最优化方法-习题解答

张彦斌

计算机学院
2014年10月20日

Contents

1 第一章最优化理论基础-P13习题1(1)、2(3)(4)、3、4

1.验证下列各集合是凸集:

(1) S={$(x_1$,$x_2)$|2x_1+x_2≥1,$x_1-2x_2\geq 1$};
需要验证:
根据凸集的定义，对任意的$x(x_1,x_2),y(y_1,y_2)\in S$及任意的实数$\lambda\in[0,1]$,都有$\lambda x+(1-\lambda)y\in S$.
即,$(\lambda x_1+(1-\lambda)y_1,\lambda x_2+(1-\lambda)y_2)\in S$
证：由$x(x_1,x_2),y(y_1,y_2)\in S$得到，

$$\left\{\begin{array}{l}2x_1+x_2\geq 1,x_1-2x_2\geq 1\\ 2y_1+y_2\geq 1,y_1-2y_2\geq 1\end{array}\right. \tag{1}$$

把(1)中的两个式子对应的左右两部分分别乘以λ和$1-\lambda$,然后再相加，即得

$$\begin{aligned}\lambda(2x_1+x_2)+(1-\lambda)(2y_1+y_2)\geq 1,\\ \lambda(x_1-2x_2)+(1-\lambda)(y_1-2y_2)\geq 1\end{aligned}\tag{2}$$

合并同类项，

$$\begin{aligned}2(\lambda x_1+(1-\lambda)y_1)+(\lambda x_2+(1-\lambda)y_2)\geq 1,\\ (\lambda x_1+(1-\lambda)y_1)-2(\lambda x_2+(1-\lambda)y_2)\geq 1\end{aligned}\tag{3}$$

证毕.

2.判断下列函数为凸（凹）函数或严格凸（凹）函数:

(3)$f(x)=x_1^2-2x_1x_2+x_2^2+2x_1+3x_2$

首先二阶导数连续可微，根据定理1.5，f在凸集上是
（I）凸函数的充分必要条件是$\nabla^2 f(x)$对一切x为半正定;
（II）严格凸函数的充分条件是$\nabla^2 f(x)$对一切x为正定。

$$\nabla^2 f(x)=\begin{pmatrix}2 & -2\\ -2 & 2\end{pmatrix}\tag{4}$$

半正定矩阵
（4）

$$\nabla^2 f(x)=\begin{pmatrix}4 & 1 & -3\\ 1 & 2 & 0\\ -3 & 0 & 4\end{pmatrix}\tag{5}$$

正定矩阵

3.证明$f(x)=\frac{1}{2}x^TGx+b^Tx$为严格凸函数当且仅当Hesse矩阵G正定。

证明：根据严格凸函数定义证明。

对任意$x\neq y$,及任意实数$\lambda\in(0,1)$都有$f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y)$.

充分性：Hesse矩阵G正定=》严格凸函数.

$f(\lambda x+(1-\lambda)y)$=$\frac{1}{2}(\lambda x+(1-\lambda)y)^TG(\lambda x+(1-\lambda)y)+b^T(\lambda x+(1-\lambda)y)$

$\lambda f(x)+(1-\lambda)f(y)$=$\lambda(\frac{1}{2}x^TGx+b^Tx)$+$(1-\lambda)(\frac{1}{2}y^TGy+b^Ty)$

$\lambda f(x)+(1-\lambda)f(y)-f(\lambda x+(1-\lambda)y)$=$\lambda(\frac{1}{2}x^TGx)$+$(1-\lambda)(\frac{1}{2}y^TGy)-$
$[\frac{1}{2}(\lambda x)^TG(\lambda x)+\frac{1}{2}(1-\lambda)y^TG(1-\lambda)y+\frac{1}{2}\lambda x^TG(1-\lambda)y+\frac{1}{2}(1-\lambda)y^TG\lambda x]$
=$\frac{1}{2}\lambda x^TG(1-\lambda)x+\frac{1}{2}(1-\lambda)y^TG\lambda y-\frac{1}{2}\lambda x^TG(1-\lambda)y-\frac{1}{2}(1-\lambda)y^TG\lambda x$

$=\frac{1}{2}\lambda x^TG(1-\lambda)(x-y)+\frac{1}{2}(1-\lambda)y^TG\lambda(y-x)$
$=\frac{1}{2}\lambda(1-\lambda)(x-y)^TG(x-y)>0$ G正定保障了严格不等式成立。

反之，必要性：严格凸函数=》Hesse矩阵G正定.

类似，当对任意$x\neq y$,及任意实数$\lambda\in(0,1)$都有$f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y)$.

$\lambda f(x)+(1-\lambda)f(y)-f(\lambda x+(1-\lambda)y)$=$\lambda(\frac{1}{2}x^TGx)$+$(1-\lambda)(\frac{1}{2}y^TGy)-[\frac{1}{2}(\lambda x)^TG(\lambda x)+\frac{1}{2}(1-\lambda)y^TG(1-\lambda)y+\frac{1}{2}\lambda x^TG(1-\lambda)y+\frac{1}{2}(1-\lambda)y^TG\lambda x]$
$=\frac{1}{2}\lambda x^TG(1-\lambda)x+\frac{1}{2}(1-\lambda)y^TG\lambda y-\frac{1}{2}\lambda x^TG(1-\lambda)y-\frac{1}{2}(1-\lambda)y^TG\lambda x$
$=\frac{1}{2}\lambda x^TG(1-\lambda)(x-y)+\frac{1}{2}(1-\lambda)y^TG\lambda(y-x)$
$=\frac{1}{2}\lambda(1-\lambda)(x-y)^TG(x-y)>0$

4.若对任意$x\in\Re^n$及实数$\theta>0$都有$f(\theta x)=\theta f(x)$,证明$f(x)$在$\Re^n$上为凸函数的充要条件是$\forall x,y\in\Re^n,f(x+y)\leq f(x)+f(y)$

证明：根据严格凸函数定义证明。

定义：对任意$x\neq y$,及任意实数$\lambda\in(0,1)$都有$f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$.

