ENGG 5501: Foundations of Optimization	**2018–19 First Term**
Homework Set 3 Solution	
Instructor: Anthony Man–Cho So	October 14, 2018

Problem 1 (15pts).

(a) **(5pts).** It suffices to show that f is convex on $\mathbb{R}_{++}$. This follows from

$$f'(x) = \ln x + 1, \quad f''(x) = \frac{1}{x} > 0 \quad \text{for all } x > 0.$$

Now, by definition, we have

$$f^*(y) = \sup_{x \in \mathbb{R}} \{yx - f(x)\} = \sup_{x>0}\{yx - x\ln x\}.$$

Upon setting the derivative of $x \mapsto yx - x\ln x$ to zero and solving for x, we see that the optimal solution to the above maximization problem is given by

$$x^* = \exp(y-1).$$

It follows that

$$f^*(y) = y\exp(y-1) - (y-1)\exp(y-1) = \exp(y-1).$$

(b) **(10pts).** Consider a fixed $x \in \mathbb{R}^n$ and let $I = \{i : f_i(x) = f(x)\}$ be the active index set at $\bar{x}$. By definition of the subdifferential, for any $i \in I$,

$$f(y) \ge f_i(y) \ge f_i(x) + (\nabla f_i(x))^T(y-x) = f(x) + (\nabla f_i(x))^T(y-x) \quad \text{for all } y \in \mathbb{R}^n.$$

It follows that $\nabla f_i(x) \in \partial f(x)$. Since $\partial f(x)$ is convex, we have

$$\partial f(x) \supseteq \operatorname{conv}\{\nabla f_i(x) : f_i(x) = f(x)\}. \tag{1}$$

To prove the converse, we proceed as follows. Let $\epsilon > 0$ be arbitrary. Since f is convex, it can be verified that the difference quotient

$$t \mapsto \frac{f(x+td) - f(x)}{t}$$

is increasing in $t > 0$. Hence, by definition of the directional derivative, we have

$$\frac{f(x+td) - f(x)}{t} \ge f'(x,d) - \epsilon$$

for any $t > 0$. Now, for each $t > 0$, define

$$I_t = \left\{ i \in I : \frac{f_i(x+td) - f(x)}{t} \ge f'(x,d) - \epsilon \right\}.$$

Note that $I_t \neq \emptyset$ for any $t > 0$. Moreover, since

$$t \mapsto \frac{f_i(x+td) - f(x)}{t} = \frac{f_i(x+td) - f_i(x)}{t} + \frac{f_i(x) - f(x)}{t}$$

is increasing in $t > 0$, we have $I_{t_1} \subseteq I_{t_2}$ for $0 < t_1 \le t_2$. It follows from the compactness of I that there exists an $i^* \in \cap_{t>0} I_t$. This yields

$$\frac{f_{i^*}(x+td) - f(x)}{t} \ge f'(x,d) - \epsilon$$

for all $t > 0$. Since $f(x) \ge f_{i^*}(x)$, we have

$$\frac{f_{i^*}(x+td) - f_{i^*}(x)}{t} \ge f'(x,d) - \epsilon,$$

which, upon taking $t \searrow 0$ and using Theorem 14(a) of Handout 2, implies that

$$\max_{s \in \operatorname{conv}\{\nabla f_i(x): f_i(x) = f(x)\}} s^T d \ge (\nabla f_{i^*}(x))^T d = f'_{i^*}(x,d) \ge f'(x,d) - \epsilon = \max_{s \in \partial f(x)} s^T d - \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have

$$\max_{s \in \operatorname{conv}\{\nabla f_i(x): f_i(x) = f(x)\}} s^T d \ge \max_{s \in \partial f(x)} s^T d.$$

Together with (1), we conclude that $\partial f(x) \subseteq \operatorname{conv}\{\nabla f_i(x) : f_i(x) = f(x)\}$, as desired.

Problem 2 (15pts).

(a) **(5pts).** First, let us show that systems (I) and (II) cannot both have solutions. Suppose to the contrary that there exist vectors $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in \mathbb{R}^m$ satisfying (I) and (II), respectively. Then, since $\bar{y} > \mathbf{0}$, $A\bar{x} \ge \mathbf{0}$, and $A\bar{x} \neq \mathbf{0}$, we have $\bar{y}^T A \bar{x} > 0$. On the other hand, since $A^T \bar{y} = \mathbf{0}$, we have $\bar{y}^T A \bar{x} = 0$. This results in a contradiction.

Now, suppose that system (I) does not have a solution. Then, by a simple scaling argument, the system

$$\text{(I}'\text{)} \quad Ax \ge \mathbf{0},\ e^T A x = 1$$

does not have a solution either. By separating the positive and negative parts of x and introducing slack variables, we see that system (I$'$) is equivalent to

$$\begin{bmatrix} A & -A & -I \\ e^T A & -e^T A & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} x^+ \\ x^- \\ s \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix},\ (x^+, x^-, s) \ge \mathbf{0}.$$

Hence, by Farkas' lemma, there exists a $\bar{z} = (\bar{u}, \bar{t}) \in \mathbb{R}^{m+1}$ such that

$$\begin{bmatrix} A^T & A^T e \\ -A^T & -A^T e \\ -I & \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{t} \end{bmatrix} \le \mathbf{0},\ \bar{t} > 0,$$

or equivalently,

$$A^T(\bar{u} + \bar{t}e) = \mathbf{0},\ \bar{u} \ge \mathbf{0},\ \bar{t} > 0.$$

Now, let $\bar{y} = \bar{u} + \bar{t}e \in \mathbb{R}^m$. Clearly, we have $A^T \bar{y} = \mathbf{0}$. Moreover, since $\bar{u} \ge \mathbf{0}$ and $\bar{t} > 0$, we have $\bar{y} \ge \bar{t}e > \mathbf{0}$. This completes the proof.

