ENGG 5501: Foundations of Optimization	**2023–24 First Term**
Solution to Midterm Examination	
Time Limit: 2 Hours	November 10, 2023

SOLVE THE FOLLOWING PROBLEMS:

Problem 1 (20pts).

(a) **(10pts).** Let $C \subseteq \mathbb{R}^n$ be a convex cone and $i_C : \mathbb{R}^n \to \{0, +\infty\}$ be the indicator function associated with C. Give an explicit expression of i_C^*, the conjugate of i_C. Show all your work.

ANSWER: By definition, we have

$$i_C^*(y) = \sup_{x\in\mathbb{R}^n} \left\{ y^T x + i_C(x) \right\} = \sup_{x\in C} y^T x.$$

Since C is a cone, we have $\alpha x \in C$ for all $\alpha > 0$ whenever $x \in C$. It follows that

$$i_C^*(y) = \begin{cases} 0 & \text{if } y^T x \le 0 \text{ for all } x \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

In particular, this shows that i_C^* is the indicator function associated with

$$-C^* = \left\{ y \in \mathbb{R}^n : y^T x \le 0 \text{ for all } x \in C \right\},$$

the negative of the dual cone of C.

(b) **(10pts).** Given a function $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a scalar $\alpha > 0$, define the function $f_\alpha : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by $f_\alpha(x) = \alpha \cdot g(x/\alpha)$. Show that $f_\alpha^*(y) = \alpha \cdot g^*(y)$, where f_α^* and g^* are the conjugates of f_α and g, respectively.

ANSWER: From the definition of f_α^*, we compute

$$\begin{aligned} f_\alpha^*(y) &= \sup_{x\in\mathbb{R}^n} \left\{ y^T x - f_\alpha(x) \right\} = \sup_{x\in\mathbb{R}^n} \left\{ y^T x - \alpha \cdot g(x/\alpha) \right\} \\ &= \alpha \cdot \sup_{x\in\mathbb{R}^n} \left\{ y^T (x/\alpha) - g(x/\alpha) \right\} \\ &= \alpha \cdot \sup_{z\in\mathbb{R}^n} \left\{ y^T z - g(z) \right\} = \alpha \cdot g^*(y), \end{aligned}$$

where the last line follows from the bijectivity of the map $x \mapsto x/\alpha$ (recall that $\alpha > 0$ by assumption).

Problem 2 (25pts).

(a) **(15pts).** Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ be a convex function and $C \subseteq \mathbb{R}^n$ be a non-empty convex set. Consider the function $g : \mathbb{R}^m \to \mathbb{R}$ defined by $g(x) = \inf_{y\in C} f(x, y)$. Suppose that $g(x) > -\infty$ for all $x \in \mathbb{R}^m$. Show that g is convex.

ANSWER: Let $x_1, x_2 \in \mathbb{R}^m$ and $\epsilon > 0$ be arbitrary. Then, there exist $y_1, y_2 \in C \subseteq \mathbb{R}^n$ such that $f(x_1, y_1) \leq g(x_1) + \epsilon$ and $f(x_2, y_2) \leq g(x_2) + \epsilon$. Now, for each $\alpha \in (0, 1)$, we have

$$\begin{aligned} g(\alpha x_1 + (1-\alpha)x_2) &= \inf_{y \in C} f(\alpha x_1 + (1-\alpha)x_2, y) \\ &\leq f(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2) \\ &\leq \alpha f(x_1, y_1) + (1-\alpha) f(x_2, y_2) \\ &\leq \alpha g(x_1) + (1-\alpha) g(x_2) + \epsilon. \end{aligned}$$

Since the above inequality holds for all $\epsilon > 0$, by letting $\epsilon \searrow 0$, we conclude that $g(\alpha x_1 + (1-\alpha)x_2) \leq \alpha g(x_1) + (1-\alpha) g(x_2)$, as desired.

(b) **(10pts).** Let $C \subseteq \mathbb{R}^n$ be a non-empty convex set. Using the result in (a), show that the function

$$\mathbb{R}^n \ni x \mapsto \text{dist}(x, C) = \inf_{y \in C} \|x - y\|_2$$

is convex.

ANSWER: We first prove that $(x, y) \mapsto \|x - y\|_2$ is convex. Let $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^n \times \mathbb{R}^n$ and $\alpha \in (0, 1)$ be arbitrary. Then, we have

$$\begin{aligned} \|(\alpha x_1 + (1-\alpha)x_2) - (\alpha y_1 + (1-\alpha)y_2)\|_2 &= \|\alpha(x_1 - y_1) + (1-\alpha)(x_2 - y_2)\|_2 \\ &\leq \alpha \|x_1 - y_1\|_2 + (1-\alpha)\|x_2 - y_2\|_2, \end{aligned}$$

as desired. Now, the convexity of $x \mapsto \text{dist}(x, C)$ follows directly from the result in (a).

Problem 3 (15pts). Consider the LP

$$\begin{array}{ll} \text{minimize} & x_1 - x_2 \\ \text{subject to} & -x_1 + x_2 \geq 1, \\ & x_1 - x_2 \geq -1, \\ & x_1, x_2 \geq 0. \end{array} \tag{1}$$

(a) **(5pts).** Show that Problem (1) has infinitely many optimal solutions.