充分条件：$\forall x,y\in\Re^n$, 有$f(x+y)\leq f(x)+f(y)$

对任意$x\neq y$,及任意实数$\lambda\in(0,1)$都有$f(\lambda x+(1-\lambda)y)\leq f(\lambda x)+f((1-\lambda)y)$

利用$f(\theta x)=\theta f(x)$,

$f(\lambda x+(1-\lambda)y)\leq f(\lambda x)+f((1-\lambda)y)$=$\lambda f(x)+(1-\lambda)f(y)$.

充分性证毕;

必要性：$f(x)$在$\Re^n$上为凸函数=》$\forall x,y\in\Re^n,f(x+y)\leq f(x)+f(y)$

根据定义有对任意$x\neq y$,及任意实数$\lambda\in(0,1)$都有$f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$.

不妨取$\lambda=\frac{1}{2}$,则
$f(\frac{1}{2}x+(1-\frac{1}{2})y)\leq\frac{1}{2}f(x)+(1-\frac{1}{2})f(y)$.

利用$f(\theta x)=\theta f(x)$,

$f(\frac{1}{2}(x+y))=\frac{1}{2}f(x+y)\leq\frac{1}{2}(f(x)+f(y))$

$\forall x,y\in\Re^n,f(x+y)\leq f(x)+f(y)$

证毕!

2 第二章线搜索算法-P27习题2、4、6

第2题:

黄金0.618算法:

function [s,phis,k,G,E]=golds(phi,a,b,delta,epsilon)

%输入：phi是目标函数，a，b是搜索区间的两个端点

% delta，epsilon分别是自变量和函数值的容许误差

%输出: s, phis 分别是近似极小点和极大值，G是$n \times 4$矩阵，

% 其第k行分别是a,p,q,b的第k次迭代值ak, pk, qk, bk,

% $E = [ds, dphi]$,分别是s和phis的误差限

%%%%%%%%%

t=(sqrt(5)-1)/2; h=b-a;

phia=feval(phi,a); phib=feval(phi,b);

p=a+(1-t)*h; q=a+t*h; phip=feval(phi,p); phiq=feval(phi,q);

k=1; G(k,:)=[a, p, q, b];

while(abs(phib-phia)>epsilon)|(h>delta)

if(phip<phiq)

b=q; phib=phiq; q=p; phiq=phip;

h=b-a; p=a+(1-t)*h; phip=feval(phi,p);

else

a=p; phia=phip; p=q; phip=phiq;

h=b-a; q=a+t*h; phiq=feval(phi,q);

end

k=k+1; G(k,:)=[a, p, q, b];

end

ds=abs(b-a); dphi=abs(phib-phia);

if(phip<=phiq)

s=p; phis=phip;

else

s=q; phis=phiq;

end

E=[ds,dphi];

运行：$[s, phis, k, G, E] = golds(inline('s^3 - 2 * s + 1'), 0, 3, 0.15, 0.01);$

结果

ak,pk,qk,bk

$$\begin{array}{cccc} 0 & 1.1459 & 1.8541 & 3.0000 \\ 0 & 0.7082 & 1.1459 & 1.8541 \\ 0 & 0.4377 & 0.7082 & 1.1459 \\ 0.4377 & 0.7082 & 0.8754 & 1.1459 \\ 0.7082 & 0.8754 & 0.9787 & 1.1459 \\ 0.7082 & 0.8115 & 0.8754 & 0.9787 \\ 0.7082 & 0.7721 & 0.8115 & 0.8754 \\ 0.7721 & 0.8115 & 0.8359 & 0.8754 \end{array} \tag{6}$$

$[s, phis, k, G, E] = golds(inline('s^3 - 2 * s + 1'), 0, 3, 0.15, 0.001);$

$\gg$ G

G =

$$\begin{matrix} 0 & 1.1459 & 1.8541 & 3.0000 \\ 0 & 0.7082 & 1.1459 & 1.8541 \\ 0 & 0.4377 & 0.7082 & 1.1459 \\ 0.4377 & 0.7082 & 0.8754 & 1.1459 \\ 0.7082 & 0.8754 & 0.9787 & 1.1459 \\ 0.7082 & 0.8115 & 0.8754 & 0.9787 \\ 0.7082 & 0.7721 & 0.8115 & 0.8754 \\ 0.7721 & 0.8115 & 0.8359 & 0.8754 \\ 0.7721 & 0.7965 & 0.8115 & 0.8359 \\ 0.7965 & 0.8115 & 0.8208 & 0.8359 \end{matrix} \tag{7}$$

第4题:

≫clear all;$[s, phis, k, ds, dphi, S] = qmin(inline('s^3 - 2 * s + 1'), 0, 3, 1e-2, 1e-4)$;

≫ s

s =

0.8165

第6题

function f=fun(x)

$f = 100 * (x(2) - x(1)^2)^2 + (1 - x(1))^2$;

function gf=gfun(x)

$gf = [-400 * (x(2) - x(1)^2) * x(1) - 2 * (1 - x(1)), 200 * (x(2) - x(1)^2)]'$;

function mk=armijo(xk,dk)

beta=0.5; sigma=0.2;

m=0; mmax=20;

while (m¡=mmax)

$if(fun(xk + beta^m * dk) <= fun(xk) + sigma * beta^m * gfun(xk)' * dk)$

mk=m; break;

end

m=m+1;

end

alpha=beta$^m k$

newxk=xk+alpha*dk

fk=fun(xk)

newfk=fun(newxk)

clear all;xk=[-1,1]';dk=[1,1]';mk=armijo(xk,dk)

alpha =

0.0020

newxk =

-0.9980
1.0020

fk =

4

newfk =

3.9956

mk =

9

3 第三章最速下降法和牛顿法P41习题1，2，3

第1题:

function f=funone(x)

$f = 3 * x(1)^2 + 2 * x(2)^2 - 4 * x(1) - 6 * x(2);$

function gf=gfunone(x)

$gf = [6 * x(1) - 4, 4 * x(2) - 6]';$

$\gg$ x0=[0,1]';[x val k]=grad('funone','gfunone',x0)

x =

0.6667

1.5000

val =

-5.8333

k =

10

第2题:

(1)牛顿法

function f=funtwo1(x)

