ENGG 5501: Foundations of Optimization	2021–22 First Term

Homework Set 4 Solution

Instructor: Anthony Man–Cho So — December 1, 2021

Problem 1 (20pts). Recall that we are given the following LP:

$$\begin{array}{ll} \text{minimize} & e^T y \\ \text{subject to} & Ax + Iy = b, \\ & x \geq \mathbf{0},\ y \geq \mathbf{0}. \end{array} \tag{1}$$

(a) **(5pts).** The dual of (1) is given by

$$\begin{array}{ll} \text{maximize} & b^T z \\ \text{subject to} & A^T z \leq \mathbf{0}, \\ & z \leq e. \end{array} \tag{2}$$

(b) **(5pts).** By assumption, we have $b \geq \mathbf{0}$. Hence, $(x, y) = (\mathbf{0}, b)$ is a feasible solution to (1). Moreover, $z = \mathbf{0}$ is a feasible solution to (2). Thus, by the LP strong duality theorem, we conclude that (1) has an optimal solution.

(c) **(10pts).** Suppose that the optimal value of (1) is zero. Let (x^*, y^*) be an optimal solution to (1). Then, since $y^* \geq \mathbf{0}$ and $e^T y^* = 0$, it follows that $y^* = \mathbf{0}$. This implies that $Ax^* = b$ and $x^* \geq \mathbf{0}$.

Conversely, suppose that there exists an $\bar{x} \in \mathbb{R}^n$ satisfying $\bar{x} \geq \mathbf{0}$ and $A\bar{x} = b$. Then, $(x, y) = (\bar{x}, \mathbf{0})$ is a feasible solution to (1) with objective value 0. Since we have $e^T y \geq 0$ for any $y \geq \mathbf{0}$, we conclude that $(x, y) = (\bar{x}, \mathbf{0})$ is in fact optimal for (1). This completes the proof.

Problem 2 (15pts). Consider the LP

$$\begin{array}{ll} \text{maximize} & e^T x \\ \text{subject to} & (I - P)^T x = \mathbf{0}, \\ & x \geq \mathbf{0} \end{array} \tag{3}$$

and its dual

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & (P - I)y \geq e. \end{array} \tag{4}$$

We claim that (4) is infeasible. Indeed, since P is a stochastic matrix, each entry of the vector Py is a convex combination of the entries of y. In particular, we have $[Py]_i \leq y_{\max}$ for $i = 1, \ldots, n$, where $y_{\max} = \max_{i \in \{1,\ldots,n\}} y_i$. However, one of the entries of $y + e$ equals $y_{\max} + 1$. This yields the desired contradiction.

It follows that (3) is either infeasible or unbounded. However, it is clear that $x = \mathbf{0}$ is feasible for (3). Hence, we conclude that (3) is unbounded, which implies that the system

$$P^T x = x, \quad x \geq \mathbf{0}, \quad x \neq \mathbf{0}$$

is solvable, as desired.

Problem 3 (25pts).

(a) **(10pts).** First, we check that C_p is convex. Let $(t_1, x^1), (t_2, x^2) \in C_p$ and $\alpha \in [0, 1]$. Then, we have

$$\|\alpha x^1 + (1-\alpha)x^2\|_p \le \alpha\|x^1\|_p + (1-\alpha)\|x^2\|_p \le \alpha t_1 + (1-\alpha)t_2,$$

where the first inequality follows from the triangle inequality for the norm $\|\cdot\|_p$. Hence, C_p is convex.

Next, we show that C_p is a cone. Let $(t, x) \in C_p$ and $\alpha > 0$. Then, we have $\|\alpha x\|_p = \alpha\|x\|_p \le \alpha t$; i.e. $\alpha(t, x) \in C_p$. It follows that C_p is a cone.

Lastly, we check that C_p is closed. Let $\{(t_k, x^k)\}_{k\ge 0}$ be a sequence in C_p such that $(t_k, x^k) \to (\bar{t}, \bar{x})$. Since the function $(t, x) \mapsto f(t, x) = \|x\|_p - t$ is continuous and $f(t_k, x^k) = \|x^k\|_p - t_k \le 0$ for all $k \ge 0$, we have

$$0 \ge \lim_{k\to\infty} f(t_k, x^k) = f\left(\lim_{k\to\infty}(t_k, x^k)\right) = f(\bar{t}, \bar{x}) = \|\bar{x}\|_p - \bar{t};$$

i.e., $(\bar{t}, \bar{x}) \in C_p$. It follows that C_p is closed.

