1. $f_0=\sum_{i=1}^{N}\|A_ix+b_i\|_2+(1/2)\|x-x_0\|_2^2$

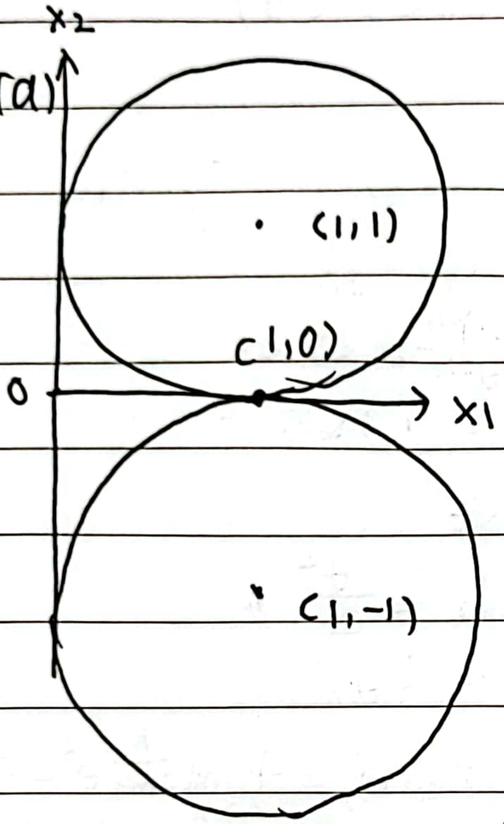
s.t. $A_ix+b_i-y_i=0$

$L(x,\nu)=\sum_{i=1}^{N}\|A_ix+b_i\|_2+(1/2)\|x-x_0\|_2^2+\sum_{i=1}^{N}\nu_i(A_ix+b_i-y_i)$

$\frac{\partial L(x,\nu)}{\partial x}=x-x_0+\sum_{i=1}^{N}A_i^T\nu_i=0 \Rightarrow x=x_0-\sum_{i=1}^{N}A_i^T\nu_i$

$g(x,\nu)=\inf_x L(x,\nu)=\sum_{i=1}^{N}\|y_i\|_2-\sum_{i=1}^{N}\nu_iy_i-\frac{1}{2}\|\sum_{i=1}^{N}A_i^T\nu_i\|_2^2+\sum_{i=1}^{N}(A_ix_0+b_i)^T\nu_i$

2.(a)



$x_1^2+x_2^2=1$

feasible set: $\{(1,0)\}$

optimal point $(1,0)$

optimal value: 1

(b) f_1: $(x_1-1)^2+(x_2-1)^2-1$, f_2: $(x_1-1)^2+(x_2+1)^2-1$

$L(x,\lambda)=x_1^2+x_2^2+\lambda_1f_1+\lambda_2f_2$

$\nabla L(x,\lambda)=2x_1+2x_2+\lambda_1(2x_1-2)+\lambda_2(2x_2+2)+\lambda_1(2x_2-2)+\lambda_2(2x_1-2)$

$$\text{KKT}\begin{cases}2x_1+2\lambda_1(x_1-1)+2\lambda_2(x_1-1)=0 & ①\\ 2x_2+2\lambda_1(x_2-1)+2\lambda_2(x_2+1)=0 & ②\\ \lambda_1((x_1-1)^2+(x_2-1)^2-1)=0\\ \lambda_2((x_1-1)^2+(x_2+1)^2-1)=0\\ \lambda_1\geq0\\ \lambda_2\geq0\\ (x_1-1)^2+(x_2-1)^2-1\leq0\\ (x_1-1)^2+(x_2+1)^2-1\leq0\end{cases}$$

$x^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} 2 = 0 & ① \\ -2\lambda_1 + 2\lambda_2 = 0 & ② \end{cases} \Rightarrow 2 \neq 0$, so there is no λ_1^*, λ_2^*

(c) $g(\lambda_1, \lambda_2) = \inf_{x_1, x_2} [x_1^2 + x_2^2 + \lambda_1 (x_1-1)^2 + \lambda_2 (x_1-1)^2 + \lambda_1 (x_2-1)^2 + \lambda_2 (x_2+1)^2 - \lambda_1 - \lambda_2]$

$= \inf_{x_1} (x_1^2 + \lambda_1 x_1^2 - 2\lambda_1 x_1 + \lambda_2 x_1^2 - 2\lambda_2 x_1) + \inf_{x_2} (x_2^2 + \lambda_1 x_2^2 - 2\lambda_1 x_2 + \lambda_2 x_2^2 + 2\lambda_2 x_2) + \lambda_1 + \lambda_2$

