Q1 IE 411 HW #3

Efe Eren Ceyani
21903359

a) $C_1 = \{\underline{x} \in \mathbb{R}^n \mid \|\underline{x}\|_2^2 = 1\}$

Counterexample: Consider $\underline{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\underline{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ for $n=2$. $\underline{x}_1, \underline{x}_2 \in C_1$.

Now, consider $\underline{x}_3 = \frac{1}{2}(\underline{x}_1 + \underline{x}_2)$, a convex combination of $\underline{x}_1$ & $\underline{x}_2$.

$$\|\underline{x}_3\|_2^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \neq 1 \Rightarrow \underline{x}_3 \notin C_1.$$

$\Rightarrow C_1$ is not convex.

b) $C_2 = \{\underline{x} \in \mathbb{R}^n \mid \max_{i \in \{1,\dots,n\}} x_i \leq 1\}$

Consider $\underline{x}, \underline{y} \in C_2$, s.t., $\begin{matrix} \max_{i \in \{1,\dots,n\}} x_i \leq 1 \\ \max_{i \in \{1,\dots,n\}} y_i \leq 1 \end{matrix}$. Now, consider $\underline{z} = \lambda \underline{x} + (1-\lambda)\underline{y}$

for $\lambda \in (0,1)$.

$$\begin{aligned} \max_{i \in \{1,\dots,n\}} z_i &= \max_{i \in \{1,\dots,n\}} (\lambda x_i + (1-\lambda) y_i) \\ &\overset{\text{max property}}{\leq} \max_{i \in \{1,\dots,n\}} (\lambda x_i) + \max_{i \in \{1,\dots,n\}} ((1-\lambda) y_i) \\ &= \lambda \max_{i \in \{1,\dots,n\}} x_i + (1-\lambda) \max_{i \in \{1,\dots,n\}} y_i \\ &\overset{\underline{x}, \underline{y} \in C_1}{\leq} \lambda + (1-\lambda) = 1 \end{aligned}$$

$\Rightarrow \max_{i \in \{1,\dots,n\}} z_i \leq 1 \Rightarrow \underline{z} \in C_2 \Rightarrow C_2$ is convex.

c) $C_3 = \{\underline{x} \in \mathbb{R}^n \mid \min_{i \in \{1,\dots,n\}} x_i \leq 1\}$

Counterexample: Consider $\underline{x}_1 = \begin{bmatrix} 99 \\ 1 \end{bmatrix}$, $\underline{x}_2 = \begin{bmatrix} 1 \\ 99 \end{bmatrix}$ for $n=2$. $\underline{x}_1, \underline{x}_2 \in C_3$.

Now consider $\underline{x}_3 = \frac{1}{2}(\underline{x}_1 + \underline{x}_2)$, a convex combination of $\underline{x}_1$ & $\underline{x}_2$.

$$\min_{i \in \{1,\dots,n\}} x_{3,i} = 50 > 1 \Rightarrow \underline{x}_3 \notin C_3$$

$\Rightarrow C_3$ is not convex.

d) $C_4 = \{\underline{x} \in \mathbb{R}^2_{++} \mid x_1 x_2 \geq 1\}$. Consider $\underline{x}, \underline{y} \in C_4$. Let $\underline{z} = \lambda \underline{x} + (1-\lambda)\underline{y}$ for $\lambda \in (0,1)$. If all components of $\underline{x}$ is greater than $\underline{y}$'s, i.e., $\underline{x} \succcurlyeq \underline{y}$, then $\underline{z} = \lambda \underline{x} + (1-\lambda)\underline{y} \succcurlyeq \underline{y} \Rightarrow z_1 z_2 \geq y_1 y_2 \geq 1$. Same result can be obtained $\underline{y} \succcurlyeq \underline{x}$ easily. However, if $\underline{x} \not\succcurlyeq \underline{y}$, i.e., one element of $\underline{x}$ is larger than $\underline{y}$'s element, and the other is smaller, then we cannot make a generalization easily.

Continued on next page.

$z_1 = \lambda x_1 + (1-\lambda) y_1$, $z_2 = \lambda x_2 + (1-\lambda) y_2$. Then,

