《数学基础》作业1

习题1. (1) 解：原问题等价于

$\min\ f_0(x) = -(x-3)^2$

s.t. $x \le 5$

$-x \le -1$

拉格朗日函数

$L(x,\lambda) = -(x-3)^2 + \lambda^T(Ax-B)$

$A = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $B = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$

设对偶问题、原问题最优解分别为 λ^*, x^*

KKT条件 $\lambda^* \ge 0$ ① $Ax^* \le B$ ② $\nabla_x L(x^*,\lambda^*) = 0 = -2(x^*-3) + \lambda^{*T}A$ ③

$\nabla_x L = -2(x-3) + \lambda^T A$

$\lambda_1^*(x^*-5) = 0$ $\lambda_2^*(-x^*+1) = 0$ ④

联立各式，可得 $x^* = 5$ 或 1

该问题满足slater条件.

(2) 解：原问题等价于

$\min\ f_0(x) = (x-3)^2$

s.t. $x \le 5$

$-x \le -1$

拉格朗日函数

$L(x,\lambda) = (x-3)^2 + \lambda^T(Ax-B)$

$A = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $B = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$

设对偶问题，原问题最优解分别为 λ^*, x^*

满足slater条件 $\nabla_x L = 2(x-3) + \lambda^T A$

KKT条件 $\lambda^* \ge 0$ ① $Ax^* \le B$ ② $\nabla_x L(x^*,\lambda^*) = 0$

$= 2(x^*-3) + \lambda^{*T} A$ ③

$\lambda_1^*(x^*-5) = 0$

$\lambda_2^*(-x^*+1) = 0$ ④ 联立可得 $x^* = 3$.

习题2. 解：原问题等价于

$\min\ f_0(x) = \frac{1}{2}\|Ax-b\|_2^2 = \frac{1}{2}(Ax-b)^T(Ax-b)$

s.t. $Gx = h$

拉格朗日函数 $L(x,\nu) = \frac{1}{2}(Ax-b)^T(Ax-b) + \nu^T(Gx-h)$

$\nabla_x L = A^TAx - A^Tb + G^T\nu$ $\nabla_x^2 L = 2A^TA \succeq 0$

$\mathrm{rank}(A^TA) = \mathrm{rank}(A) = n$ A^TA 满秩对称矩阵

由 $\nabla_x^2 L \succeq 0$ 可得 $L(x,\nu)$ 关于x为凸函数

令 $\nabla_x L = \vec{0}$ 可得 $A^TAx = A^Tb - G^T\nu$ 即 $x = (A^TA)^{-1}(A^Tb - G^T\nu)$

将此代入，得 对偶函数 $g(\nu) = \inf_x L(x,\nu) =$

$= \frac{1}{2}b^Tb - \nu^T h + \frac{1}{2}(\nu^TG - b^TA)\cdot(A^TA)^{-1}(A^Tb - G^T\nu)$

对偶问题 $\max\ g(\nu)$ $\nabla_\nu g = -h + G(A^TA)^{-1}\cdot A^Tb - G\cdot(A^TA)^{-1}G^T\nu$

$\nabla_\nu^2 g = -G(A^TA)^{-1}G^T \preceq 0$ 故 g为关于ν的凹函数

令 $\nabla_\nu g = \vec{0}$ 可得 对偶问题最优解 ν^* 满足

$$G(A^TA)^{-1}G^T\nu^* = -h + G(A^TA)^{-1}\cancel{b^TA}\,A^Tb \quad (i)$$

$\longrightarrow G\cdot(A^TA)^{-1}A^Tb$!

此问题满足slater条件，KKT条件为：

$Gx^* = h$ ①　　$\nabla_x L(x^*,\nu^*) = \vec{0} = \cancel{2}A^TAx^* - \cancel{2}A^Tb + G^T\nu^*$ ②

$$x^* = (A^TA)^{-1}(A^Tb - \tfrac{1}{2}G^T\nu^*) \quad (ii)$$

(i)、(ii)即为所求表达式。

×××　×××　×××　×××

其实仅从做题角度，该题没那么复杂，直接上KKT条件，

便可得到 $x^* = (A^TA)^{-1}(A^Tb - G^T\nu^*)$ 再结合 $Gx^* = h$

可构造 $\begin{bmatrix} A^TA & G^T \\ G & 0 \end{bmatrix}\begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} A^Tb \\ h \end{bmatrix}$ 解这个线性方程组即可。