(b) **(15pts).** Let us establish the following more general result: For any $p, q \ge 1$ satisfying $1/p + 1/q = 1$, we have $C_p^* = C_q$. Recall that by definition,

$$C_p^* = \left\{(t', x') \in \mathbb{R} \times \mathbb{R}^n : tt' + x^T x' \ge 0 \text{ for all } (t, x) \in C_p\right\}.$$

Let $(t', x') \in C_q$ be arbitrary. For any $(t, x) \in C_p$, we compute

$$\begin{aligned} tt' + x^T x' &\ge \|x\|_p \cdot \|x'\|_q + x^T x' && \text{(since } (t,x) \in C_p \text{ and } (t',x') \in C_q\text{)} \\ &\ge \sum_{j=1}^n |x_j x_j'| + x^T x' && \text{(by Hölder's inequality)} \\ &\ge 0; \end{aligned}$$

i.e. $(t', x') \in C_p^*$. It follows that $C_q \subseteq C_p^*$.

Now, suppose that $(t', x') \notin C_q$; i.e. $\|x'\|_q > t'$. Our goal is to show that there exists a pair $(t, x) \in C_p$ satisfying $tt' + x^T x' < 0$, so that $(t', x') \notin C_p^*$. Consider the following two cases:

<u>**Case 1**</u>: $p > 1$.

Define $u = (\|x'\|_q^{q/p}, \tilde{x}') \in \mathbb{R} \times \mathbb{R}^n$, where

$$\tilde{x}_j' = -\text{sgn}(x_j') \cdot |x_j'|^{q-1} \qquad \text{for } j = 1, \ldots, n$$

and for any $x \in \mathbb{R}$,

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

We claim that $u \in C_p$. Indeed, we have

$$\|\tilde{x}'\|_p^p = \sum_{j=1}^n |x_j'|^{p(q-1)} = \sum_{j=1}^n |x_j'|^q = \|x'\|_q^q,$$

where we use the fact that $p(q-1) = p+q-p = q$. Moreover, we compute

$$(t', x')^T u = t' \cdot \|x'\|_q^{q/p} - \sum_{j=1}^{n} \mathrm{sgn}(x'_j) \cdot |x'_j|^{q-1} x'_j < \|x'\|_q^{q/p+1} - \sum_{j=1}^{n} |x'_j|^q = 0,$$

where we use the fact that $(p+q)/p = q$. It follows that $(t', x') \notin C_p^*$.

<u>Case 2</u>: $p = 1$.

Let $j^* = \min\{k : |x'_k| = \max_{1 \le j \le n} |x'_j|\}$ and define $\tilde{x}'$ by

$$\tilde{x}'_j = \begin{cases} -\mathrm{sgn}(x'_j) & \text{if } j = j^*, \\ 0 & \text{otherwise.} \end{cases}$$

Then, we have $u = (1, \tilde{x}') \in C_1$. Moreover, we compute

$$(t', x')^T u = t' - \mathrm{sgn}(x'_{j^*}) x'_{j^*} < \|x'\|_\infty - \|x'\|_\infty = 0,$$

which implies that $(t', x') \notin C_p^*$.

(c) **[Extra Credit (10pts).]** We claim that $\mathrm{int}(C_p) = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : \|x\|_p < t\}$. Suppose that $(t, x) \in \mathrm{int}(C_p)$. By definition, there exists a $\delta > 0$ such that $(t, x) + \delta(s, y) \in C_p$ for any $(s, y) \in \mathbb{R} \times \mathbb{R}^n$ satisfying $\|(s, y)\|_2 \le 1$. Take $(s, y) = (-1, \mathbf{0})$. Then, we have $(t, x) + \delta(-1, \mathbf{0}) \in C_p$, which implies that $\|x\|_p \le t - \delta < t$.

Conversely, suppose that $(t, x) \in \mathbb{R} \times \mathbb{R}^n$ satisfies $\|x\|_p < t$. Since the function $(t, x) \mapsto f(t, x) = \|x\|_p - t$ is continuous, if we take $\epsilon = (t - \|x\|_p)/2 > 0$, then there exists a $\delta > 0$ such that whenever $(t', x') \in \mathbb{R} \times \mathbb{R}^n$ satisfies $\|(t', x') - (t, x)\|_2 \le \delta$, we have $|f(t', x') - f(t, x)| \le \epsilon$ (see Section 3.1 of Handout C for the definition of continuity of functions). This implies that $f(t', x') \le f(t, x) + \epsilon = -\epsilon < 0$, or equivalently, $(t', x') \in C_p$. It follows that $(t, x) \in \mathrm{int}(C_p)$.

Problem 4 (15pts).

