，$\mathbf{b}\in\Re^n$，求二次函数 $f(\mathbf{x})=\dfrac{1}{2}\mathbf{x}^T\mathbf{G}\mathbf{x}+\mathbf{b}^T\mathbf{x}$ 的一阶和二阶导数。

解： $f(\mathbf{x})=\frac{1}{2}\mathbf{x}^T\mathbf{G}\mathbf{x}+\mathbf{b}^T\mathbf{x}$ ($\mathbf{G}$为$n\times n$的对称矩阵)

$$\begin{cases}\nabla f(\mathbf{x}) = \mathbf{Gx} + \mathbf{b} \\ \nabla^2 f(\mathbf{x}) = \mathbf{G}\end{cases}$$

推导过程：

$$f(\mathbf{x}) = \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} G_{j,k} x_j x_k + \sum_{j=1}^{n} b_j x_j$$

唯一包含变量x_i的项为：

$$\frac{1}{2}G_{ii}x_i^2 + \frac{1}{2}\sum_{j\neq i} G_{ji}\, x_j x_i + \frac{1}{2}\sum_{k\neq i} G_{ik}\, x_i x_k + b_i x_i$$

对x_i求偏导得：

$$\frac{\partial f}{\partial x_i} = G_{ii}x_i + \frac{1}{2}\sum_{j\neq i} G_{ji}\, x_j + \frac{1}{2}\sum_{k\neq i} G_{ik}\, x_k + b_i = \sum_{j=1}^{n} G_{ij}x_j + b_i = (\mathbf{Gx} + \mathbf{b})_i$$

(因为 $G_{ij} = G_{ji}$)

因此 $\nabla f(\mathbf{x}) = \mathbf{Gx} + \mathbf{b}$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = G_{ij}$$

所以 $\nabla^2 f(x) = \mathbf{G}$

6. **[5.6]** 假设$\mathbf{G} \in \Re^{n\times n}$是正定矩阵，$\mathbf{b} \in \Re^n$，对二次函数

$f(\mathbf{x}) = \dfrac{1}{2}\mathbf{x}^T\mathbf{Gx} + \mathbf{b}^T\mathbf{x}$，证明其沿射线$\mathbf{x}_k + \alpha_k \mathbf{p}_k$的一维精确线搜索极小值为

$\alpha_k = -\dfrac{\nabla f_k^T \mathbf{p}_k}{\mathbf{p}_k^T \mathbf{G} \mathbf{p}_k}$。

解： 令：

$$\Phi(\alpha) = f(\boldsymbol{x_k} + \alpha \boldsymbol{p_k})$$

则需证明$\Phi(\alpha)$的极小值点为$-\dfrac{\nabla f_k^T \boldsymbol{p_k}}{\boldsymbol{p_k^T G p_k}}$

求导得：

$$\begin{aligned}\Phi'(\alpha) &= \nabla f(\boldsymbol{x_k} + \alpha\boldsymbol{p_k})^T \cdot \boldsymbol{p_k} = (\boldsymbol{G}(\boldsymbol{x_k} + \alpha\boldsymbol{p_k}) + \boldsymbol{b})^{\boldsymbol{T}}\boldsymbol{p_k} \\ &= (\boldsymbol{Gx_k} + \boldsymbol{b})^{\boldsymbol{T}}\boldsymbol{p_k} + \alpha\boldsymbol{p_k^T G^T p_k} = \nabla f(\boldsymbol{x_k})^T\boldsymbol{p_k} + \alpha\boldsymbol{p_k^T G^T p_k}\end{aligned}$$

$$\Phi''(\alpha) = \boldsymbol{p_k^T G^T p_k} > 0\left(\boldsymbol{G}\text{正定}\right)$$

因此令一阶导数等于 0 即得：

$$\alpha_k = -\frac{\nabla f(\boldsymbol{x_k})^T \boldsymbol{p_k}}{\boldsymbol{p_k^T G^T p_k}} = -\frac{\nabla f_k^T \boldsymbol{p_k}}{\boldsymbol{p_k^T G p_k}}$$

7. **[5.11]** 约束极值问题：

$$\begin{aligned} \min \quad & (x_1-2)^2 + x_2^2 \\ \text{s.t.} \quad & \begin{cases} x_1 - x_2^2 \ge 0 \\ -x_1 + x_2 \ge 0 \end{cases} \end{aligned}$$

检验 $\mathbf{x}^{(1)} = (0,0)^T$ 和 $\mathbf{x}^{(2)} = (1,1)^T$ 是否为 K-T 点。

解： $\nabla f(\mathbf{x}) = \begin{bmatrix} 2(x_1-2) \\ 2x_2 \end{bmatrix}, \nabla c_1(\mathbf{x}) = \begin{bmatrix} 1 \\ -2x_2 \end{bmatrix}, \nabla c_2(\mathbf{x}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

✓ 验证 $\mathbf{x}^{(1)}$。由于 $\nabla f\left(\mathbf{x}^{(1)}\right) = \begin{bmatrix} -4 \\ 0 \end{bmatrix}, \nabla c_1\left(\mathbf{x}^{(1)}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$，所以代入

$\nabla f\left(\mathbf{x}^{(1)}\right) - \lambda_1 \nabla c_1\left(\mathbf{x}^{(1)}\right) - \lambda_2 \nabla c_2\left(\mathbf{x}^{(1)}\right) = 0$，得：$\lambda_1 = -4,\ \lambda_2 = 0$。因为 $\lambda_1 < 0$，所以 $\mathbf{x}^{(1)}$ 不是 K-T 点。

✓ 验证 $\mathbf{x}^{(2)}$。由于 $\nabla f\left(\mathbf{x}^{(2)}\right) = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \nabla c_1\left(\mathbf{x}^{(2)}\right) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$，所以代入

$\nabla f\left(\mathbf{x}^{(2)}\right) - \lambda_1 \nabla c_1\left(\mathbf{x}^{(2)}\right) - \lambda_2 \nabla c_2\left(\mathbf{x}^{(2)}\right) = 0$，得：$\lambda_1 = 0,\ \lambda_2 = 2 \ge 0$，所以 $\mathbf{x}^{(2)}$ 是 K-T 点。