作业 #5

（提交日期：2023/12/12）

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. 假设$\mathbf{G}\in\Re^{n\times n}$是正定矩阵，$\mathbf{b}\in\Re^{n}$，求二次函数$f(\mathbf{x})=\frac{1}{2}\mathbf{x}^T\mathbf{G}\mathbf{x}+\mathbf{b}^T\mathbf{x}$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)

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

推导过程：

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

唯一包含变量x_i的项为：

$$\frac{1}{2}G_{ii}x_i^2+\frac{1}{2}\sum_{j\neq i}G_{ji}\,x_jx_i+\frac{1}{2}\sum_{k\neq i}G_{ik}\,x_ix_k+b_ix_i$$

对x_i求偏导得：

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

(因为 $G_{ij}=G_{ji}$)

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

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

所以 $\nabla^2 f(x)=\mathbf{G}$

5. **[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$f(\mathbf{x})=\frac{1}{2}\mathbf{x}^T\mathbf{G}\mathbf{x}+\mathbf{b}^T\mathbf{x}$，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$\alpha_k=-\frac{\nabla f_k^T\mathbf{p}_k}{\mathbf{p}_k^T\mathbf{G}\mathbf{p}_k}$。

解：令：

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

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

求导得：

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

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

因此令一阶导数等于 0 即得：

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

6. **[5.7]** 用黄金分割法求函数

$$f\left(x\right) = e^{-x} + x^2$$

在区间$\left[0,1\right]$上的近似极小点，要求缩短后的区间长度$L \le 0.2$。

解：计算列表如下。

k	a_k	b_k	$b_k - a_k$	λ_k	μ_k	$f\left(\lambda_k\right)$	$f\left(\mu_k\right)$
1	0	1	1	0.382	0.618	0.828	0.921
2	0	0.618	0.618	0.236	0.382	0.845	0.828
3	0.236	0.618	0.382	0.382	0.472	0.828	0.847
4	0.236	0.472	0.236	0.236	0.382	0.8281	0.8284
5	0.236	0.382	0.146				

当$k=5$时，近似极小点所在区间为$\left[0.236, 0.382\right]$，区间长度小于 0.2。近似极小点可取为：$\mathbf{x}^* \approx \frac{0.236+0.382}{2} = 0.309$。

（不列表，计算如下）$a_0 = 0, b_0 = 1$时，

$x_1 = 0.382 \times (1-0) = 0.382, x_1{}' = 0.618 \times (1-0) = 0.618.$

$f(x_1) = e^{-0.382} + 0.382^2 = 0.828, f(x_1{}') = e^{-0.618} + 0.618^2 = 0.921.$

由$f(x_1) < f(x_1{}')$得：

$a_1 = a_0 = 0, b_1 = x_1' = 0.618, x_2' = x_1 = 0.382,$

$x_2 = 0 + (1-0.618) \times (0.618 - 0) = 0.236$

$f(x_2) = e^{-0.236} + 0.236^2 = 0.845, f(x_2{}') = e^{-0.382} + 0.382^2 = 0.828.$

由$f(x_2) > f(x_2{}')$得：

$a_2 = x_2 = 0.236, b_2 = b_1 = 0.618, x_2' = x_3 = 0.382,$

$x_3{}' = 0.236 + 0.618 \times (0.618 - 0.236) = 0.472$

$f(x_3{}') = e^{-0.472} + 0.472^2 = 0.847, f(x_3) = e^{-0.382} + 0.382^2 = 0.828.$

由$f(x_3) < f(x_3{}')$得：

$a_3 = a_2 = 0.236, b_3 = x_3' = 0.472, x_4' = x_3 = 0.382,$

$x_4 = 0.236 + (1-0.618) \times (0.472 - 0.236) = 0.326$

$f(x_4) = e^{-0.326} + 0.326^2 = 0.8281, f(x_4') = e^{-0.382} + 0.382^2 = 0.8284.$

由$f(x_4) < f(x_4{}')$得：

$a_4 = a_3 = 0.236, b_4 = x_4' = 0.382$

$b_4 - a_4 = 0.146 < 0.2$

所以区间[0.236,0.382]为所求区间

$x^* \approx \dfrac{0.236 + 0.382}{2} = 0.309.$

7. **[5.11]** 约束极值问题：

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

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。