数学基础 11th

习题一

(1) $\max \quad f(x) = (x-3)^2 \qquad 1 \le x \le 5$

$\Downarrow$

$\min \quad f(x) = -(x-3)^2 \qquad g_1(x) = 1-x \le 0$

$g_2(x) = x-5 \le 0$

$$L(x, \lambda_1, \lambda_2) = -(x-3)^2 + \lambda_1(1-x) + \lambda_2(x-5)$$

KKT 条件 $\dfrac{\partial L(x, \lambda_1, \lambda_2)}{\partial (x)} = -2(x-3) - \lambda_1 + \lambda_2 = 0$

$\lambda_1 g_1(x) = \lambda_1(1-x) = 0$

$\lambda_2 g_2(x) = \lambda_2(x-5) = 0$

$g_1(x) \le 0$

$g_2(x) \le 0$

$\lambda_1 \ge 0$

$\lambda_2 \ge 0$

求解：

$\lambda_1 = \lambda_2 = 0 \Rightarrow x = 3$

λ_1, λ_2 都 $\ne 0 \Rightarrow x=1$ 且 $x=5$ 且 $x=3$ ×

$\lambda_1 = 0 \quad \lambda_2 \ne 0 \Rightarrow x = 5, \ \lambda_2 = 4$

$\lambda_2 = 0, \ \lambda_1 \ne 0 \Rightarrow x = 1, \ \lambda_1 = 4$

$x = 1$ 或 5 时 $\min f(x) = -4$

$\max f(x) = 4$

(2) min $f(x)=(x-3)^2$ $1\le x\le 5$

$g_1(x)=1-x\le 0$

$g_2(x)=x-5\le 0$

$$L(x,\lambda_1,\lambda_2)=(x-3)^2+\lambda_1(1-x)+\lambda_2(x-5)$$

KKT条件

$$\frac{\partial L(x,\lambda_1,\lambda_2)}{\partial(x)}=2(x-3)-\lambda_1+\lambda_2=0$$

$\lambda_1 g_1(x)=\lambda_1(1-x)=0$

$\lambda_2 g_2(x)=\lambda_2(x-5)=0$

$g_1(x)\le 0$

$g_2(x)\le 0$

$\lambda_1\ge 0$

$\lambda_2\ge 0$

求解:

