SI251 - Convex Optimization, 2024 Spring

Homework 1

Name: Zhou Shouchen

Student ID: 2021533042

Due 23:59 (CST), Mar. 27, 2024

1 Convex sets

1. Please prove that the following sets are convex:

   1) $S = \{x \in \boldsymbol{R}^m \mid \mid p(t) \mid \leq 1 \text{ for } |t| \leq \pi/3\}$, where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$. (5 pts)

   2) (Ellipsoids) $\left\{x | \sqrt{(x-x_c)^T P(x-x_c)} \leq r\right\} \quad (x_c \in \mathbb{R}^n, r \in \mathbb{R}, P \succeq 0)$. (5 pts)

   3) (Symmetric positive semidefinite matrices) $S_+^{n\times n} = \left\{P \in S^{n\times n} | P \succeq 0\right\}$. (5 pts)

   4) The set of points closer to a given point than a given set, i.e.,

   $$\left\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\right\},$$

   where $S \in R^n$. (5 pts)

(1) For a fixed $t \in [-\frac{\pi}{3}, \frac{\pi}{3}]$, we could know that $\cos t, \cos 2t, \cdots, \cos mt$ are certein constants, so $p(t)$ is a linear function of $\boldsymbol{x}$.
Since $|p(t)| \leq 1 \Leftrightarrow -1 \leq p(t) \leq 1$.
So let $S_t = \{\boldsymbol{x} | -1 \leq x_1 \cos t + \cdots + x_n \cos nt \leq 1\}$.
Since $p(t)$ is linear function of $\boldsymbol{x}$, so S_t the interaction of two half spaces, which is a convex set.
And we could know that

$$S = \bigcap_{-\frac{\pi}{3} \leq t \leq \frac{\pi}{3}} S_t$$

From the theorem, we could know that the intersection of convex sets is also a convex set, so S is a convex set.

So above all, we have proved that S is a convex set.

(2) Let S be the Ellipsoids set, and we could know that $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, we have $(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_1 - \boldsymbol{x}_c) \leq r^2$ and $(\boldsymbol{x}_2 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_2 - \boldsymbol{x}_c) \leq r^2$.
And $\forall \theta \in [0,1]$, since $P \succeq 0$, so P is symmetric, so we have

$$\begin{aligned}
&[(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) - \boldsymbol{x}_c]^T P[(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) - \boldsymbol{x}_c] \\
&= \theta^2(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_1 - \boldsymbol{x}_c) + 2\theta(1-\theta)(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_2 - \boldsymbol{x}_c) + (1-\theta)^2(\boldsymbol{x}_2 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_2 - \boldsymbol{x}_c) \\
&\leq \theta^2 r^2 + (1-\theta)^2 r^2 + 2\theta(1-\theta)(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_2 - \boldsymbol{x}_c)
\end{aligned}$$

And since $P \succeq 0$, so P could be decomposition as $P = Q^T \Lambda Q$, so $P^{\frac{1}{2}} = Q^T \Lambda^{\frac{1}{2}} Q$, i.e. $P^{\frac{1}{2}} \succeq 0$.
So

$$\begin{aligned}
(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_2 - \boldsymbol{x}_c) &= (\boldsymbol{x}_1 - \boldsymbol{x}_c)^T (P^{\frac{1}{2}})^T (P^{\frac{1}{2}})(\boldsymbol{x}_2 - \boldsymbol{x}_c) \\
&= [P^{\frac{1}{2}}(\boldsymbol{x}_1 - \boldsymbol{x}_c)]^T [P^{\frac{1}{2}}(\boldsymbol{x}_2 - \boldsymbol{x}_c)] \\
&\leq \|[P^{\frac{1}{2}}(\boldsymbol{x}_1 - \boldsymbol{x}_c)]\|_2 \cdot \|P^{\frac{1}{2}}(\boldsymbol{x}_2 - \boldsymbol{x}_c)\|_2
\end{aligned}$$

Since $x_1 \in S$, so

$$(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_1 - \boldsymbol{x}_c) = (\boldsymbol{x}_1 - \boldsymbol{x}_c)^T (P^{\frac{1}{2}})^T [P^{\frac{1}{2}}(\boldsymbol{x}_1 - \boldsymbol{x}_c)] = \|P^{\frac{1}{2}}(\boldsymbol{x}_1 - \boldsymbol{x}_c)\|_2^2 \leq r^2$$

i.e.

