DATA130026.01 Optimization
Solution of Assignment 1

1. Use the definition of convex function ($\lambda f(x)+(1-\lambda)f(y) \geq f(\lambda x+(1-\lambda)y)$) to show that the quadratic function $\frac{1}{2}x^T Ax + b^T x + c$ is convex if and only if $A \succeq 0$.

   Solution. $\forall x, y$ and $0 \leq \lambda \leq 1$,

$$
\begin{aligned}
& f(x) = \frac{1}{2}x^T Ax + b^T x + c \ \text{ is convex} \\
\Leftrightarrow \quad & \lambda f(x) + (1-\lambda)f(y) \geq f\big(\lambda x + (1-\lambda)y\big) \\
\Leftrightarrow \quad & \lambda(\frac{1}{2}x^T Ax + b^T x + c) + (1-\lambda)(\frac{1}{2}y^T Ay + b^T y + c) \geq \\
& \frac{1}{2}\big(\lambda x + (1-\lambda)y\big)^T A\big(\lambda x + (1-\lambda)y\big) + b^T\big(\lambda x + (1-\lambda)y\big) + c \\
\Leftrightarrow \quad & \lambda(1-\lambda)x^T Ax + \lambda(1-\lambda)y^T Ay - \lambda(1-\lambda)x^T Ay - \lambda(1-\lambda)y^T Ax \geq 0 \\
\Leftrightarrow \quad & x^T Ax + y^T Ay - x^T Ay - y^T Ax \geq 0 \\
\Leftrightarrow \quad & (x-y)^T A(x-y) \geq 0 \\
\Leftrightarrow \quad & A \succeq 0
\end{aligned}
$$

2. Show that both the second order cone, i.e., $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\|_2 \leq t\}$, and the semidefinite cone, i.e., $\{Z \in \mathbb{S}^n : Z \succeq 0\}$, are convex cones.

   Solution.

   (a) For any $(x_1, t_1), (x_2, t_2) \in \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\|_2 \leq t\}$ and $0 \leq \theta \leq 1$, we have

$$\|\theta x_1 + (1-\theta)x_2\|_2 \leq \theta\|x_1\|_2 + (1-\theta)\|x_2\|_2 \leq \theta t_1 + (1-\theta)t_2,$$

   (The first inequality is the triangle inequality, and the second inequality follows from the definition of the second order cone.) so that $\big(\theta x_1 + (1-\theta)x_2, \theta t_1 + (1-\theta)t_2\big) \in \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\|_2 \leq t\}$, which means the second order cone is convex.

   (b) For any $A, B \in \{Z \in \mathbb{S}^n : Z \succeq 0\}$ and $0 \leq \theta \leq 1$, since $\forall x \in \mathbb{R}^n$, $x^T Ax \geq 0, x^T Bx \geq 0$, we obtain

$$x^T\big(\theta A + (1-\theta)B\big)x = \theta x^T Ax + (1-\theta)x^T Bx \geq 0.$$

   So we have $\theta A + (1-\theta)B \in \{Z \in \mathbb{S}^n : Z \succeq 0\}$, which means the semidefinite cone is convex.

3. Let $a, b \in \mathbb{R}^n (a \neq b)$. For what values of μ ($\mu > 0$) is the set

$$S_\mu = \{x \in \mathbb{R}^n : \|x-a\|_2 \leq \mu\|x-b\|_2\}$$

   convex?

   Solution. Since $\mu > 0$, we have

$$
\begin{aligned}
S_\mu &= \{x \in \mathbb{R}^n : \|x-a\|_2 \leq \mu\|x-b\|_2\} \\
&= \{x \in \mathbb{R}^n : \|x-a\|_2^2 \leq \mu^2\|x-b\|_2^2\} \\
&= \{x \in \mathbb{R}^n : (1-\mu^2)x^T x - 2(a-\mu^2 b)^T x + (a^T a - \mu^2 b^T b) \leq 0\}.
\end{aligned}
$$

(a) If $\mu = 1$, it is a halfspace, which is convex;

(b) If $0 < \mu < 1$, it is a ball

$$\{x \in \mathbb{R}^n : (x - x_0)^T (x - x_0) \le R^2\},$$

with center x_0 and radius R given by

$$x_0 = \frac{a - \mu^2 b}{1 - \mu^2}, \quad R = \frac{\mu}{1 - \mu^2}\|a - b\|_2,$$

which is convex;

(c) If $\mu > 1$, it is the whole space outside the ball

$$\{x \in \mathbb{R}^n : (x - x_0)^T (x - x_0) \ge R^2\},$$

(the ball also has center x_0 and radius R,) which is not convex.