$f = 4 * x(1)^2 + x(2)^2 - 8 * x(1) - 4 * x(2);$

function gf=gfuntwo1(x)

$gf = [8 * x(1) - 8, 2 * x(2) - 4]';$

x0=[0,1]';[x val k]=grad('funtwo1','gfuntwo1',x0)

x =

1

2

val =

-8

k =

2

（2）阻尼牛顿法

function He=Hesstwo(x)

n=length(x);

He=zeros(n,n);

He=[8, 0;

0, 2];

≫ x0=[0,1]';[x val k]=dampnm('funtwo1','gfuntwo1','Hesstwo',x0)

x =

1

2

val =

-8

k =

1

第3题.

function f=fun(x)

$f = (x(1)-2)^4 + (x(1)-2*x(2))^2;$

function gf=gfun(x)

$gf = [4*(x(1)-2)^3 + 2*(x(1)-2*x(2)), -4*(x(1)-2*x(2))]';$

≫clear all;

≫x0=[0 3]';[v,val,k]=grad('fun','gfun',x0) x =

2.0139

1.0070

val =

3.7685e-008

k =

2111

4 第四章共轭梯度法P51习题1，3，6(1)

1.证明向量$\alpha_1 = (1,0)^T$和$\alpha_2 = (3,-2)^T$关于矩阵

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \tag{8}$$

共轭. 验证$\alpha_1^T A\alpha_2 = 0$.

3.设$f(x) = \frac{1}{2}x^T Hx + b^T x$,其中

$$H = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \tag{9}$$

(1)证明$d_0 = (1,0)^T$与$d_1 = (-1,2)^T$关于H共轭;

(2)以$x_0 = (0,0)^T$为初始点,d_0和d_1为搜索方向，用精确线搜索求f的极小点.

验证(1)$d_0^T Hd_1 = 0$. (2)首先，$g(\vec{X}) = \nabla f(\vec{X}) = H\vec{X} + b =$

$$\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \end{pmatrix} \tag{10}$$

用定理4.1，也就是算法4.1产生的迭代序列，则每一步迭代点x_{k+1}都是$f(x)$在x_0和方向$d_0, d_1, \ldots, d_k$所张成的线性流形，$S_k = \{x | x = x_0 + \sum_{i=0}^{k} \alpha_i d_i, \forall \alpha_i\}$ 中的极小点，特别地，$x_n = x^* = -G^{-1}b$是问题的唯一极小点.

精确线搜索得到步长因子α_k具有如下性质，$g_{k+1}^T d_k = 0$

$$\begin{cases} \vec{X_{k+1}} = \vec{X_k} + \alpha_k d_k \\ g_{k+1}^T d_k = 0 \end{cases} \tag{11}$$

$\vec{X_{k+1}} = \vec{X_k} + \alpha_k d_k$,即$\vec{X_1} = \vec{X_0} + \alpha_0 d_0$;

$g(\vec{X_1}) = g_1 = g(\vec{X_1}) = H\vec{X_1} + b$,;$g_1^T d_0 = 0$

$\alpha_0 = -3/4, \vec{X_1} = (-3/4, 0)^T, f(\vec{X_1}) = -9/8$;
同理，利用(11)迭代，即

$$\begin{cases} \vec{X_2} = \vec{X_1} + \alpha_1 d_1 \\ g_2^T d_1 = 0 \end{cases} \tag{12}$$

$\alpha_1 = -1/4; \vec{X_2} = (-1/2, -1/2)^T, f(\vec{X_2}) = -3/2$,

$f(\vec{X_2}) < f(\vec{X_1})$,

定理4.1保证了极小点为$\vec{X_2} = (-1/2, -1/2)^T$

6.(1)$f(x) = 4x_1^2 + 4x_2^2 - 4x_1x_2 - 12x_2$,取初始点$x_0 = (-0.5, 1)^T$;
$g(x) = \nabla f(x) = Gx + b, G(x) = \nabla^2 f(x) = G$;

共轭方向的构造过程，取初始方向$d_0 = -g_0$,令$x_1 = x_0 + \alpha_0 d_0$,其中$\nabla f(x_1)^T d_0 = g_1^T d_0 = 0$,在$x_1$处，用f在$x_1$的负梯度方向$-g_1$与$d_0$的组合来生成$d_1$，即$d_1 = -g_1 + \beta_0 d_0$,然后选取系数$\beta_0$，使得$d_1$与$d_0$关于G共轭，即令$d_1^T G d_0 = 0$确定$\beta_0$.因此，$\beta_0 = \frac{g_1^T G d_0}{d_0^T G d_0}$,$g_1 - g_0 = G(x_1 - x_0) = \alpha_0 G d_0$,

利用定理4.1可知$g_2^T d_i = 0 (i = 0, 1)$
计算过程:

$$g(x) = Gx + b = \begin{pmatrix} 8 & -4 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ -12 \end{pmatrix} \tag{13}$$

$$G = \begin{pmatrix} 8 & -4 \\ -4 & 8 \end{pmatrix} \tag{14}$$

$$d_0 = -g(x0) = -Gx - b = -\begin{pmatrix} 8 & -4 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} -0.5 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ -12 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix} \tag{15}$$

$$x_1 = x_0 + \alpha_0 d_0 = \begin{pmatrix} -0.5 \\ 1 \end{pmatrix} + \alpha_0 \begin{pmatrix} 8 \\ 2 \end{pmatrix} = \begin{pmatrix} 8\alpha_0 - 0.5 \\ 2\alpha_0 + 1 \end{pmatrix} \tag{16}$$

$\nabla f(x_1)^T d_0 = g_1^T d_0 = 0$

$$g_1^T d_0 = \left[\begin{pmatrix} 8 & -4 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} 8\alpha_0 - 0.5 \\ 2\alpha_0 + 1 \end{pmatrix} + \begin{pmatrix} 0 \\ -12 \end{pmatrix} \right]^T \begin{pmatrix} 8 \\ 2 \end{pmatrix} = 0 \tag{17}$$