(b) **(10pts).** The dual of the given LP is

$$\begin{array}{ll} \text{maximize} & 9y_2 + 6y_3 \\ \text{subject to} & y_1 + 2y_2 + y_3 \leq -3, \\ & 2y_1 - 2y_2 - y_3 \leq 1, \\ & -y_1 + 3y_2 + 2y_3 \leq 3, \\ & y_1 + 3y_2 - y_3 \leq -1. \end{array}$$

It is clear that $\bar{x}$ is feasible for the given LP. Hence, by the LP strong duality theorem, $\bar{x}$ is an optimal solution to the given LP if and only if the following system has a solution:

$$\left.\begin{array}{rcrcrcr} y_1 & + & 2y_2 & + & y_3 & \leq & -3, \\ 2y_1 & - & 2y_2 & - & y_3 & \leq & 1, \\ -y_1 & + & 3y_2 & + & 2y_3 & \leq & 3, \\ y_1 & + & 3y_2 & - & y_3 & \leq & -1, \end{array}\right\} \text{dual feasibility}$$

$$\left.\begin{array}{rcrcrcr} y_1 & + & 2y_2 & + & y_3 & = & -3, \\ 2y_1 & - & 2y_2 & - & y_3 & = & 1, \\ -y_1 & + & 3y_2 & + & 2y_3 & = & 3. \end{array}\right\} \text{complementarity}$$

Upon solving the complementarity conditions, we get

$$\bar{y}_1 = -\frac{2}{3}, \quad \bar{y}_2 = -7, \quad \bar{y}_3 = \frac{35}{3}.$$

Since $(\bar{y}_1, \bar{y}_2, \bar{y}_3)$ also satisfies

$$y_1 + 3y_2 - y_3 \leq -1,$$

we conclude that $\bar{x}$ is optimal for the given LP.

Problem 3. We show that (i)⇒(iii)⇒(ii)⇒(i).

Step 1: (i)⇒(iii). Suppose that P contains a recession direction $d \in \mathbb{R}^n$. Then, for any $x_0 \in P$, we have $x(\lambda) = x_0 + \lambda d \in P$ for all $\lambda \geq 0$. Let $i \in \{1, \ldots, n\}$ be an index such that $d_i \neq 0$. Then, we have

$$\|x(\lambda)\|_2 \geq |(x_0)_i + \lambda d_i| \to \infty \qquad \text{as } \lambda \to \infty.$$

It follows that P is unbounded.

Step 2: (iii)⇒(ii). We establish the contrapositive. Suppose that the system

$$Ad = \mathbf{0}, \;\; d \geq \mathbf{0}, \;\; d \neq \mathbf{0}$$

has no solution. Then, by Gordan's theorem (Corollary 2 of Handout 3), there exists a $v \in \mathbb{R}^m$ such that $A^T v \geq e$. Now, let $x \in P$ be arbitrary. Since $Ax = b$ and $x \geq \mathbf{0}$, we have

$$b^T v = x^T A^T v \geq e^T x = \|x\|_1.$$

In particular, we see that P is bounded.

Step 3: (ii)$\Rightarrow$(i). Suppose that $d \in \mathbb{R}^n$ satisfies

$$Ad = \mathbf{0}, \ \ d \geq \mathbf{0}, \ \ d \neq \mathbf{0}.$$

Let $x_0 \in P$ and $\lambda \geq 0$ be arbitrary. Consider the point $x(\lambda) = x_0 + \lambda d \in \mathbb{R}^n$. Since $Ax_0 = b$ and $x_0 \geq \mathbf{0}$, we have

$$Ax(\lambda) = Ax_0 + \lambda Ad = b, \ \ x(\lambda) = x_0 + \lambda d \geq \mathbf{0};$$

i.e., $x(\lambda) \in P$. It follows that P contains the recession direction d.

Problem 4. Suppose that there exists a $d \in C \setminus \{\mathbf{0}\}$ satisfying $c^T d < 0$. Then, we have $\lambda d \in C$ for any $\lambda > 0$, which implies that $v^* = -\infty$.

Conversely, suppose that $v^* = -\infty$. Let $A \in \mathbb{R}^{m \times n}$ be the $m \times n$ matrix whose i-th row is a_i^T, where $i = 1, \ldots, m$. By scaling if necessary, there exists an $\bar{x} \in C$ such that $c^T \bar{x} = -1$. This implies that the polyhedron

$$P = \left\{ x \in \mathbb{R}^n : a_i^T x \geq 0 \text{ for } i = 1, \ldots, m, \, c^T x = -1 \right\}$$

is non-empty. Since $\mathbf{0}$ is a basic feasible solution of C, there exist n vectors in the collection $\{a_1, \ldots, a_m\}$ that are linearly independent. Hence, by Theorem 3 of Handout 3, P has at least one extreme point, say $d \in \mathbb{R}^n$. Note that there are n linearly independent active constraints at d. Moreover, since $c^T d = -1$, we have $d \neq \mathbf{0}$. Thus, there are $n - 1$ linearly independent constraints of the form $a_i^T x \geq 0$ that are active at d. This completes the proof.