SI251 - Convex Optimization, Spring 2021
Homework 1

Due on Mar 25, 2021, before class

I. Convex Set

1. Describe the dual cone for each of the following cones.

(1) $K = \{0\}$. (5 points)

(2) $K = \mathbb{R}^2$. (5 points)

(3) $K = \{(x_1, x_2) \mid |x_1| \le x_2\}$. (5 points)

(4) $K = \{(x_1, x_2) \mid x_1 + x_2 = 0\}$. (5 points)

Solution:

(1) $K^* = \{y \mid yx \ge 0 \text{ for all } x \in K\} = \{y \mid 0 \ge 0\} = \mathbb{R}$.

(2) $K^* = \{\mathbf{0}\}$. To see this, we need to identify the values of $\boldsymbol{y} \in \mathbb{R}^2$ for which $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} \ge 0$ for all $\boldsymbol{x} \in \mathbb{R}^2$. But given any $\boldsymbol{y} \ne \mathbf{0}$, consider the choice $\boldsymbol{x} = -\boldsymbol{y}$, for which we have $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} = -\|\boldsymbol{y}\|_2^2 < 0$. So the only possible choice is $\boldsymbol{y} = \mathbf{0}$ (which indeed satisfies $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} \ge \mathbf{0}$ for all $\boldsymbol{x} \in \mathbb{R}^2$).

(3) $K^* = \{(x_1, x_2) \mid |x_1| \le x_2\}$. (This cone is self-dual.)

(4)

$$\begin{aligned} K^* &= \{(y_1, y_2) \mid x_1 y_1 + x_2 y_2 \ge 0 \text{ for all } \boldsymbol{x} \in K\} \\ &= \{(y_1, y_2) \mid x_1(y_1 - y_2) \ge 0 \text{ for all } x_1\} \\ &= \{(y_1, y_2) \mid y_1 = y_2\} \end{aligned} \tag{1}$$

2. *Hyperbolic sets.* Show that the *hyperbolic set* $\left\{\mathbf{x} \in \mathbb{R}_+^2 \mid x_1 x_2 \ge 1\right\}$ is convex. As a generalization, show that $\left\{\mathbf{x} \in \mathbb{R}_+^n \mid \prod_{i=1}^n x_i \ge 1\right\}$ is convex. *Hint.* If $a, b \ge 0$ and $0 \le \theta \le 1$, then $a^\theta b^{1-\theta} \le \theta a + (1-\theta) b$. (15 points)

Solution:

(1) Let $\mathbf{z} = \theta \mathbf{x} + (1-\theta)\mathbf{y}$ where $\mathbf{x}, \mathbf{y} \in \left\{\mathbf{x} \in \mathbb{R}_+^2 \mid x_1 x_2 \ge 1\right\}$ and $\theta \in [0,1]$, then $\mathbf{z} \in \mathbb{R}_+^2$. We need to prove that $(\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) \ge 1$. When $\mathbf{x} \succeq \mathbf{y}$, it means that $x_1 \ge y_1$ and $x_2 \ge y_2$, then $\theta x_1 + (1-\theta) y_1 \ge y_1$ and $\theta x_2 + (1-\theta) y_2 \ge y_2$,

$$\text{so } (\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) \ge y_1 y_2 \ge 1 \,.$$

When $\mathbf{y} \succeq \mathbf{x}$, it means that $y_1 \ge x_1$ and $y_2 \ge x_2$, then $\theta x_1 + (1-\theta) y_1 \ge x_1$ and $\theta x_2 + (1-\theta) y_2 \ge x_2$,

$$\text{so } (\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) \ge x_1 x_2 \ge 1 \,.$$

Besides, when $\mathbf{x} \not\succeq \mathbf{y}$, then $(y_1 - x_1)(y_2 - x_2) < 0$.

$$\begin{aligned} &(\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) \\ &= \theta x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta) x_1 y_2 + \theta(1-\theta) x_2 y_1 \\ &= \theta x_1 x_2 + (1-\theta) y_1 y_2 - \theta(1-\theta)(y_1 - x_1)(y_2 - x_2) \\ &\ge 1 \end{aligned}$$

