Special Topics in Management Science 26:711:685

Convex Analysis and Optimization

Fall 2023 Rutgers University Prof. Eckstein

Solutions to Homework 5 / Take-Home Midterm

1. We first consider the case $h(x) < 0$. Since h is continuous, it follows that there exists $\epsilon > 0$ such that $h(x') < 0$ for all x' such that $\|x' - x\| < \epsilon$. Therefore $B(x, \epsilon) \subset L(0, h)$, where $B(x, \epsilon)$ denotes the open ball

$$B(x, \epsilon) \doteq \{ x' \in \mathbb{R}^n \mid \|x' - x\| < \epsilon \}.$$

We can then deduce that $F_{L(0,h)}(x) \supseteq \{ \beta(x' - x) \mid x' \in B(x, \epsilon), \beta \geq 0 \} = \mathbb{R}^n$, and we have $F_{L(0,h)}(x) = \mathbb{R}^n$. Therefore, using the convexity of $L(0, h)$, we have $N_{L(0,h)}(x) = \big(F_{L(0,h)}(x)\big)^* = (\mathbb{R}^n)^* = \{0\}$.

If $h(x) > 0$, then $x \notin L(0, h)$, and so $N_{L(0,h)}(x) = \emptyset$, since the normal cone of set is always empty at points outside the set.

It remains to consider the the case $h(x) = 0$. Again, we will attempt to determine the cone of feasible directions $F_{L(0,h)}(x)$. Consider any direction $d \in \mathbb{R}^n$. If $\langle \nabla h(x), d \rangle > 0$, then $h(x + \alpha d) > h(x) = 0$ for all sufficiently small $\alpha > 0$.[1] Therefore $d \notin F_{L(0,h)}(x)$. On the other hand, suppose that $\langle \nabla h(x), d \rangle < 0$. Then for all sufficiently small $\alpha > 0$, we must have $h(x + \alpha d) < h(x) = 0$ and thus $x + \alpha d \in L(0, h)$. Therefore, $d \in F_{L(0,h)}(x)$. Combining these two cases, we conclude that

$$\{ d \in \mathbb{R}^n \mid \langle \nabla h(x), d \rangle < 0 \} \subseteq F_{L(0,h)}(x) \subseteq \{ d \in \mathbb{R}^n \mid \langle \nabla h(x), d \rangle \leq 0 \}. \qquad (1)$$

Let L be the set on the left and R the set on the right in (1), so that $L \subseteq F_{L(0,h)}(x) \subseteq R$. Using the assumed-to-exist point $\bar{x}$ with $h(\bar{x}) < 0$, we have by convexity of h that

$$0 > h(\bar{x}) \geq h(x) + \langle \nabla h(x), \bar{x} - x \rangle = \langle \nabla h(x), \bar{x} - x \rangle.$$

So, the direction $d = \bar{x} - x$ has the property $\langle \nabla h(x), d \rangle < 0$, meaning that

$$d \in L \quad \Rightarrow \quad L \neq \emptyset \quad \Rightarrow \quad \nabla h(x) \neq 0.$$

Therefore,

$$\operatorname{cl} L = \operatorname{cl} \{ d \in \mathbb{R}^n \mid \langle \nabla h(x), d \rangle < 0 \} = \{ d \in \mathbb{R}^n \mid \langle \nabla h(x), d \rangle \leq 0 \}, = R.$$

so it is safe to conclude from (1) that

$$\operatorname{cl} F_{L(0,h)}(x) = R = \{ d \in \mathbb{R}^n \mid \langle \nabla h(x), d \rangle \leq 0 \}.$$

[1] In this situation, one can easily use the convexity of the function show that $h(x + \alpha d) > h(x) = 0$ for $\alpha > 0$, but this fact is not necessary for the rest of the proof.

(Note that without the assumption about $\bar{x}$, it is possible that $\nabla h(x) = 0$, so $L = \emptyset$ and $R = \mathbb{R}^n$, and it would not be possible to conclude anything useful from (1).)

