Special Topics in Management Science 26:711:685

Convex Analysis and Optimization

Fall 2013 Rutgers University Prof. Eckstein

Solutions to Homework 3

1. (a) Consider any $y = (y^1, \ldots, y^m) \in C_1^* \times \cdots \times C_m^*$. Then for any $x = (x^1, \ldots, x^m) \in C_1 \times \cdots \times C_m$, we have $\langle x^i, y^i \rangle \le 0$ for $i = 1, \ldots, m$, since $x^i \in C_i$ and $y^i \in C_i^*$. Thus,

$$\langle x, y \rangle = \sum_{i=1}^{m} \langle x^i, y^i \rangle \le 0,$$

because each term in the sum is nonpositive. Since $x \in C_1 \times \cdots \times C_m$ was arbitrary, $y \in (C_1 \times \cdots \times C_m)^*$. Since $y \in C_1^* \times \cdots C_m^*$ was arbitrary, $C_1^* \times \cdots \times C_m^* \subseteq (C_1 \times \cdots \times C_m)^*$.

For the reverse inclusion, consider any $y = (y^1, \ldots, y^m) \notin C_1^* \times \cdots \times C_m^*$. This means that for at least one $i \in \{1, \ldots, m\}$, we have $y^i \notin C_i^*$. This in turn means that there exists $x^i \in C_i$ such that $\langle x^i, y^i \rangle > 0$. Consider now the vector $x = (0, \ldots, 0, x^i, 0, \ldots, 0)$, where the nonzero entry is in the i^{th} position. Since the C_j, $j \neq i$, are cones, they contain 0, and so $x \in C_1 \times \cdots \times C_m$. But then we have $\langle x, y \rangle = \langle x^i, y^i \rangle > 0$, so $y \notin (C_1 \times \cdots \times C_m)^*$. Thus, in view of the above inclusion, $C_1^* \times \cdots \times C_m^* = (C_1 \times \cdots \times C_m)^*$.

(b) This result may be proved relatively compactly as follows:

$$\begin{aligned}
y \in \bigcap_{i \in I} C_i^* \qquad &\Leftrightarrow \qquad (\forall\, i \in I) \quad y \in C_i^* \\
&\Leftrightarrow \qquad (\forall\, i \in I)\ (\forall\, x_i \in C_i) \quad \langle x_i, y \rangle \le 0 \\
&\Leftrightarrow \qquad \left(\forall\, y \in \bigcup_{i \in I} C_i \right) \quad \langle x, y \rangle \le 0 \\
&\Leftrightarrow \qquad y \in \left(\forall\, y \in \bigcup_{i \in I} C_i \right)^* .
\end{aligned}$$

(c) Consider $y \in C_1^* \cap C_2^*$, and any $x \in C_1 + C_2$. Now, we must have $x = x^1 + x^2$, where $x^1 \in C_1$ and $x_2 \in C_2$. Since $y \in C_1^*$, we have $\langle x^1, y \rangle \le 0$, and since $y \in C_2^*$, we also have $\langle x^2, y \rangle \le 0$. Therefore, $\langle x, y \rangle = \langle x^1 + x^2, y \rangle = \langle x^1, y \rangle + \langle x^2, y \rangle \le 0 + 0 = 0$, and $y \in (C_1 + C_2)^*$. Thus, $C_1^* \cap C_2^* \subseteq (C_1 + C_2)^*$. Next, consider $y \notin C_1^* \cap C_2^*$. Then for either $i = 1$ or $i = 2$, we have $y \notin C_i^*$. There must then exist some $x \in C_i$ with $\langle x, y \rangle > 0$. But this x is also in $C_1 + C_2$ (just add the vector 0 from the other cone), and so $y \notin (C_1 + C_2)^*$. So, $C_1^* \cap C_2^* = (C_1 + C_2)^*$.