In summary, when $0 < \mu \le 1$, the set S_μ is convex.

4. Let $C \in \mathbb{R}^n$ be a nonempty convex set. For each $x \in C$ define the normal cone of C at x by

$$N_C(x) = \{w \in \mathbb{R}^n : w^T(y - x) \le 0 \text{ for all } y \in C\},$$

and define $N_C(x) = \varnothing$ when $x \notin C$. Show that $N_C(x)$ is convex and closed. Particularly, when $x \in \text{int}(C)$, we have $N_C(x) = \{0\}$.

Solution.

(a) Show that $N_C(x)$ is convex.
Method 1 Since

$$\begin{aligned} N_C(x) &= \{w \in \mathbb{R}^n : w^T(y - x) \le 0 \text{ for all } y \in C\} \\ &= \bigcap_{y \in C} \{w \in \mathbb{R}^n : w^T(y - x) \le 0\}, \end{aligned}$$

which means $N_C(x)$ can be expressed as the intersection of convex half spaces of $\mathbb{R}^n$, so that $N_C(x)$ is convex.
Method 2 For any $w_1, w_2 \in N_C(x)$ and $y \in C$, we have

$$w_1^T(y - x) \le 0, \quad w_2^T(y - x) \le 0.$$

Then $\forall \lambda \in [0, 1]$,

$$\begin{aligned} (\lambda w_1 + (1 - \lambda) w_2)^T (y - x) &= \lambda w_1^T(y - x) + (1 - \lambda) w_2^T(y - x) \\ &\le 0. \end{aligned}$$

Therefore $\lambda w_1 + (1 - \lambda) w_2 \in N_C(x)$, which means $N_C(x)$ is convex.

(b) Show that $N_C(x)$ is closed.
Method 1 Since

$$\begin{aligned} N_C(x) &= \{w \in \mathbb{R}^n : w^T(y-x) \le 0 \text{ for all } y \in C\} \\ &= \bigcap_{y \in C} \{w \in \mathbb{R}^n : w^T(y-x) \le 0\}, \end{aligned}$$

which means $N_C(x)$ can be expressed as the intersection of closed half spaces of $\mathbb{R}^n$, so that $N_C(x)$ is closed.
Method 2 For any w_0, if $\exists \{w_i\}_{i=1}^{\infty} \in N_C(x)$ s.t. $w_0 = \lim_{i\to\infty} w_i$, then $\forall y \in C$, we have

$$\begin{aligned} & w_i^T(y-x) \le 0, \qquad \forall i \in \mathbb{N}_+ \\ \Rightarrow \quad & w_0^T(y-x) = \lim_{i\to\infty} w_i^T(y-x) \le 0 \\ \Rightarrow \quad & w_0 \in N_C(x). \end{aligned}$$

Therefore $N_C(x)$ contains all of its limit points, which means it is closed.

(c) Show that $N_C(x) = \{0\}$ when $x \in \operatorname{int}(C)$.
Proof by contradiction. Obviously $\mathbf{0} \in N_C(x)$, now suppose that $\exists\ \widetilde{w} \neq \mathbf{0}$ s.t. $\widetilde{w} \in N_C(x)$.
Since $x \in \operatorname{int}(C)$, there exists $\delta > 0$ s.t. $O(x,\delta) \subset C$. Let $y = x + \frac{\delta}{2}\frac{\widetilde{w}}{\|\widetilde{w}\|_2} \in O(x,\delta) \subset C$, we have

$$\widetilde{w}^T(y-x) = \frac{\delta}{2}\frac{\widetilde{w}^T\widetilde{w}}{\|\widetilde{w}\|_2} = \frac{\delta}{2}\|\widetilde{w}\|_2 > 0,$$

This leads to a contradiction. Therefore, $N_C(x)$ contains only one element $\mathbf{0}$.