(2) Let $\mathbf{z} = \theta \mathbf{x} + (1-\theta)\mathbf{y}$ where $\mathbf{x}, \mathbf{y} \in \left\{\mathbf{x} \in \mathbb{R}_+^n \mid \prod_{i=1}^n x_i \ge 1\right\}$ and $\theta \in [0,1]$, then $\mathbf{z} \in \mathbb{R}_+^n$. According to the hint, we have

$$\prod_{i=1}^n (z_i) = \prod_{i=1}^n (\theta x_i + (1-\theta) y_i) \ge \prod_{i=1}^n \left(x_i^\theta y_i^{1-\theta}\right) = \left(\prod_{i=1}^n x_i\right)^\theta \left(\prod_{i=1}^n y_i\right)^{1-\theta} \ge 1$$

So $\left\{\mathbf{x} \in \mathbb{R}_+^n \mid \prod_{i=1}^n x_i \ge 1\right\}$ is convex.

3. *Solution set of a quadratic inequality.* Let $C \subseteq \mathbb{R}^n$ be the solution set of a quadratic inequality,

$$C = \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + c \leq 0\right\},$$

with $\mathbf{A} \in \mathbb{S}^n, \mathbf{b} \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

(1) Show that C is convex if $\mathbf{A} \succeq 0$. (5 points)

(2) Show that the intersection of C and the hyperplane defined by $\mathbf{g}^{\mathrm{T}}\mathbf{x}+h=0$ (where $\mathbf{g} \neq \mathbf{0}$) is convex if $\mathbf{A} + \lambda\mathbf{g}\mathbf{g}^{\mathrm{T}} \succeq \mathbf{0}$ for some $\lambda \in \mathbb{R}$. (10 points)

Solution:

(1) As we know a set is convex if and only if its intersection with an arbitrary line $\{\hat{\mathbf{x}} + t\mathbf{v} \mid t \in \mathbb{R}\}$ is convex. Insert $\hat{\mathbf{x}} + t\mathbf{v}$ into $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + c \leq 0$

$$\begin{aligned}
&(\hat{\mathbf{x}} + t\mathbf{v})^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}} + t\mathbf{v} + \mathbf{b}^{\mathrm{T}}\hat{\mathbf{x}} + t\mathbf{v} + c \\
&= \left(\mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v}\right)t^2 + \left(\mathbf{b}^{\mathrm{T}}\mathbf{v} + 2\hat{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\mathbf{v}\right)t + c + \mathbf{b}^{\mathrm{T}}\hat{\mathbf{x}} + \hat{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}} \\
&= \alpha t^2 + \beta t + \gamma \leq 0
\end{aligned}$$

The intersection of C and the arbitrary line is the set defined as $\left\{\hat{\mathbf{x}} + t\mathbf{v} \mid \alpha t^2 + \beta t + \gamma \leq 0\right\}$. Since $\mathbf{A} \succeq 0$, then $\mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v} \geq 0$, so the above set is a continuous and closed line segment which is convex.

(2) The intersection of C and the hyperplane defined by $\mathbf{g}^{\mathrm{T}}\mathbf{x}+h=0$ (where $\mathbf{g} \neq \mathbf{0}$) is $\left\{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + c \leq 0 \mid \mathbf{g}^{\mathrm{T}}\mathbf{x} + h = 0\right\}$. Then we take a point $\hat{\mathbf{x}}$ in the above set, then $\mathbf{g}^{\mathrm{T}}\hat{\mathbf{x}}+h=0$. Insert an arbitrary line $\hat{\mathbf{x}}+t\mathbf{v}$ into above set: $I = \left\{\hat{\mathbf{x}} + t\mathbf{v} \mid \alpha t^2 + \beta t + \gamma \leq 0, \mathbf{g}^{\mathrm{T}}\mathbf{v}t = 0\right\}, \alpha = \mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v}, \beta = \mathbf{b}^{\mathrm{T}}\mathbf{v}+2\hat{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\mathbf{v}, \gamma = c + \mathbf{b}^{\mathrm{T}}\hat{\mathbf{x}} + \hat{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}}$. When $\mathbf{g}^{\mathrm{T}}\mathbf{v} \neq 0$, then $t = 0, I = \{\hat{\mathbf{x}} \mid \gamma \leq 0\}$. I is convex whether γ is greater than 0 or not.

