Linear and Convex Optimization Homework 03

Qiu Yihang, 2021/10/14-10/17

1. ***Proof:***

We prove the proposition by contradiction.

Assume M is not a convex set, **i.e.**

$$\exists \boldsymbol{x_1}, \boldsymbol{x_2} \in M, \exists \theta \in [0,1] \textbf{ s.t. } \boldsymbol{z} \triangleq \theta \boldsymbol{x_1} + \bar{\theta}\boldsymbol{x_2} \notin M \text{ (where } \bar{\theta} = 1-\theta).$$

By definition of M, we have $\boldsymbol{x_1}, \boldsymbol{x_2} \in S$ and $f(\boldsymbol{x_1}) = f(\boldsymbol{x_2}) < f(\boldsymbol{z})$ (otherwise, $\boldsymbol{z} \in M$).

Since S is convex, $\boldsymbol{z} \in S$.

Since f is a convex function, $f(\boldsymbol{z}) = f(\theta\boldsymbol{x_1} + \bar{\theta}\boldsymbol{x_2}) \le \theta f(\boldsymbol{x_1}) + \bar{\theta} f(\boldsymbol{x_2}) = f(\boldsymbol{x_1})$.

Therefore, $f(\boldsymbol{z}) \le f(\boldsymbol{x_1}) < f(\boldsymbol{z})$. **Contradiction!**

Thus, M is a convex set. *Qed.* ■

2. ***Proof:***

We prove the proposition by contradiction.

Since f is a convex function, $\forall \theta \in [0,1]$, $f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) \le \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$.

Assume $\exists \theta_1 \in [0,1]$ **s.t.** $f(\theta_1\boldsymbol{x} + \overline{\theta_1}\boldsymbol{y}) < \theta_1 f(\boldsymbol{x}) + \overline{\theta_1} f(\boldsymbol{y})$. Obviously $\theta_1 \notin \{0,1,\theta_0\}$.

Case 1. $\theta_1 \in (0, \theta_0)$.

We have $\theta_0\boldsymbol{x} + \overline{\theta_0}\boldsymbol{y} = \frac{\overline{\theta_0}}{\overline{\theta_1}}(\theta_1\boldsymbol{x} + \overline{\theta_1}\boldsymbol{y}) + \frac{\theta_0\overline{\theta_1} - \overline{\theta_0}\theta_1}{\overline{\theta_1}}\boldsymbol{x}$.

Since $\frac{\overline{\theta_0}}{\overline{\theta_1}} + \frac{\theta_0\overline{\theta_1} - \overline{\theta_0}\theta_1}{\overline{\theta_1}} = \frac{\theta_0\overline{\theta_1} + \overline{\theta_0}(1-\theta_1)}{\overline{\theta_1}} = \frac{\theta_0\overline{\theta_1} + \overline{\theta_0}\overline{\theta_1}}{\overline{\theta_1}} = 1$,

$$\theta_0\boldsymbol{x} + \overline{\theta_0}\boldsymbol{y} \text{ is a convex combination of } \theta_1\boldsymbol{x} + \overline{\theta_1}\boldsymbol{y} \text{ and } \boldsymbol{x}.$$

Let $\varphi = \frac{\overline{\theta_0}}{\overline{\theta_1}} \in (0,1)$ (since $\theta_1 \in (0,\theta_0), \theta_0 \in (0,1)$). Then $\bar{\varphi} = 1 - \varphi = \frac{\theta_0\overline{\theta_1} - \overline{\theta_0}\theta_1}{\overline{\theta_1}} \in (0,1)$.

Since f is a convex function,

$$\begin{aligned} f(\theta_0\boldsymbol{x} + \overline{\theta_0}\boldsymbol{y}) &\le \varphi f(\theta_1\boldsymbol{x} + \overline{\theta_1}\boldsymbol{y}) + \bar{\varphi} f(\boldsymbol{x}) \\ &< (\varphi\theta_1 + \bar{\varphi}) f(\boldsymbol{x}) + \varphi\overline{\theta_1} f(\boldsymbol{y}) = \theta_0 f(\boldsymbol{x}) + \overline{\theta_0} f(\boldsymbol{y}). \end{aligned}$$

However, $f(\theta_0\boldsymbol{x} + \overline{\theta_0}\boldsymbol{y}) = \theta_0 f(\boldsymbol{x}) + \overline{\theta_0} f(\boldsymbol{y})$.