5. *Supporting hyperplanes.*

(a) Express the closed convex set $\{\boldsymbol{x} \in \mathbb{R}_+^n \mid x_1x_2 \ge 1\}$ as an intersection of halfspaces.

(b) Let $C = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\|_\infty \le 1\}$, the l_∞ norm unit $\mathbb{R}^n$ and let $\hat{\boldsymbol{x}}$ be a point in the boundary of C. Identify the supporting hyperplanes of C at $\hat{\boldsymbol{x}}$ explicitly.

Solution.

(a) For any point $\hat{\boldsymbol{x}} := \left(t, \frac{1}{t}, \hat{x_3}, \ldots, \hat{x_n}\right)$ in the boundary of $\{\boldsymbol{x} \in \mathbb{R}_+^n \mid x_1x_2 \ge 1\}$, the supporting hyperplane can be expressed as $\boldsymbol{a}^T(\boldsymbol{x} - \hat{\boldsymbol{x}}) = 0$, where

$$\boldsymbol{a} = \left(\frac{1}{t}, t, 0, \ldots, 0\right)$$

is the normal vector of the tangent plane at $\hat{\boldsymbol{x}}$. By supporting hyperplane theorem, the halfspace satisfies

$$\begin{aligned} & \boldsymbol{a}^T(\boldsymbol{x} - \hat{\boldsymbol{x}}) \ge 0 \\ \Rightarrow\ & \frac{1}{t}(x_1 - t) + t(x_2 - \frac{1}{t}) \ge 0 \\ \Rightarrow\ & x_1 + t^2 x_2 \ge 2t. \end{aligned}$$

Obviously $t > 0$, so the closed convex set can be expressed as

$$\bigcap_{t>0} \{\boldsymbol{x} \in \mathbb{R}^n \mid x_1 + t^2 x_2 \geq 2t\}.$$

(b) Denote $\boldsymbol{a} := (a_1, a_2, \ldots, a_n)$ and $\hat{\boldsymbol{x}} := (\hat{x_1}, \hat{x_2}, \ldots, \hat{x_n})$, then the supporting hyperplanes of C at $\hat{x}$ can be expressed as

$$\boldsymbol{a}^T(\boldsymbol{x} - \hat{\boldsymbol{x}}) = 0 \qquad (\boldsymbol{a} \neq \boldsymbol{0}).$$

By supporting hyperplane theorem, for any $\boldsymbol{x}$, the halfspace satisfies

$$\begin{aligned} &\boldsymbol{a}^T(\boldsymbol{x} - \hat{\boldsymbol{x}}) \leq 0 \\ \Leftrightarrow\ &\sum_{i=1}^{n} a_i(x_i - \hat{x}_i) \leq 0 \\ \Leftrightarrow\ &a_i(x_i - \hat{x}_i) \leq 0, \quad \forall i. \end{aligned}$$

Since $\|x\|_\infty \leq 1$,

i. When $\hat{x}_i = 1$, we have $x_i - \hat{x}_i \leq 0$, which means $a_i \geq 0$.
ii. When $\hat{x}_i = -1$, we have $x_i - \hat{x}_i \geq 0$, which means $a_i \leq 0$.
iii. When $-1 < \hat{x}_i < 1$, the sign of $x_i - \hat{x}_i$ is uncertain, which means $a_i = 0$.

Then the supporting hyperplanes of C at $\hat{\boldsymbol{x}}$ can be expressed as

$$\boldsymbol{a}^T(\boldsymbol{x} - \hat{\boldsymbol{x}}) = 0 \qquad (\boldsymbol{a} \neq \boldsymbol{0}),$$

where

$$\begin{cases} a_i \geq 0, & \text{if } \hat{x}_i = 1 \\ a_i \leq 0, & \text{if } \hat{x}_i = -1 \\ a_i = 0, & \text{if } \hat{x}_i \in (-1, 1). \end{cases}$$