2. There is a nearly trivial proof based on the polar cone theorem: if there does not exist an $a \in K^*$ with $\langle a, z \rangle > 0$, then $\langle y, z \rangle \le 0$ for all $y \in K^*$, which means that

$y \in K^{**}$. Since K is nonempty, convex, and closed, the polar cone theorem asserts that $K^{**} = K$, so $z \in K$, contradicting the hypothesis.

There are several alternative approaches using more basic principles. Here is one: since K is a closed convex set, there exists some $a \in \mathbb{R}^n \backslash \{0\}$ and $b \in \mathbb{R}$ such that $\langle a, z \rangle > b$ and $\langle a, x \rangle \leq b$ for all $x \in K$. Since $0 \in K$, we must have $0 = \langle a, 0 \rangle \leq b$, and thus $\langle a, z \rangle > 0$. Furthermore, if there were any $x \in K$ such that $\langle a, x \rangle > 0$, then for sufficiently large $\alpha > 0$ we would have $\alpha x \in K$ and $\langle a, \alpha x \rangle = \alpha \langle a, x \rangle > b$, a contradiction. So we know $\langle a, x \rangle \leq 0$ for all $x \in K$, that is, $a \in K^*$.

3. Referring to an earlier result from class, convexity of $C_1 + C_2$ can be established by showing it is a cone that is closed under addition. To show that $C_1 + C_2$ is a cone, consider an arbitrary $z \in C_1 + C_2$ and $\alpha \geq 0$. If we can show that $\alpha z \in C_1 + C_2$, then $C_1 + C_2$ is a cone. Now, z must be of the form $x + y$, where $x \in C_1$ and $y \in C_2$. Since $x \in C_1$, $\alpha \geq 0$, and C_1 is a cone, $\alpha x \in C_1$. Similarly, $\alpha y \in C_2$. Therefore $C_1 + C_2 \ni \alpha x + \alpha y = \alpha(x + y) = \alpha z$. So, $C_1 + C_2$ is a cone.

   Now consider any $u, v \in C_1 + C_2$. If we can show that $u + v \in C_1 + C_2$, then the $C_1 + C_2$ is closed under addition and the proof will be complete. By construction, we must have $u = u^1 + u^2$ and $v = v^1 + v^2$, where $u^1, v^1 \in C_1$ and $u^2, v^2 \in C_2$. So

$$u + v = (u^1 + u^2) + (v^1 + v^2) = (u^1 + v^1) + (u^2 + v^2).$$

   Since C_1 is convex, it is closed under addition, hence $u^1 + v^1 \in C_1$. Similarly, $u^2 + v^2 \in C_2$. So $u + v$ is the sum of a vector in C_1 and a vector in C_2, that is, $u + v \in C_1 + C_2$.

4. Later I will present my original solution, corresponding to the hint. But first, a majority of students had a more elegant solution based on using the polar cone theorem one additional time. This approach first notes that

$$\big(\operatorname{cl}(C_1^* + C_2^*)\big)^* = (C_1^* + C_2^*)^* = C_1^{**} \cap C_2^{**} = C_1 \cap C_2,$$

   where the three equalities hold because, respectively,

   - $(\operatorname{cl} X)^* = X^*$ for any set X, as proved in class when we first introduced the notion of a polar
   - The result of problem 1(c)
   - By the polar cone theorem, since C_1 and C_2 are closed and convex.

   The sum of two convex cones is convex (see the previous problem) so $C_1^* + C_2^*$ is convex and $\operatorname{cl}(C_1^* + C_2^*)$ is closed and convex. From above, we have $C_1 \cap C_2 = \big(\operatorname{cl}(C_1^* + C_2^*)\big)^*$, and taking the polar of both sides and applying the polar cone theorem, we obtain

$$(C_1 \cap C_2)^* = \big(\operatorname{cl}(C_1^* + C_2^*)\big)^{**} = \operatorname{cl}(C_1^* + C_2^*).$$

   Here is a longer, alternative proof based more closely on the hint: take any $y \in C_1^* + C_2^*$. Then $y = y^1 + y^2$, where $y^1 \in C_1^*$ and $y_2 \in C_2^*$. Now consider any $x \in C_1 \cap C_2$. Since $x \in C_1$ and $y^1 \in C_1^*$, we have $\langle x, y^1 \rangle \leq 0$. Since we also have $x \in C_2$, we similarly have