Contradiction!

Case 2. $\theta_1 \in (\theta_0, 1)$.

The proof is similar as Case 1.

In conclusion, for the same $\boldsymbol{x}, \boldsymbol{y}$, $f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$ holds for any $\theta \in [0,1]$.

Qed. ■

3.(a) ***Solution:***

$\underline{f(\boldsymbol{x})\text{ is convex.}}$ The proof is as follows.

$f(\boldsymbol{x})$ can be rewritten in the following form: $f(\boldsymbol{x}) = \left(x_1 + \frac{1}{2}x_3\right)^2 + \left(x_2 + \frac{1}{2}x_3\right)^2$.

First we prove that $\left(\theta\left(u_1 + \frac{1}{2}u_2\right) + \bar{\theta}\left(v_1 + \frac{1}{2}v_2\right)\right)^2 \le \theta\left(u_1 + \frac{1}{2}u_2\right)^2 + \bar{\theta}\left(v_1 + \frac{1}{2}v_2\right)^2$

$$\Leftrightarrow\ (\theta - \theta^2)\left(\left(u_1 + \frac{1}{2}u_2\right)^2 + \left(v_1 + \frac{1}{2}v_2\right)^2 - 2\left(u_1 + \frac{1}{2}u_2\right)\left(v_1 + \frac{1}{2}v_2\right)\right) \ge 0$$

$$\Leftrightarrow\ (\theta - \theta^2)\left(u_1 + \frac{1}{2}u_2 - v_1 - \frac{1}{2}v_2\right)^2 \ge 0. \qquad (\theta \in [0,1] \Rightarrow \theta - \theta^2 \ge 0)$$

Let $\boldsymbol{x} = (x_1, x_2, x_3), \boldsymbol{y} = (y_1, y_2, y_3)$.

$f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) - \theta f(\boldsymbol{x}) - \bar{\theta} f(\boldsymbol{y})$

$$= \left(\theta\left(x_1 + \frac{1}{2}x_3\right) + \bar{\theta}\left(y_1 + \frac{1}{2}y_3\right)\right)^2 - \theta\left(x_1 + \frac{1}{2}x_3\right)^2 - \bar{\theta}\left(y_1 + \frac{1}{2}y_3\right)^2$$
$$+ \left(\theta\left(x_2 + \frac{1}{2}x_3\right) + \bar{\theta}\left(y_2 + \frac{1}{2}y_3\right)\right)^2 - \theta\left(x_2 + \frac{1}{2}x_3\right)^2 - \bar{\theta}\left(y_2 + \frac{1}{2}y_3\right)^2$$
$$\ge 0 + 0 = 0$$

Therefore, $f(\boldsymbol{x})$ is convex. ■

Another Proof: $\nabla f = (2x_1 + x_3, 2x_2 + x_3, x_1 + x_2 + x_3)$, $\nabla^2 f = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

$$2 \ge 0, 1 \ge 0, \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 \ge 0, \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1 \ge 0, \begin{vmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0 \ge 0 \Rightarrow \nabla^2 f \succcurlyeq \mathbf{0}.$$

By Theorem, $f(\boldsymbol{x})$ is convex. □

(b) Solution:

$\underline{f(\boldsymbol{x})\text{ is convex.}}$ The proof is as follows.

$$\nabla f = \left(-\frac{1}{x_1^2 x_2}, -\frac{1}{x_1 x_2^2}\right), \ \nabla^2 f = \begin{pmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{pmatrix}.$$

Since $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2_{++}$,

$$\frac{2}{x_1^3 x_2} > 0, \frac{2}{x_1 x_2^3} > 0, \begin{vmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{vmatrix} = \frac{3}{x_1^4 x_2^4} > 0.$$

i.e. $\nabla^2 f \succ \mathbf{0}$.