$\lambda_1=\lambda_2=0 \Rightarrow x=3$

λ_1,λ_2都$\neq 0 \Rightarrow x=1$ 且 $x=5$ 且 $x=3$ ✗

$\lambda_1=0$ $\lambda_2\neq 0 \Rightarrow x=5,\ \lambda_2=-4$ ✗

$\lambda_2=0$ $\lambda_1\neq 0 \Rightarrow x=1,\ \lambda_1=-4$ ✗

$\therefore$ 当 $x=3$ 时 $\min f(x)=0$ $(\lambda_1=\lambda_2=0)$

习题2：

$$\min \|Ax-b\|_2^2 \qquad Gx=h$$

A $m\times n$

x $n\times 1$

b $m\times 1$

$Ax-b$ $m\times 1$

$$\|Ax-b\|_2^2 = (Ax-b)^T(Ax-b)$$

$$L(x,v) = (Ax-b)^T(Ax-b) + v^T(Gx-h)$$

$$= (x^TA^T-b^T)(Ax-b)+v^T(Gx-h)$$

$$= x^TA^TAx - x^TA^Tb - b^TAx + b^Tb + v^TGx - v^Th$$

$$(x^TA^Tb)^T = b^TAx$$

$$= x^TA^TAx + (v^TG-b^TA)x + b^Tb - v^Th - \underbrace{x^TA^Tb}_{1\times 1}$$

$$= x^TA^TAx + (G^Tv-2A^Tb)^Tx - v^Th$$

$$\frac{\partial L(x,v)}{\partial (x)} = A^TAx + A^TAx + (v^TG-b^TA)^T - A^Tb$$

$$= 2A^TAx + G^Tv - A^Tb - A^Tb = 2A^TAx - 2A^Tb + G^Tv$$

KKT 条件

$$Gx^* = h \quad ① \qquad 2A^TAx - 2A^Tb + G^Tv^* = 0 \quad ②$$

$$2A^TAx^* = 2A^Tb - G^Tv^*$$

$$A^TAx^* = A^Tb - \frac{1}{2}G^Tv^*$$

$\because A \in R^{m\times n}$ $\quad rank(A)=n$ $\quad \therefore A^TA$ 可逆

$$x^* = (A^TA)^{-1}(A^Tb - \frac{1}{2}G^Tv^*) \quad ③$$

③代入①

$$G(A^TA)^{-1}(A^Tb - \frac{1}{2}G^Tv^*) = h$$

$$G(A^TA)^{-1}A^Tb - \frac{1}{2}G(A^TA)^{-1}G^Tv^* = h$$

$$G(A^TA)^{-1}G^Tv^* = 2(G(A^TA)^{-1}A^Tb - h)$$

$rank(G)=p$ $\quad \therefore G(A^TA)^{-1}G^T$ 可逆

$$\therefore \boxed{v^* = 2(G(A^TA)^{-1}G^T)^{-1}(G(A^TA)^{-1}A^Tb-h) \quad ④}$$

④代入③

$$\therefore \boxed{x^* = (A^TA)^{-1}(A^Tb - G^T(G(A^TA)^{-1}G^T)^{-1}(G(A^TA)^{-1}A^Tb-h))}$$

习题3、

$$\|A\|_2 = \max_{\|X\|_2=1,\ X\in R^{n\times 1}} \|AX\|_2$$

$$\|A\|_2^2 = \left(\max_{\|X\|_2=1,\ X\in R^{n\times 1}} \|AX\|_2\right)^2 = \max_{\|X\|_2=1,\ X\in R^{n\times 1}} \|AX\|_2^2$$

$\|X\|_2=1 = \sqrt{X^TX}=1$ 等价于 $X^TX=1 \Leftrightarrow \|X\|_2^2=1$

$\therefore\ L(X,\lambda) = \|AX\|_2^2 - \lambda(\|X\|_2^2-1) = X^TA^TAX + \lambda(X^TX-1)$

$$\frac{\partial L(X,\lambda)}{\partial (X)} = A^TAX + A^TAX - \lambda(X+X) = 0 \quad ①$$

$$\frac{\partial L(X,\lambda)}{\partial (\lambda)} = X^TX-1 = 0$$

$\cancel{2}A^TAX = \cancel{2}\lambda X \qquad X^TX=1 \qquad \|X\|_2=1 \quad \therefore X\neq 0$

$\therefore$ λ是A^TA的特征值，X为特征向量

$$\therefore \max_{\|X\|_2=1,\ X\in R^{n\times 1}} \|AX\|_2^2 = \max_{\|X\|_2=1,\ X\in R^{n\times 1}} X^TA^TAX = \max_{\|X\|_2=1,\ X\in R^{n\times 1}} X^T\lambda X$$

$$= \max_{\|X\|_2=1,\ X\in R^{n\times 1}} \lambda X^TX$$

$\because$ 取max时 $X^TX=1$

$\therefore$ 原式 $= \max \lambda$ （λ是A^TA的特征值）

$\therefore\ \|A\|_2 = \max_{\|X\|_2=1,\ X\in R^{n\times 1}} \|AX\|_2$ 的平方，即 $\max_{\|X\|_2=1,\ X\in R^{n\times 1}} \|AX\|_2^2$ 是A^TA特征值的最大值