$\alpha_0 = 17/104, x_1 = (21/26, 69/52)^T \approx (0.80769, 1.32692)^T$

$g_1 = (15/13, -60/13)^T$

$\beta_0 = \frac{g_1^T G d_0}{d_0^T G d_0} = 225/676 \approx 0.33284$

$d_1 = -g_1 + \beta_0 d_0 = (3315/2197, 23205/4394)^T \approx (1.5088757, 5.281065)^T$;

$x_2 = x_1 + \alpha_1 d_1; \nabla f(x_2)^T d_1 = g_2^T d_1 = 0$

$$x_2 = \begin{pmatrix} (255\alpha_1)/169 + 21/26 \\ (1785\alpha_1)/338 + 69/52 \end{pmatrix} \tag{18}$$

$$g_2 = \begin{pmatrix} 15/13 - (1530\alpha_1)/169 \\ (6120\alpha_1)/169 - 60/13 \end{pmatrix} \tag{19}$$

由此可以求出$\alpha_1 = 0.127450980392157$;

极值点为$X_2 = (1, 2)^T$;

5 第五章拟牛顿法P73-2

2.DFP程序算法调用

极值点$x = (-0.2203 \times 10^{-6}, -0.1599 \times 10^{-6})$; 极小值val=$1.2527 \times 10^{-13}$

附程序:

function [x,val,k]=dfp(fun,gfun,x_0)

% 功能：用DFP算法求解无约束问题：min f(x)

% 输入：x_0是初始点，fun，gfun分别是目标函数及其梯度

%输出：x,val分别是近似最优点和最优值，k是迭代次数。

maxk=1e5; % 给出最大迭代次数

ρ=0.55;σ=0.4;ϵ=1e-5;

k=0;n=length(x_0);

%Hk=inv(feval('Hess',x_0));

%Hk=eye(n);

```
Hk=[2 1;1 1];
while(k<maxk)
    gk=feval(gfun,x0);% 计算梯度
    if(norm(gk)< ϵ),break;end % 检验终止准则
    dk=-Hk*gk;% 计算搜索方向
    m=0;mk=0;
    while(m<20) % 用Armijo搜索求步长
        if(feval(fun,x0+rho^m*dk) < feval(fun,x0)+sigma*rho^m*gk'*dk)
            mk=m;break;
        end
        m=m+1;
    end
    %DFP校正
    x = x0 + rho^mk*dk;
    sk=x-x0;yk=feval(gfun,x)-gk;
    if(sk'*yk>0)
        Hk=Hk-(Hk*yk*yk'*Hk)/(yk'*Hk*yk)+(sk*sk')/(sk'*yk);
    end
    k=k+1; x0=x;
end
val=feval(fun,x0);
```

（I）当

$$H_0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \tag{20}$$

[x,val,k]=dfp('fun','gfun',[1,-1]')

x =

1.0e-006 *

-0.220306134442640

-0.159928197216675

val =

1.252658776679855e-013

k =

4

（II）当采用%Hk=inv(feval('Hess',x0));

[x,val,k]=dfp('fun','gfun',[1,-1]')

x =

0

0

val =

0

k =

1

6 第六章信赖域方法P86-8

8(1)

>>gk=[-6 -3]';Bk=[4 -4;-4 8];

>> dta=1;

```
>> [d,val,lam,k]=trustq(gk,Bk,dta)

d =

    0.870281791219574
    0.492554154744547
val =

   -5.928777686124834
lam =

   5.158202203432865
k =

    5
>> dta=2;

[d,val,lam,k]=trustq(gk,Bk,dta)

d =

    1.726569382044938

    1.009434577568092

val =

    -10.321239036609670

lam =

    1.813689513237923

k =

    7

>> dta=5;

[d,val,lam,k]=trustq(gk,Bk,dta)

d =

    3.749999980155628
```

```
    2.249999987787719

val =

    -14.624999999999998

lam =

    8.078453007598365e-009

k =

    4
```

(2)

```
>> gk=[1 -3 -2]';Bk=[3 -1 2;-1 2 0;2 0 4];

dta=1;

[d,val,lam,k]=trustq(gk,Bk,dta)

d =

    -0.262643366009954

    0.837433127446609

    0.479295543075525

val =

    -2.501140183861169

lam =

    1.268746535391740

k =

    7

>> dta=2;

[d,val,lam,k]=trustq(gk,Bk,dta)
```

```
d =

    -0.333333333333382

    1.333333333329036

    0.666666666665635

val =

    -2.833333333333333

lam =

    6.261736529506079e-012

k =

    5

>> dta=5;

[d,val,lam,k]=trustq(gk,Bk,dta)

d =

    -0.333333333333433

    1.333333333180722

    0.666666666628600

val =

    -2.833333333333333

lam =

    2.286320834416492e-010

k =

    4
```

7 第七章非线性最小二乘问题P98-1，2，6

1. 设有非线性方程组

$$\begin{aligned} f_1(x) &= x_1^3 - 2x_2^2 - 1 = 0 \\ f_2(x) &= 2x_1 + x_2 - 2 = 0 \end{aligned} \tag{21}$$

(1)列出求解这个方程组的非线性最小二乘问题的数学模型；

最小二乘问题的数学表达式：$min_{x\in R^n} f(x) = \frac{1}{2}\|F(x)\| = \frac{1}{2}\sum_{i=1}^{m} f_i^2(x)$

(2)写出求解该问题的高斯-牛顿法迭代公式的具体形式：

$$J_k = F'(x(k)) = (\nabla F_1(x(k)), \cdots, \nabla F_m(x(k)))^T = \begin{pmatrix} 3x_{1,k}^2 & -4x_{2,k} \\ 2 & 1 \end{pmatrix} \tag{22}$$

$d_k^{GN} = -[J_k^T J_k]^{-1} J_k^T F(x_k) =$

$$\left[\begin{pmatrix} 3x_{1,k}^2 & 2 \\ -4x_{2,k} & 1 \end{pmatrix}\begin{pmatrix} 3x_{1,k}^2 & -4x_{2,k} \\ 2 & 1 \end{pmatrix}\right]^{-1}\begin{pmatrix} 3x_{1,k}^2 & 2 \\ -4x_{2,k} & 1 \end{pmatrix}\begin{pmatrix} x_{1,k}^3 - 2x_{2,k}^2 - 1 \\ 2x_{1,k} + x_{2,k} - 2 \end{pmatrix} \tag{23}$$