By Theorem, we know $f(\boldsymbol{x})$ is a convex function. ■

(c) ***Solution:***

$f(\boldsymbol{x})$ is neither convex nor concave. The proof is as follows.

Let $\boldsymbol{x} = (1,3), \boldsymbol{y} = (3,1), \theta = \frac{1}{2}$. Then $\bar{\theta} = \frac{1}{2}, \theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} = (2,2)$.

$f(\boldsymbol{x}) = 9, f(\boldsymbol{y}) = 3, f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = 8 > 6 = \frac{1}{2}(9+3) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$.

Thus, $f(\boldsymbol{x})$ is not convex.

Let $\boldsymbol{x} = (1,1), \boldsymbol{y} = (1,3), \theta = \frac{1}{2}$. Then $\bar{\theta} = \frac{1}{2}, \theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} = (1,2)$.

$f(\boldsymbol{x}) = 1, f(\boldsymbol{y}) = 9, f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = 4 < 5 = \frac{1}{2}(1+9) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$.

Thus, $f(\boldsymbol{x})$ is not concave.

In conclusion, we know $f(\boldsymbol{x})$ is neither convex nor concave. ■

(d) ***Solution:***

$f(\boldsymbol{x})$ is neither convex nor concave. The proof is as follows.

Let $\boldsymbol{x} = (1,1), \boldsymbol{y} = (3,3), \theta = \frac{1}{2}$. Then $\bar{\theta} = \frac{1}{2}, \theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} = (2,2)$.

$f(\boldsymbol{x}) = 1, f(\boldsymbol{y}) = \sqrt{3}, f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = \sqrt{2} = \frac{2\sqrt{2}}{2} > \frac{1}{2}\left(1+\sqrt{3}\right) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$.

Thus, $f(\boldsymbol{x})$ is not convex.

Let $\boldsymbol{x} = (1,3), \boldsymbol{y} = (3,1), \theta = \frac{1}{2}$. Then $\bar{\theta} = \frac{1}{2}, \theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} = (2,2)$.

$f(\boldsymbol{x}) = \frac{1}{\sqrt{3}}, f(\boldsymbol{y}) = 3, f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = \sqrt{2} = \frac{6\sqrt{2}}{6} < \frac{9+\sqrt{3}}{6} = \frac{1}{2}\left(\frac{1}{\sqrt{3}} + 3\right) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$.

Thus, $f(\boldsymbol{x})$ is not concave.

In conclusion, we know $f(\boldsymbol{x})$ is neither convex nor concave. ■

(e) ***Solution:***

CASE 01. $\alpha = 0$.

We have $f(\boldsymbol{x}) = x_2$. $\nabla f = (0,1)$, $\nabla^2 f = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{0}$.

Thus, $f(\boldsymbol{x})$ is both convex and concave.

CASE 02. $\alpha = 1$.

We have $f(\boldsymbol{x}) = x_1$. $\nabla f = (1,0)$, $\nabla^2 f = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{0}$.

Thus, $f(\boldsymbol{x})$ is both convex and concave.

CASE 03. $0 < \alpha < 1$.

$$\nabla f = (\alpha x_1^{\alpha-1} x_2^{1-\alpha}, (1-\alpha) x_1^{\alpha} x_2^{-\alpha}),$$

$$\nabla^2 f = \begin{pmatrix} \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} & \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \\ \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} & \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1} \end{pmatrix}.$$

Since $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}_{++}^2$,

$$\alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} < 0, \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1} < 0,$$

$$\begin{vmatrix} \alpha(\alpha-1) x_1^{\alpha-2} x_2^{1-\alpha} & \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \\ \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} & \alpha(\alpha-1) x_1^{\alpha} x_2^{-\alpha-1} \end{vmatrix}$$

$$= \alpha(\alpha-1) x_1^{\alpha-2} x_2^{-\alpha-1} (x_2^2 + x_1^2 + 2x_1x_2)$$

$$= \alpha(\alpha-1) x_1^{\alpha-2} x_2^{-\alpha-1} (x_1 + x_2)^2 < 0$$

i.e. $\nabla^2 f \prec \mathbf{0}$.

