## POM-HW5-Jasenv@CC98

## **Solution 1:**

$$\min_{x \in R^n} f(x) = rac{1}{2} r(x)^{ ext{T}} r(x) = rac{1}{2} \sum_{i=1}^m \left[ r_i(x) 
ight]^2$$

By derivation, we get the gradient and hessian matrix of f(x).

$$abla f(x) = 
abla r(x)^{\mathrm{T}} r(x) = \sum_{i=1}^m r_i(x) 
abla r_i(x)$$

$$abla^2 f(x) = 
abla r(x)^{\mathrm{T}} 
abla r(x) + \sum_{i=1}^m r_i(x) 
abla^2 r_i(x) = M(x) + S(x)$$

So equivalently, this problem reduces to prove

$$S(x) = \sum_{i=1}^m r_i(x) 
abla^2 r_i(x) 
ightarrow O, x 
ightarrow x^*$$

For this specific problem:

$$r_1(x) = x_2 - x_1^2, r_2(x) = 1 - x_2$$

then,

$$abla^2 r_1(x) = \left(egin{array}{cc} -2 & 0 \ 0 & 0 \end{array}
ight), 
abla^2 r_2(x) = \left(egin{array}{cc} 0 & 0 \ 0 & 0 \end{array}
ight)$$

SO,

$$egin{aligned} S(x) &= \left(x_2 - x_1^2
ight) \cdot \left(egin{array}{cc} -2 & 0 \ 0 & 0 \end{array}
ight) + \left(1 - x_2
ight) \left(egin{array}{cc} 0 & 0 \ 0 & 0 \end{array}
ight) \ &= \left(egin{array}{cc} -2 \left(x_2 - x_1^2
ight) & 0 \ 0 & 0 \end{array}
ight) 
ightarrow O, x 
ightarrow x^* = \left(1,1
ight)^T \end{aligned}$$

## **Solution 2:**

Define two functions:

$$egin{aligned} x(\mu) &= -ig(A^TA + \mu Iig)^{-1}A^Tr, \quad \delta(x) = \|Ax + r\|^2. \end{aligned}$$

For  $x(\mu)$ ,

$$x'(\mu) = \left(A^TA + \mu I
ight)^{-2}A^Tr$$

For  $\delta(x)$ , use the known condition:  $A^TAx = -A^Tr - \mu x$ 

$$egin{aligned} \delta(x) &= (Ax+r)^T(Ax+r) = x^TA^TAx + 2r^TAx + r^Tr \ &
abla \delta(x) &= 2\left(A^TAx + A^Tr
ight) = 2\left(-A^Tr - \mu x + A^Tr
ight) = -2\mu x \end{aligned}$$

By the chain principle of composition function's derivation,  $\mu>0$ 

$$egin{aligned} \delta'(x(\mu)) &= (
abla \delta(x))^T x'(\mu) = -2 \mu x^T ig(A^T A + \mu Iig)^{-2} A^T r \ &= 2 \mu x^T ig(A^T A + \mu Iig)^{-1} ig\left[-ig(A^T A + \mu Iig)^{-1} A^T r
ight] \ &= 2 \mu x^T ig(A^T A + \mu Iig)^{-1} x > 0 \end{aligned}$$

So,  $\delta(x(\mu))$  is the increasing function of  $\mu$ ,  $\mu_{1}>\mu_{2}>0$ ,  $\delta\left(x\left(\mu_{2}\right)\right)<\delta\left(x\left(\mu_{1}\right)\right)$ ,  $i.\,e.\,\left\Vert Ax_{2}+r\right\Vert ^{2}<\left\Vert Ax_{1}+r\right\Vert ^{2}.$