**ENGG 5501: Foundations of Optimization** **2021–22 First Term**

# Homework Set 3 Solution

Instructor: Anthony Man–Cho So October 26, 2021

**Problem 1 (20pts).**

(a) **(5pts).** Let $x \in C$ be arbitrary. Then, we have $\mathbb{I}_C(x) = 0$. By definition, we have $s \in \partial \mathbb{I}_C(x)$ iff

$$\mathbb{I}_C(y) \geq \mathbb{I}_C(x) + s^T(y-x) = s^T(y-x) \quad \text{for all } y \in \mathbb{R}^n.$$

Note that if $y \notin C$, then $\mathbb{I}_C(y) = +\infty$, and the above inequality holds automatically. On the other hand, if $y \in C$, then $\mathbb{I}_C(y) = 0$, and the above inequality reduces to $s^T(y-x) \leq 0$. It follows that

$$\partial \mathbb{I}_C(x) = \{s \in \mathbb{R}^n : s^T(y-x) \leq 0 \text{ for all } y \in C\}.$$

(b) **(15pts).** Let $x \in \mathbb{R}^n_+$ be given. By the result in (a), we have

$$\partial \mathbb{I}_{\mathbb{R}^n_+}(x) = \left\{ s \in \mathbb{R}^n : s^T(y-x) \leq 0 \text{ for all } y \in \mathbb{R}^n_+ \right\}.$$

Let $I = \{i \in \{1,\ldots,n\} : x_i = 0\}$ be the active index set associated with $x$. By definition, we have $x_i > 0$ whenever $i \notin I$. Suppose that $s \in \mathbb{R}^n$ satisfies $s \leq \mathbf{0}$ and $x^T s = 0$. Since

$$0 = x^T s = \sum_{i \in I} x_i s_i + \sum_{i \notin I} x_i s_i = \sum_{i \notin I} x_i s_i,$$

we have $s_i = 0$ whenever $i \notin I$. This implies that

$$s^T(y-x) = \sum_{i \in I} s_i(y_i - x_i) + \sum_{i \notin I} s_i(y_i - x_i) = \sum_{i \in I} s_i y_i \leq 0$$

for all $y \in \mathbb{R}^n_+$; i.e., $s \in \partial \mathbb{I}_{\mathbb{R}^n_+}(x)$. Conversely, suppose that $s \in \partial \mathbb{I}_{\mathbb{R}^n_+}(x)$; i.e., $s^T(y-x) \leq 0$ for all $y \in \mathbb{R}^n_+$. Observe that for $i = 1,\ldots,n$, if we take $y = x + e_i \geq \mathbf{0}$ (here, $e_i$ is the $i$-th basis vector), then we get $0 \geq s^T(y-x) = s_i$. This shows that $s \leq \mathbf{0}$. Moreover, if we take $y = \mathbf{0}$, then we get $x^T s \geq 0$. Since $x \geq \mathbf{0}$ and $s \leq \mathbf{0}$, this implies that $x^T s = 0$. It follows that $s \in \mathbb{R}^n$ satisfies $s \leq \mathbf{0}$ and $x^T s = 0$, as desired.

**Problem 2 (20pts).**

(a) **(10pts).** Let $x, y \in \mathbb{R}^2$ and $\alpha \in (0,1)$ be arbitrary with $x \neq y$. Consider the following cases:

<u>Case 1</u>: $\|x\|_2 \leq 1$ and $\|y\|_2 \leq 1$.

In this case, since $x \neq y$ and $\alpha \in (0,1)$, we have

$$\|\alpha x + (1-\alpha)y\|_2^2 = \alpha\|x\|_2^2 + (1-\alpha)\|y\|_2^2 - \alpha(1-\alpha)\|x-y\|_2^2 < 1.$$

It follows that $f(\alpha x + (1-\alpha)y) = 0 \leq \alpha f(x) + (1-\alpha) f(y)$.

<u>Case 2</u>: $\|x\|_2 > 1$.

In this case, since $f(x) = +\infty$, we trivially have $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$.

Since $x$ and $y$ are interchangeable, we conclude from the above that $f$ is convex, as desired.