(3)初始点取为$x_0 = (2,2)^T$,迭代三次：

迭代公式：

$X_{k+1} = X_k + d_k^{GN}$

$X_1 = X_0 + d_0^{GN}$=

3.107142857142859

3.785714285714287

$X_2 = X_1 + d_1^{GN}$=

5.157431640715118

7.685136718569831

$X_3 = X_2 + d_2^{GN}$=

8.766682264589718

16.466635470820520

2.
解答：（1）测得的t_1,t_2和y一共5组数据，分别代入关系式

$$y=\frac{x_1x_3t_1}{1+x_1t_1+x_2t_2} \tag{24}$$

$$\begin{cases} 0.13=\frac{x_1x_3}{1+x_1+x_2} \\ 0.22=\frac{2x_1x_3}{1+2x_1+x_2} \\ 0.08=\frac{x_1x_3}{1+x_1+2x_2} \\ 0.13=\frac{2x_1x_3}{1+2x_1+2x_2} \\ 0.19=\frac{0.1x_1x_3}{1+0.1x_1} \end{cases} \tag{25}$$

$$\begin{cases} F_1(x)=x_1x_3-0.13(1+x_1+x_2) \\ F_2(x)=2x_1x_3-0.22(1+2x_1+x_2) \\ F_3(x)=x_1x_3-0.08(1+x_1+2x_2) \\ F_4(x)=2x_1x_3-0.13(1+2x_1+2x_2) \\ F_5(x)=0.1x_1x_3-0.19(1+0.1x_1) \end{cases} \tag{26}$$

(1)最小二乘问题模型表示为$min_{x\in R^n}f(x)=\frac{1}{2}\|F(x)\|=\frac{1}{2}\sum_{i=1}^{m}F_i^2(x)$

(2)高斯牛顿迭代公式的具体公式为:

$d_k^{GN}=-[J_k^TJ_k]^{-1}J_k^TF(x_k)$

6.利用LM方法的matlab程序求解$minf(x)=\frac{1}{2}\sum_{i=1}^{5}r_i^2(x)$
其中

$$\begin{cases} r_1(x)=x_1^2+x_2^2+x_3^2-1 \\ r_2(x)=x_1+x_2+x_3-1 \\ r_3(x)=x_1^2+x_2^2+(x_3-2)^2-1 \\ r_4(x)=x_1+x_2-x_3+1 \\ r_5(x)=x_1^3+3x_2^2+(5x_3-x_1+1)^2-36t \end{cases} \tag{27}$$

t为参数，可取$t=0.5,1,5$等，注意当$t=1$时，$x^*=(0,0,1)^T$是全局极小点，这时问题为零残量，比较不同参数的计算效果。

function [x,val,k]=lmm(Fk,JFk,x0)

%功能: 用L-M方法求解非线性方程组: F(x)=0

%输入: x0是初始点, Fk, JFk 分别是求F(xk)及F'(xk)的函数

%输出: x, val分别是近似解及——F(xk)——的值, k是迭代次数.

maxk=1000; %给出最大迭代次数

$\rho = 0.55; \sigma = 0.4; \mu_k = norm(feval(Fk, x0));$

k=0; epsilon=1e-6; n=length(x0);

while(k<maxk)

fk=feval(Fk,x0); %计算函数值

jfk=feval(JFk,x0); %计算Jacobi阵

gk=jfk'*fk;

$dk = -(jfk'*jfk + \mu_k * eye(n)) \backslash gk;$ %解方程组Gk*dk=-gk, 计算搜索方向

if(norm(gk)¡epsilon), break; end %检验终止准则

m=0; mk=0;

while(m<20) % 用Armijo搜索求步长

$newf = 0.5 * norm(feval(Fk, x0 + \rho^m * dk))^2;$

$oldf = 0.5 * norm(feval(Fk, x0))^2;$

$if(newf < oldf + sigma * \rho^m * gk' * dk)$

mk=m; break;

end

m=m+1;

end

$x0 = x0 + \rho^m k * dk;$

muk=norm(feval(Fk,x0));

k=k+1;

end

x=x0;

val=0.5*μ_k^2;

%gval=norm(gfun(x));

%% 目标函数

(I)t=0.5

function y=Fk(x)

$y(1) = x(1)^2 + x(2)^2 + x(3)^2 - 1;$

$y(2) = x(1) + x(2) + x(3) - 1;$

$y(3) = x(1)^2 + x(2)^2 + (x(3) - 2)^2 - 1;$

$y(4) = x(1) + x(2) - x(3) + 1;$

$y(5) = x(1)^3 + 3 * x(2)^2 + (5 * x(3) - x(1) + 1)^2 - 36 * 0.5;$

$y = y(:);$

%%%% Jacobi 阵%%%%%%function JF=JFk(x)

JF=[2*x(1), 2*x(2), 2*x(3);

1, 1, 1;

2*x(1),2*x(2),2*(x(3)-2);

1, 1 -1;

$3 * x(1)^2 - 2 * (5 * x(3) - x(1) + 1), 6 * x(2), 10 * (5 * x(3) - x(1) + 1)];$

>>$x0 = [1, 1, 1]'; [x, val, k] = lmm('Fk',' JFk', x0)$

x =

0.339361063668441

-0.200183578804671

0.714384339944574

val =

0.486062168183995

k =

9

(II)t=1;注意，这里$x^* = (0,0,1)^T$是全局极小点，这时问题为零残量。
$>>clearall; x0 = [1,1,1]'; [x, val, k] = lmm('Fk','JFk', x0)$

x =

-0.000000000000080

0.000000000000087

0.999999999999985

val =

2.815888304992978e-027

k =

8

(III)t=5;