$\frac{\partial \inf_{x_1} (x_1^2 + \lambda_1 x_1^2 - 2\lambda_1 x_1 + \lambda_2 x_1^2 - 2\lambda_2 x_1)}{\partial x_1} = 2x_1 + 2\lambda_1 x_1 + 2\lambda_2 x_1 - 2\lambda_1 - 2\lambda_2 = 0$

$x_1 = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2}$

Similarly, $x_2 = \frac{\lambda_1 - \lambda_2}{1 + \lambda_1 + \lambda_2}$

$g(\lambda_1, \lambda_2) = -\frac{(\lambda_1 + \lambda_2)^2}{1 + \lambda_1 + \lambda_2} - \frac{(\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} + \lambda_1 + \lambda_2$

the dual problem: maximize $g(\lambda_1, \lambda_2) = -\frac{(\lambda_1 + \lambda_2)^2}{1 + \lambda_1 + \lambda_2} - \frac{(\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} + \lambda_1 + \lambda_2$

subject to. $\lambda_1 \geq 0$, $\lambda_2 \geq 0$

to check whether it's strong duality, we can first get the maximum of $g(\lambda_1, \lambda_2)$, since $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, we can let $t = \lambda_1 + \lambda_2$, $t \geq 0$, then $g(t, \lambda_1, \lambda_2)$ is maximum if $\lambda_1 = \lambda_2$ $\left(-\frac{(\lambda_1 - \lambda_2)^2}{1+t} \; (t \geq 0)\right) \Rightarrow -\frac{t^2}{1+t} + t = \frac{t}{1+t} < 1$,

while $f_0(x)$ can be 1 and 1 is not optimal,

So $\min f_0(x) < 1$, Therefore, there exists a t that $\max g(t) = \frac{t}{1+t} > \min f_0(x)$. So the strong duality doesn't hold

3. $L(x,\nu) = \|Ax-b\|_2^2 + \nu^T(Gx-h)$

$$\frac{\partial L}{\partial x} = 2A^TAx - 2A^Tb + G^T\nu = 0$$

$$\text{KKT}: \begin{cases} 2A^TAx - 2A^Tb + G^T\nu = 0 \\ Gx - h = 0 \end{cases}$$

$$\begin{cases} 2A^TAx^* - 2A^Tb + G^T\nu^* = 0 \\ Gx^* - h = 0 \end{cases}$$

$$\Downarrow \quad \text{rank}(G) = p$$

$$\begin{cases} \nu^* = (G^T)^{-1}(-2A^TAG^{-1}\cdot h + 2A^Tb) \\ x^* = G^{-1}\cdot h \end{cases}$$

4. l_1-norms: minimize $|(x_1-b_1)|+\dots+|(x_n-b_n)|$

solution: median of $b_1, b_2 \dots b_n$

l_2-norms: minimize $\sqrt{(x_1-b_1)^2+\dots+(x_n-b_n)^2}$

solution: $\frac{\sum_{i=1}^{n} b_i}{n}$

l_∞-norm: minimize $\max\{|x_1-b_1|, \dots, |x_n-b_n|\}$

solution: $\cancel{\frac{b_{min}+b_{max}}{2}}$ $\cancel{\text{和}}$ $\frac{\max\{b_i\}-\min\{b_i\}}{2}$

5. Lagrangian: $L(x,r,\nu) = \sum_{i=1}^{m} \phi(r_i) + \nu^T(Ax-b-r)$

$g(r,\nu) = \inf_x L(x,r,\nu) = \sum_{i=1}^{m} \phi(r_i) - \nu^T b - \nu^T r \ (\nu^T A = 0)$

$$\therefore g(r,\nu) = \begin{cases} \sum_{i=1}^{m} \phi(r_i) - \nu^T b - \sum_{i=1}^{m} \nu_i r_i & (\nu^T A = 0) \\ -\infty & \text{otherwise} \end{cases}$$

$$\inf_r g(r,\nu) = -\nu^T b + \inf_r \left(\sum_{i=1}^{m} \phi(r_i) - \sum_{i=1}^{m} \nu_i r_i\right)$$

$$= -\nu^T b + \sum_{i=1}^{m} \inf_{r_i} (\phi(r_i) - \nu_i r_i)$$

$$= -\nu^T b - \sum_{i=1}^{m} \sup_{r_i} (\nu_i r_i - \phi(r_i))$$

$$= -\nu^T b - \sum_{i=1}^{m} \phi^*(\nu_i)$$

$$\therefore g(\nu) = \begin{cases} -\nu^T b - \sum_{i=1}^{m} \phi^*(\nu_i), & \nu^T A = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$\Rightarrow$ maximize $\quad -\nu^T b - \sum_{i=1}^{m} \phi^*(\nu_i)$

s.t. $\quad \nu^T A = 0$

6.(1) $x^{(0)}=1$, $p^*=\log 2$, $f'(x)=\frac{e^x-e^{-x}}{e^x+e^{-x}}$, $f''(x)=\frac{4e^{2x}}{(e^{2x}+1)^2}$