习题3. 证明：原问题等价于

$$\max_x \tfrac{1}{2}x^TA^TAx$$
$$s.t.\ \ x^Tx = 1$$

拉格朗日函数

$L(x,\nu) = \frac{1}{2}x^TA^TAx + \nu(x_1^2 + \cdots x_n^2 - 1)$

$\nabla_x L = 2A^TAx + 2\nu\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (2A^TA + 2\nu)x$

令 $\nabla_x L = \vec{0}$ 可得 $A^TAx = -\nu x$

$\max x^TA^TAx \Rightarrow \max x^T(-\nu x) \Rightarrow \max \cancel{x} -\nu\cdot x^Tx = -\nu$

而 $-\nu$ 是 A^TA 的特征值 故所求为 A^TA 最大的特征值

* 其实更严谨的思路是 设对偶问题最优解为 ν^*，原问题最优解为 x^*

$\nabla_x L = (2A^TA + 2\nu)x$　　$\nabla_x L(x^*,\nu^*) = 0 = (2A^TA + 2\nu^*)x^*$

该问题满足slater条件　　$A^TAx^* = -\nu^*x^*$

可以看到 $(-\nu^*)$ 是 A^TA 对应的特征值 x^* 为其对应的特征向量

$\nabla_x L = \vec{0}$ 可得 $A^TAx = -\nu x$ 代入 $L(x,\nu)$ 得对偶函数

$$g(\nu) = -\nu x^Tx + \nu\cdot x^Tx - \nu = -\nu$$

解对偶问题 $\max g(\nu)$ 得 $-\nu^*$ 即为 A^TA 最大的特征值

习题4. 证明：$\min \|x\|_2$ $\Rightarrow$ $\min \frac{1}{2}\|x\|_2^2 = \frac{1}{2}x^Tx$

$s.t.\ Ax = b$　　　　$s.t.\ Ax = b$

解：

$L(x,\nu) = \frac{1}{2}x^Tx + \nu^T(Ax - b)$

$\nabla_x^2 L = 2I \succeq 0$　L为关于x的凸函数

$\nabla_x L = x + A^T\nu$

令 $\nabla_x L = \vec{0}$ 可得 $x = -A^T\nu$ 代入L中，可得对偶函数

$g(\nu) = \inf_x L(x,\nu) = \frac{1}{2}\nu^T AA^T\nu - \nu^T AA^T \nu - \nu^T b$

$= -\frac{1}{2}\nu^T AA^T \nu - \nu^T b$

对偶问题 $\max_\nu g$

$\nabla_\nu g = -AA^T\nu - b \qquad \nabla_\nu^2 g \preceq 0 \qquad g$ 为关于 ν 的凹函数

令 $\nabla_\nu g = \vec{0}$ 可得 $AA^T\nu^* = -b$ （ν^*, x^* 分别为对偶问题、原问题最优解）

该问题满足slater条件 原问题的KKT条件中 $\nabla_x L(x^*, \nu^*)$

$= x^* + A^T\nu^* = \vec{0} \qquad x^* = -A^T\nu^* = -A^T(-1)\cdot(AA^T)^{-1}b$

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

即为所求。

习题5、解：$\nabla f(x) = 2\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 2x \qquad \nabla f(x^{(0)}) = 2\begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix} \qquad \|\nabla f(x^{(0)})\|_2^2 > \varepsilon$

$\varepsilon = 0.001$

令 $x^{(1)} = x^{(0)} - \lambda_0 \nabla f(x^{(0)})$ 代入，可得 $x^{(1)} = \begin{bmatrix} 2-4\lambda_0 \\ 2-4\lambda_0 \\ 1-2\lambda_0 \end{bmatrix}$

$f(x^{(1)}) = 9(1-2\lambda_0)^2$ 使其最小，得 $\lambda_0 = \frac{1}{2}$

$x^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \nabla f(x^{(1)}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \|\nabla f(x^{(1)})\|_2^2 < \varepsilon$

故所求点为 $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

习题6、解：取一个~~特别~~无穷小量 $\varepsilon \to 0$ 题中已固定步长

$\nabla f = \begin{bmatrix} 2x_1 - 2 \\ 32x_2 - 64 \end{bmatrix} \qquad x^{(0)} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \qquad \|\nabla f(x^{(0)})\|_2^2 > \varepsilon$

