225040374 丁方孙草. HW5.

1. Solution. Denote all the sets in each question as S.

(a) Given S_1 and S_2 are convex:

$\forall x_{11}, x_{12} \in S_1$. $\forall \lambda_1 \in [0,1]$ $\lambda_1 x_{11} + (1-\lambda_1) x_{12} \in S_1$.

$\forall x_{21}, x_{22} \in S_2$. $\forall \lambda_2 \in [0,1]$ $\lambda_2 x_{21} + (1-\lambda_2) x_{22} \in S_2$

$\Rightarrow \forall y_1, y_2 \in S$. $\exists y_{11}, y_{21} \in S_1$. $y_{21}, y_{22} \in S_2$

s.t. $y_1 = y_{11} + y_{12}$. $y_2 = y_{21} + y_{22}$

$\forall \lambda \in [0,1]$ $\lambda y_1 + (1-\lambda) y_2 = \lambda y_{11} + (1-\lambda) y_{21} + \lambda y_{21} + (1-\lambda) y_{22}$

Denote $z_1 = \lambda y_{11} + (1-\lambda) y_{21}$

$z_2 = \lambda y_{21} + (1-\lambda) y_{22}$

$\Rightarrow z_1 \in S_1$. $z_2 \in S_2$ Then $z_1 + z_2 \in S$.

So S is convex.

(b). For $ax^2+2x+\frac{1}{a}$. $\Delta = 2^2-4\cdot a\cdot\frac{1}{a}=0$.

Given $a>0$. $\forall x\in\mathbb{R}$. $ax^2+2x+\frac{1}{a}\geq 0$.

$\mathbb{R}$ is convex i.e. $\forall x_1.x_2\in\mathbb{R}$. $\lambda x_1+(1-\lambda)x_2\in\mathbb{R}$

Then the set is convex

(c). f is a convex function :

$\forall x_1.x_2.\ \forall\lambda\in[0,1].\ f(\lambda x_1+(1-\lambda)x_2)\leq\lambda f(x_1)+(1-\lambda)f(x_2)$

$\Rightarrow \forall s_1=(y_1,x_1),\ s_2=(y_2,x_2)\in S$

$y_1\geq f(x_1).\quad y_2\geq f(x_2)$

$\lambda y_1+(1-\lambda)y_2\geq\lambda f(x_1)+(1-\lambda)f(x_2)\geq f(\lambda x_1+(1-\lambda)x_2)$

Then $\lambda s_1+(1-\lambda)s_2=(\lambda y_1+(1-\lambda)y_2,\ \lambda x_1+(1-\lambda)x_2)\in S$.

$\Rightarrow S$ is convex.

(d). Similar to (c).

$\forall s_1 = (y_1, x_1),\ s_2 = (y_2, x_2) \in S \quad \forall \lambda \in [0,1].$

$y_1 = f(x_1).\quad y_2 = f(x_2)$

$$\lambda y_1 + (1-\lambda) y_2 = \lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)$$

Then S is convex only when f is linear function

(e) Define $B = \frac{1}{2}(A + A^T)$. $\Rightarrow S = \{x \in \mathbb{R}^3;\ x^T B x \leq 0\}$.

$$B = \begin{pmatrix} 2 & -1 & 4 \\ -1 & 3 & 2 \\ 4 & 2 & -5 \end{pmatrix}$$

sequential principal minors of B are

$M_1 = 2 > 0$

$M_2 = 5 > 0$

$M_3 = \mathrm{Det}(B) = -97 < 0$

So B is a indefinite matrix. S is not convex.

2. Solution

(a) $\forall x_1, x_2,\ \forall \lambda \in [0,1]$.

$$f(\lambda x_1 + (1-\lambda) x_2) = \sum_{i=1}^{d} |\lambda x_1 + (1-\lambda) x_2|$$

$$\leq \sum_{i=1}^{d} |\lambda x_1| + |(1-\lambda) x_2|$$

$$= \lambda \sum_{i=1}^{d} |x_1| + (1-\lambda) \sum_{i=1}^{d} |x_2|$$

$$= \lambda f(x_1) + (1-\lambda) f(x_2)$$

So $f(x)$ is convex.

(b). $\forall x_1, x_2 \in \mathbb{R}^d,\ \forall \lambda \in [0,1]$.

$$f(\lambda x_1 + (1-\lambda) x_2) = \sum_{i=1}^{d} f_i(\lambda x_1 + (1-\lambda) x_2)$$

$$\leq \sum_{i=1}^{d} \lambda f_i(x_1) + (1-\lambda) f_i(x_2)$$

$$= \lambda f(x_1) + (1-\lambda) f(x_2)$$

So $f(x)$ is convex.