$>>clearall; x0 = [1,1,1]'; [x, val, k] = lmm('Fk','JFk', x0)$
x =

-0.490713830929549

0.103144026198463

2.384345136824180

val =

14.450411547247533

k =

14

8 第八章最优性条件P112-1，2,5,6

1.验证$\bar{x}=(2,1)^T$是否为下列最优化问题的KT点:

$$\begin{array}{ll} min & f(x)=(x_1-3)^2+(x_2-2)^2 \\ s.t. & x_1^2+x_2^2\leq 5, \\ & x_1+2x_2=4, \\ & x_1,x_2\geq 0. \end{array} \tag{28}$$

验证：计算

$$\nabla f(\bar{x})=\left[\begin{array}{c} 2(x_1-3) \\ 2(x_2-2) \end{array}\right]\Bigg|_{x=\bar{x}}=\left[\begin{array}{c} -2 \\ -2 \end{array}\right],\nabla h(\bar{x})=\left[\begin{array}{c} 1 \\ 2 \end{array}\right] \tag{29}$$

$$\nabla g_1(\bar{x})=\left[\begin{array}{c} -2x_1 \\ -2x_2 \end{array}\right]=\left[\begin{array}{c} -4 \\ -2 \end{array}\right],\nabla g_2(\bar{x})=\left[\begin{array}{c} 1 \\ 0 \end{array}\right],\nabla g_3(\bar{x})=\left[\begin{array}{c} 0 \\ 1 \end{array}\right] \tag{30}$$

令

$\nabla f(\bar{x})-\bar{\mu}\nabla h(\bar{x})-\bar{\lambda_1}\nabla g_1(\bar{x})=0$

即

$$\left[\begin{array}{c} -2 \\ -2 \end{array}\right]-\bar{\mu}\left[\begin{array}{c} 1 \\ 2 \end{array}\right]-\bar{\lambda_1}\left[\begin{array}{c} -4 \\ -2 \end{array}\right]-\bar{\lambda_2}\left[\begin{array}{c} 1 \\ 0 \end{array}\right]-\bar{\lambda_3}\left[\begin{array}{c} 0 \\ 1 \end{array}\right]=0 \tag{31}$$

令$\bar{\lambda_2}=0,\bar{\lambda_3}=0$,解得$\bar{\mu}=-\frac{2}{3},\bar{\lambda_1}=\frac{1}{3}$

所以

$$\left\{\begin{array}{l} \nabla f(\bar{x})-\bar{\mu}\nabla h(\bar{x})-\sum_{i=1}^{3}\bar{\lambda}_i\nabla g_i(\bar{x})=0 \\ \bar{\lambda}_i g_i(\bar{x})=0,\bar{\lambda}_i\geq 0,i=1,2,3 \end{array}\right. \tag{32}$$

这表明$\bar{x}$是KT点，$(\bar{x},(\bar{\mu},\bar{\lambda}))$是KT对，其中$\bar{\mu}=-\frac{2}{3},\bar{\lambda}=(\frac{1}{3},0,0)^T$.

2.对于最优化问题:

$$\begin{array}{ll} min & f(x)=4x_1-3x_2 \\ s.t. & -(x_1-3)^2+x_2+1\geq 0, \\ & 4-x_1-x_2\geq 0, \\ & x_2+7\geq 0. \end{array} \tag{33}$$

求满足KT条件的点。

解：类似第1题

$$\nabla f(\bar{x}) = \begin{bmatrix} 4 \\ -3 \end{bmatrix}\Bigg|_{x=\bar{x}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}, \nabla h(\bar{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{34}$$

$$\nabla g_1(\bar{x}) = \begin{bmatrix} -2(\bar{x}_1 - 3) \\ 1 \end{bmatrix}, \nabla g_2(\bar{x}) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \nabla g_3(\bar{x}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{35}$$

令

$$\begin{cases} \nabla f(\bar{x}) - \bar{\mu}\nabla h(\bar{x}) - \sum_{i=1}^{3} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0 \\ \bar{\lambda}_i g_i(\bar{x}) = 0, \bar{\lambda}_i \geq 0, i = 1, 2, 3 \end{cases} \tag{36}$$

即:

$$\begin{cases} \begin{bmatrix} 4 \\ -3 \end{bmatrix} - \bar{\lambda}_1 \begin{bmatrix} -2(\bar{x}_1 - 3) \\ 1 \end{bmatrix} - \bar{\lambda}_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \bar{\lambda}_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \\ \bar{\lambda}_1(-(\bar{x}_1 - 3)^2 + \bar{x}_2 + 1) = 0 \\ \bar{\lambda}_2(4 - \bar{x}_1 - \bar{x}_2) = 0 \\ \bar{\lambda}_3(\bar{x}_2 + 7) = 0 \\ \bar{\lambda}_i \geq 0, i = 1, 2, 3 \end{cases} \tag{37}$$

取$\bar{\lambda}_3 = 0$

$4 - \bar{x}_1 - \bar{x}_2 = 0 => \bar{x}_2 = 4 - \bar{x}_1$

代入$-(\bar{x}_1 - 3)^2 + \bar{x}_2 + 1 = 0 =>$

$-(\bar{x}_1 - 3)^2 + 4 - \bar{x}_1 + 1 = 0 =>$

$\bar{x}_1 = 1$或$\bar{x}_1 = 4$

当$\bar{x}_1 = 4$时，$\bar{x}_2 = 0$,$\bar{\lambda}_1 = -7/3, \bar{\lambda}_2 = 2/3$,不满足$\bar{\lambda}_i \geq 0$舍去;

当$\bar{x}_1 = 1$时，$\bar{x}_2 = 3$,$\bar{\lambda}_1 = 7/3, \bar{\lambda}_2 = 16/3$,满足$\bar{\lambda}_i \geq 0$;

5.利用KT条件推出线性规划

$$\begin{array}{ll} min & z = c^T x \\ s.t. & Ax \leq b, \\ & x \geq 0, \end{array} \tag{38}$$

的最优化条件。

解:

$$\begin{cases} \nabla f(x) - \sum_{i=1}^{2} \lambda_i \nabla g_i(x) = 0 \\ \lambda_i g_i(x) = 0, \lambda_i \geq 0, i = 1, 2 \end{cases} \tag{39}$$