$x^{(1)} = x^{(0)} - \lambda \cdot \nabla f(x^{(0)}) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} - \frac{1}{100}\begin{bmatrix} 2 \\ 32 \end{bmatrix} = \begin{bmatrix} 1.98 \\ 2.68 \end{bmatrix}$

$\|\nabla f(x^{(1)})\|_2^2 > \varepsilon \qquad x^{(2)} = x^{(1)} - \lambda \nabla f(x^{(1)})$

$= \begin{bmatrix} 1.98 \\ 2.68 \end{bmatrix} - \frac{1}{100}\begin{bmatrix} 1.96 \\ 21.76 \end{bmatrix} = \begin{bmatrix} 1.9604 \\ 2.4624 \end{bmatrix}$

```
'''

迭代 1 次后的函数值是 0.5483225357352826  此时在[-0.41604038    0.13139973]
迭代 2 次后的函数值是 0.0015494982987473648  此时在[-0.00177646 -0.02265656]
迭代 3 次后的函数值是 5.219734036052285e-06  此时在[-0.00011553 -0.00131399]
迭代 4 次后的函数值是 1.7546426992072408e-08  此时在[-6.74554262e-06 -7.61794296e-05]
迭代 5 次后的函数值是 5.897435933282354e-11  此时在[-3.91229201e-07 -4.41645324e-06]
最优值是
[-3.91229201e-07 -4.41645324e-06]

'''
import numpy as np

def hessian(x):
    return np.array([ [6-2*x[1] , (-2)*x[0] ],
                      [(-2)*x[0] , 6 ] ])
def f1(x):
    return   np.array([ 6*x[0]-2*x[0]*x[1],6*x[1]-x[0]*x[0] ])

def f(x):
    return 3*(x[0]**2)+3*(x[1]**2)-((x[0]**2)*(x[1]))

def wolfe(x,dir,t,c1=0.5,c2=0.618):
  #看看当前是否满足 wolfe 条件
  #当前在 x  方向 dir  步长 t
  #系数 c1 c2
  #要保证  0<c1<c2<1
  cond1=( f(x)+c1*t*np.dot(f1(x),dir)>=f(x+t*dir) )
  cond2=( np.dot(f1(x+t*dir),dir)>=c2*np.dot(f1(x),dir) )
  return (cond1 and cond2)

def newton_method(x0,eps=1e-8):
    x=x0
    dir=(np.dot(np.linalg.inv(hessian(x)),f1(x)))*(-1) #方向
    iter=0 #迭代次数
    redu=0.9 #缩小系数

    iter1=0

    while float(np.dot((-1)*dir,f1(x)))>eps and iter<6666:
        iter=iter+1
        iter1=0

        #使用迭代法非精确线搜索步长直到可能值满足 wolfe 条件
        t=20 #初始步长
        while(wolfe(x,dir,t)==0 and iter1<=100):
            t=redu*t
            iter1=iter1+1
```

```
        x=x+t*dir
        print("迭代{}次后的函数值是{}  此时在{}".format(iter,f(x),x))
        dir=(np.dot(np.linalg.inv(hessian(x)),f1(x)))*(-1)
        if iter==6666:
            print('迭代了太多次没有结果\n')
    return x

if __name__=="__main__":
    x0=np.array([1.5,1.5])
    xans=newton_method(x0)
    print("最优值在")
    print(xans)
```

习题8、经过与助教学长的协商，该题有误，第个交叉项应为 $-2x_1x_2$.

则原问题：用DFP法求二次函数 $f(x)=3x_1^2+x_2^2-2x_1x_2-4x_1$ 的极小点.