By Theorem, we know $f(\boldsymbol{x})$ is a concave function.

In conclusion, $f(\boldsymbol{x})$ is $\begin{cases} \text{concave.} & \alpha \in (0,1) \\ \text{both convex and concave.} & \alpha = 0 \ or \ \alpha = 1 \end{cases}$ ■

4. ***Proof:***

For any $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2, \boldsymbol{y} = (y_1, y_2) \in \mathbb{R}^2, \theta \in [0,1]$, let $\bar{\theta} = 1 - \theta$.

Since $f_1(x)$ and $f_2(x)$ are strictly convex, we have $\theta f_i(x_i) + \bar{\theta} f_i(y_i) > f_i(\theta x_i + \bar{\theta} y_i), i \in \{1,2\}$.

Thus,

$$\begin{aligned} f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) &= f_1(\theta x_1 + \bar{\theta} y_1) + f_2(\theta x_2 + \bar{\theta} y_2) \\ &< \theta f_1(x_1) + \bar{\theta} f_1(y_1) + \theta f_2(x_2) + \bar{\theta} f_2(y_2) \\ &= \theta\big(f_1(x_1) + f_2(x_2)\big) + \bar{\theta}\big(f_1(y_1) + f_2(y_2)\big) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y}). \end{aligned}$$

Therefore, $f(x_1, x_2)$ is a strictly convex function. *Qed.* ■

Now we prove $f(x_1, x_2) = x_1^2 + x_2^4$ is strictly convex.

We know $f_1(x) = x^2, f_2(x) = x^4$ are both strictly convex.

From the conclusion proved above, we know $f(x_1, x_2) = f_1(x_1) + f_2(x_2)$ is strictly convex. ■

5. ***Proof:***

First we prove the sufficiency. Assume f is convex.

For any $\boldsymbol{x}, \boldsymbol{y} \in C$. Let $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{x}, g(t) = \nabla f(\boldsymbol{x} + t\boldsymbol{d})^T \boldsymbol{d}$. Thus, $g'(t) = \boldsymbol{d}^T \nabla^2 f(\boldsymbol{x} + t\boldsymbol{d}) \boldsymbol{d}$.

Since f is convex, $\nabla^2 f \succcurlyeq \mathbf{0}$.

Since C is convex, $\boldsymbol{x} + t\boldsymbol{d} \in C$, $g'(t) = \boldsymbol{d}^T \nabla^2 f(\boldsymbol{x} + t\boldsymbol{d}) \boldsymbol{d} \geq \mathbf{0}$, **i.e.** $g(t)$ is increasing on $[0,1]$.

Thus, $[g(0)-g(1)](0-1)\geq 0$, **i.e.**

$$-(\nabla f(\boldsymbol{x})^T\boldsymbol{d}-\nabla f(\boldsymbol{y})^T\boldsymbol{d})=\left(\nabla f(\boldsymbol{x})-\nabla f(\boldsymbol{y})\right)^T(-\boldsymbol{d})\geq \boldsymbol{0},$$

i.e.

$$\langle \nabla f(\boldsymbol{x})-\nabla f(\boldsymbol{y}),\boldsymbol{x}-\boldsymbol{y}\rangle\geq \boldsymbol{0},\qquad \forall \boldsymbol{x},\boldsymbol{y}\in C.$$

Now we prove the necessity. Assume $\forall x,y\in C,\langle \nabla f(\boldsymbol{x})-\nabla f(\boldsymbol{y}),\boldsymbol{x}-\boldsymbol{y}\rangle\geq \boldsymbol{0}$.

Let $\boldsymbol{d}=\boldsymbol{x}-\boldsymbol{y},h(t)=f(\boldsymbol{y}+t\boldsymbol{d})$. Thus, $h'(t)=\nabla f(\boldsymbol{y}+t\boldsymbol{d})^T\boldsymbol{d}$.

