浙江工业大学《最优化方法》期末试卷

(2011 ~ 2012 第一学期)

一、 (10 分) 设多元函数 $f:\mathbf{R}^n\to\mathbf{R}^1$ 是凸函数，试证复合函数 $h(x)=e^{f(x)}, x\in\mathbf{R}^n$ 是凸函数。

答:

先证 e^x 是凸函数，将 $\nabla^2 e^x = e^x$ 看作 1×1 的矩阵，显然 $\nabla^2 e^x$ 正定，则 e^x 是凸函数，且单调递增. $\forall x, y\in\mathbf{R}^n, \lambda\in[0,1]$，则

$$
\begin{aligned}
h[\lambda x+(1-\lambda)y] &= \exp\{f[\lambda x+(1-\lambda)y]\}\\
&\le \exp[\lambda f(x)+(1-\lambda)f(y)] \quad (f\text{是凸函数，以及}e^x\text{单调递增})\\
&\le \lambda\exp[f(x)]+(1-\lambda)\exp[f(y)] \quad (e^x\text{是凸函数})\\
&= \lambda h(x)+(1-\lambda)f(y)
\end{aligned}
$$

所以 $h(x)=e^{f(x)}, x\in\mathbf{R}^n$ 是凸函数。

二、 (10 分)

(1) 试列举所学过的精确线搜索方法。

(2) 证明精确线搜索满足 $g_{k+1}^{\mathrm{T}}d_k=0$，并依此说明最速下降法收敛缓慢的原因。

答:

(1) 0.618 法、Fibonacci 法、二分法。

(2) 精确线搜索下，令 $\varphi(\alpha)=f(x_k+\alpha d_k)$，所以

$$
\left.\frac{\mathrm{d}\varphi(\alpha)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_k} = \nabla f(x_k+\alpha_k d_k)^{\mathrm{T}}d_k = \nabla f(x_{k+1})^{\mathrm{T}}d_k = g_{k+1}^{\mathrm{T}}d_k = 0,
$$

证毕.

最速下降法中 $d_k=-g_k$，所以有

$$
g_{k+1}^{\mathrm{T}}g_k = d_{k+1}^{\mathrm{T}}d_k = 0,
$$

所以相邻两次搜索方向是相互直交的，若目标函数等值线接近一个椭球，则会产生锯齿现象，越接近极小点，步长越小，前进越慢，故最速下降法收敛缓慢.

三、 (10 分) 拟牛顿法有哪些优点？

答：

(1) 仅需一阶导数，无需二阶导数；
(2) $B_k(H_k)$ 正定，方向具有下降性质；
(3) 需 $O(n^2)$ 次乘法运算；
(4) 搜索方向相互共轭，有二次终止性；
(5) 超线性收敛性；
(6) 是在椭球范数 $||\cdot||_{B_k}$ 意义下的最速下降法

四、 (12 分) 试用最优性条件求解如下无约束优化问题的局部最优解：

$$(P_1): \min_{x\in\mathbf{R}^2} f(x) = 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3$$

答：

$$\nabla f = \begin{pmatrix} 4x_1 - 2x_2 + 6x_1^2 \\ 2x_2 - 2x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \nabla^2 f = \begin{pmatrix} 12x_1 + 4 & -2 \\ -2 & 2 \end{pmatrix}$$

解得 $x^{(1)} = (0,0)^{\mathrm{T}}, x^{(2)} = \left(-\dfrac{1}{3}, -\dfrac{1}{3}\right)$，代入 $\nabla^2 f$ 得：

$\nabla^2 f(x^{(1)}) = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix}$，正定，$x^{(1)}$ 是局部极小点；

$\nabla^2 f(x^{(2)}) = \begin{pmatrix} 0 & -2 \\ -2 & 2 \end{pmatrix}$，不定，$x^{(2)}$ 是鞍点。

五、 (16 分) 取初始点 $x^{(0)} = (2,4)^{\mathrm{T}}$，用 FR 共轭梯度法求解

$$(P_2): \min_{x\in\mathbf{R}^2} f(x) = 3x_1^2 + x_2^2 - 2x_1x_2 - 4x_1,$$

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

答：

$$G = \begin{pmatrix} 6 & -2 \\ -2 & 2 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 0 \end{pmatrix},$$

所以

(1) $g_0 = Gx^{(0)} - b = (0,4)^{\mathrm{T}}$，$d_0 = -g_0 = (0,-4)^{\mathrm{T}}$，$\alpha_0 = -\dfrac{g_0^{\mathrm{T}} d_0}{d_0^{\mathrm{T}} G d_0} = \dfrac{1}{2}$；

(2) $x^{(1)}=x^{(0)}+\alpha_0 d_0=(2,2)^{\mathrm{T}}$, $g_1=Gx^{(1)}-b=(4,0)^{\mathrm{T}}$, $\beta_0=\dfrac{g_1^{\mathrm{T}}g_1}{g_0^{\mathrm{T}}g_0}=1$, $d_1=-g_1+\beta_0 d_0=(-4,-4)^{\mathrm{T}}$，

$\alpha_1=-\dfrac{g_1^{\mathrm{T}}d_1}{d_1^{\mathrm{T}}Gd_1}=\dfrac{1}{4}$，$x^{(2)}=x^{(1)}+\alpha_1 d_1=(1,1)^{\mathrm{T}}$，$g_2=Gx^{(2)}-b=(0,0)^{\mathrm{T}}$。