(c). $\forall x_1, x_2 \in \mathbb{R}^d$. $\forall \mu \in [0,1]$.

$$f(\mu x_1 + (1-\mu)x_2) \leq f(\mu x_1) + f((1-\mu)x_2)$$
$$= |\mu| f(x_1) + |1-\mu| f(x_2)$$
$$= \mu f(x_1) + (1-\mu) f(x_2).$$

So $f(x)$ is convex.

(d). $f(x)$ is symmetric by axis y.

Consider $g(x) = x^{3/2}$. $x \geq 0$.

$$\frac{dg(x)}{dx} = \frac{3}{2} x^{\frac{1}{2}} \geq 0. \qquad \frac{d^2 g(x)}{dx^2} = \frac{3}{4} x^{-\frac{1}{2}} \geq 0.$$

$g(x)$ is convex so does $f(x)$.

(e). $x + \sqrt{1+x^2} \geq 0 \Longleftrightarrow x \in \mathbb{R}$.

$$\frac{df(x)}{dx} = \frac{1}{\sqrt{1+x^2}} + 2x. \qquad \frac{d^2 f(x)}{dx^2} = 2 - x(1+x^2)^{-\frac{3}{2}}$$

$$2 - x(1+x^2)^{-\frac{3}{2}} \geq 0 \iff \frac{x}{(1+x^2)^{3/2}} \leq 2$$

$$\iff x \leq 2(1+x^2)^{3/2}$$

$$\iff x^2 \leq 4(1+x^2)^3$$

$\forall x \in \mathbb{R}$. $x^2 \leq 4(1+x^2)$ So $\frac{d^2 f(x)}{dx^2} \geq 0$.

$f(x)$ is convex.

3. Solution.

(a). $f(\beta_0, \beta) = \sum_{i=1}^{n} (\beta_0 + \beta^T x_i - y_i)^2$

$$= \sum_{i=1}^{n} \beta_0^2 + 2\beta_0(\beta^T x_i - y_i) + (\beta^T x_i - y_i)^2$$

$$= n\beta_0^2 + 2\beta_0\left(\sum_{i=1}^{n} \beta^T x_i - \sum_{i=1}^{n} y_i\right) + \sum_{i=1}^{n} (\beta^T x_i - y_i)^2$$

$$= n\beta_0^2 + \sum_{i=1}^{n} (\beta^T x_i - y_i)^2$$

For $n\beta_0^2 \geq 0$. to minimize $f(x)$. β_0^* must be 0.

(b). To minimize $f(x)$. $\beta_0^* = 0$. β^* minimizes $(\beta' x_i - y_i)^2$.

$$\tilde{f}(\beta_0, \beta) = \sum_{i=1}^{n} (\beta_0 + \beta^T x_i' - y_i')^2$$

$$= \sum_{i=1}^{n} \beta_0^2 + 2\beta_0 (\beta^T x_i' - y_i') + (\beta^T x_i' - y_i')^2$$

$$= n\beta_0^2 + 2\beta_0 \left(\sum_{i=1}^{n} \beta^T x_i' - \sum_{i=1}^{n} y_i' \right) + \sum_{i=1}^{n} (\beta^T x_i' - y_i')^2$$

we have: $\sum_{i=1}^{n} \beta^T x_i' = n\beta^T q$. $\sum_{i=1}^{n} y_i' = nr$

$$\Rightarrow \tilde{f}(\beta_0, \beta) = n\beta_0^2 + 2n\beta_0(\beta^T q - r) + \sum_{i=1}^{n} (\beta^T x_i' - y_i')^2$$

to minimize $g(\beta_0) = n\beta_0^2 + 2n\beta_0(\beta^T q - r)$

$$\beta_0 = -\frac{2n(\beta^T q - r)}{2n} = -\beta^T q + r.$$

$\Rightarrow \beta_0 = \beta_0^* - \beta^T q + r$ $\qquad g(\beta_0) = -n\beta_0^2 = -n(\beta^T q - r)^2$.

with best β_0.

$$\tilde{f}(\beta_0, \beta) = \sum_{i=1}^{n} (\beta^T x_i' - y_i')^2 - (\beta^T g - r)^2$$

$$= \sum_{i=1}^{n} (\beta^T x_i' - y_i' - \beta^T g + r)(\beta^T x_i' - y_i' + \beta^T g - r)$$

$$= \sum_{i=1}^{n} (\beta^T x_i - y_i)(\beta^T x_i - y_i + 2\beta^T g - 2r)$$

$$= \sum_{i=1}^{n} (\beta^T x_i - y_i)^2 + 2\beta_0 (\beta^T x_i - y_i)$$

$$= \sum_{i=1}^{n} (\beta^T x_i - y_i)^2 + 2\beta_0 \left(\sum_{i=1}^{n} \beta^T x_i - \sum_{i=1}^{n} y_i \right)$$

$$= \sum_{i=1}^{n} (\beta^T x_i - y_i)^2$$

As β^* minimizes $(\beta' x_i - y_i)^2$. here $\beta = \beta^*$.