For any $\boldsymbol{x},\boldsymbol{y}\in C$, since C is convex, $\forall \alpha_1,\alpha_2\in[0,1],\alpha_1\boldsymbol{x}+\overline{\alpha_1}\boldsymbol{y}\in \boldsymbol{C},\alpha_2\boldsymbol{x}+\overline{\alpha_2}\boldsymbol{y}\in \boldsymbol{C}$

(where $\overline{\alpha_i}=1-\alpha_i,i\in\{1,2\}$).

Let $\alpha_1<\alpha_2$.

Thus,

$$\begin{aligned}&\langle \nabla f(\alpha_1\boldsymbol{x}+\overline{\alpha_1}\boldsymbol{y})-\nabla f(\alpha_2\boldsymbol{x}+\overline{\alpha_2}\boldsymbol{y}),(\alpha_1\boldsymbol{x}+\overline{\alpha_1}\boldsymbol{y})-(\alpha_2\boldsymbol{x}+\overline{\alpha_2}\boldsymbol{y})\rangle\\&\qquad=(\alpha_1-\alpha_2)\left(\nabla f(\alpha_1\boldsymbol{x}+\overline{\alpha_1}\boldsymbol{y})-\nabla f(\alpha_2\boldsymbol{x}+\overline{\alpha_2}\boldsymbol{y})\right)^T\boldsymbol{d}\geq 0\end{aligned}$$

i.e.

$$\nabla f(\alpha_1\boldsymbol{x}+\overline{\alpha_1}\boldsymbol{y})^T\boldsymbol{d}\leq \nabla f(\alpha_2\boldsymbol{x}+\overline{\alpha_2}\boldsymbol{y})^T\boldsymbol{d},\qquad \forall\ \alpha_1,\alpha_2\in[0,1],\alpha_1<\alpha_2.$$

i.e.

$$h'(\alpha_1)\leq h'(\alpha_2),\qquad \forall\ \alpha_1,\alpha_2\in[0,1],\alpha_1<\alpha_2.$$

By **<u>Lagrange Mean Value Theorem</u>**, for any $\theta\in(0,1)$, exist $\varphi_1\in(0,\theta),\varphi_2\in(\theta,1)$ **s.t.**

$$\frac{h(\theta)-h(0)}{\theta-0}=h'(\varphi_1),\frac{h(1)-h(\theta)}{1-\theta}=h'(\varphi_2).$$

Since $1>\varphi_2>\theta>\varphi_1>0$, we have $h'(\varphi_1)\leq h'(\varphi_2)$.

Thus,

$$\begin{aligned}\frac{h(\theta)-h(0)}{\theta}\leq\frac{h(1)-h(\theta)}{\bar\theta}&\Leftrightarrow \bar\theta h(\theta)-\bar\theta h(0)\leq \theta h(1)-\theta h(\theta)\\&\Leftrightarrow\ h(\theta)\leq\bar\theta h(0)+\theta h(1)\end{aligned}$$

i.e.

$$f(\theta\boldsymbol{x}+\bar\theta\boldsymbol{y})\leq\bar\theta f(\boldsymbol{y})+\theta f(\boldsymbol{x}).\qquad \forall \boldsymbol{x},\boldsymbol{y}\in C,\forall\theta\in(0,1).$$

Considering when $\theta\in\{0,1\}$, obviously $f(\theta\boldsymbol{x}+\bar\theta\boldsymbol{y})=\bar\theta f(\boldsymbol{y})+\theta f(\boldsymbol{x})$, we have

$$f(\theta\boldsymbol{x}+\bar\theta\boldsymbol{y})\leq\bar\theta f(\boldsymbol{y})+\theta f(\boldsymbol{x}).\qquad \forall \boldsymbol{x},\boldsymbol{y}\in C,\forall\theta\in[0,1].$$

Therefore, f is convex.

In conclusion, f is convex **<u>iff.</u>** $\langle \nabla f(\boldsymbol{x})-\nabla f(\boldsymbol{y}),\boldsymbol{x}-\boldsymbol{y}\rangle\geq \boldsymbol{0}$, $\forall \boldsymbol{x},\boldsymbol{y}\in C$.

Qed. ■