ENGG 5501: Foundations of Optimization	2025–26 First Term

Homework Set 2 Solution

Instructor: Anthony Man–Cho So — November 3, 2025

Problem 1 (10pts). 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 $f(x_1, y_1) \le g(x_1) + \epsilon$ and $f(x_2, y_2) \le g(x_2) + \epsilon$. Now, for each $\alpha \in (0,1)$, we have

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

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

Problem 2 (10pts).

(a) **(10pts).** By definition, for every $x \in C$, we have $i_C(x) = 0$ and

$$\partial i_C(x) = \{ w \in \mathbb{R}^n : i_C(y) \ge w^T(y - x) \text{ for all } y \in \mathbb{R}^n \}.$$

The desired expression for $\partial i_C(x)$ follows by noting that $i_C(y) = 0$ if $y \in C$ and $i_C(y) = +\infty$ otherwise. The set $\partial i_C(x)$ is known as the *normal cone* to C at x.

(b) **(10pts).** In general, we do not have $\partial(i_{C_1} + i_{C_2})(x) = \partial i_{C_1}(x) + \partial i_{C_2}(x)$. Indeed, consider $C_1 = \{y \in \mathbb{R}^2 : \|y - (-1,0)\|_2 \le 1\}$, $C_2 = \{y \in \mathbb{R}^2 : \|y - (1,0)\|_2 \le 1\}$, and $x = (0,0)$. By the result in (a), we have $\partial i_{C_1}((0,0)) = \{w \in \mathbb{R}^2 : w^T y \le 0 \text{ for all } y \in C_1\}$. Note that every $y \in C_1$ can be expressed as $y = (-1,0) + v$ for some $v \in \mathbb{R}^2$ with $\|v\|_2 \le 1$. Thus, we have $w^T y \le 0$ for all $y \in C_1$ if and only if $w_1(v_1 - 1) + w_2 v_2 \le 0$ for all $v = (v_1, v_2)$ satisfying $\|v\|_2 \le 1$. This is only possible when $w_1 \ge 0$ and $w_2 = 0$, as $-1 \le v_1, v_2 \le 1$. In other words, we have

$$\partial i_{C_1}((0,0)) = \{(w_1, 0) \in \mathbb{R}^2 : w_1 \ge 0\}.$$

A similar argument shows that

$$\partial i_{C_2}((0,0)) = \{(w_1, 0) \in \mathbb{R}^2 : w_1 \le 0\}.$$

It follows that $\partial i_{C_1}((0,0)) + \partial i_{C_2}((0,0)) = \{(w_1, 0) \in \mathbb{R}^2 : w_1 \in \mathbb{R}\}$. On the other hand, since $i_{C_1}(\cdot) + i_{C_2}(\cdot) = i_{C_1 \cap C_2}(\cdot) = i_{\{(0,0)\}}(\cdot)$, we have $\partial(i_{C_1} + i_{C_2})((0,0)) = \mathbb{R}^2$.

Problem 3 (20pts).

(a) **(10pts).** Given $\rho \ge 0$, let $g : \mathbb{R}^n \to \mathbb{R}$ be given by $g(x) = f(x) + \frac{\rho}{2}\|x\|_2^2$. We compute

$$\begin{aligned} &\alpha\|x\|_2^2 + (1-\alpha)\|y\|_2^2 - \|\alpha x + (1-\alpha)y\|_2^2 \\ &= \alpha\|x\|_2^2 + (1-\alpha)\|y\|_2^2 - \alpha^2\|x\|_2^2 - 2\alpha(1-\alpha)x^T y - (1-\alpha)^2\|y\|_2^2 \\ &= \alpha(1-\alpha)(\|x\|_2^2 + \|y\|_2^2 - 2x^T y) \\ &= \alpha(1-\alpha)\|x - y\|_2^2. \end{aligned}$$

Hence, we have

$$
\begin{aligned}
& f \text{ is } \rho\text{-convex} \\
\iff & f(\alpha x+(1-\alpha)y) \le \alpha f(x)+(1-\alpha)f(y)+\frac{\rho\alpha(1-\alpha)}{2}\|x-y\|_2^2, \quad \forall x,y\in\mathbb{R}^n;\ \alpha\in[0,1] \\
\iff & f(\alpha x+(1-\alpha)y) \\
& \le \alpha f(x)+(1-\alpha)f(y)+\frac{\rho}{2}\left(\alpha\|x\|_2^2+(1-\alpha)\|y\|_2^2-\|\alpha x+(1-\alpha)y\|_2^2\right), \ \forall x,y\in\mathbb{R}^n;\ \alpha\in[0,1] \\
\iff & g(\alpha x+(1-\alpha)y) \le \alpha g(x)+(1-\alpha)g(y), \quad \forall x,y\in\mathbb{R}^n;\ \alpha\in[0,1] \\
\iff & g \text{ is convex.}
\end{aligned}
$$