① $\frac{f'(x^{(0)})}{f''(x^{(0)})}=1.8134$, $x^{(1)}=x^{(0)}-\frac{f'(x^{(0)})}{f''(x^{(0)})}=-0.8134$

$f(x^{(1)})=1.4324$

② $\frac{f'(x^{(1)})}{f''(x^{(1)})}=-1.2228$, $x^{(2)}=x^{(1)}-\frac{f'(x^{(1)})}{f''(x^{(1)})}=0.4094$

$f(x^{(2)})=1.2997$

③ $x^{(3)}=x^{(2)}-\frac{f'(x^{(2)})}{f''(x^{(2)})}=0.0473$

$f(x^{(3)})=1.0815$

$x^{(0)}=1.1$, $x^{(1)}=x^{(0)}-\frac{f'(x^{(0)})}{f''(x^{(0)})}=-1.129$

$f(x^{(1)})=1.512$

$x^{(2)}=x^{(1)}-\frac{f'(x^{(1)})}{f''(x^{(1)})}=1.234$

$f(x^{(2)})=1.535$

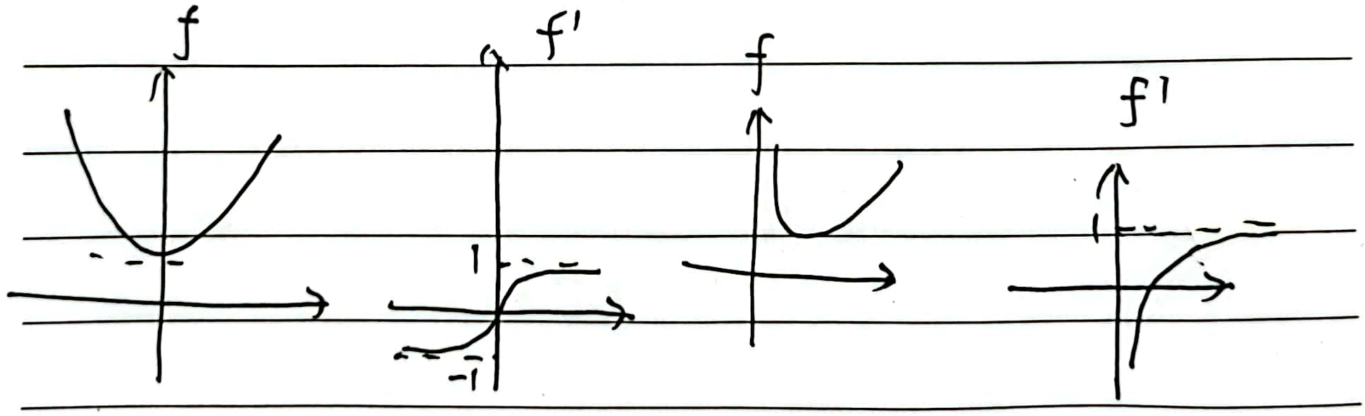
$x^{(3)}=x^{(2)}-\frac{f'(x^{(2)})}{f''(x^{(2)})}=-1.695$

$f(x^{(3)})=1.622$

(2) $f(x)=-\log x+x$, $f'(x)=-\frac{1}{x}+1$, $f''(x)=\frac{1}{x^2}$

$p^*=1$

$x^{(0)}=3$, $x^{(1)}=x^{(0)}-\frac{f'(x^{(0)})}{f''(x^{(0)})}=-3$, not in D



(1) (2)

7. let $f_0(x) = f(x) + (Ax-b)^T Q (Ax-b)$

$\nabla^2 f_0(x) = \nabla^2 f(x) + 2A^TQA$, $\nabla f_0(x) = \nabla f(x) + 2A^TQAx - 2A^TQb$

$$\begin{pmatrix} \nabla^2 f(x) + 2A^TQA & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x_{nt} \\ w \end{pmatrix} = \begin{pmatrix} -\nabla f(x) - 2A^TQAx + 2A^TQb \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x_{nt}' \\ w' \end{pmatrix} = \begin{pmatrix} -\nabla f(x) \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{cases} w' = w + 2QAx - 2Qb \\ \Delta x_{nt} = \Delta x_{nt}' \end{cases}$$

Yes