$$\|P^{\frac{1}{2}}(\boldsymbol{x}_1 - \boldsymbol{x}_c)\| \le r$$

Similarly, we have

$$\|P^{\frac{1}{2}}(\boldsymbol{x}_2 - \boldsymbol{x}_c)\| \le r$$

So

$$\begin{aligned}(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_2 - \boldsymbol{x}_c) &\le \|[P^{\frac{1}{2}}(\boldsymbol{x}_1 - \boldsymbol{x}_c)]\|_2 \cdot \|P^{\frac{1}{2}}(\boldsymbol{x}_2 - \boldsymbol{x}_c)\|_2 \\ &\le r \cdot r \\ &= r^2\end{aligned}$$

So

$$\begin{aligned}&[(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) - \boldsymbol{x}_c]^T P[(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) - \boldsymbol{x}_c] \\ \le\ &\theta^2 r^2 + (1-\theta)^2 r^2 + 2\theta(1-\theta)(\boldsymbol{x}_1 - \boldsymbol{x}_c)^T P(\boldsymbol{x}_2 - \boldsymbol{x}_c) \\ \le\ &\theta^2 r^2 + (1-\theta)^2 r^2 + 2\theta(1-\theta) r^2 \\ =\ &r^2\end{aligned}$$

So $\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in S$.

So above all, we have proved that $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S, \forall\theta \in [0,1]$, we have $\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in S$. So S i.e. the Ellipsoids is a convex set.

(3) $\forall A, B \in S_+^{n\times n}$, we have $A^T = A, B^T = B$, and $\forall \boldsymbol{y} \in \mathbb{R}^n, \boldsymbol{y}^T A\boldsymbol{y} \ge 0, \boldsymbol{y}^T B\boldsymbol{y} \ge 0$.
So $\forall\theta \in [0,1]$, we have

$$(\theta A + (1-\theta)B)^T = \theta A^T + (1-\theta)B^T = \theta A + (1-\theta)B$$

And

$$\boldsymbol{y}^T(\theta A + (1-\theta)B)\boldsymbol{y} = \theta\boldsymbol{y}^T A\boldsymbol{y} + (1-\theta)\boldsymbol{y}^T B\boldsymbol{y} \ge 0$$

So $\theta A + (1-\theta)B$ is symmetric and semi-positive defined.
So $\theta A + (1-\theta)B \in S_+^{n\times n}$.

So above all, we have proved that $S_+^{n\times n}$ is a convex set.

(4) Let $\mathcal{C} = \Big\{x \mid \|x - x_0\|_2 \le \|x - y\|_2 \text{ for all } y \in S\Big\}$.
$\forall \boldsymbol{x} \in \mathcal{C}$, and for a fixed $\boldsymbol{y}$, we have

$$\begin{aligned}\|\boldsymbol{x} - \boldsymbol{x}_0\|_2 &\le \|\boldsymbol{x} - \boldsymbol{y}\|_2 \\ \|\boldsymbol{x} - \boldsymbol{x}_0\|_2^2 &\le \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 \\ (\boldsymbol{x} - \boldsymbol{x}_0)^T(\boldsymbol{x} - \boldsymbol{x}_0) &\le (\boldsymbol{x} - \boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y}) \\ \boldsymbol{x}^T(\boldsymbol{x}_0 - \boldsymbol{y}) &\ge \frac{1}{2}(\boldsymbol{x}_0^T\boldsymbol{x}_0 - \boldsymbol{y}^T\boldsymbol{y})\end{aligned}$$

From the definition, we know that for a fixed $\boldsymbol{y}$, $\boldsymbol{x}^T(\boldsymbol{x}_0 - \boldsymbol{y}) \ge \frac{1}{2}(\boldsymbol{x}_0^T\boldsymbol{x}_0 - \boldsymbol{y}^T\boldsymbol{y})$ is a half-space $S_{\boldsymbol{y}}$.
So $\forall \boldsymbol{y} \in S$, we could see that $\mathcal{C} = \bigcap_{\boldsymbol{y}\in S} S_{\boldsymbol{y}}$.
And since each $S_{\boldsymbol{y}}$ is a half-space, which is a convex set. And from the theorem we have known, that the intersection of convex sets is also a convex set, so $\mathcal{C}$ is a convex set.