So the statement holds for $f(\beta_0, \beta)$ and $\tilde{f}(\beta_0, \beta)$.

4. Solution

(a) ① Objective function $f(x_1, x_2)$

$$\frac{\partial}{\partial x_1} f(x_1, x_2) = 6x_1 + 2x_2$$

$$\frac{\partial}{\partial x_2} f(x_1, x_2) = 2x_1 + 8x_2$$

$$\Rightarrow \nabla^2 f(x_1, x_2) = \begin{pmatrix} 6 & 2 \\ 2 & 8 \end{pmatrix} \text{ is p.d.}$$

$\Rightarrow$ Objective function is convex.

② Constraint $g(x_1, x_2) = (1 + x_1^2 + x_2^2)^{1/2} + x_1$

$$\frac{\partial}{\partial x_1} g(x_1, x_2) = x_1 (1 + x_1^2 + x_2^2)^{-1/2} + 1$$

$$\frac{\partial}{\partial x_2} g(x_1, x_2) = x_2 (1 + x_1^2 + x_2^2)^{-1/2}$$

$$\nabla g(x_1, x_2) = \begin{pmatrix} (1+x_1^2+x_2^2)^{-1/2} - x_1^2 (1+x_1^2+x_2^2)^{-3/2}, & -x_1 x_2 (1+x_1^2+x_2^2)^{-3/2} \\ -x_1 x_2 (1+x_1^2+x_2^2)^{-3/2}, & (1+x_1^2+x_2^2)^{-1/2} - x_2^2 (1+x_1^2+x_2^2)^{-3/2} \end{pmatrix}$$

$$= (1+x_1^2+x_2^2)^{-3/2} \begin{pmatrix} 1+x_2^2 & -x_1x_2 \\ -x_1x_2 & 1+x_1^2 \end{pmatrix}$$

For $A = \begin{pmatrix} 1+x_2^2 & -x_1x_2 \\ -x_1x_2 & 1+x_1^2 \end{pmatrix}$

$M_1 = 1+x_2^2 > 0.$

$M_2 = (1+x_1^2)(1+x_2^2) - x_1^2x_2^2 = 1+x_1^2+x_2^2 > 0.$

$\Rightarrow$ $g(x_1, x_2)$ is convex.

So this is a convex problem.

(b). Lagrangian function:

$L(x_1, x_2, \lambda) = 3x_1^2+4x_2^2+2x_1x_2+\lambda[(1+x_1^2+x_2^2)^{1/2}+x_1-2]$

KKT condition: ① $(1+x_1^2+x_2^2)^{1/2}+x_1-2 \le 0.$

② $\lambda \geq 0.$

③ $\nabla_1 L = 6x_1+2x_2+\lambda[x_1(1+x_1^2+x_2^2)^{-1/2}+1] = 0.$

$\nabla_2 L = 2x_1+8x_2+\lambda x_2(1+x_1^2+x_2^2)^{-1/2} = 0.$

④ $\lambda[(1+x_1^2+x_2^2)^{1/2}+x_1-2]=0.$

$x_1 \nabla_1 L=0.$

$x_2 \nabla_2 L=0.$

(c). 1). $\lambda=0$. $\nabla_1 L=6x_1+2x_2=0.$ $\Rightarrow (x_1, x_2, \lambda)=(0,0,0).$

$\nabla_2 L=2x_1+8x_2=0.$

2). $\lambda>0$. $(1+x_1^2+x_2^2)^{1/2}+x_1-2=0.$

$\nabla_1 L=6x_1+2x_2+\lambda[x_1(1+x_1^2+x_2^2)^{-1/2}+1]=0.$

$\nabla_2 L=2x_1+8x_2+\lambda x_2(1+x_1^2+x_2^2)^{-1/2}=0.$

Denote $y=(1+x_1^2+x_2^2)^{-1/2}>0.$

$\Rightarrow \nabla_1 L=6x_1+2x_2+\lambda(x_1 y+1)=0.$

$\nabla_2 L=2x_1+8x_2+\lambda x_2 y=0.$

Also. $1/y+x_1-2=0.$ $y=\frac{1}{x_1-2}$

The equations do not have solutions.

$\Rightarrow$ KKT point is $(0,0)$