习题4

$$\max \quad f(x) = \frac{1}{2}\|x\|_2^2$$
$$\text{s.t} \quad Ax = b$$

$$L(x,\lambda) = \frac{1}{2}\|x\|_2^2 - \lambda^T(Ax-b) = \frac{1}{2}x^Tx - \lambda^T(Ax-b)$$

$$\frac{\partial L}{\partial x} = x - A^T\lambda = 0$$
$$x = A^T\lambda$$

$$g(\lambda) = \frac{1}{2}\lambda^T A A^T\lambda - \lambda^T(AA^T\lambda - b) = \frac{1}{2}\lambda^T A A^T \lambda - \lambda^T A A^T \lambda + \lambda^T b$$
$$= -\frac{1}{2}\lambda^T A A^T \lambda + \lambda^T b$$

$$\frac{\partial g}{\partial \lambda} = -AA^T\lambda + b = 0$$

$\because A \in R^{m\times n}$, $\text{rank}(A) = m$ $\therefore AA^T$ 可逆

$$\therefore \lambda = (AA^T)^{-1}b$$

$\therefore$ x 满足 $Ax=b$ 的最小二范数解

$$x = A^T(AA^T)^{-1}b$$

扫描全能王 创建

习题五：

由题意：

$f(x) = x^T x$ ，$\nabla_x f(x) = 2x$ ，设最速下降法的步长为 λ

$$f(x - \lambda \nabla_x f(x)) = (x - \lambda \nabla_x f(x))^T (x - \lambda \nabla_x f(x))$$
$$= x^T x - 2\lambda \nabla_x f(x)^T x + \lambda^2 \nabla_x f(x)^T \nabla_x f(x)$$

在 $x - \lambda \nabla_x f(x)$ 方向上，使最小的 λ 满足

$$\frac{\partial f(x - \lambda \nabla_x f(x))}{\partial \lambda} = -2\nabla_x f(x)^T x + 2\lambda \nabla_x f(x)^T \nabla_x f(x)$$

$$\lambda = \frac{\nabla_x f(x)^T x}{\nabla_x f(x)^T \nabla_x f(x)} = \frac{1}{2}$$

所以

$$x^{(1)} = x^{(0)} - \frac{1}{2}\nabla_x f(x^{(0)}) = (0,0,0)^T$$

$$f(x^{(1)}) = 0$$

$$x^{(2)} = x^{(1)} - \frac{1}{2}\nabla_x f(x^{(1)}) = (0,0,0)^T$$

$$f(x^{(2)}) = 0$$

同理 $f(x^{(n)}) = 0 \ (n > 0)$ ，因此当 $|f(x^{(n+1)}) - f(x^{(n)})| = 0 < 0.001$ 时，迭代终止

习题 6

$\lambda = 0.01$ 固定步长

$$\min f(x) = (x_1-1)^2 + 16(x_2-2)^2 \qquad x^{(0)} = (2,3)^T$$

$$\nabla f(x) = [2(x_1-1), 32(x_2-2)]^T$$

$$\nabla f(x^{(0)}) = (2, 32)^T$$

$$x^{(1)} = x^{(0)} - 0.01\nabla f(x^{(0)}) = (2,3)^T - 0.01(2,32)^T = \left(\frac{99}{50}, \frac{67}{25}\right)^T$$

$$\nabla f(x^{(1)}) = \left(\frac{49}{25}, \frac{544}{25}\right)^T$$

$$x^{(2)} = x^{(1)} - 0.01\nabla f(x^{(1)}) = \left(\frac{99}{50}, \frac{67}{25}\right)^T - 0.01\left(\frac{49}{25}, \frac{544}{25}\right)^T = \left(\frac{4901}{2500}, \frac{1539}{625}\right)^T$$

$$\nabla f(x^{(2)}) = \left(1.9208, \frac{9248}{625}\right)^T$$

习题 7

$$\min f(x) = 3x_1^2 + 3x_2^2 - x_1^2 x_2 \qquad x^{(0)} = (1.5, 1.5)^T$$

$$g(x) = \nabla f(x) = (6x_1 - 2x_1x_2, 6x_2 - x_1^2)^T$$

Hessian 矩阵

$$\begin{pmatrix} 6-2x_2 & -2x_1 \\ -2x_1 & 6 \end{pmatrix}$$

```
初始点为:
[1.5 1.5]

第 1 次迭代结果:
[-3.75 -2.25]

第 2 次迭代结果:
[ 0.625 -3.125]

第 3 次迭代结果:
[0.31901607 0.00135752]
```