When $\mathbf{g}^{\mathrm{T}}\mathbf{v} = 0, I = \left\{\hat{\mathbf{x}} + t\mathbf{v} \mid \alpha t^2 + \beta t + \gamma \leq 0\right\}$. Since $\mathbf{v}^{\mathrm{T}}\left(\mathbf{A} + \lambda\mathbf{g}\mathbf{g}^{\mathrm{T}}\right)\mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v} + \lambda\mathbf{v}^{\mathrm{T}}\mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{v} =$

$\mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v} = \alpha \succeq 0, I$ is convex. According to the fact that a set is convex if and only if its intersection with an arbitrary line is convex, the intersection of C and the hyperplane defined by $\mathbf{g}^{\mathrm{T}}\mathbf{x} + h = 0$ (where $\mathbf{g} \neq \mathbf{0}$) is convex.

II. Convex Function

1. Determine the convexity (i.e., convex, concave, or neither) of the following functions.

(1) $f(x_1, x_2) = 1/(x_1 x_2)$ on $\mathbb{R}^2_{++}$ (5 points)

(2) $f(x_1, x_2) = x_1/x_2$ on $\mathbb{R}^2_{++}$. (5 points)

(3) $f(x_1, x_2) = x_1^2/x_2$ on $\mathbb{R} \times \mathbb{R}_{++}$. (5 points)

(4) $f(x_1, x_2) = \sqrt{x_1 x_2}$ on $\mathbb{R}^2_{++}$. (5 points)

Solution:

(1)

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{bmatrix} \succeq \mathbf{0}.$$

The Hessian of f is positive semidefinite, hence f is <u>convex</u>.

(2)

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}.$$

The Hessian of f is indefinite, hence f is <u>neither convex nor concave</u>.

(3) Method 1:

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2^2} \\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix} \succeq \mathbf{0}.$$

The Hessian of f is positive semidefinite, hence f is <u>convex</u>.
Method 2: The f is quadratic-over-linear function, and hence is <u>convex</u>.

(4)

$$\nabla^2 f(x_1,x_2) = -\frac{\sqrt{x_1x_2}}{4}\begin{bmatrix} \frac{1}{x_1^2} & -\frac{1}{x_1x_2} \\ -\frac{1}{x_1x_2} & \frac{1}{x_2^2} \end{bmatrix} \preceq \mathbf{0}.$$

The Hessian of f is negative semidefinite, hence f is concave.

2. *Convex hull of functions.* Suppose g and h are convex functions, bounded below, with $\mathbf{dom}\ g = \mathbf{dom}\ h = \mathbb{R}^n$. The convex hull function of g and h is defined as

$$f(\mathbf{x}) = \inf\{\theta g(\mathbf{y}) + (1-\theta)h(\mathbf{z}) \mid \theta\mathbf{y} + (1-\theta)\mathbf{z} = \mathbf{x},\ 0 \le \theta \le 1\},$$

where the infimum is over θ, $\mathbf{y}$, $\mathbf{z}$. Show that the convex hull of h and g is convex. Describe $\mathbf{epi}\ f$ in terms of $\mathbf{epi}\ g$ and $\mathbf{epi}\ h$. (10 points)

Solution:

Since g and h are convex functions, $\mathbf{epi}\ g$ and $\mathbf{epi}\ h$ are convex set.

$$\begin{aligned} \mathbf{epi}\ g =& \{(\mathbf{y}, t_1) \in \mathbb{R}^{n+1} \mid \mathbf{y} \in \mathbf{dom}\ g,\ g(\mathbf{y}) \le t_1\}, \\ \mathbf{epi}\ h =& \{(\mathbf{z}, t_2) \in \mathbb{R}^{n+1} \mid \mathbf{z} \in \mathbf{dom}\ h,\ h(\mathbf{z}) \le t_2\}. \end{aligned}$$

For $0 \le \theta \le 1$ and $\theta\mathbf{y} + (1-\theta)\mathbf{z} = \mathbf{x}$, we have

$$\theta g(\mathbf{y}) + (1-\theta)h(\mathbf{z}) \le \theta t_1 + (1-\theta)t_2,$$

i.e., $f(\mathbf{x}) \le t$, where $t = \theta t_1 + (1-\theta)t_2$. Thus

$$\mathbf{epi}\ f = \mathbf{conv}\ (\mathbf{epi}\ g \cup \mathbf{epi}\ h),$$

i.e., $\mathbf{epi}\ f$ is the convex hull of the union of the epigraphs of g and h. This shows that f is convex.