2. (15 pts) For a given norm $\|\cdot\|$ on $\boldsymbol{R}^n$, the dual norm, denoted $\|\cdot\|_*$, is defined as

$$\|y\|_* = \sup_{x\in\boldsymbol{R}^n} \{y^T x \mid \|x\| \le 1\}.$$

Show that the dual of Euclidean norm is the Euclidean norm, i.e., $\sup\limits_{x\in\boldsymbol{R}^n} \{z^T x \mid \|x\|_2 \le 1\} = ||z||_2$.

$$\begin{aligned}
&\boldsymbol{z}^T\boldsymbol{x} \\
&\le \|\boldsymbol{z}\|_2\|\boldsymbol{x}\|_2 \quad \text{(Cauchy-Schwarz inequality)} \\
&\le \|\boldsymbol{z}\|_2 \qquad\qquad (\|\boldsymbol{x}\|_2 \le 1)
\end{aligned}$$

If and only if $\|\boldsymbol{x}\|_2 = 1$ and $\cos\langle\boldsymbol{z},\boldsymbol{x}\rangle = 1$ will take the equation condition.

So above all, we have proved that $\sup\limits_{\boldsymbol{x}\in\boldsymbol{R}^n} \{\boldsymbol{z}^T\boldsymbol{x} \mid \|\boldsymbol{x}\|_2 \le 1\} = \|\boldsymbol{z}\|_2$.

3. (15 pts) Define a norm cone as

$$\mathcal{C} \equiv \left\{(x,t) : x \in \mathbb{R}^d, t \geq 0, \|x\| \leq t\right\} \subseteq \mathbb{R}^{d+1}$$

Show that the norm cone is convex by using the definition of convex sets.

$\forall (\boldsymbol{x}_1, t_1), (\boldsymbol{x}_2, t_2) \in \mathcal{C}$, we have $\|\boldsymbol{x}_1\| \leq t_1, \|\boldsymbol{x}_2\| \leq t_2, t_1 \geq 0, t_2 \geq 0$.
And $\forall \theta \in [0,1]$, we have

$$\begin{aligned}
&\|\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2\| \\
&\leq \|\theta \boldsymbol{x}_1\| + \|(1-\theta)\boldsymbol{x}_2\| \\
&= \theta\|\boldsymbol{x}_1\| + (1-\theta)\|\boldsymbol{x}_2\| \\
&= \theta t_1 + (1-\theta)t_2
\end{aligned}$$

Also, since $t_1, t_2 \geq 0$, so $\theta t_1 + (1-\theta)t_2 \geq 0$.

So $\theta(\boldsymbol{x}_1, t_1) + (1-\theta)(\boldsymbol{x}_2, t_2) = (\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \theta t_1 + (1-\theta)t_2) \in \mathcal{C}$.

So above all, we have proved that $\mathcal{C}$ is a convex set.

2 Convex functions

4. (18 pts) Let $C \subset \mathbb{R}^n$ be convex and $f : C \to R^\star$. Show that the following statements are equivalent:

   (a) epi(f) is convex.

   (b) For all points $x_i \in C$ and $\{\lambda_i | \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1, i = 1, 2, \cdots, n\}$, we have
   $$f\Big(\sum_{i=1}^n \lambda_i x_i\Big) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

   (c) For $\forall x, y \in C$ and $\lambda \in [0, 1]$,
   $$f\Big((1-\lambda)x + \lambda y\Big) \leq (1-\lambda)f(x) + \lambda f(y).$$

- $(a) \Rightarrow (c)$
  $\forall \boldsymbol{x}, \boldsymbol{y} \in C$, we have $(\boldsymbol{x}, f(\boldsymbol{x})), (\boldsymbol{y}, f(\boldsymbol{y})) \in \text{epi}(f)$.
  From (a), we have known that epi(f) is convex, so $\forall \lambda \in [0, 1]$, we have
  $$((1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y}, (1-\lambda)f(\boldsymbol{x}) + \lambda f(\boldsymbol{y})) \in \text{epi}(f)$$
  which means that
  $$f((1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y}) \leq (1-\lambda)f(\boldsymbol{x}) + \lambda f(\boldsymbol{y})$$
  So $(a) \Rightarrow (c)$ has been proved.