(b) **(10pts).** Let us take $f : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ to be the function given in the problem and specify

$$f(x_1, x_2) = \begin{cases} 0 & \text{if } \|(x_1, x_2)\|_2 = 1 \text{ and } x_2 < 0, \\ +\infty & \text{if } \|(x_1, x_2)\|_2 = 1 \text{ and } x_2 \geq 0. \end{cases}$$

We claim that this furnishes an example of a convex function whose epigraph is not closed. Indeed, observe that

$$\begin{aligned} \text{epi}(f) &= \{(x, t) \in \mathbb{R}^2 \times \mathbb{R} : f(x) \leq t\} \\ &= \{(x, t) \in \mathbb{R}^2 \times \mathbb{R} : \|x\|_2 < 1,\, t \geq 0\} \cup \{(x, t) \in \mathbb{R}^2 \times \mathbb{R} : \|x\|_2 = 1,\, x_2 < 0,\, t \geq 0\}. \end{aligned}$$

In particular, we see that the sequence $\{(0, 1 - 1/k, 0)\}_{k \geq 1}$ belongs to epi$(f)$, but its limit $(0, 1, 0)$ does not belong to epi$(f)$.

**Problem 3 (15pts).** It is clear that $f$ is differentiable at any $x \in \mathbb{R} \setminus \{\pm 1\}$. Now, note that

$$\lim_{t \to 0^+} \frac{f(1+t) - f(1)}{t} = \lim_{t \to 0^+} \frac{\frac{1}{2}(1+t)^2 - |1+t| - (-\frac{1}{2})}{t} = 0$$

and

$$\lim_{t \to 0^-} \frac{f(1+t) - f(1)}{t} = \lim_{t \to 0^-} \frac{-\frac{1}{2} - (-\frac{1}{2})}{t} = 0.$$

It follows that $f$ is differentiable at $x = 1$. A similar argument shows that $f$ is differentiable at $x = -1$. Hence, the function $f$ is in fact differentiable on $\mathbb{R}$. Moreover, it is straightforward to verify that $f$ attains its minimum at $x \in \{\pm 1\}$. This implies that $f(x) \geq -\frac{1}{2}$ for all $x \in \mathbb{R}$.

By Theorem 9 of Handout 2, it remains to show that $f(y) \geq f(x) + f'(x)(y - x)$ for all $x, y \in \mathbb{R}$. We consider two cases:

Case 1: $|x| \leq 1$.
Since $f(x) = -\frac{1}{2}$, $f'(x) = 0$, and $f(y) \geq -\frac{1}{2}$ for all $y \in \mathbb{R}$, we have $f(y) \geq f(x) + f'(x)(y - x)$ for all $x$ satisfying $|x| \leq 1$ and $y \in \mathbb{R}$.

Case 2: $|x| > 1$.
It is easy to verify that $f'(x) = x - \text{sgn}(x)$, where

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{otherwise.} \end{cases}$$

Now, for any $y \in \mathbb{R}$, we have

$$f(x) + f'(x)(y - x) = \frac{1}{2}x^2 - |x| + (x - \text{sgn}(x))(y - x) = -\frac{1}{2}x^2 + (x - \text{sgn}(x))y,$$

where we use the fact that $|x| = \text{sgn}(x) \cdot x$. Consider the following sub-cases:

Case 2a: $|y| \leq 1$.
We have

$$-\frac{1}{2}x^2 + (x - \text{sgn}(x))y \leq -\frac{1}{2}x^2 + |x| - 1 = -\frac{1}{2}(|x| - 1)^2 - \frac{1}{2} \leq -\frac{1}{2} = f(y).$$

Case 2b: $|y| > 1$.

We compute

$$-\frac{1}{2}x^2 + (x - \operatorname{sgn}(x))y = -\frac{1}{2}x^2 + xy - \frac{1}{2}y^2 + \left(\frac{1}{2}y^2 - |y|\right) + |y| - \operatorname{sgn}(x) \cdot y. \tag{1}$$

If $\operatorname{sgn}(x) = \operatorname{sgn}(y)$, then $|y| - \operatorname{sgn}(x) \cdot y = 0$, which implies that

$$-\frac{1}{2}x^2 + (x - \operatorname{sgn}(x))y = -\frac{1}{2}(x-y)^2 + \frac{1}{2}y^2 - |y| \le \frac{1}{2}y^2 - |y| = f(y).$$

Otherwise, we have $|y| - \operatorname{sgn}(x) \cdot y = 2|y|$. Since $|x| > 1$ and $xy < 0$, we also have $xy < -|y|$. It follows from (1) that

$$-\frac{1}{2}x^2 + (x - \operatorname{sgn}(x))y < \left(-\frac{1}{2} + |y| - \frac{1}{2}y^2\right) + \left(\frac{1}{2}y^2 - |y|\right) = -\frac{1}{2}(|y|-1)^2 + \frac{1}{2}y^2 - |y| \le f(y).$$

In summary, Case 2 shows that $f(y) \ge f(x) + f'(x)(y-x)$ for all $x$ satisfying $|x| > 1$ and $y \in \mathbb{R}$.