$$
\begin{aligned}
z_1 z_2 &= (\lambda x_1 + (1-\lambda) y_1)(\lambda x_2 + (1-\lambda) y_2) \\
&= \lambda^2 x_1 x_2 + (1-\lambda)^2 y_1 y_2 + \lambda(1-\lambda)(x_1 y_2 + x_2 y_1) \\
&= (\lambda - \lambda(1-\lambda)) x_1 x_2 + ((1-\lambda) - \lambda(1-\lambda)) y_1 y_2 + \lambda(1-\lambda)(x_1 y_2 + x_2 y_1) \\
&= \lambda x_1 x_2 + (1-\lambda) y_1 y_2 - \lambda(1-\lambda)(x_1 x_2 + y_1 y_2 - x_1 y_2 - x_2 y_1) \\
&= \lambda x_1 x_2 + (1-\lambda) y_1 y_2 \underbrace{- \lambda(1-\lambda) \underbrace{(y_1 - x_1)(y_2 - x_2)}_{<0 \text{ because one of them is } >0 \text{ and other is } <0 \text{ due to } \underline{x} \not\geq \underline{y} \text{ assumption.}}}_{>0} \\
&\geq \lambda x_1 x_2 + (1-\lambda) y_1 y_2 \\
&\geq \lambda \cdot 1 + (1-\lambda) \cdot 1 \\
&= 1 \quad \Rightarrow z_1 z_2 \geq 1 \text{ for all cases}
\end{aligned}
$$

$\Rightarrow C_4$ is convex.

Q2 Using definitions, $\forall \underline{x}, \underline{y} \in C \subseteq \mathbb{R}^n$, $\lambda \in (0,1)$,

$$f(\lambda \underline{x} + (1-\lambda)\underline{y}) \leq \lambda f(\underline{x}) + (1-\lambda) f(\underline{y})$$

$$g(\lambda \underline{x} + (1-\lambda)\underline{y}) < \lambda g(\underline{x}) + (1-\lambda) g(\underline{y})$$

Let $h(\underline{x}) = f(\underline{x}) + g(\underline{x})$. Then, $\forall \underline{x}, \underline{y} \in C \subseteq \mathbb{R}^n$, $\lambda \in (0,1)$

$$
\begin{aligned}
h(\lambda \underline{x} + (1-\lambda)\underline{y}) &= f(\lambda \underline{x} + (1-\lambda)\underline{y}) + g(\lambda \underline{x} + (1-\lambda)\underline{y}) \\
&\leq \lambda f(\underline{x}) + (1-\lambda) f(\underline{y}) + g(\lambda \underline{x} + (1-\lambda)\underline{y}) \\
&< \lambda f(\underline{x}) + (1-\lambda) f(\underline{y}) + \lambda g(\underline{x}) + (1-\lambda) g(\underline{y}) \\
&= \lambda (f(\underline{x}) + g(\underline{x})) + (1-\lambda)(f(\underline{y}) + g(\underline{y})) \\
&= \lambda h(\underline{x}) + (1-\lambda) h(\underline{y})
\end{aligned}
$$

$\Rightarrow h = f + g$ is strictly convex over $C \subseteq \mathbb{R}^n$.

Q3 $g: \mathbb{R}_+ \to \mathbb{R}$, $g(x) = x^p$, $x \in \mathbb{R}_+$, $p > 1$.

First, notice that $\mathbb{R}_+$ is an open convex set. We can apply second order conditions.

$$g(x) = x^p, \quad g'(x) = p x^{p-1}, \quad g''(x) = \underbrace{p(p-1)}_{>0} \underbrace{x^{p-2}}_{>0}$$

$$g''(x) > 0, \ \forall x \in \mathbb{R}_+ \Rightarrow g \text{ is a convex function.}$$

Q4 $f: \mathbb{R}^n \to \mathbb{R}$, $f(\underline{x}) = \log\left(\sum_{i=1}^{n} e^{x_i}\right)$. Consider $\underline{x} = \underline{1}_n$, $\underline{y} = 2 \cdot \underline{1}_n$, where $\underline{1}_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_n$.

$$f(\underline{x}) = \log\left(\sum_{i=1}^{n} e^1\right) = \log(ne) = \log(n) + \log(e) = \log(n) + 1$$

$$f(\underline{y}) = \log\left(\sum_{i=1}^{n} e^2\right) = \log(ne^2) = \log(n) + \log(e^2) = \log(n) + 2$$

For $\lambda \in (0,1)$,

$$①: \lambda f(\underline{x}) + (1-\lambda) f(\underline{y}) = \lambda \log(n) + \lambda + (1-\lambda)\log(n) + (1-\lambda)2$$
$$= \log(n) + \lambda + 2(1-\lambda)$$

$$②: f(\lambda \underline{x} + (1-\lambda)\underline{y}) = \log\left(\sum_{i=1}^{n} e^{(\lambda + 2(1-\lambda))}\right) = \log(ne^{(\lambda + 2(1-\lambda))})$$
$$= \log(n) + \log(e^{\lambda + 2(1-\lambda)})$$
$$= \log(n) + \lambda + 2(1-\lambda)$$

Since expressions ① & ② are equal to each other, f cannot be a strictly convex function.