- $(c) \Rightarrow (a)$
  $\forall \boldsymbol{x}, \boldsymbol{y} \in C$, and $\forall (\boldsymbol{x}, t_1), (\boldsymbol{y}, t_2) \in \text{epi}(f)$, we have $t_1 \geq f(\boldsymbol{x}), t_2 \geq f(\boldsymbol{y})$.
  And $\forall \lambda \in [0, 1]$, we have
  $$\begin{aligned} & f((1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y}) \\ \leq\ & (1-\lambda)f(\boldsymbol{x}) + \lambda f(\boldsymbol{y}) \\ \leq\ & (1-\lambda)t_1 + \lambda t_2 \end{aligned}$$
  So $((1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y}, (1-\lambda)t_1 + \lambda t_2) \in \text{epi}(f)$, so epi$(f)$ is convex.
  So $(c) \Rightarrow (a)$ has been proved.

- $(b) \Rightarrow (c)$
  Let $n = 2, \lambda_1 = 1 - \lambda, \lambda_2 = \lambda$, then we have
  $$f((1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y}) \leq (1-\lambda)f(\boldsymbol{x}) + \lambda f(\boldsymbol{y})$$
  So $(b) \Rightarrow (c)$ has been proved.

- $(c) \Rightarrow (b)$
  We can use induction to prove this.
  When $n = 2$, let $\lambda_1 = 1 - \lambda, \lambda_2 = \lambda$, then we have
  $$f\Big(\sum_{i=1}^n \lambda_i \boldsymbol{x}_i\Big) \leq \sum_{i=1}^n \lambda_i f(\boldsymbol{x}_i)$$

And since f is convex, so we have $\sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i \in C$.

Suppose when $n = k$, (b) holds.
i.e. $\forall \boldsymbol{y}_i \in C$ and $\{\nu_i | \nu_i \geq 0, \sum_{i=1}^{k} \nu_i = 1, i = 1, 2, \cdots, k\}$, we have

$$f\Big(\sum_{i=1}^{k} \nu_i \boldsymbol{y}_i\Big) \leq \sum_{i=1}^{k} \nu_i f(\boldsymbol{y}_i)$$

And also suppose that $\boldsymbol{z} = \sum_{i=1}^{k} \nu_i f(\boldsymbol{y}_i) \in C$.

Then for $n = k+1$, $\forall \boldsymbol{x} \in C$, we have:
$\forall \lambda \in [0,1]$, since $\boldsymbol{x}, \boldsymbol{z} \in C$, so we can get that

$$\begin{aligned}
f((1-\lambda)\boldsymbol{x} + \lambda \boldsymbol{z}) &\leq (1-\lambda) f(\boldsymbol{x}) + \lambda f(\boldsymbol{z}) && \text{(from (c))} \\
&= (1-\lambda) f(\boldsymbol{x}) + \lambda f(\sum_{i=1}^{k} \nu_i f(\boldsymbol{y}_i)) \\
&\leq (1-\lambda) f(\boldsymbol{x}) + \lambda \left(\sum_{i=1}^{k} \nu_i f(\boldsymbol{y}_i) \right) && \text{(the assuption when } n = k) \\
&= (1-\lambda) f(\boldsymbol{x}) + \sum_{i=1}^{k} (\lambda \nu_i) f(\boldsymbol{y}_i)
\end{aligned}$$

Let $\lambda_{k+1} = 1 - \lambda, \boldsymbol{x}_{k+1} = \boldsymbol{x}, \lambda_i = \lambda \nu_i, \boldsymbol{x}_i = \boldsymbol{y}_1$, where $i = 1, 2, \cdots, k$.
Since we have $\sum_{i=1}^{k} \nu_i = 1$,
so $\sum_{i=1}^{n} \lambda_i = (1-\lambda) + \lambda \left(\sum_{i=1}^{k} \nu_i \right) = 1, \lambda_i \geq 0$.
So we have proved that when $n = k+1$, (b) holds.
Also, since $\boldsymbol{x}, \boldsymbol{z} \in C$, and C is a convex set, so $\forall \lambda \in [0,1], (1-\lambda)\boldsymbol{x} + \lambda \boldsymbol{z} \in C$, i.e. $\sum_{i=1}^{k+1} \lambda_i \boldsymbol{x}_i \in C$
So $\forall n \geq 2$, we have proved that $(c) \Rightarrow (b)$.

Since we have proved that $(a) \Leftrightarrow (c)$ and $(b) \Leftrightarrow (c)$, so we could say that (a),(b),(c) three statements are equivalent.