习题8

$\min\ f(x) = 3x_1^2 + x_2^2 - 2x_1x_2 - 4x_1$ （题目有误，不是 $-2x_1^2x_2$ 是 $-2x_1x_2$）

给定 $x^{(0)} = (-2, 4)^T$ 取 $\bar{H}^{(0)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

求 $f(x)$ 的极小点 即求 $\frac{1}{2}f(x)$ 的极小点

即求 $\frac{1}{2}f(x) \Rightarrow \frac{3}{2}x_1^2 + \frac{1}{2}x_2^2 - x_1x_2 - 2x_1$ 的极小点

$$\nabla f(x) = [(3x_1 - x_2 - 2), (x_2 - x_1)]^T$$
$$\nabla f(x^{(0)}) = (-12, 6)^T$$

$$p^{(0)} = -\bar{H}^{(0)}\nabla f(x^{(0)}) = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -12 \\ 6 \end{pmatrix} = \begin{pmatrix} 12 \\ -6 \end{pmatrix}$$

利用一维搜索，即 $\min_\lambda f(x^{(0)} + \lambda p^{(0)})$，$\lambda_0 = \frac{5}{17}$

$$x^{(1)} = x^{(0)} + \lambda_0 p^{(0)} = \begin{pmatrix} -2 \\ 4 \end{pmatrix} + \frac{5}{17}\begin{pmatrix} 12 \\ -6 \end{pmatrix} = \left(\frac{26}{17}, \frac{38}{17}\right)^T$$
$$\nabla f(x^{(1)}) = \left(\frac{6}{17}, \frac{12}{17}\right)^T$$

$$\Delta x^{(0)} = x^{(1)} - x^{(0)} = \left(\frac{26}{17}, \frac{38}{17}\right)^T - (-2, 4)^T = \left(\frac{60}{17}, -\frac{30}{17}\right)^T$$
$$\Delta g^{(0)} = \nabla f(x^{(1)}) - \nabla f(x^{(0)}) = \left(\frac{6}{17}, \frac{12}{17}\right)^T - (-12, 6)^T = \left(\frac{210}{17}, -\frac{90}{17}\right)^T$$

更新 Hessian 逆矩阵为

$$\bar{H}^{(1)} = \bar{H}^{(0)} + \frac{\Delta x^{(0)}(\Delta x^{(0)})^T}{(\Delta g^{(0)})^T \Delta x^{(0)}} - \frac{\bar{H}^{(0)}(\Delta g^{(0)})^T \bar{H}^{(0)}}{(\Delta g^{(0)})^T \bar{H}^{(0)} \Delta g^{(0)}}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\left(\frac{60}{17}, -\frac{30}{17}\right)^T\left(\frac{60}{17}, -\frac{30}{17}\right)}{\left(\frac{210}{17}, -\frac{90}{17}\right)\left(\frac{60}{17}, -\frac{30}{17}\right)^T} - \frac{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\left(\frac{210}{17}, -\frac{90}{17}\right)^T\left(\frac{210}{17}, -\frac{90}{17}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}{\left(\frac{210}{17}, -\frac{90}{17}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\left(\frac{210}{17}, -\frac{90}{17}\right)^T}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{17}\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} - \frac{1}{58}\begin{pmatrix} 49 & -21 \\ -21 & 9 \end{pmatrix} = \frac{1}{986}\begin{pmatrix} 385 & 241 \\ 241 & 891 \end{pmatrix}$$

$$p^{(1)} = -\bar{H}^{(1)}\nabla f(x^{(1)}) = -\frac{1}{986}\begin{pmatrix} 385 & 241 \\ 241 & 891 \end{pmatrix}\begin{pmatrix} \frac{6}{17} \\ \frac{12}{17} \end{pmatrix} = -\begin{pmatrix} \frac{9}{29} \\ \frac{21}{29} \end{pmatrix}$$

