浙江工业大学《最优化方法》期末试卷

（ 2019 ~ 2020 第 一 学 期 ）

学院、班级________________ 学号________________ 姓名______________

一	二	三	四	五	六	七	八	九	总分

一、（10 分）证明 $f(x)=x_1e^{-(x_1+x_2)}$ 在任何区域都不是凸函数。

$$\frac{\partial f}{\partial x_1}=(1-x_1)e^{-(x_1+x_2)},\quad \frac{\partial f}{\partial x_2}=-x_1e^{-(x_1+x_2)}$$ ---2分

$$\frac{\partial^2 f}{\partial x_1^2}=(x_1-2)e^{-(x_1+x_2)},\quad \frac{\partial^2 f}{\partial x_1\partial x_2}=\frac{\partial^2 f}{\partial x_2\partial x_1}=(x_1-1)e^{-(x_1+x_2)}$$

---6分

$$\frac{\partial^2 f}{\partial x_2^2}=x_1e^{-(x_1+x_2)}$$

---8分

$$\nabla^2 f(x)=e^{-(x_1+x_2)}\begin{pmatrix} x_1-2 & x_1-1 \\ x_1-1 & x_1 \end{pmatrix} \qquad \begin{vmatrix} x_1-2 & x_1-1 \\ x_1-1 & x_1 \end{vmatrix}=-1<0$$

故 $\nabla^2 f(x)$ 在任何区域都不是半正定的，$f(x)$ 在任何区域都非凸。

---10分

二、（10 分）求出函数

$$f(x_1,x_2)=x_1x_2(3-x_1-x_2)$$

的所有稳定点，并判断它们是极大值点、极小值点还是鞍点。

$$\nabla f(x)=\begin{pmatrix} 3x_2-2x_1x_2-x_2^2 \\ 3x_1-x_1^2-2x_1x_2 \end{pmatrix}=0$$

$\Rightarrow$ 稳定点 $x^{(1)}=\begin{pmatrix}0\\0\end{pmatrix}$，$x^{(2)}=\begin{pmatrix}1\\1\end{pmatrix}$，$x^{(3)}=\begin{pmatrix}3\\0\end{pmatrix}$，$x^{(4)}=\begin{pmatrix}0\\3\end{pmatrix}$ ----4分

$$\nabla^2 f(x)=\begin{pmatrix} -2x_2 & 3-2x_1-2x_2 \\ 3-2x_1-2x_2 & -2x_1 \end{pmatrix}$$ ---6分

$\nabla^2 f(x^{(1)})=\begin{pmatrix}0 & 3\\ 3 & 0\end{pmatrix}$ 不定，$x^{(1)}$ 为鞍点；

$\nabla^2 f(x^{(2)})=\begin{pmatrix}-2 & -1\\ -1 & -2\end{pmatrix}$ 负定，$x^{(2)}$ 为极大值点；

$\nabla^2 f(x^{(3)})=\begin{pmatrix}0 & -3\\ -3 & -6\end{pmatrix}$ 不定，$x^{(3)}$ 为鞍点；

$\nabla^2 f(x^{(4)})=\begin{pmatrix}-6 & -3\\ -3 & 0\end{pmatrix}$ 不定，$x^{(4)}$ 为鞍点。 ----10分

三、（12 分）考虑无约束优化问题

$$\min f(x_1,x_2)=(1-x_1)^2+2(x_2-x_1^2)^2,$$

选取初始点 $x^{(0)}=\left(\frac{1}{2},0\right)^{\mathrm{T}}$，利用最速下降法和带步长因子的牛顿法，选择精确线搜索，各迭代一步，并判断哪个方法得到的是该问题的最优解。

$$\nabla f(x)=\begin{pmatrix}-2(1-x_1)-8(x_2-x_1^2)x_1\\ 4(x_2-x_1^2)\end{pmatrix},\quad \nabla^2 f(x)=\begin{pmatrix}16x_1^2-8(x_2-x_1^2)+2 & -8x_1\\ -8x_1 & 4\end{pmatrix}$$ ···4分

$\nabla f(x^{(0)})=\begin{pmatrix}0\\-1\end{pmatrix}$, $\nabla^2 f(x^{(0)})=\begin{pmatrix}8&-4\\-4&4\end{pmatrix}$ 最速下降方向 $d_S=\begin{pmatrix}0\\1\end{pmatrix}$ ···5分

$\alpha_0=\arg\min\limits_{\alpha\geq 0} f(x^{(0)}+\alpha d_S)=\arg\min\limits_{\alpha\geq 0}\left[\frac{1}{4}+2\left(\alpha-\frac{1}{4}\right)^2\right]=\frac{1}{4}$ ···7分