所以极小点为 $(1,1)^{\mathrm{T}}$，极小值为 -2。

六、 (18 分) 对于约束最优化问题：

$$(P_3):\quad \begin{array}{ll} \min & (x_1-1)^2+(x_2-1)^2 \\ \text{s.t.} & 1-x_1^2-x_2^2\geq 0, \\ & 1-x_1-x_2\geq 0. \end{array}$$

(1) 写出 (P_3) 的 KKT 条件；

(2) 求出 (P_3) 的 KKT 点，并判断它 (们) 是否为全局最优解 (写出理由)。

答:

(1) KKT 条件如下：

$$\begin{cases} \begin{pmatrix} 2x_1-2 \\ 2x_2-2 \end{pmatrix}=\lambda_1\begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}+\lambda_2\begin{pmatrix} -1 \\ -1 \end{pmatrix} \\ 1-x_1^2-x_2^2\geq 0 \\ 1-x_1-x_2\geq 0 \\ \lambda_1,\lambda_2\geq 0 \\ \lambda_1(1-x_1^2-x_2^2)=0 \\ \lambda_2(1-x_1-x_2)=0 \end{cases}$$

(2) 分情况讨论：

(a) $\lambda_1=0,\lambda_2=0$，则 $x_1=x_2=1$，不可行；

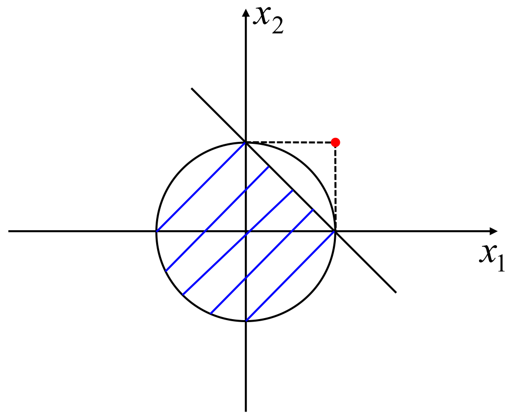
(b) $\lambda_1=0, 1-x_1-x_2=0$，则 $x_1=x_2=\dfrac{1}{2},\lambda_2=1$，可行，是 KKT 点；

(c) $\lambda_2=0, 1-x_1^2-x_2^2=0$，则 $x_1=x_2=\dfrac{\sqrt{2}}{2}$(不可行) 或 $x_1=x_2=-\dfrac{\sqrt{2}}{2},\lambda_1=1+\sqrt{2}$(可行，是 KKT 点)；

(d) $1-x_1^2-x_2^2=0, 1-x_1-x_2=0$，则 $x_1=1,x_2=0,\lambda_1=-1,\lambda_2=2$，不可行。

所以 KKT 点是 $\left(\dfrac{1}{2},\dfrac{1}{2}\right)^{\mathrm{T}}$ 和 $\left(-\dfrac{\sqrt{2}}{2},-\dfrac{\sqrt{2}}{2}\right)^{\mathrm{T}}$。

利用图解法判断是否为全局最优解：



几何意义就是求蓝色区域内的点到点 $(1,1)^{\mathrm{T}}$ 的最短距离的平方，故全局最优解为 $\left(\dfrac{1}{2},\dfrac{1}{2}\right)^{\mathrm{T}}$。

七、 (12 分) 利用直接消去法求解如下二次规划问题：

$$(P_4):\quad \begin{array}{ll} \min\limits_{x\in\mathbf{R}^3} & q(x)=x_1^2+2x_2^2+3x_3^2 \\ \text{s.t.} & x_1+x_2+x_3=10. \end{array}$$

答：

拉格朗日函数为 $L(x,\lambda)=x_1^2+2x_2^2+3x_3^2-\lambda(x_1+x_2+x_3-10)$，则

$$\begin{cases} \frac{\partial L}{\partial x_1}=2x_1-\lambda=0, \\ \frac{\partial L}{\partial x_2}=4x_2-\lambda=0, \\ \frac{\partial L}{\partial x_3}=6x_3-\lambda=0, \\ \frac{\partial L}{\partial \lambda}=x_1+x_2+x_3-10=0 \end{cases} \qquad \Rightarrow x^*=\begin{pmatrix} \frac{60}{11} \\ \frac{30}{11} \\ \frac{20}{11} \end{pmatrix},\lambda^*=\frac{120}{11}.$$

八、 (12 分) 对于非线性约束最优化问题

$$(P_5):\quad \begin{array}{ll} \min & x_2-2x_1 \\ \text{s.t.} & x_2-x_1^2\geq 0, \\ & x_1+1\geq 0. \end{array}$$

写出其对数障碍函数 $P(x;\mu)$，并利用对数障碍函数方法和一阶最优化条件求解 (P_5)。

答：

对数障碍函数为

$$P(x;\mu) = x_2 - 2x_1 - \mu\log(x_2 - x_1^2) - \mu\log(x_1 + 1),$$

所以

$$\begin{aligned}\frac{\partial P}{\partial x_1} &= -2 - \mu\frac{-2x_1}{x_2 - x_1^2} - \mu\frac{1}{x_1 + 1} = 0,\\ \frac{\partial P}{\partial x_2} &= 1 - \mu\frac{1}{x_2 - x_1^2} = 0,\end{aligned}$$

解得

$$x_1 = \sqrt{\frac{2+\mu}{2}}, x_2 = \frac{2+3\mu}{2},$$

令 $\mu \to 0^+$，则

$$x_1 \to 1, x_2 \to 1,$$

所以极小点为 $(1,1)^{\mathrm{T}}$。