$\langle x, y^2 \rangle \le 0$. We then calculate $\langle x, y \rangle = \langle x, y^1 + y^2 \rangle = \langle x, y^1 \rangle + \langle x, y^2 \rangle \le 0 + 0 = 0$. We then conclude that $C_1^* + C_2^* \subseteq (C_1 \cap C_2)^*$. In general, we may not know that $C_1^* + C_2^*$ is closed, but since $(C_1 \cap C_2)^*$ must be a closed set and contains $C_1^* + C_2^*$, it certainly also contains $\mathrm{cl}(C_1^* + C_2^*)$.

For the opposite inclusion, consider some $z \notin \mathrm{cl}(C_1^* + C_2^*)$. Because $\mathrm{cl}(C_1^* + C_2^*)$ is a closed convex cone, we may invoke problem 2 to conclude that there exists some $a \in [\mathrm{cl}(C_1^* + C_2^*)]^*$ such that $\langle a, z \rangle > 0$. Now, as shown in class, $(\mathrm{cl}\, X)^* = X^*$ for any set X, so

$$
\begin{aligned}
a &\in (C_1^* + C_2^*)^* \\
&= C_1^{**} \cap C_2^{**} \qquad \text{[by problem 1(c)]} \\
&= C_1 \cap C_2 \qquad \text{[by polar cone theorem, since } C_1, C_2 \text{ are closed convex cones]}.
\end{aligned}
$$

Since $\langle a, z \rangle > 0$ and $a \in C_1 \cap C_2$, $z \notin (C_1 \cap C_2)^*$. We conclude that $(C_1 \cap C_2)^* = \mathrm{cl}(C_1^* + C_2^*)$.

5. Note that in this problem I deliberately avoided the somewhat tricky issue of whether the cone $P = A^\top C^*$ is closed. It does *not* follow that because P is the linear (or continuous) image of a (possibly unbounded) closed set, it is closed. This reasoning only works for *pre*images. If C and hence C^* is finitely generated (as is always the case in $n = 2$ dimensions), then $P = A^\top C^*$ must also be finitely generated and hence closed. However, for $n \ge 3$ it is possible to devise examples in which C^* is a closed convex cone but $P = A^\top C^*$ is not closed for some matrices A. Proposition 1.5.8 on page 65 of the Bertsekas textbook gives some conditions sufficient to rule out such situations.

   (a) Suppose $\{x^k\} \subseteq K$ is a convergent sequence whose limit is x. Then $Ax^k \to Ax$, and since $Ax^k \in C$ for all k and C is closed, we have $Ax \in C$. Therefore, $x \in K$, and we can deduce that K is a closed set.

   Take any $x \in K$ and $\alpha \ge 0$. Then $A(\alpha x) = \alpha Ax \in C$, because $Ax \in C$ and C is a cone. This establishes that K is a cone.

   Finally, take any $x^1, x^2 \in K$. Then $A(x^1 + x^2) = Ax^1 + Ax^2 \in C$ because $Ax^1, Ax^2 \in C$ and C is a convex cone. Since we already established that K is a cone, it must be convex.

   (b) Suppose $z \in P$ and $\alpha \ge 0$. Then $z = A^\top y$ for some $y \in C^*$; since C^* is a cone, $\alpha y \in C^*$ and thus $A^\top(\alpha y) = \alpha A^\top y = \alpha z \in P$. This establishes that P is cone. It follows that $\mathrm{cl}\, P$ is also a cone (I omit the proof, but it is very simple).

   Consider any $x \in K$ and $z \in P$. Then $z = A^\top y$ for some $y \in C^*$, and $\langle x, z \rangle = \langle x, A^\top y \rangle = \langle Ax, y \rangle \le 0$, the last inequality following from $Ax \in C$ and $y \in C^*$. This establishes that $P \subseteq K^*$. Since K^* must be a closed set and it contains P, we also have $\mathrm{cl}\, P \subseteq K^*$.