$x^{(1)}=\begin{pmatrix}1/2\\1/4\end{pmatrix}$. 牛顿方向 $d_N=-\begin{pmatrix}8&-4\\-4&4\end{pmatrix}^{-1}\begin{pmatrix}0\\-1\end{pmatrix}=\begin{pmatrix}1/4\\1/2\end{pmatrix}$ ···9分

$\alpha_0=\arg\min\limits_{\alpha\geq 0} f(x^{(0)}+\alpha d_N)=\arg\min\limits_{\alpha\geq 0}\frac{1}{128}\left[8(2-\alpha)^2+(2-\alpha)^4\right]=2$, ···11分

$x^{(1)}=\begin{pmatrix}1\\1\end{pmatrix}$. $\nabla f\left(\begin{pmatrix}1\\1\end{pmatrix}\right)=\begin{pmatrix}0\\0\end{pmatrix}$, $\nabla^2 f\left(\begin{pmatrix}1\\1\end{pmatrix}\right)=\begin{pmatrix}18&-8\\-8&4\end{pmatrix}$ 正定，故

牛顿法得到的 $\begin{pmatrix}1\\1\end{pmatrix}$ 为该问题的最优解.

···12分

四、（10 分）考虑约束最优化问题：

$$\begin{aligned}\min\ & f(x)=x_1^2+x_2^2-4x_1+4\\ s.t.\ & c_1(x)=x_1-x_2+2\geq 0,\\ & c_2(x)=-x_1^2+x_2-1\geq 0,\\ & c_3(x)=x_1\geq 0,\\ & c_4(x)=x_2\geq 0.\end{aligned}$$

（1）验证该问题为凸规划问题；

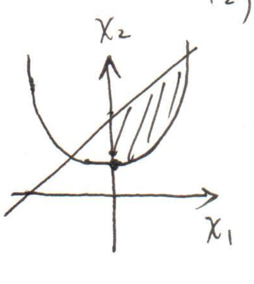
（2）给出可行点 $\bar{x}=(0,1)^{\mathrm{T}}$ 处的线性化可行方向 $d=(d_1,d_2)^{\mathrm{T}}$ 应满足的条件（用 d_1,d_2 的表达式给出）；并判断 $\bar{x}$ 处的线性化可行方向集合与可行方向集合是否相等。

(1) $\nabla^2 f(x)=\begin{pmatrix}2&\\&2\end{pmatrix}$ 正定，$f(x)$ 为凸函数.

c_1, c_3, c_4 为线性函数，对 $\nabla^2 c_2(x)=\begin{pmatrix}-2&0\\0&0\end{pmatrix}$ 半负定. $c_2(x)$ 为凹函数.

故该问题为凸规划. ···4分

(2) 点 $\bar{x}$ 处的起作用约束为 c_2, c_3, $\nabla c_2=\begin{pmatrix}-2x_1\\1\end{pmatrix}$ $\nabla c_3=\begin{pmatrix}1\\0\end{pmatrix}$



$\begin{cases}d^{\mathrm{T}}\nabla c_2(\bar{x})\geq 0\\ d^{\mathrm{T}}\nabla c_3(\bar{x})\geq 0\\ d\neq 0\end{cases}\Rightarrow d_1\geq 0,\ d_2\geq 0$ 且 $d\neq 0$ 为线性化可行方向应满足条件. ···8分

由图知，$\bar{x}$ 处的方向 $\begin{pmatrix}d_1\\0\end{pmatrix}$ 当 $d_1>0$ 时 为线性化可行方向. 但不为可行方向

故两集合不相等. ···10分

五、（12 分）利用逐次插值逼近法求解一维极小化问题

$$\min_{\alpha\geq 0}\varphi(\alpha)=-4\alpha^3+10\alpha^2-5\alpha+1,$$

（1）取插值点 $\alpha_1=0$，$\alpha_2=1$，利用插值条件 $\varphi(\alpha_1),\varphi'(\alpha_1),\varphi'(\alpha_2)$ 构造二次插值多项式，求出插值多项式的极小点 $\bar{\alpha}$，为保证每步迭代原函数在两个插值点处的导数值异号，下一步迭代的两个插值点是哪些？

（2）仍取 $\alpha_1=0$，利用插值条件 $\varphi(\alpha_1),\varphi'(\alpha_1),\varphi''(\alpha_1)$ 构造二次插值多项式，则插值多项式的极小点 $\tilde{\alpha}$ 又会是多少？