Q5

a) Consider $f(\underline{x}) = \|\underline{x}\|_2$, $\forall \underline{x}, \underline{y} \in \mathbb{R}^n$, $\lambda \in (0,1)$,

$$f(\lambda \underline{x} + (1-\lambda)\underline{y}) = \|\lambda \underline{x} + (1-\lambda)\underline{y}\|_2$$
$$\leq \|\lambda \underline{x}\|_2 + \|(1-\lambda)\underline{y}\|_2$$
$$= \lambda \|\underline{x}\|_2 + (1-\lambda)\|\underline{y}\|_2$$
$$= \lambda f(\underline{x}) + (1-\lambda) f(\underline{y}) \Rightarrow f(\underline{x}) = \|\underline{x}\|_2 \text{ is convex.}$$

Now, consider $f(\underline{x}) = \|\underline{x}\|_2^4$. Taking the fourth power is a non-decreasing function here because $\|\underline{x}\|_2 \geq 0$. Hence, $f(\underline{x}) = \|\underline{x}\|_2^4$ is also a convex function.

Proof: Let g be non-decreasing and convex. Then,

$$g(f(\lambda \underline{x} + (1-\lambda)\underline{y})) \leq g(\lambda f(\underline{x}) + (1-\lambda) f(\underline{y})) \leq \lambda g(f(\underline{x})) + (1-\lambda) g(f(\underline{y}))$$

Hence $g \circ f$ is also convex.

Another solution is more direct:

$$f(\lambda \underline{x} + (1-\lambda)\underline{y}) = \|\lambda \underline{x} + (1-\lambda)\underline{y}\|_2^4 \quad \lambda \in (0,1)$$
$$\leq (\|\lambda \underline{x}\|_2 + \|(1-\lambda)\underline{y}\|_2)^4$$
$$= (\lambda \|\underline{x}\|_2 + (1-\lambda)\|\underline{y}\|_2)^4 \leq \lambda \|\underline{x}\|_2^4 + (1-\lambda)\|\underline{y}\|_2^4 = \lambda f(\underline{x}) + (1-\lambda) f(\underline{y})$$

b) $f(\underline{x}) = (2x_1^2 + 3x_2^2)\left(\frac{x_1^2}{2} + \frac{x_2^2}{3}\right)$, $f: \mathbb{R}^2 \to \mathbb{R}$.

$$f(\underline{x}) = x_1^4 + x_2^4 + \frac{13x_1^2x_2^2}{6}$$

$$\nabla f(\underline{x}) = \begin{bmatrix} 4x_1^3 + \frac{13x_1x_2^2}{3} \\ 4x_2^3 + \frac{13x_1^2x_2}{3} \end{bmatrix}$$

$$\nabla^2 f(\underline{x}) = \begin{bmatrix} 12x_1^2 + \frac{13x_2^2}{3} & \frac{26x_1x_2}{3} \\ \frac{26x_1x_2}{3} & 12x_2^2 + \frac{13x_1^2}{3} \end{bmatrix}$$

Notice that leading determinants of $\nabla^2 f(\underline{x})$ are always non-negative; hence, by the second order condition, $f(\underline{x})$ is a convex function.

c) First, consider $g(\underline{x}) = \sqrt{x_1^2 + x_2^2 + 20x_3^2 - x_1x_2 - 4x_2x_3 + 1}$

$$g(\underline{x}) = \sqrt{x_1^2 + x_2^2 + 20x_3^2 - \tfrac{1}{2}x_1x_2 - \tfrac{1}{2}x_1x_2 - 2x_2x_3 - 2x_2x_3 + 1}$$

$$= \sqrt{x_1(x_1 - \tfrac{1}{2}x_2 + 0) + x_2(-\tfrac{1}{2}x_1 + x_2 - 2x_3) + x_3(0 - 2x_2 + 20x_3) + 1}$$

$= \sqrt{\underline{x}^T A \underline{x} + 1}$, where $A = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & -2 \\ 0 & -2 & 20 \end{bmatrix}$ which is positive definite from leading determinants.