$\nabla g_1(x) = -A,$

$\nabla g_2(x) = I,$

其拉格朗日函数为

$L(x, \lambda_1, \lambda_2) = c^T x - \lambda_1^T (b - Ax) - \lambda_2^T x$

对上述函数关于x求极小. 令

$\nabla_x L(x, \lambda_1, \lambda_2) = c - \lambda_2 + A^T \lambda_1 = 0,$

由(39)$\lambda_2 g_2(x) = \lambda_2 x = 0,$

令$\lambda_2 = 0,$

因此最优性条件为:

$$\begin{cases} c + A^T \lambda_1 = 0 \\ \lambda_1 (b - Ax) = 0, \lambda_1 \geq 0 \end{cases} \tag{40}$$

6.设二次规划

$$\begin{array}{ll} min & f(x) = \frac{1}{2} x^T H x + c^T x \\ s.t. & Ax = b, \end{array} \tag{41}$$

其中H为n阶对称正定矩阵，矩阵A行满秩，求其最优解并说明解的唯一性。

解:

首先写出该问题的拉格朗日函数为

$L(x, \lambda) = \frac{1}{2} x^T H x + c^T x - \lambda^T (Ax - b).$

对上述函数关于x求极小. 由于H对称正定, 故函数$L(x, \lambda)$关于x为凸函数. 令

$\nabla_x L(x, \lambda) = Hx + c - A^T \lambda = 0,$

H对称正定，以及等式约束条件$Ax = b,$

$Hx + c - A^T\lambda = 0,$

$x + H^{-1}c - H^{-1}A^T\lambda = 0,$

$Ax + AH^{-1}c - AH^{-1}A^T\lambda = 0,$

$b + AH^{-1}c - AH^{-1}A^T\lambda = 0,$

H对称正定，A行满秩，因此，$AH^{-1}A^T$可逆（需要简单证明），

$\lambda = (AH^{-1}A^T)^{-1}(b + AH^{-1}c),$

因此有拉格朗日乘子的唯一性解，

也就有了最优解$x = -H^{-1}c + H^{-1}A^T\lambda$的唯一性。

9 第九章罚函数法P132，1-(1)、2-(1)、3-(3),6

1-(1):用外罚函数法求解下列约束优化问题:

$$\begin{array}{ll} min & f(x) = -x_1 - x_2 \\ s.t. & x_1^2 + x_2^2 = 1, \end{array} \tag{42}$$

解:

由等式约束得$x_2 = \pm\sqrt{1 - x_1^2}$, 代入目标函数得到一个无约束的单变量极小化问题

$min\phi(x_1) = -x_1 \pm \sqrt{1 - x_1^2}$

其全局极小点为$x_1 = \sqrt{\frac{1}{2}}$,从而得到原问题的全局极小点为$(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$.

现在要使构造的罚函数P(x),满足

$$P(x) \begin{cases} = 0, & x_1^2 + x_2^2 - 1 = 0 \\ > 0, & x_1^2 + x_2^2 - 1 \neq 0, \end{cases} \tag{43}$$

只要令$\bar{P}(x) = (x_1^2 + x_2^2 - 1)^2$即可. 现在考察目标函数和上述罚函数的组合

$P(x, \sigma) = f(x) + \bar{P}(x) = -x_1 - x_2 + \sigma\bar{P}(x)$

其中$\sigma > 0$是充分大的正数，称为罚因子（罚参数）。求这个组合函数的极小点. 由

$\frac{\partial P(x,\sigma)}{\partial x_1} = \frac{\partial P(x,\sigma)}{\partial x_2} = 0$,

得

$$\begin{cases} -1+4\sigma x_1(x_1^2+x_2^2-1) &= 0 \\ -1+4\sigma x_2(x_1^2+x_2^2-1) &= 0 \end{cases} \tag{44}$$

由此可得$x_1 = x_2 \neq 0$，因此$x_1(2x_1^2-1) = \frac{1}{4\sigma}$,当$\sigma \to \infty$,$x_1 = 0$(舍去)和$x_1 = \sqrt{\frac{1}{2}}$。所以$x_1 = x_2 = \sqrt{\frac{1}{2}}$,$minf(x) = -\sqrt{2}$.

2-(1).用内点法求解下列约束优化问题:

（1）

$$\begin{array}{ll} min & f(x) = x_1 + x_2 \\ s.t. & -x_1^2 + x_2 \leq 0, \\ & x_1 \geq 0; \end{array} \tag{45}$$

————————————更正————————————

$$\begin{array}{ll} min & f(x) = x_1 + x_2 \\ s.t. & -x_1^2 + x_2 \geq 0, \\ & x_1 \leq 0; \end{array} \tag{46}$$

————————————————————————————

解:

令$g_1(x) = x_1^2 - x_2, g_2(x) = x_1$,给出增广目标函数为

$H(x,\tau) = x_1 + x_2 - \tau(ln(x_1^2 - x_2) + ln(x_1))$

令

$$\begin{cases} \frac{\partial H}{\partial x_1} &= 1 - \frac{2\tau x_1}{x_1^2 - x_2} - \frac{\tau}{x_1} = 0 \\ \frac{\partial H}{\partial x_2} &= 1 + \frac{\tau}{x_1^2 - x_2} = 0 \end{cases} \tag{47}$$

$x_1^2 - x_2 = -\tau$,$1 + 2x_1 - \frac{\tau}{x_1} = 0$,

$x_1 = \frac{-1 \pm \sqrt{1+8\tau}}{4}$,

$\tau \to 0$,$x_1 = 0$或$x_1 = -\frac{1}{2}$

$x_1^2 - x_2 = -\tau$

$x_2 = 0$或$x_2 = \frac{1}{4}$

当$x_1 = -\frac{1}{2}$，$x_2 = \frac{1}{4}$时min f(x)=$-\frac{1}{4}$.