再由一维搜索 $\min_\lambda f(x^{(1)} + \lambda p^{(1)})$，得

$\lambda_1 = \frac{29}{17}$ $\nabla f(x^{(2)}) = (0, 0)^T$

$x^{(2)} = (1, 1)^T$ 为极小点

$$x^{(2)} = x^{(1)} + \lambda_1 p^{(1)} = \begin{pmatrix} \frac{26}{17} \\ \frac{38}{17} \end{pmatrix} + \frac{29}{17}\begin{pmatrix} -\frac{9}{29} \\ -\frac{21}{29} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

习题9:

构造罚函数

$$P(X,M)=\frac{1}{3}(x_1+1)^3+x_2+M\{[\min(0,-(1-x_1))]^2+[\min(0,x_2)]^2\}$$

$$\frac{\partial P}{\partial x_1}=(x_1+1)^2+2M[\min(0,\ x_1-1)]\ \cancel{+2M[\min(0}$$

$$\frac{\partial P}{\partial x_2}=1+2M[\min(0,x_2)]$$

对于不满足约束条件的点, $X=(x_1,x_2)^T$, 有 $1-x_1>0$, $x_2<0$

$x_1<1$, $x_2<0$

$$\frac{\partial P}{\partial x_1}=2 \qquad \frac{\partial P}{\partial x_2}=0 \qquad 1+2Mx_2=0$$

$$x_2=-\frac{1}{2M}$$

$$(x_1+1)^2+2M(x_1-1)=2$$

$$x_1^2+2x_1+1+2Mx_1-2M=2$$

$$x_1^2+(2+2M)x_1-2M-1=0$$

$$x_1=\frac{-(2+2M)\pm2\sqrt{M^2+4M+2}}{2}=-(1+M)\pm\sqrt{M^2+4M+2}=\sqrt{M^2+4M+2}-1-M$$

$M=1$: $X=(-2+\sqrt{7}\approx0.645,\ -\frac{1}{2})$

$M=2$: $X=(\sqrt{14}-3\approx0.741,\ -\frac{1}{4})$

$M=3$: $X=(\sqrt{23}-4\approx0.795,\ -\frac{1}{6})$

$\Downarrow$ $M\to+\infty$时 近似

$(1,0)$

习题 10

$1-x_1 \le 0 \qquad -x_2 \le 0$

$x_1 \ge 1 \qquad x_2 \ge 0$

构造倒数障碍函数

$$\bar{P}(x,r) = \frac{1}{3}(x_1+1)^3 + x_2 + \frac{r}{x_1-1} + \frac{r}{x_2}$$

$$\frac{\partial \bar{P}}{\partial x_1} = (x_1+1)^2 + r\cdot(-1)\frac{1}{(x_1-1)^2} = 0$$

$$\frac{\partial \bar{P}}{\partial x_2} = 1 + r\cdot(-1)\cdot\frac{1}{x_2^2} = 0$$

$$(x_1+1)^2 = \frac{r}{(x_1-1)^2} \qquad 1 = \frac{r}{x_2^2}$$

$$(x_1^2-1)^2 = r \qquad r = x_2^2 \qquad x_2 = \sqrt{r}$$

$$x_1^2 - 1 = \sqrt{r}$$

$$x_1^2 = \sqrt{r}+1$$

$$x_1 = \sqrt{\sqrt{r}+1}$$

∴ 如此得最优解

$$x_{min} = \lim_{r\to 0}\left(\sqrt{1+\sqrt{r}},\ \sqrt{r}\right)^T = (1,0)^T$$