5. (14 pts) Monotone Mappings. A function $\psi : \boldsymbol{R}^n \to \boldsymbol{R}^n$ is called monotone if for all $x, y \in \boldsymbol{dom}\psi$,

$$(\psi(x) - \psi(y))^T (x - y) \geq 0$$

Suppose $f : \boldsymbol{R}^n \to \boldsymbol{R}^n$ is a differentiable convex function. Show that its gradient ∇f is monotone. Is the converse true, i.e., is every monotone mapping the gradient of a convex function?

1. If f is a differentiable convex function, we could know that $\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{R}^n$,

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T (\boldsymbol{y} - \boldsymbol{x})$$
$$f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y})^T (\boldsymbol{x} - \boldsymbol{y})$$

Add these two inequality, we will get that

$$\begin{aligned} \nabla f(\boldsymbol{x})^T (\boldsymbol{y} - \boldsymbol{x}) + \nabla f(\boldsymbol{y})^T (\boldsymbol{x} - \boldsymbol{y}) &\leq 0 \\ \nabla f(\boldsymbol{x})^T (\boldsymbol{x} - \boldsymbol{y}) - \nabla f(\boldsymbol{y})^T (\boldsymbol{x} - \boldsymbol{y}) &\geq 0 \\ (\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}))^T (\boldsymbol{x} - \boldsymbol{y}) &\geq 0 \end{aligned}$$

So we have proved that a differentiable convex function's gradient ∇f is monotone.

2. Suppose that $\psi(x_1, x_2) = (-x_2, x_1)$.
Let $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$, then

$$\begin{aligned} (\psi(\boldsymbol{x}) - \psi(\boldsymbol{y}))^T (\boldsymbol{x} - \boldsymbol{y}) &= [(-x_2, x_1) - (-y_2, y_1)]^T [(x_1, x_2) - (y_1, y_2)] \\ &= (-x_2 + y_2, x_1 - y_1)^T [(x_1 - y_1, x_2 - y_2)] \\ &= (x_1 - y_1) \cdot (x_2 - y_2) + (x_1 - y_1) \cdot [-(x_2 - y_2)] \\ &= 0 \\ &\geq 0 \end{aligned}$$

So our constructed $\psi(\boldsymbol{x})$ is monotone.
Let f be the primitive function of ψ, then $\dfrac{\partial^2 f}{\partial x_1 \partial x_2} = -1, \dfrac{\partial^2 f}{\partial x_2 \partial x_1} = 1$.

But for a differentiable convex function f, it must have $\dfrac{\partial^2 f}{\partial x_1 \partial x_2} = \dfrac{\partial^2 f}{\partial x_2 \partial x_1}$.

So for a monotone fuction, it may not a gradient of a differentianle convex function.

So above all, differentiable convex function's gradient ∇f is monotone, but its converse is not true.

6. (18 pts) Please determine whether the following functions are convex, concave or none of those, and give a detailed explanation for your choice.

   1)
   $$f_1(x_1, x_2, \cdots, x_n) = \begin{cases} -(x_1x_2\cdots x_n)^{\frac{1}{n}}, & \text{if } \ x_1, \cdots, x_n > 0 \\ \infty & \text{otherwise;} \end{cases}$$

   2) $f_2(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, where $0 \le \alpha \le 1$, on $\mathbb{R}^2_{++}$;

   3) $f_3(x, u, v) = -\log(uv - x^T x)$ on $dom f = \{(x, u, v) | uv > x^T x, \quad u, v > 0\}$.

(1) We could see that f_1 is twice continuously differentiable over dom f_1, and $\forall \boldsymbol{x} \in$ dom f_1, its Hessian matrix is that:

$$\nabla^2 f_1(\boldsymbol{x}) = \frac{-(x_1x_2\cdots x_n)^{\frac{1}{n}}}{n^2} \begin{bmatrix} \frac{1-n}{x_1^2} & \frac{1}{x_1x_2} & \cdots & \frac{1}{x_1x_n} \\ \frac{1}{x_2x_1} & \frac{1-n}{x_2^2} & \cdots & \frac{1}{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_nx_1} & \frac{1}{x_1x_2} & \cdots & \frac{1-n}{x_n^2} \end{bmatrix}$$