解：取 $x_0=\begin{bmatrix}-2\\4\end{bmatrix}$ $\quad \bar{H}_0=\begin{bmatrix}1&0\\0&1\end{bmatrix}$ $\quad \nabla f(x)=\begin{bmatrix}6x_1-2x_2-4\\2x_2-2x_1\end{bmatrix}$

$\nabla f(x_0)=\begin{bmatrix}-24\\12\end{bmatrix}$ $\quad p_0=\cancel{-\bar{H}_0\cdot x_0=}$ $-\bar{H}_0\cdot\nabla f(x_0)=\begin{bmatrix}24\\-12\end{bmatrix}$

精确一维线搜索 $\min f(x_0+\lambda p_0)$ 可得 $\lambda_0=\frac{5}{34}$

$x_1=x_0+\lambda_0 p_0=\begin{bmatrix}\frac{26}{17}\\\frac{38}{17}\end{bmatrix}$ $\quad \nabla f(x_1)=\begin{bmatrix}\frac{12}{17}\\\frac{24}{17}\end{bmatrix}$

$\Delta x_0=x_1-x_0=\begin{bmatrix}\frac{60}{17}\\-\frac{30}{17}\end{bmatrix}$ $\quad \Delta g_0=\nabla f(x_1)-\nabla f(x_0)=\begin{bmatrix}\frac{420}{17}\\-\frac{180}{17}\end{bmatrix}$

$$\bar{H}_1=\bar{H}_0+\frac{\Delta x_0(\Delta x_0)^T}{(\Delta g_0)^T\Delta x_0}-\frac{\bar{H}_0(\Delta g_0)(\Delta g_0)^T(\bar{H}_0)}{(\Delta g_0)^T\bar{H}_0(\Delta g_0)}$$

$$=\begin{bmatrix}1&0\\0&1\end{bmatrix}+\begin{bmatrix}\frac{2}{17}&-\frac{1}{17}\\-\frac{1}{17}&\frac{1}{34}\end{bmatrix}-\begin{bmatrix}\frac{49}{58}&-\frac{21}{58}\\-\frac{21}{58}&\frac{9}{58}\end{bmatrix}=\begin{bmatrix}\frac{269}{986}&\frac{299}{986}\\\frac{299}{986}&\frac{431}{493}\end{bmatrix}$$

$p_1=-\bar{H}_1\cdot\nabla f(x_1)=\begin{bmatrix}-\frac{18}{29}\\-\frac{42}{29}\end{bmatrix}$ $\quad$ 再一维线搜索 $\min f(x_1+\lambda p_1)$
精确

可得 $\lambda_1=\frac{29}{34}$

$x_2=x_1+\lambda_1 p_1=\begin{bmatrix}1\\1\end{bmatrix}$ $\quad \nabla f(x_2)=\begin{bmatrix}0\\0\end{bmatrix}$ $\quad$ 故 $x_2=\begin{bmatrix}1\\1\end{bmatrix}$ 为极小点

习题9. $\min f_0(x)=\frac{1}{3}(x_1+1)^3+x_2$ $x^{(0)}=\begin{bmatrix}2\\0\end{bmatrix}$ 迭代两步

s.t. $f_1(x)=1-x_1\le 0$

$f_2(x)=-x_2\le 0$

解：罚函数 $P(x,M)=\cancel{\ast}\frac{1}{3}(x_1+1)^3+x_2+M([\min(0,x_1-1)]^2+[\min(0,x_2)]^2)$

~~对于不满足约束条件的点~~ $\cancel{1-x_1>0}$ $\cancel{-x_2>0}$

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

~~故~~ $\cancel{P(x,M)=\frac{1}{3}(x_1+1)^3+x_2}$

$\frac{\partial P}{\partial x_2}=1+2M\min(x_2,0)$ 对于不满足约束条件的点来说

$1-x_1>0$ $-x_2<0$ 令 $\frac{\partial P}{\partial x_1}=0$ $\frac{\partial P}{\partial x_2}=0$

可得 $x=[-(1+M)+\sqrt{M^2+4M},\ -\frac{1}{2M}]^T$ 或 $[-(1+M)-\sqrt{M^2+4M},\ -\frac{1}{2M}]^T$

$\downarrow$

$M\to\infty$ 时该极限不存在。(舍)

令 $M\to+\infty$ 可得原问题的最优解为 $\begin{bmatrix}1\\0\end{bmatrix}$

习题10 解：倒数障碍函数

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

$\frac{\partial P}{\partial x_1}=(x_1+1)^2-\frac{r}{(x_1-1)^2}=0$ $\frac{\partial P}{\partial x_2}=1-\frac{r}{x_2^2}=0$

联立，可得 $x_1=\sqrt{1+\sqrt{r}}$ $x_2=\sqrt{r}$

可得此问题最优解为 $x=\lim_{r\to 0}(\sqrt{1+\sqrt{r}},\sqrt{r})^T=\begin{bmatrix}1\\0\end{bmatrix}$