ANSWER: The first two constraints imply that $x_1 - x_2 = -1$. It follows that every point in the set $\mathcal{X} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 - x_2 = -1;\ x_1, x_2 \geq 0\}$ is optimal for Problem (1). Since the set $\mathcal{X}$ is convex and has at least two distinct points, it has infinitely many points.

(b) **(10pts).** Write down the dual of (1) and show that it also has infinitely many solutions.

ANSWER: The dual of (1) is given by

$$\begin{array}{ll} \text{maximize} & y_1 - y_2 \\ \text{subject to} & -y_1 + y_2 \geq 1, \\ & y_1 - y_2 \geq -1, \\ & y_1, y_2 \geq 0. \end{array} \tag{2}$$

Again, the first two constraints imply that $y_1 - y_2 = -1$. It follows that every point in the set $\mathcal{Y} = \{(y_1, y_2) \in \mathbb{R}^2 : y_1 - y_2 = -1;\ y_1, y_2 \geq 0\}$ is optimal for Problem (2), and the set $\mathcal{Y}$ has infinitely many points.

Problem 4 (40pts). Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ be given. Consider the polyhedron $P = \{x \in \mathbb{R}^n : Ax \le b\}$. For any $x \in P$, define

$$
\begin{aligned}
N(x) &= \{u \in \mathbb{R}^n : u^T(y - x) \le 0 \text{ for all } y \in P\}, \\
Q(x) &= \{A^T z \in \mathbb{R}^n : z^T(b - Ax) = 0,\ z \in \mathbb{R}^m_+\}.
\end{aligned}
$$

(a) **(10pts).** Show that $Q(x) \subseteq N(x)$.

ANSWER: Let $u \in Q(x)$. Then, we have $u = A^T z$ for some $z \in \mathbb{R}^m$ satisfying $z^T(b-Ax) = 0$ and $z \ge \mathbf{0}$. Now, for each $y \in P$, we get $z^T A(y - x) \le z^T(b - Ax) = 0$, where the inequality follows from $z \ge \mathbf{0}$ and $Ay \le b$.

(b) **(15pts).** Let $u \in \mathbb{R}^n$ be such that $u \notin Q(x)$. Furthermore, let $I(x) = \{i : a_i^T x = b_i\}$, where a_i^T is the i-th row of A. Show that there exists a $w \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ satisfying $w^T a_i \le 0$ for all $i \in I(x)$.

ANSWER: If $I(x) = \emptyset$, then every $w \in \mathbb{R}^n\backslash\{\mathbf{0}\}$ satisfies the desired conclusion vacuously. Hence, we may assume that $I(x) \ne \emptyset$. Since $Q(x)$ is a non-empty closed convex set, by the separation theorem and the definition of $Q(x)$, there exists a $w \in \mathbb{R}^n\backslash\{\mathbf{0}\}$ such that

$$
\theta^* \triangleq \max\left\{w^T A^T z : z^T(b - Ax) = 0,\ z \in \mathbb{R}^m_+\right\} < w^T u. \tag{3}
$$

Since $\{z \in \mathbb{R}^m : z^T(b - Ax) = 0,\ z \ge \mathbf{0}\}$ is a cone containing the origin and the right-hand side of (3) is finite, we have $\theta^* = 0$. Now, for each $i \in I(x)$, the i-th basis vector $e_i \in \mathbb{R}^m$ satisfies $e_i^T(b - Ax) = 0$ and $e_i \ge \mathbf{0}$. It follows that $w^T A^T e_i = w^T a_i \le \theta^* = 0$ for all $i \in I(x)$.

(c) **(15pts).** Show that for some $\epsilon > 0$, we have $x + \epsilon w \in P$, where w is the vector found in (b). Hence, conclude that $u \notin N(x)$.

ANSWER: Consider the following cases:

Case 1: $i \in I(x)$.

The result in (b) implies that $a_i^T(x + \epsilon w) \le b_i$ for all $\epsilon > 0$.

Case 2: $i \notin I(x)$.

Since $x \in P$, we have $a_i^T x < b_i$. If $w^T a_i \le 0$, then we have $a_i^T(x + \epsilon w) < b_i$ for all $\epsilon > 0$. On the other hand, upon letting

$$
\bar{\epsilon} \triangleq \left(\min_{i: i \notin I(x),\, w^T a_i > 0} (b_i - a_i^T x) \right) \Big/ \left(\max_{i: i \notin I(x),\, w^T a_i > 0} w^T a_i \right) > 0,
$$

we see that if $w^T a_i > 0$, then $a_i^T(x + \epsilon w) \le b_i$ for all $\epsilon \in (0, \bar{\epsilon}]$.

Putting the above two cases together, we conclude that $x + \bar{\epsilon} w \in P$.

Finally, using (3), we have $u^T((x + \bar{\epsilon} w) - x) = \bar{\epsilon} w^T u > 0$. It follows that $u \notin N(x)$.