(b) **(10pts).** Given $x\in\mathbb{R}^n$ and $\rho\ge 0$, we have $-\|x\|_1+\frac{\rho}{2}\|x\|_2^2=\sum_{i=1}^n\left(-|x_i|+\frac{\rho}{2}x_i^2\right)$. Thus, it suffices to understand whether the univariate function $\mathbb{R}\ni x\mapsto -|x|$ is ρ-convex for some $\rho\ge 0$. Towards that end, for each $\rho\ge 0$, let $g_\rho:\mathbb{R}\to\mathbb{R}$ be the function given by $g_\rho(x)=-|x|+\frac{\rho}{2}x^2$. Clearly, the function g_0 is non-convex. Moreover, for $\rho>0$, we have

$$
g_\rho(0)=g_\rho\left(\frac{\rho^{-1}+(-\rho^{-1})}{2}\right)=0>-\frac{1}{2\rho}=\frac{1}{2}g_\rho(\rho^{-1})+\frac{1}{2}g_\rho(-\rho^{-1}).
$$

It follows that g_ρ is non-convex for all $\rho\ge 0$.

Problem 4 (30pts). The systems (I) and (II) cannot be simultaneously solvable. Indeed, suppose that $\bar{x}\in\mathbb{R}^n$ solves (I) and $\bar{y}\in\mathbb{R}^m$ solves (II). Then, since $\bar{y}>\mathbf{0}$, $A\bar{x}\le\mathbf{0}$, and $A\bar{x}\ne\mathbf{0}$, we have $\bar{y}^TA\bar{x}<0$. On the other hand, since $\bar{x}\ge\mathbf{0}$ and $A^T\bar{y}\ge\mathbf{0}$, we have $\bar{y}^TA\bar{x}\ge 0$. This results in a contradiction.

Suppose that (I) is not solvable. Then, by a simple scaling argument, we see that

$$
(\mathrm{I}') \quad Ax\le\mathbf{0},\ e^TAx=-1,\ x\ge\mathbf{0}
$$

is not solvable either. By introducing slack variables, we see that (I′) is equivalent to

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

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

$$
\begin{bmatrix} A^T & A^Te \\ I & \mathbf{0} \end{bmatrix}\begin{bmatrix} \bar{u} \\ \bar{t} \end{bmatrix}\ge\mathbf{0},\ \bar{t}>0,
$$

or equivalently,

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

Now, let $\bar{y}=\bar{u}+\bar{t}e\in\mathbb{R}^m$. Clearly, we have $A^T\bar{y}\ge\mathbf{0}$. Moreover, since $\bar{u}\ge\mathbf{0}$ and $\bar{t}>0$, we have $\bar{y}\ge\bar{t}e>\mathbf{0}$. It follows that (II) is solvable, as desired.

Problem 5 (20pts). Suppose that $(x^*,y^*)\in\mathbb{R}_+^n\times\mathbb{R}^m$ is a Nash equilibrium of the game. We first show that $Ax^*=b$. Suppose there exists an index i such that $(b-Ax^*)_i\ne 0$. For any $\alpha>0$, let $y_\alpha=\alpha\cdot\mathrm{sgn}((b-Ax^*)_i)\cdot e_i\in\mathbb{R}^m$. Since $L(x^*,y_\alpha)\le L(x^*,y^*)$, we have

$$
L(x^*,y_\alpha)=c^Tx^*+y_\alpha^T(b-Ax^*)=c^Tx^*+\alpha|(b-Ax^*)_i|\le L(x^*,y^*)=c^Tx^*+(y^*)^T(b-Ax^*).
$$

Since the above inequality holds for all $\alpha > 0$ and the rightmost quantity is fixed, we obtain a contradiction by letting $\alpha \nearrow +\infty$. It follows that $Ax^* = b$. This, together with the assumption that $x^* \in \mathbb{R}^n_+$, implies that x^* is feasible for the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \\ & x \in \mathbb{R}^n_+. \end{array} \tag{1}$$

Now, observe that for any feasible solution $\bar{x}$ to Problem (1), we have $L(x^*, y^*) \leq L(\bar{x}, y^*)$, which implies that $c^T x^* \leq c^T \bar{x}$. Hence, we conclude that x^* is optimal for Problem (1).