——————————————更正——————————————————————

解:

令$g_1(x)=-(x_1^2-x_2), g_2(x)=-x_1$,给出增广目标函数为

$H(x,\tau)=x_1+x_2-\tau(ln(-x_1^2+x_2)+ln(-x_1))$

令

$$\left\{\begin{array}{lll}\frac{\partial H}{\partial x_1} & = & 1+\frac{2\tau x_1}{-x_1^2+x_2}-\frac{\tau}{x_1}=0 \\ \frac{\partial H}{\partial x_2} & = & 1-\frac{\tau}{-x_1^2+x_2}=0\end{array}\right. \tag{48}$$

$x_1^2-x_2=-\tau$,$1+2x_1-\frac{\tau}{x_1}=0$,

$x_1=\frac{-1\pm\sqrt{1+8\tau}}{4}$,

$\tau\to 0$,$x_1=0$或$x_1=-\frac{1}{2}$

$x_1^2-x_2=-\tau$

$x_2=0$或$x_2=\frac{1}{4}$

当$x_1=-\frac{1}{2}$，$x_2=\frac{1}{4}$时min f(x)=$-\frac{1}{4}$.

———————————————————————————————————-

3-(1).用乘子法求解下列问题:

（1）

$$\begin{array}{ll}min & f(x)=x_1^2+x_2^2 \\ s.t. & x_1\geq 0;\end{array} \tag{49}$$

解:

PHR算法:

我们回到一般约束优化问题(9.28,9.33)（书上）, 我们来构造求解(49) 的乘子法. 此时, 增广拉格朗日函数为

$\psi(x,\mu,\lambda,\sigma)=f(x)-\sum_{i=1}^{l}\mu_i h_i(x)+\frac{\sigma}{2}\sum_{i=1}^{l}h_i^2(x)+\frac{1}{2\sigma}\sum_{i=1}^{m}([min\{0,\sigma g_i(x)-\lambda_i\}]^2-\lambda_i^2)$

乘子迭代公式为

$(\mu_{k+1})_i=(\mu_k)_i-\sigma h_i(x_k), i=1,2,\cdot,l,$

$(\lambda_{k+1})_i = max\{0, (\lambda_k)_i - \sigma g_i(x_k)\}, i = 1, 2, \cdot, m.$

令

$\beta_k = (\sum_{i=1}^{l} h_i^2(x_k) + \sum_{i=1}^{m}[min\{g_i(x_k), \frac{(\lambda_k)_i}{\sigma}\}]^2)^{\frac{1}{2}}$

则终止准则为$\beta_k \le \varepsilon$

———————————-重要———————————

$\psi(x, \lambda, \sigma) = f(x) + \frac{1}{2\sigma}([min\{0, \sigma x_1 - \lambda_1\}]^2 - \lambda_1^2)$

令

$$\begin{cases} \frac{\partial \psi}{\partial x_1} = 2x_1 = 0, if(\sigma x_1 - \lambda_1) < 0 \\ \quad\quad = 2x_1 + (\sigma x_1 - \lambda_1), if(\sigma x_1 - \lambda_1) > 0 \\ \frac{\partial \psi}{\partial x_2} = 2x_2 = 0 \end{cases} \tag{50}$$

——-数值方法角度——————————————

取初始点$x_0 = (0, 0)^T, \lambda_1 = 1, \sigma_1 = 2, \varepsilon = 1e - 5$ $x_1 = x_2 = 0$,min f(x)=0

$x_1 = \frac{\lambda_1}{2+\sigma}, x_2 = 0$;或者$x_1 = 0, x_2 = 0$

$\sigma \to \infty$,$x_1 = 0, x_2 = 0$;min f(x)=0

6.略。

10　第十一章二次规划习题11 P178-1（1），5

1.用拉格朗日方法求解下列二次规划问题：（1）

$$\begin{array}{ll} min & f(x) = 2x_1^2 + x_2^2 + x_1x_2 - x_1 - x_2, \\ s.t. & x_1 + x_2 = 1; \end{array} \tag{51}$$

首先写出该问题的拉格朗日函数为

$L(x, \lambda) = \frac{1}{2}x^T Hx + c^T x - \lambda(AX - 1).$

$$H = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}, c = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, A = (1, 1), \tag{52}$$

对上述函数关于x求极小. 由于H对称正定, 故函数$L(x,\lambda)$关于x为凸函数. 令

$\nabla_x L(x,\lambda) = Hx + c - A^T\lambda = 0,$

H对称正定，以及等式约束条件$Ax = 1$,

$$\begin{pmatrix} H & -A^T \\ -A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} -c \\ -1 \end{pmatrix}, \tag{53}$$

$$\begin{pmatrix} 4 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \lambda \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \tag{54}$$

解得

$$\begin{pmatrix} x_1 \\ x_2 \\ \lambda \end{pmatrix} = \begin{pmatrix} 1/4 \\ 3/4 \\ 3/4 \end{pmatrix}, \tag{55}$$

5.设$A \in R^{m\times n}$行满秩，$a \in R^n$,证明二次规划问题

$$\begin{array}{ll} min & \frac{1}{2}(x-a)^T(x-a), \\ s.t. & Ax = b; \end{array} \tag{56}$$

的解以及相应的拉格朗日乘子分别为:

$x^* = a + A^T(AA^T)^{-1}(b - Aa)$,$\lambda^* = (AA^T)^{-1}(b - Aa)$

证明:

二次规划问题等价于

$$\begin{array}{ll} min & \frac{1}{2}x^T Hx - a^T x + \frac{1}{2}a^T a, \\ s.t. & Ax = b; \end{array} \tag{57}$$

其中H单位矩阵E,

对上述函数关于x求极小. 由于H对称正定, 故函数$L(x,\lambda)$关于x为凸函数. 令

$\nabla_x L(x,\lambda) = Hx - a - A^T\lambda = 0,$

H对称正定，以及等式约束条件$Ax = b$,

$Hx - a - A^T\lambda = 0,$

$x + H^{-1}(-a) - H^{-1}A^T\lambda = 0,$

$Ax + AH^{-1}(-a) - AH^{-1}A^T\lambda = 0,$

$b + AH^{-1}(-a) - AH^{-1}A^T\lambda = 0,$

其中H单位矩阵E，A行满秩，因此，AA^T可逆（需要简单证明），

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

因此有拉格朗日乘子的唯一性解，

也就有了最优解$x = -H^{-1}(-a) + H^{-1}A^T\lambda = a + A^T(AA^T)^{-1}(b - Aa)$的唯一性。