Now, from the convexity of $L(0, h)$, we have

$$N_{L(0,h)}(x) = \left(F_{L(0,h)}(x)\right)^* = \left(\operatorname{cl} F_{L(0,h)}(x)\right)^* = \{d \in \mathbb{R}^n \mid \langle \nabla h(x), d \rangle \le 0\}^*.$$

Using homework 3, problem 5(b) with $A = \nabla h(x)^\top$ and $C = \mathbb{R}_-$, we conclude that

$$\begin{aligned} N_{L(0,h)}(x) &= \{d \mid Ad \in C\}^* \\ &= \operatorname{cl}\{A^\top \alpha \mid \alpha \in C^*\} \\ &= \operatorname{cl}\{\nabla h(x)\alpha \mid \alpha \in \mathbb{R}_+\} \\ &= \{\alpha h(x) \mid \alpha \ge 0\}, \end{aligned}$$

where we can drop the closure operation because the set involved is a single ray and thus clearly closed (it is also clearly finitely generated).

2. Let $Z = \{x \in \mathbb{R}^n \mid Ax = b\}$, and define the following convex funtions

$$\begin{aligned} f_0(x) &= f(x) \\ f_j(x) &= \delta_{L(0,h_j)}(x) = \begin{cases} 0, & \text{if } h_j(x) \le 0 \\ +\infty, & \text{if } h_j(x) > 0 \end{cases} \qquad\qquad j = 1, \ldots, r \\ f_{r+1}(x) &= \delta_Z(x) = \begin{cases} 0, & \text{if } Ax = b \\ +\infty, & \text{if } Ax \ne b \end{cases} \\ f_{r+2}(x) &= \delta_X(x) = \begin{cases} 0, & \text{if } x \in X \\ +\infty, & \text{if } x \notin X, \end{cases} \end{aligned}$$

Note that $(f_0 + \cdots + f_{r+2})(x) = +\infty$ if x violates any of the constraints in (1), and otherwise $(f_0 + \cdots + f_{r+2})(x) = f(x)$. Therefore, solving (1) is equivalent to minimizing $f_0 + \cdots + f_{r+2}$ over $\mathbb{R}^n$, or equivalently solving

$$0 \in \partial(f_0 + \cdots + f_{r+2})(x).$$

Note that

$$\begin{aligned} \operatorname{ri}\operatorname{dom} f_0 &= \operatorname{ri}\operatorname{dom} f \\ \operatorname{ri}\operatorname{dom} f_j &= \{x \in \mathbb{R}^n \mid h_j(x) < 0\} \qquad j = 1, \ldots, r \qquad \text{(as proved in class)} \\ \operatorname{ri}\operatorname{dom} f_{r+1} &= \operatorname{ri} Z = Z \qquad\qquad\qquad\qquad\qquad \text{(since } Z = \operatorname{aff} Z\text{)} \\ \operatorname{ri}\operatorname{dom} f_{r+2} &= \operatorname{ri} X. \end{aligned}$$

The point $\overline{x}$ stipulated in the assumption lies in all these sets, and so we have

$$\operatorname{ri}\operatorname{dom} f_0 \cap \operatorname{ri}\operatorname{dom} f_1 \cap \cdots \cap \operatorname{ri}\operatorname{dom} f_{r+2} \ne \emptyset,$$

The many-function version of the Rockefellar-Moreau theorem then guarantees that

$$\partial(f_0 + \cdots + f_{r+2})(x) = \partial f_0(x) + \cdots + \partial f_{r+2}(x)$$

for all $x \in \mathbb{R}^n$, and so a necessary and sufficient condition for x^* to be optimal is

$$0 \in \partial f_0(x^*) + \cdots + \partial f_{r+2}(x^*).$$

We proved in problem 1 that

$$\partial f_j(x) = \{\mu_j \nabla h_j(x) \mid \mu_j \geq 0,\ \mu_j h_j(x) = 0\},$$

and in class that

$$\partial f_{r+1}(x) = \begin{cases} \{A^\top \lambda \mid \lambda \in \mathbb{R}^m\}, & \text{if } Ax = b \\ \emptyset & \text{if } Ax \neq b. \end{cases}$$

We also know that $\partial f_{r+2}(x) = \partial \delta_X(x) = N_X(x)$. We thus have that $0 \in \partial f_0(x^*) + \cdots + \partial f_{r+2}(x^*)$ if and only if it satisfies all the constraints in (1) and

$$0 \in \partial f(x^*) + \sum_{j=1}^{r} \{\mu_j \nabla h_j(x^*) \mid \mu_j \geq 0,\ \mu_j h_j(x^*) = 0\} + \{A^\top \lambda \mid \lambda \in \mathbb{R}^m\} + N_X(x).$$