Using Cholesky decomposition, $A = BB^T$, where B is a lower triangular matrix,

$$= \sqrt{\underline{x}^T B B^T \underline{x} + 1}$$

$\underline{y} = B^T\underline{x}$

$= \sqrt{\underline{y}^T\underline{y} + 1} = \sqrt{\|\underline{y}\|_2^2 + 1} = \|\hat{\underline{y}}\|_2$, where $\hat{\underline{y}} = \begin{bmatrix} \underline{y} \\ 1 \end{bmatrix}$

For $\underline{x}, \underline{y} \in \mathbb{R}^3$, $\lambda \in (0,1)$, and $\hat{\underline{x}} = \begin{bmatrix} \underline{x} \\ 1 \end{bmatrix}$, $\hat{\underline{y}} = \begin{bmatrix} \underline{y} \\ 1 \end{bmatrix}$

$$\begin{aligned} g(\lambda\underline{x} + (1-\lambda)\underline{y}) &= \|\lambda\hat{\underline{x}} + (1-\lambda)\hat{\underline{y}}\|_2 \\ &\leq \|\lambda\hat{\underline{x}}\|_2 + \|(1-\lambda)\hat{\underline{y}}\|_2 \\ &= \lambda\|\hat{\underline{x}}\|_2 + (1-\lambda)\|\hat{\underline{y}}\|_2 \\ &= \lambda g(\underline{x}) + (1-\lambda)g(\underline{y}) \end{aligned}$$

$\Rightarrow g(\underline{x})$ is a convex function.

Second, consider $h(\underline{x}) = (x_1^2 + x_2^2 + x_1 + x_2 + 2)^2$

$$h(\underline{x}) = (x_1^2 + x_2^2 + x_1 + x_2 + 2)^2$$
$$= \left((x_1 + \tfrac{1}{2})^2 + (x_2 + \tfrac{1}{2})^2 + \tfrac{3}{2} \right)^2$$

Let $\hat{\underline{x}} = \underline{x} + \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}$ (because we are in $\mathbb{R}^3$)

$$= \left(\hat{x}_1^2 + \hat{x}_2^2 + \tfrac{3}{2} \right)^2$$

Let $\tilde{\underline{x}} = \begin{bmatrix} \hat{\underline{x}} \\ \sqrt{\frac{3}{2}} \end{bmatrix}$, $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (From previous parts, we know that this is convex, but let's prove it again).

$$= \|A\tilde{\underline{x}}\|_2^4$$

Now, $\lambda \in (0,1)$, $\underline{x}, \underline{y} \in \mathbb{R}^3$, $h_2(\underline{x}) = \|A\tilde{\underline{x}}\|_2$

$$h_2(\lambda \underline{x} + (1-\lambda)\underline{y}) \underset{\lambda \in (0,1)}{=} \|\lambda A\tilde{\underline{x}} + (1-\lambda)A\tilde{\underline{y}}\|_2$$
$$\leq \lambda \|A\tilde{\underline{x}}\|_2 + (1-\lambda)\|A\tilde{\underline{y}}\|_2$$
$$= \lambda h(\underline{x}) + (1-\lambda) h(\underline{y})$$

$h_2(\underline{x}) = \|A\tilde{\underline{x}}\|_2$ is convex $\Rightarrow h(\underline{x}) = \|A\tilde{\underline{x}}\|_2^4$ is also convex because taking fourth power is convex (shown in part (a)).

Ultimately, $f(\underline{x}) = \max(g(\underline{x}), h(\underline{x}))$ is composed of two convex functions. We must show that maximum of two convex functions is also convex.

We know that for $\underline{x}, \underline{y} \in \mathbb{R}^3$ & $\lambda \in (0,1)$

$$g(\lambda \underline{x} + (1-\lambda)\underline{y}) \leq \lambda g(\underline{x}) + (1-\lambda) g(\underline{y})$$
$$h(\lambda \underline{x} + (1-\lambda)\underline{y}) \leq \lambda h(\underline{x}) + (1-\lambda) h(\underline{y})$$

$\Rightarrow$

$$f(\lambda \underline{x} + (1-\lambda)\underline{y}) = \max\{ g(\lambda \underline{x} + (1-\lambda)\underline{y}), h(\lambda \underline{x} + (1-\lambda)\underline{y}) \} \leq \max\{ \lambda g(\underline{x}) + (1-\lambda) g(\underline{y}), \lambda h(\underline{x}) + (1-\lambda) h(\underline{y}) \}$$
$$\leq \max\{ \lambda g(\underline{x}) + \lambda h(\underline{x}) \} + \max\{ (1-\lambda) g(\underline{y}) + (1-\lambda) h(\underline{y}) \}$$
$$= \lambda \max\{ g(\underline{x}), h(\underline{x}) \} + (1-\lambda) \max\{ g(\underline{y}), h(\underline{y}) \}$$
$$= \lambda f(\underline{x}) + (1-\lambda) f(\underline{y})$$

$\Rightarrow f$ is a convex function $\Rightarrow$ maximum of two convex functions is convex

$\Rightarrow f(\underline{x}) = \max\left\{ \sqrt{x_1^2 + x_2^2 + 20x_3^2 - x_1x_2 - 4x_2x_3 + 1}, (x_1^2 + x_2^2 + x_1 + x_2 + 2)^2 \right\}$ is a convex function over $\mathbb{R}^3$.