Next, we show that $c - A^T y^* \in \mathbb{R}^n_+$. Suppose there exists an index i such that $(c - A^T y^*)_i < 0$. For any $\alpha > 0$, let $x_\alpha = \alpha \cdot e_i \in \mathbb{R}^n_+$. Since $L(x^*, y^*) \leq L(x_\alpha, y^*)$, we have

$$L(x^*, y^*) = c^T x^* + (y^*)^T (b - Ax^*) \leq L(x_\alpha, y^*) = c^T x_\alpha + (y^*)^T (b - Ax_\alpha) = \alpha (c - A^T y^*)_i + b^T y^*.$$

Since the above inequality holds for all $\alpha > 0$ and the leftmost quantity is fixed, we obtain a contradiction by letting $\alpha \nearrow +\infty$. It follows that $c - A^T y^* \in \mathbb{R}^n_+$; i.e., y^* is feasible for the dual LP

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y \in \mathbb{R}^n_+, \\ & y \in \mathbb{R}^m. \end{array} \tag{2}$$

Now, since $L(x^*, y^*) \leq L(\mathbf{0}, y^*)$ and $Ax^* = b$, we have $c^T x^* \leq b^T y^*$. Moreover, since x^* is feasible for (1) and y^* is feasible for (2), by the LP weak duality theorem, we have $c^T x^* = b^T y^*$. It follows that y^* is optimal for Problem (2).

Conversely, suppose that x^* is an optimal solution to (1) and y^* is an optimal solution to (2). Then, since $Ax^* = b$, for any $y \in \mathbb{R}^m$, we have $L(x^*, y) = L(x^*, y^*)$. Moreover, since $Ax^* = b$ and $c - A^T y^* \in \mathbb{R}^n_+$, for any $x \in \mathbb{R}^n_+$, we have

$$L(x, y^*) = (c - A^T y^*)^T x + b^T y^* \geq b^T y^* = c^T x^* = L(x^*, y^*),$$

where the second-to-last equality follows from the LP strong duality theorem. This completes the proof.

Problem 6 (20pts).

(a) **(10pts).** The integer programming formulation of the problem is given by

$$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{j=1}^{m} c_j x_j & \\ \text{subject to} & \displaystyle\sum_{j: i \in \mathcal{S}_j} x_j \geq 1 & \text{for } i = 1, \ldots, n, \\ & x_j \in \{0, 1\} & \text{for } j = 1, \ldots, m. \end{array} \tag{3}$$

Indeed, given a feasible solution $\bar{x} \in \{0,1\}^m$ to Problem (3), let $\bar{\mathscr{S}} = \{\mathcal{S}_j : \bar{x}_j = 1, j \in \{1, \ldots, m\}\}$ be the corresponding sub-collection. Then, the objective function gives the total cost of the subsets in $\bar{\mathscr{S}}$. Moreover, observe that the union of the subsets in $\bar{\mathscr{S}}$ equals $\mathcal{U}$ if and only if for each element $i \in \mathcal{U}$, there exists a subset $\mathcal{S}_j \in \bar{\mathscr{S}}$ such that $i \in \mathcal{S}_j$. This yields the first set of constraints in Problem (3).

The LP relaxation of Problem (3) is given by

$$
\begin{array}{lll}
\text{minimize} & \displaystyle\sum_{j=1}^{m} c_j x_j & \\
\text{subject to} & \displaystyle\sum_{j:i\in\mathcal{S}_j} x_j \geq 1 & \text{for } i = 1,\ldots,n, \\
& x_j \geq 0 & \text{for } j = 1,\ldots,m.
\end{array} \tag{4}
$$

Note that if x^* is an optimal solution to Problem (4), then we necessarily have $x^* \leq e$. This follows from the non-negativity of the cost c and the observation that the vector $\bar{x}' \in [0,1]^m$ defined by $\bar{x}'_j = \min\{\bar{x}_j, 1\}$ for $j = 1,\ldots,m$ is feasible for Problem (4) whenever the vector $\bar{x}$ is feasible for Problem (4).

(b) **(10pts).** Let y_i be the dual variable corresponding to the constraint $\sum_{j:i\in\mathcal{S}_j} x_j \geq 1$, where $i = 1,\ldots,n$. Note that the first set of constraints in Problem (4) takes the form $Ax \geq e$, where $A \in \{0,1\}^{n\times m}$ and the (i,j)-th entry of A equals 1 if and only if $i \in \mathcal{S}_j$. It follows that the dual of Problem (4) is given by

$$
\begin{array}{lll}
\text{maximize} & \displaystyle\sum_{i=1}^{n} y_i & \\
\text{subject to} & \displaystyle\sum_{i\in\mathcal{S}_j} y_i \leq c_j & \text{for } j = 1,\ldots,m, \\
& y_i \geq 0 & \text{for } i = 1,\ldots,n.
\end{array}
$$