（3）简要对比上述两个方法的优缺点。

$\varphi'(\alpha)=-12\alpha^2+20\alpha-5$

(1) $\varphi(\alpha_1)=1,\ \varphi'(\alpha_1)=-5,\ \varphi'(\alpha_2)=3.$

$$\begin{cases} a\alpha_1^2+b\alpha_1+c=\varphi(\alpha_1)=1 \\ 2a\alpha_1+b=\varphi'(\alpha_1)=-5 \\ 2a\alpha_2+b=\varphi'(\alpha_2)=3 \end{cases} \Rightarrow \begin{cases} a=4 \\ b=-5 \\ c=1 \end{cases} \Rightarrow \bar{\alpha}=\frac{-b}{2a}=\frac{5}{8}.$$

$\varphi'(\bar{\alpha})=\frac{45}{16}>0$，故下一步迭代取插值点 $\alpha_1,\bar{\alpha}$. ---5分

(2) $\varphi''(\alpha)=-24\alpha+20\quad \varphi''(\alpha_1)=20.$

$$\begin{cases} a\alpha_1^2+b\alpha_1+c=1 \\ 2a\alpha_1+b=-5 \\ 2a=20 \end{cases} \Rightarrow \begin{cases} a=10 \\ b=-5 \\ c=1 \end{cases} \Rightarrow \tilde{\alpha}=\frac{-b}{2a}=\frac{1}{4}.$$

--- 10分

(3). (1)为割线法. 仅用一阶导数. 收敛阶次为1.618.

(2)为牛顿法. 需用二阶导数. 某些问题二阶导较难获得.

收敛阶为2阶. 比割线法快. ---12分

六、（10 分）利用对数障碍函数法求解

$$\min\ x_1x_2$$
$$s.t.\quad -2x_1+x_2+3\geq 0.$$

$P(x;\mu)=x_1x_2-\mu\log(-2x_1+x_2+3)$ ---2分

$$\begin{cases} \dfrac{\partial P}{\partial x_1}=x_2+\dfrac{2\mu}{-2x_1+x_2+3}=0 \\ \dfrac{\partial P}{\partial x_2}=x_1-\dfrac{\mu}{-2x_1+x_2+3}=0 \end{cases}$$ ---4分

解得

$$\begin{cases} x_1(\mu)=\dfrac{3+\sqrt{9-16\mu}}{8} \\ x_2(\mu)=-\dfrac{3+\sqrt{9-16\mu}}{4} \end{cases}$$

或

$$\begin{cases} x_1(\mu)=\dfrac{3-\sqrt{9-16\mu}}{8} \\ x_2(\mu)=-\dfrac{3-\sqrt{9-16\mu}}{4} \end{cases}$$ ---8分

令 $\mu\to 0^+$ 得

$$x^*=\begin{pmatrix}\frac{3}{4}\\ -\frac{3}{2}\end{pmatrix}$$ 或 $$x^*=\begin{pmatrix}0\\0\end{pmatrix}$$ ---10分

七、（14 分）（1）取初始点 $x^{(0)}=(6,3)^{\mathrm{T}}$，用 FR 共轭梯度法求解

$$\min_{x\in R^2} f(x)=\frac{1}{2}x_1^2+x_2^2,$$

其中 FR 公式：$\beta_{k-1}=\dfrac{g_k^{\mathrm{T}}g_k}{g_{k-1}^{\mathrm{T}}g_{k-1}}$。

（2）与最速下降法和牛顿法相比，共轭梯度法有哪些优点？试从计算量、存储量、收敛速度等方面讨论。

(1) $f(x)=\frac{1}{2}x^{T}Gx \qquad G=\begin{pmatrix}1 & \\ & 2\end{pmatrix}$

$x^{(0)}=\begin{pmatrix}6\\3\end{pmatrix}$, $g_0=Gx^{(0)}=\begin{pmatrix}6\\6\end{pmatrix}$, $d_0=-g_0=\begin{pmatrix}-6\\-6\end{pmatrix}$ ……3分

$\alpha_0=-\dfrac{d_0^{T}g_0}{d_0^{T}Gd_0}=\dfrac{2}{3}$, $x^{(1)}=x^{(0)}+\alpha_0 d_0=\begin{pmatrix}2\\-1\end{pmatrix}$ ……5分

$g_1=Gx^{(1)}=\begin{pmatrix}2\\-2\end{pmatrix}$, $\beta_0=\dfrac{g_1^{T}g_1}{g_0^{T}g_0}=\dfrac{1}{9}$, ……7分