This means in turn that there must exist $\mu^* \in \mathbb{R}^r$ and $\lambda^* \in \mathbb{R}^m$ with $\mu^* \geq 0$, $\mu_j^* h_j(x^*) = 0$ for $j = 1 \ldots, r$, and

$$0 \in \partial f(x^*) + \sum_{j=1}^{r} \mu_j^* \nabla h_j(x^*) + A^\top \lambda^* + N_X(x). \tag{2}$$

Considering that all the μ_j^* are nonnegative, while all $h_j(x^*)$ must be nonpositive at a feasible solution, the complementary slackness condition that $\mu_j^* h_j(x^*) = 0$ for all $j = 1, \ldots, r$ is the same as the constraint $\sum_{j=1}^{r} \mu_j^* h_j(x^*) = 0$. Coupling this observation with (2) and feasibility with respect to all the constraints, we obtain the set of conditions to be proved.

3. (a) Since K is a convex set, we know that for $x \in K$ the cone of feasible directions $F_K(x)$ takes the form

$$F_K(x) = \{\beta(w - x) \mid w \in K, \beta \geq 0\}$$

Next, we claim that

$$\{\beta(w - x) \mid z \in K, \beta \geq 0\} = \{z - \alpha x \mid z \in K, \alpha \geq 0\}. \tag{3}$$

To establish (3), let L be the set on its left and R the set on its right. Take any point $t = \beta(w - x) \in L$, where $w \in K$ and $\beta \geq 0$; it can be written as $t = \beta w - \beta x = z - \beta x$, where we set $z = \beta x$; this choice of z is in K since K is a cone and $\beta \geq 0$. Since $t = z - \beta x$, it is clear that $t \in R$, and thus we have shown $L \subseteq R$.

To prove the reverse identity, take any point $u = z - \alpha x \in R$, where $z \in K$ and $\alpha \geq 0$. If $\alpha > 0$, then we can take $\beta = \alpha$ and $w = \frac{1}{\alpha} z \in K$ (because K is a cone) in the definition of L and write $L \ni \beta(w - x) = \alpha(\frac{1}{\alpha} z - x) = z - \alpha x = u$. The other possibility is that $\alpha = 0$, and thus $u = z \in K$. Then, since $x, u \in K$ and K is a convex cone, $x + u \in K$. Setting $w = x + u \in K$ and $\beta = 1$ in the definition of L, we have $L \ni \beta(w - x) = (x + u) - x = u$. Thus, $L \supseteq R$ and in fact $L = R$. Thus, we have established

$$F_K(x) = \{ z - \alpha x \mid z \in K, \alpha \geq 0 \}.$$

(b) To now obtain expression for $N_K(x)$, we use the result proved in Homework 3, problem 1(c), showing that $(C_1 + C_2)^* = C_1^* \cap C_2^*$. In this case, we let $C_1 = K$ and $C_2 = \{ -\alpha x \mid \alpha \geq 0 \} = \{ \alpha(-x) \mid \alpha \geq 0 \}$, so that $C_1 + C_2 = F_K(x)$. We then get $C_1^* = K^*$ and, using Homework 3, question 5(c) with $A = x^\top$ and therefore $A^\top = x$, that $C_2^* = \{ y \in \mathbb{R}^n \mid \langle -x, y \rangle \leq 0 \} = \{ y \in \mathbb{R}^n \mid \langle x, y \rangle \geq 0 \}$. Thus, we have

$$N_K(x) = [F_K(x)]^* = K^* \cap \{ y \in \mathbb{R}^n \mid \langle x, y \rangle \geq 0 \} = \{ y \in K^* \mid \langle x, y \rangle = 0 \},$$

the last equality following because $x \in K$ and $y \in K^*$ imply that $\langle x, y \rangle \leq 0$.