Combining the conclusions of Cases 1 and 2, we conclude that $f$ is convex.

**Problem 4 (15pts).** Yes. By assumption, the polyhedron $P$ contains the constraints

$$\begin{cases} x_i & \ge & 0 & \text{for } i \in I, \\ x_i & \le & 0 & \text{for } i \notin I, \end{cases}$$

where $I \subseteq \{1, \ldots, n\}$. We claim that $P$ does not contain a line, which would then imply the desired conclusion. Suppose that this is not the case. Then, there exist $x_0 \in P$ and $d \ne \mathbf{0}$ such that $x_0 + \alpha d \in P$ for all $\alpha \in \mathbb{R}$. Let $j \in \{1, \ldots, n\}$ be such that $d_j \ne 0$. If $j \in I$, then $(x_0 + \alpha d)_j < 0$ as $\alpha \searrow -\infty$, which contradicts the hypothesis that $(x_0 + \alpha d)_j \ge 0$ for all $\alpha \in \mathbb{R}$. One can derive a similar contradiction for the case where $j \notin I$. Hence, the claim is established.

**Problem 5 (30pts).**

(a) **(15pts).** First, we show that (I) and (II) cannot be simultaneously solvable. Suppose to the contrary that $\bar{x}$ (resp. $\bar{y}$) solves (I) (resp. (II)). Then, on one hand, we have $\bar{y}^T A\bar{x} \ge 0$ because $\bar{x} \ge \mathbf{0}$ and $A^T\bar{y} \ge \mathbf{0}$. On the other hand, we have $\bar{y}^T A\bar{x} < 0$ because $A\bar{x} < \mathbf{0}$ and $\mathbf{0} \ne \bar{y} \ge \mathbf{0}$. This results in a contradiction.

Next, observe that (I) is solvable iff the system

$$(\mathrm{I}') \quad Ax + s = -e, \;\; (x, s) \ge \mathbf{0}$$

is solvable. Indeed, if $(\bar{x}, \bar{s})$ is a solution to (I'), then $A\bar{x} \le -e < \mathbf{0}$ and $\bar{x} \ge \mathbf{0}$, which implies that $\bar{x}$ is a solution to (I). Conversely, if $\bar{x}$ is a solution to (I), then there exists a $\theta > 0$ such that $A\bar{x} \le -\theta e$. By letting $\tilde{x} = \bar{x}/\theta$ and $\tilde{s} = -e - A\tilde{x}$, we have $A\tilde{x} + \tilde{s} = -e$ and $(\tilde{x}, \tilde{s}) \ge \mathbf{0}$, which implies that $(\tilde{x}, \tilde{s})$ is a solution to (I').

Now, if (I') is not solvable, then by Farkas' lemma, the system

$$(\mathrm{II}') \quad A^T w \le \mathbf{0}, \;\; w \le \mathbf{0}, \;\; -e^T w > 0$$

is solvable. However, it is clear that if $\bar{w}$ is a solution to (II'), then $\bar{y} = -\bar{w}$ is a solution to (II). This completes the proof.

(b) **(15pts).** Again, we show that (I) and (II) cannot be simultaneously solvable. Suppose to the contrary that $\bar{x}$ (resp. $\bar{y}$) solves (I) (resp. (II)). Then, on one hand, we have $\bar{x}^T A^T \bar{y} = 0$ because $A^T \bar{y} = \mathbf{0}$. On the other hand, we have $\bar{x}^T A^T \bar{y} > 0$ because $A\bar{x} \geq \mathbf{0}$, $A\bar{x} \neq \mathbf{0}$, and $\bar{y} > \mathbf{0}$. This results in a contradiction.

Next, observe that (I) is solvable iff the system

$$(\mathrm{I}') \quad Ax \geq \mathbf{0}, \;\; e^T Ax > 0$$

is solvable. Indeed, if $\bar{x}$ is a solution to (I), then there exists an $i \in \{1, \ldots, m\}$ such that $(A\bar{x})_i > 0$. Since $A\bar{x} \geq \mathbf{0}$, we see that $e^T A\bar{x} \geq (A\bar{x})_i > 0$; i.e., $\bar{x}$ is a solution to (I'). Conversely, if $\bar{x}$ is a solution to (I'), then it is clear that $\bar{x}$ is also a solution to (I), as $e^T A\bar{x} > 0$ implies that $A\bar{x} \neq \mathbf{0}$.

Now, if (I') is not solvable, then by Farkas' lemma, the system

$$(\mathrm{II}') \quad -A^T w = A^T e, \;\; w \geq \mathbf{0}$$

is solvable. It follows that $\bar{y} = w + e$ is a solution to (II).