$d_1=-g_1+\beta_0 d_0=\begin{pmatrix}-8/3\\ \frac{4}{3}\end{pmatrix}$ $\quad \alpha_1=-\dfrac{d_1^{T}g_1}{d_1^{T}Gd_1}=\dfrac{3}{4}$ ……9分

$x^{(2)}=x^{(1)}+\alpha_1 d_1=\begin{pmatrix}0\\0\end{pmatrix}$, $g_2=Gx^{(2)}=\begin{pmatrix}0\\0\end{pmatrix}$. ……11分

(2) 共轭梯度法仅需利用一阶导数信息，但克服了最速下降法收敛慢的缺点，又避免了牛顿法需要存储和计算Hesse矩阵并求逆的缺点，是解大型非线性最优化问题的首选方法。

……14分

八、（12 分）求原点$\begin{pmatrix}0\\0\end{pmatrix}$到凸集$S=\{x|x_1+x_2-4\geq 0,\ 2x_1+x_2-5\geq 0\}$的最小距离，将该问题写成二次规划问题，列出其 KKT 条件并进行求解。

可写成如下二次规划

$$\min\ x_1^2+x_2^2$$

$$\text{s.t}\quad x_1+x_2-4\geq 0$$

$$2x_1+x_2-5\geq 0. \quad \text{---2分}$$

KKT条件：

$$\begin{cases}\begin{pmatrix}2x_1\\2x_2\end{pmatrix}=\lambda_1\begin{pmatrix}1\\1\end{pmatrix}+\lambda_2\begin{pmatrix}2\\1\end{pmatrix}\\ x_1+x_2-4\geq 0\\ 2x_1+x_2-5\geq 0\\ \lambda_1\geq 0,\ \lambda_2\geq 0. \quad \text{---6分}\\ \lambda_1(x_1+x_2-4)=0\\ \lambda_2(2x_1+x_2-5)=0\end{cases}$$

(a) $\lambda_1=0,\lambda_2=0 \Rightarrow x_1=x_2=0$ 不可行

(b) $\lambda_1=0\quad 2x_1+x_2-5=0,\ \begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}2\\1\end{pmatrix}\quad \lambda_2=2.$ 不可行.

(c) $x_1+x_2-4=0,\ \lambda_2=0,\ \begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}2\\2\end{pmatrix}$ 可行 $\lambda_1=4>0$、为KKT点.

(d) $x_1+x_2-4=0.\ 2x_1+x_2-5=0.\ \begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}1\\3\end{pmatrix}\Rightarrow\lambda_2=-4<0$. 违反乘子非负.

---10分

故KKT点为$\begin{pmatrix}2\\2\end{pmatrix}$. 又该问题为凸规划. 故$\begin{pmatrix}2\\2\end{pmatrix}$为全局最优解.

最小距离为 $2\sqrt{2}$. ---12分

九、（10 分）考虑二次函数极小化问题

$$\min f(x)=\frac{1}{2}x^{\mathrm{T}}Ax+b^{\mathrm{T}}x,$$

其中 A 为对称正定矩阵。设 $\bar{x}$ 是 $f(x)$ 的极小点，p 是 A 的属于特征值 λ 的特征向量，$x^{(1)}=\bar{x}+p$，

证明：（1）$\nabla f(x^{(1)})=\lambda p$.

（2）如果用最速下降法，从 $x^{(1)}$ 出发，采用精确线搜索，则一步可达最优解 $\bar{x}$ 。

(1) $\nabla f(x)=Ax+b$，$\nabla^2 f(x)=A$

由 A 对称正定，故 $\bar{x}=-A^{-1}b$ 满足 $\nabla f(\bar{x})=0$ 为 $f(x)$ 的极小点。 ……2分

$\nabla f(x^{(1)})=Ax^{(1)}+b=A(\bar{x}+p)+b=Ap=\lambda p$

……4分

(2). 最速下降方向 $d_s=-\nabla f(x^{(1)})=-\lambda p$. ……6分

精确线搜索步长 $\alpha=-\dfrac{d_s^T\nabla f(x^{(1)})}{d_s^TAd_s}$

$$=-\frac{-\lambda p^T\cdot\lambda p}{-\lambda p^TA(-\lambda p)}$$

$$=\frac{\lambda^2p^Tp}{\lambda^2p^TAp}=\frac{p^Tp}{p^T\lambda p}=\frac{1}{\lambda}$$

……8分

$x^{(2)}=x^{(1)}+\alpha d_s=\bar{x}+p+\frac{1}{\lambda}\cdot(-\lambda p)=\bar{x}$,

即一步可达最优解。 ……10分