(c) First, let us compute the cone of feasible directions $F_Z(x)$ to Z at some $x \in Z$. The members of $F_Z(x)$ are vectors $d \in \mathbb{R}^n$ such that for all sufficiently small $\delta > 0$, $x + \delta d \in Z$, that is,

$$A(x + \delta d) - b \in K \qquad \Leftrightarrow \qquad (Ax - b) + \delta A d \in K.$$

In other words, $d \in F_Z(x)$ if and only if $Ad \in F_K(Ax - b)$. Note that Z is convex because it is the preimage of the convex set K under an affine mapping. Since Z is convex,

$$\begin{aligned}
N_Z(x) &= [F_Z(x)]^* \\
&= \{ d \in \mathbb{R}^n \mid Ad \in F_K(Ax - b) \}^* \\
&= \operatorname{cl} \{ A^\top \lambda \mid \lambda \in [F_K(Ax - b)]^* \} && \text{(homework 3, problem 5(b))} \\
&= \operatorname{cl} \{ A^\top \lambda \mid \lambda \in N_K(Ax - b) \} && \text{(since K is convex)} \\
&= \operatorname{cl} \{ A^\top \lambda \mid \lambda \in K^*, \langle Ax - b, \lambda \rangle = 0 \} && \text{(by part (a)).}
\end{aligned}$$

(d) Suppose that $A\bar{x} - b \in \operatorname{ri} K$. We would like to show that $\bar{x} \in \operatorname{ri} Z$, and hence that $\operatorname{ri} Z \supseteq \{ x \in \mathbb{R}^n \mid Ax - b \in \operatorname{ri} K \}$. Since $A\bar{x} - b \in \operatorname{ri} K$, we know that for any $w \in K$, there exists some $\delta > 0$ such that $A\bar{x} - b + \delta\big(w - (A\bar{x} - b)\big) \in K$. In particular, if we take any $x \in Z$, then $w = Ax - b \in K$ and so there exists $\delta > 0$ such that

$$\begin{aligned}
& & A\bar{x} - b + \delta\big((Ax - b) - (A\bar{x} - b)\big) &\in K \\
&\Leftrightarrow & A\bar{x} - b + \delta A(x - \bar{x}) &\in K \\
&\Leftrightarrow & A\big(\bar{x} + \delta(x - \bar{x})\big) - b &\in K \\
&\Rightarrow & \bar{x} + \delta(x - \bar{x}) &\in Z.
\end{aligned}$$

The prolongation principle thus asserts that $\bar{x} \in \operatorname{ri} Z$. Note that the same chain of reasoning does not work in reverse, because there may be members of K which are not of the form $Ax - b$. Thus, having $\bar{x} \in \operatorname{ri} Z$ does not necessarily imply that $A\bar{x} - b \in \operatorname{ri} K$.

(e) Using the notation of part (b), the problem is simply to minimize $f(x)$ over $x \in Z$. Defining $g = \delta_Z$, this problem is equivalent to minimizing $f(x) + g(x)$, or equivalently finding x^* such that $\partial(f + g)(x^*) \ni 0$. Now, since $A\bar{x} - b \in \operatorname{ri} K$, the previous part of this problem asserts that $\bar{x} \in \operatorname{ri} Z = \operatorname{ri} \operatorname{dom} g$, and since we also have $\bar{x} \in \operatorname{ri} \operatorname{dom} f$, the Rockafeller-Moreau theorem asserts that $\partial(f + g)(x) = \partial f(x) + \partial g(x)$ for all $x \in \mathbb{R}^n$. Therefore, x^* is an optimal solution to the stated optimization problem if and only if $0 \in \partial f(x^*) + \partial g(x^*)$

Now, in part (c) we proved that

$$\begin{aligned} N_Z(x) &= \operatorname{cl} \{ A^\top \lambda \mid \lambda \in K^*, \langle Ax - b, \lambda \rangle = 0 \} \\ &= \operatorname{cl} \left(A^\top K^* \cap \{ A^\top \lambda \mid \langle Ax - b, \lambda \rangle = 0 \} \right). \end{aligned}$$

The second set within the closure operation above is a linear image of a linear subspace, and hence also a linear subspace and thus closed. Since $A^\top K^*$ is also closed by assumption, the intersection of the two sets above is closed and we can drop the the "cl" operation, yielding

$$\partial g(x) = N_Z(x) = \{ A^\top \lambda \mid \lambda \in K^*, \langle Ax - b, \lambda \rangle = 0 \}.$$

Plugging the above formula for $N_Z(x)$ into the condition $0 \in \partial f(x^*) + \partial g(x^*)$, we obtain the Karush-Kuhn-Tucker conditions

$$\partial f(x^*) + A^\top \lambda^* \ni 0 \qquad \lambda^* \in K^* \qquad \langle Ax^* - b, \lambda^* \rangle = 0.$$