(a) **(15pts).** Since $Q \in \mathcal{S}_+^n$, there exists a $Q^{1/2} \in \mathcal{S}_+^n$ such that $Q = Q^{1/2} Q^{1/2}$. Observe that

$$\begin{aligned} x^T Q x \le t \quad &\Longleftrightarrow \quad \left(t - \frac{1}{4}\right)^2 + \|Q^{1/2} x\|_2^2 \le \left(t + \frac{1}{4}\right)^2 \\ &\Longleftrightarrow \quad \left\| \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q^{1/2} \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} - \begin{bmatrix} 1/4 \\ \mathbf{0} \end{bmatrix} \right\|_2 \le t + \frac{1}{4} \\ &\Longleftrightarrow \quad \begin{bmatrix} 1 & \mathbf{0} \\ 1 & \mathbf{0} \\ \mathbf{0} & Q^{1/2} \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} - \begin{bmatrix} -1/4 \\ 1/4 \\ \mathbf{0} \end{bmatrix} \in \mathcal{Q}^{n+2}. \end{aligned}$$

This shows that X is SOC-representable.

(b) **[Extra Credit (10pts).]** Although t is not necessarily non-negative, observe that

$$
\begin{aligned}
t \le \sqrt{x_1 x_2} \quad &\iff \quad t \le \tau, \quad 0 \le \tau \le \sqrt{x_1 x_2} \\
&\iff \quad t \le \tau, \quad \tau \ge 0, \quad \tau^2 \le \left[\frac{(x_1+x_2)^2}{4} - \frac{(x_1-x_2)^2}{4} \right] \\
&\iff \quad t \le \tau, \quad \tau \ge 0, \quad \left\| \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} \tau \\ x_1 \\ x_2 \end{bmatrix} \right\|_2 \le \frac{x_1+x_2}{2} \\
&\iff \quad \left\{ \begin{array}{l} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tau \\ t \\ x_1 \\ x_2 \end{bmatrix} \in \mathcal{Q}^2, \\[2ex] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tau \\ t \\ x_1 \\ x_2 \end{bmatrix} \in \mathcal{Q}^2, \\[2ex] \begin{bmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} \tau \\ t \\ x_1 \\ x_2 \end{bmatrix} \in \mathcal{Q}^3. \end{array} \right.
\end{aligned}
$$

Hence, X is SOC-representable.

Problem 5 (20pts).

(a) **(5pts).** Let $y \in K_1^* + K_2^*$. Then, we can write $y = u + v$, where $u \in K_1^*$ and $v \in K_2^*$. Now, for any $z \in K_1 \cap K_2$, we have $\langle u, z \rangle \ge 0$ and $\langle v, z \rangle \ge 0$. It follows that $\langle y, z \rangle \ge 0$. Since this holds for any $y \in K_1^* + K_2^*$ and $z \in K_1 \cap K_2$, we conclude that $K_1^* + K_2^* \subseteq (K_1 \cap K_2)^*$.

(b) **(10pts).** First, we show that $K_1^* + K_2^*$ is convex. Let $y^1, y^2 \in K_1^* + K_2^*$ and $\alpha \in [0,1]$. Since we can write $y^1 = u^1 + v^1$ and $y^2 = u^2 + v^2$ for some $u^1, u^2 \in K_1^*$ and $v^1, v^2 \in K_2^*$, we have

$$\alpha y^1 + (1-\alpha) y^2 = (\alpha u^1 + (1-\alpha) u^2) + (\alpha v^1 + (1-\alpha) v^2).$$

By Proposition 2(a) of Handout 5, the sets K_1^*, K_2^* are convex. This gives $\alpha u^1 + (1-\alpha) u^2 \in K_1^*$ and $\alpha v^1 + (1-\alpha) v^2 \in K_2^*$. It follows that $\alpha y^1 + (1-\alpha) y^2 \in K_1^* + K_2^*$, as desired.

Next, we show that $K_1^* + K_2^*$ is a cone. Let $y \in K_1^* + K_2^*$ and write $y = u + v$, where $u \in K_1^*$ and $v \in K_2^*$. By Proposition 2(a) of Handout 5, for any $\alpha > 0$, we have $\alpha u \in K_1^*$ and $\alpha v \in K_2^*$. It follows that $\alpha y \in K_1^* + K_2^*$ for all $\alpha > 0$, as desired.

(c) **[Extra Credit (15pts).]** For any $u \in E$, define $\|u\|^2 = \langle u, u \rangle$. Let $\{z^k\}_{k \ge 0}$ be a sequence in $K_1^* + K_2^*$ such that $z^k \to z$. The proof proceeds in four steps:

Step 1: (Boundedness of $\{z^k\}_{k\geq 0}$). Fix an $\epsilon > 0$. Since $z^k \to z$, there exists an index $K \geq 0$ such that $\|z^k - z\| \leq \epsilon$ for all $k \geq K$. Let $M = \max\{\|z\| + \epsilon, \|z^1\|, \dots, \|z^{K-1}\|\}$. Then, we have $\|z^k\| \leq M$ for all $k \geq 0$. This shows that $\{z^k\}_{k\geq 0}$ is bounded.