3. Show that the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are convex.

(1) The difference between the maximum and minimum value of a polynomial on a given interval, as a function of its coefficients:

$$f(\mathbf{x}) = \sup_{t\in[a,b]} p(t) - \inf_{t\in[a,b]} p(t), \text{ where } p(t) = x_1 + x_2 t + x_3 t^2 + \cdots + x_n t^{n-1}.$$

a, b are real constants with $a < b$. (5 points)

Solution:

The $\sup_{t\in[a,b]} p(t)$ is the supremum of a family of linear functions of x, so it is convex. The $\inf_{t\in[a,b]} p(t)$ is the infimum of a family of linear functions, so it is concave. Therefore, f is the difference of a convex and a concave function, which is convex.

(2) The 'exponential barrier' of a set of inequalities:

$$f(\mathbf{x}) = \sum_{i=1}^{m} e^{-1/f_i(\mathbf{x})},\ \mathbf{dom}\ f = \{\mathbf{x} \mid f_i(\mathbf{x}) < 0, i = 1, \ldots, m\}.$$

The functions $f_i(\mathbf{x})$ are convex. (5 points)

Solution:

$h(u) = e^{1/u}$ is convex and decreasing on $\mathbb{R}_{++}$:

$$h' = -\frac{1}{u^2}e^{1/u}, \quad h''(u) = \frac{2}{u^3}e^{1/u} + \frac{1}{u^4}e^{1/u}.$$

Therefore, the composition $h(-f_i(\mathbf{x})) = e^{-1/f_i(\mathbf{x})}$ is convex since $f_i(\mathbf{x})$ is convex. Since the sum of convex functions is still convex, $f(\mathbf{x})$ is convex.

(3) The function

$$f(\mathbf{x}) = \inf_{\alpha>0} \frac{g(\mathbf{y} + \alpha\mathbf{x}) - g(\mathbf{y})}{\alpha}$$

if g is convex and $\mathbf{y} \in \mathbf{dom}\ g$. (It can be shown that this is the directional derivative of g at $\mathbf{y}$ in the direction $\mathbf{x}$.) (10 points)

Solution:

Method 1: The original problem can be written as

$$f(\mathbf{x}) = \inf_{t>0} t \left(g(\mathbf{y} + \frac{1}{t}\mathbf{x}) - g(\mathbf{y}) \right),$$

which is the infimum over t of the perspective of the convex function $g(\mathbf{y}+\mathbf{x}) - g(\mathbf{y})$. Therefore, $f(\mathbf{x})$ is convex.

Method 2:

$$\frac{\partial}{\partial \alpha} \frac{g(\mathbf{y}+\alpha\mathbf{x}) - g(\mathbf{y})}{\alpha} = \frac{\alpha \nabla g(\mathbf{y}+\alpha\mathbf{x})^{\mathrm{T}}\mathbf{x} - g(\mathbf{y}+\alpha\mathbf{x}) + g(\mathbf{y})}{\alpha^2}.$$

Since g is convex, we have

$$g(\mathbf{y}) \geq g(\mathbf{y}+\alpha\mathbf{x}) - \alpha \nabla g(\mathbf{y}+\alpha\mathbf{x})^{\mathrm{T}}\mathbf{x}.$$

Therefore, we can get

$$\frac{\partial}{\partial \alpha} \frac{g(\mathbf{y}+\alpha\mathbf{x}) - g(\mathbf{y})}{\alpha} \geq 0.$$

$\frac{g(\mathbf{y}+\alpha\mathbf{x}) - g(\mathbf{y})}{\alpha}$ is the increasing function with respect to α. Therefore, we have

$$\begin{aligned} f(\mathbf{x}) =& \inf_{\alpha>0} \frac{g(\mathbf{y}+\alpha\mathbf{x}) - g(\mathbf{y})}{\alpha} \\ =& \lim_{\alpha\to 0} \frac{g(\mathbf{y}+\alpha\mathbf{x}) - g(\mathbf{y})}{\alpha} \\ =& \nabla g(\mathbf{y})^{\mathrm{T}}\mathbf{x}. \end{aligned}$$

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