And $\forall \boldsymbol{x} \in$ dom $f_1, \boldsymbol{y} \in \mathbb{R}^n$, we have

$$\begin{aligned} \boldsymbol{y}^T \nabla^2 f_1(x) \boldsymbol{y} &= \frac{-(x_1x_2\cdots x_n)^{\frac{1}{n}}}{n^2} \cdot \left[\sum_{i=1}^{n} \frac{y_i^2(1-n)}{x_i^2} + \sum_{i \neq j} \frac{y_iy_j}{x_ix_j} \right] \\ &= \frac{-(x_1x_2\cdots x_n)^{\frac{1}{n}}}{n^2} \cdot \left[\left(\sum_{i=1}^{n} \frac{y_i}{x_i} \right)^2 - n \sum_{i=1}^{n} \left(\frac{y_i}{x_i} \right)^2 \right] \end{aligned}$$

From the multivariate mean inequality, we could know that

$$\left(\frac{\sum_{i=1}^n a_i}{n} \right)^2 \le \frac{1}{n} \sum_{i=1}^{n} a_i^2 \Rightarrow \left(\sum_{i=1}^{n} a_i \right)^2 \le n \sum_{i=1}^{n} a_i^2$$

So we could know that

$$\left[\left(\sum_{i=1}^{n} \frac{y_i}{x_i} \right)^2 - n \sum_{i=1}^{n} \left(\frac{y_i}{x_i} \right)^2 \right] \le 0$$

And since $\frac{-(x_1x_2\cdots x_n)^{\frac{1}{n}}}{n^2} < 0$, so we could know that

$$\forall \boldsymbol{x} \in \mathrm{dom} f_1, \boldsymbol{y} \in \mathbb{R}^n, \boldsymbol{y}^T \nabla^2 f_1(x) \boldsymbol{y} \ge 0$$

So above all, $f_1(x_1, \cdots, x_n)$ is convex.

(2) The Hessian of f_2 is that:

$$\begin{aligned} \nabla^2 f_2(x) &= \begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^{\alpha}x_2^{-\alpha-1} \end{bmatrix} \\ &= \alpha(\alpha-1)x_1^{\alpha}x_2^{1-\alpha} \begin{bmatrix} \frac{1}{x_1^2} & -\frac{1}{x_1x_2} \\ -\frac{1}{x_1x_2} & \frac{1}{x_2^2} \end{bmatrix} \end{aligned}$$

So $\forall \boldsymbol{y} = (y_1, y_2)$, we have

$$\boldsymbol{y}^T \nabla^2 f_2(\boldsymbol{x}) \boldsymbol{y} = \alpha(\alpha - 1) x_1^\alpha x_2^{1-\alpha} \left(\frac{y_1}{x_1} - \frac{y_2}{x_2} \right)^2 \leq 0$$

So above all, $f_2(x_1, x_2)$ is concave.

(3) $f_3(\boldsymbol{x}, u, v) = -\log(uv - \boldsymbol{x}^T\boldsymbol{x}) = -\log\left(u(v - \frac{\boldsymbol{x}^T\boldsymbol{x}}{u})\right) = -\log u - \log\left(v - \frac{\boldsymbol{x}^T\boldsymbol{x}}{u}\right)$.

From we have known about the perspective: if $f(\boldsymbol{x})$ is convex, then its perspective $g(\boldsymbol{x}, t) = tf\left(\frac{\boldsymbol{x}}{t}\right)$ is convex.

Since $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x}$ is convex, so $g(\boldsymbol{x}, t) = \frac{\boldsymbol{x}^T\boldsymbol{x}}{t}$ is convex.

And since v is affine, $-\frac{\boldsymbol{x}^T\boldsymbol{x}}{t}$ is concave, so $\left(v - \frac{\boldsymbol{x}^T\boldsymbol{x}}{u}\right)$ is concave.

Since $h(x) = -\log x$ is convex and non-increasing, $\left(v - \frac{\boldsymbol{x}^T\boldsymbol{x}}{u}\right)$ is concave, so from the composition with scalar functions, we could know that $-\log\left(v - \frac{\boldsymbol{x}^T\boldsymbol{x}}{u}\right)$ is convex.

Also, since $-\log u$ is convex, so $f_3(\boldsymbol{x}, u, v) = -\log u - \log\left(v - \frac{\boldsymbol{x}^T\boldsymbol{x}}{u}\right)$ is convex.

So above all, $f_3(\boldsymbol{x}, u, v)$ is convex.