Step 2: (Properties of u). Let $y \in K_1^* \setminus \{\mathbf{0}\}$. By definition, we have $\langle u, y\rangle \geq 0$. We claim that $\langle u, y\rangle > 0$. Indeed, since we have $u \in \mathrm{int}(K_1)$, there exists an $\epsilon > 0$ such that $u + \epsilon w \in K_1$ for any $w \in E$ satisfying $\|w\| \leq 1$. By taking $w = -y/\|y\|$, we see that if $\langle u, y\rangle = 0$, then $\langle u + \epsilon w, y\rangle = -\epsilon \cdot \|y\| < 0$. This contradicts the fact that $u + \epsilon w \in K_1$ and $y \in K_1^*$. Hence, we conclude that $\langle u, y\rangle > 0$ for all $y \in K_1^* \setminus \{\mathbf{0}\}$. Since we also have $u \in \mathrm{int}(K_2)$, by the same argument, we conclude that $\langle u, y\rangle > 0$ for all $y \in K_2^* \setminus \{\mathbf{0}\}$.

Step 3: (Bounded sequences in K_1^* and K_2^*). For any $k \geq 0$, since $z^k \in K_1^* + K_2^*$, there exist $a^k \in K_1^*$ and $b^k \in K_2^*$ such that $z^k = a^k + b^k$. By the Cauchy-Schwarz inequality and the result in Step 1, we have

$$\langle u, a^k\rangle + \langle u, b^k\rangle = \langle u, z^k\rangle \leq \|u\| \cdot \|z^k\| \leq M \cdot \|u\|.$$

Moreover, the definitions of u, a^k, and b^k imply that $\langle u, a^k\rangle \geq 0$ and $\langle u, b^k\rangle \geq 0$. It follows that $\langle u, a^k\rangle \leq M \cdot \|u\|$ and $\langle u, b^k\rangle \leq M \cdot \|u\|$ for all $k \geq 0$. This implies that whenever $a^k \neq \mathbf{0}$, we have

$$M \cdot \|u\| \geq \langle u, a^k\rangle = \|a^k\| \cdot \left\langle u, \frac{a^k}{\|a^k\|} \right\rangle \geq \|a^k\| \cdot \inf_{y \in K_1^* : \|y\|=1} \langle u, y\rangle. \tag{5}$$

By the result in Step 2, we have $\langle u, y\rangle > 0$ for any $y \in K_1^* \setminus \{\mathbf{0}\}$. Since K_1^* is closed by Proposition 2(a) of Handout 5 and $\{y \in E : \|y\| = 1\}$ is compact, the set $\{y \in K_1^* : \|y\| = 1\} = K_1^* \cap \{y \in E : \|y\| = 1\}$ is compact. Hence, we can invoke Weierstrass' theorem to deduce that $\inf_{y \in K_1^* : \|y\|=1} \langle u, y\rangle > 0$. This, together with (5), shows that $\{a^k\}_{k\geq 0}$ is bounded. By the same argument, we deduce that $\{b^k\}_{k\geq 0}$ is bounded.

Step 4: (Completing the proof). Since both $\{a^k\}_{k\geq 0}$ and $\{b^k\}_{k\geq 0}$ are bounded, by passing to subsequences if necessary, we may assume that $a^k \to a$ and $b^k \to b$. Since $a^k \in K_1^*$ for all $k \geq 0$ and K_1^* is closed by Proposition 2(a) of Handout 5, we have $a \in K_1^*$. A similar argument yields $b \in K_2^*$. Since $z^k \to z$ by assumption, we conclude that $z = a + b \in K_1^* + K_2^*$, as desired.

(d) **(10pts).** Suppose that $y \notin K_1^* + K_2^*$. By the results in (b), (c), and Problem 3 of the Midterm Examination, there exists a $b \in E$ such that $\langle b, u\rangle \leq 0 < \langle b, y\rangle$ for all $u \in K_1^* + K_2^*$. Since $\mathbf{0} \in K_1^* \cap K_2^*$, we have $\langle b, u\rangle \leq 0$ for all $u \in K_1^*$ and $\langle b, u\rangle \leq 0$ for all $u \in K_2^*$. This, together with Proposition 2(b) of Handout 5, implies that $-b \in (K_1^*)^* = K_1$ and $-b \in (K_2^*)^* = K_2$; i.e., $-b \in K_1 \cap K_2$. Since $\langle -b, y\rangle < 0$, we conclude that $y \notin (K_1 \cap K_2)^*$. This completes the proof.