   For the reverse inclusion, consider any $z \notin \mathrm{cl}\, P$. From problem 2, there must exist $q \in (\mathrm{cl}\, P)^* = P^*$ such that $\langle q, z \rangle > 0$. Since $q \in P^*$, we have $\langle q, w \rangle \le 0$ for all $w \in P$, that is, $\langle q, A^\top y \rangle \le 0$ for all $y \in C^*$. Equivalently, $\langle Aq, y \rangle \le 0$ for all $y \in C^*$, which means $Aq \in C^{**} = C$ (because C is a closed convex cone). From the definition of K, this means $q \in K$, and since $\langle q, z\rangle > 0$, we therefore have $z \notin K^*$. Thus, we have $\operatorname{cl} P = K^*$.

(c) We have just established $\operatorname{cl} P = K^*$. We then have

$$P^* = (\operatorname{cl} P)^* = K^{**} = K.$$

The justifications for the above three equalities are, respectively,

- $X^* = (\operatorname{cl} X)^*$ for any set X
- Taking the polar of both sides in $\operatorname{cl} P = K^*$
- The polar cone theorem, since K is closed and convex.

6. (a) We start by considering the case $n = 1$, in which case $K = [0, \infty)$. Then $K^* = \{y \in \mathbb{R} \mid xy \leq 0 \; \forall\, x \in [0, \infty)\} = (-\infty, 0] = -K$. For $n > 1$, we note that $K = [0, \infty)^n$, and so problem 1(a) implies that $K^* = ([0, \infty)^n)^* = ([0, \infty)^*)^n = (-\infty, 0]^n = -K$.

(b) First, we note that

$$\begin{aligned} -K &= \{-(x, w) \in \mathbb{R}^n \times \mathbb{R} \mid w \geq \|x\|\} \\ &= \{(y, z) \in \mathbb{R}^n \times \mathbb{R} \mid -z \geq \|-y\|\} \\ &= \{(y, z) \in \mathbb{R}^n \times \mathbb{R} \mid z \leq -\|-y\|\} \\ &= \{(y, z) \in \mathbb{R}^n \times \mathbb{R} \mid z \leq -\|y\|\}. \end{aligned}$$

Now consider any $(x, w) \in K$ and $(y, z) \in -K$. Then

$$\begin{aligned} \langle (x, w), (y, z)\rangle &= \langle x, y\rangle + wz \\ &\leq \|x\|\,\|y\| + wz && \text{[by the Cauchy-Schwarz inequality]} \\ &\leq \|x\|\,\|y\| - \|x\|\,\|y\| && \text{[since } w \geq \|x\| \text{ and } z \leq -\|y\|\text{]} \\ &= 0. \end{aligned}$$

Thus, we conclude that $-K \subseteq K^*$. Now consider any $(u, v) \in (\mathbb{R}^n \times \mathbb{R}) \backslash (-K)$. Then we must have $v > -\|u\|$. We now distinguish two cases:

- $u = 0$: In this case, $v > 0$. Considering the vector $(0, 1) \in \mathbb{R}^n \times \mathbb{R}$, we have $\langle (0, 1), (u, v)\rangle = v > 0$. Since $(0, 1) \in K$, we conclude that $(u, v) \notin K^*$. Note: many students forgot to consider this case.
- $u \neq 0$: Consider the vector $(u, \|u\|) \in K$. Then

$$\begin{aligned} \langle (u, v), (u, \|u\|)\rangle &= \|u\|^2 + \|u\| \cdot v \\ &> \|u\|^2 - \|u\|^2 && \text{[because } v > -\|u\| \text{ and } \|u\| \neq 0\text{]} \\ &= 0. \end{aligned}$$

  Since $(u, \|u\|) \in K$, we conclude that $(u, v) \notin K^*$.

Combining these two cases and that we have already proved $-K \subseteq K^*$, we must have $-K = K^*$.