Special Topics in Operations Research 16:711:611

Convex Analysis and Optimization

Spring 2009 Rutgers University Prof. Eckstein

Solutions to Homework 6

1. (a) The problem has the form $\min_{x \in \mathbb{R}^n} f(x) + g(Ax)$, where we let $A = I$ (the identity matrix) and $g = \delta_K$, the indicator function of the cone K. The corresponding Fenchel dual problem has the form $\max_{y \in \mathbb{R}^m} -\widehat{f}(-A^\top y) - \widehat{g}(y)$; in this case, since $A = I$ (and hence $m = n$) and $g = \delta_K$, we have the dual problem

$$\max_{y \in \mathbb{R}^n} -\widehat{f}(-y) - \widehat{\delta_K}(y).$$

So, we need to determine the form of the function $\widehat{\delta_K}$:

$$\widehat{\delta_K}(y) = \sup_{y \in \mathbb{R}^n} \{\langle x, y \rangle - \delta_K(x)\} = \sup_{x \in K} \{\langle x, y \rangle\}.$$

If $y \in K^*$, then $\langle x, y \rangle \leq 0$ for all $x \in K$, and the supremum above is 0. Otherwise, there exists some $z \in K$ such that $\langle z, y \rangle > 0$, and by taking $x = \alpha z \in K$ with $\alpha \to +\infty$, we obtain that the supremum is $-\infty$. We conclude that $\widehat{\delta_K} = \delta_{K^*}$. Thus, the dual problem has the following equivalent forms:

$$\begin{aligned} & \max_{x \in \mathbb{R}^n} \{-\widehat{f}(-y) - \delta_{K^*}(y)\} \\ \sim \quad & \max_{y \in K^*} \{\widehat{f}(-y)\} \\ \sim \quad & \max_{\lambda \in -K^*} \widehat{f}(\lambda), \end{aligned}$$

where the last form follows by substituting $\lambda = -y$.

(b) We note that $\operatorname{ri} \operatorname{dom} \delta_K = \operatorname{ri} K$, and so $\operatorname{ri} \operatorname{dom} f \cap \operatorname{ri} \operatorname{dom} \delta_K \neq \emptyset$, and thus the Rockafellar-Moreau theorem assert that for all $x \in \mathbb{R}^n$ we have

$$\partial(f + \delta_K)(x) = \partial f(x) + \partial \delta_K(x) = \partial f(x) + N_K(x).$$

Thus, a necessary and sufficient condition for x^* to be optimal for the primal problem $\min f(x) + \delta_K(x)$ is

$$0 \in \partial(f + \delta_K)(x^*) \quad \Leftrightarrow \quad 0 \in \partial f(x^*) + N_K(x^*).$$

This condition holds if and only if there exists $y^* \in N_K(x^*)$ such that $-y^* \in \partial f(x)$. Keeping in mind the assignments $A = I$ and $g = \delta_K$, these conditions are exactly the same as the sufficient conditions established in class, namely, that $-A^\top y^* \in \partial f(x^*)$ and $y^* \in \partial g(Ax^*)$.

(c) The conclusion is immediate from the observation that if V is a linear subspace, then $-V^* = -V^\perp = V^\perp$.

(d) $L(x,\lambda) = \inf_{u\in\mathbb{R}^m}\{F(x,u) - \langle\lambda,u\rangle\} = \inf_{u:\, x-u\in K}\{f(x) - \langle\lambda,u\rangle\}$. Making a change of variables $d = x-u$, hence $u = x-d$, we then obtain

$$\begin{aligned} L(x,\lambda) &= \inf_{d\in K}\{f(x) - \langle\lambda, x-d\rangle\} \\ &= \inf_{d\in K}\{f(x) - \langle\lambda, x\rangle + \langle\lambda, d\rangle\} \\ &= f(x) - \langle\lambda, x\rangle + \inf_{d\in K}\{\langle\lambda, d\rangle\} \end{aligned}$$

It is easily seen that the last infimum above is 0 if $\lambda \in -K^*$, and is otherwise $-\infty$; however, if $f(x) = +\infty$, then the first infimum and thus $L(x,\lambda)$ are $+\infty$. Therefore we may express the Lagrangian as

$$L(x,\lambda) = \begin{cases} +\infty, & \text{if } x \notin \operatorname{dom} f, \\ f(x) - \langle\lambda, x\rangle, & \text{if } x \in \operatorname{dom} f,\ \lambda \in -K^* \\ -\infty, & \text{if } x \in \operatorname{dom} f,\ \lambda \notin -K^*. \end{cases}$$

We then obtain the dual function by minimizing over x, which effectively eliminates the case $x \notin \operatorname{dom} f$:

$$\begin{aligned} q(\lambda) = F^*(0,\lambda) &= \inf_{x\in\mathbb{R}^n}\{L(x,\lambda)\} \\ &= \begin{cases} \inf_{x\in\mathbb{R}^n}\{f(x) - \langle\lambda, x\rangle\}, & \text{if } \lambda \in -K^* \\ -\infty, & \text{otherwise} \end{cases} \\ &= \begin{cases} -\widehat{f}(\lambda), & \text{if } \lambda \in -K^* \\ -\infty, & \text{otherwise,} \end{cases} \end{aligned}$$

which implies exactly the same dual problem as obtained by the Fenchel duality framework.

(e) First, we would like to show that F is closed, proper, and convex.

To show that F is closed, note that

$$\begin{aligned} \operatorname{epi} F &= \{(x,u,z) \mid x \in \mathbb{R}^n,\ x-u \in K,\ z \geq f(x)\} \\ &= \{(x,u,z) \mid (x,z) \in \operatorname{epi} f,\ x-u \in K\}. \end{aligned}$$

Take any sequence $\{(x^k, u^k, z_k)\} \subseteq \operatorname{epi} F$ converging to some limit $(x^\infty, u^\infty, z_\infty)$. Then $\{(x^k, z_k)\} \subseteq \operatorname{epi} f$, and since f is closed, $\lim_{k\to\infty}(x^k, z_k) = (x^\infty, z_\infty) \in \operatorname{epi} f$. We also have $x^k - u^k \to x^\infty - u^\infty$, and since $x^k - u^k \in K$ for all k and K is closed, it follows that $x^\infty - u^\infty \in K$. Using the second expression for $\operatorname{epi} F$ above, we then conclude that $(x^\infty, u^\infty, z_\infty) \in \operatorname{epi} F$, and so $\operatorname{epi} F$ is closed.

Next, to show F is proper, we use that f is proper, and hence that there exists $x \in \mathbb{R}^n$ such that $f(x) \in \mathbb{R}$. Since K is nonempty, we can then select any $d \in K$ and set $u = x - d$, implying $x - u = d \in K$. We then have $F(x,u) = f(x) \in \mathbb{R}$, and F must be proper.

Finally, we need to show F is convex. Take any $(x,u),(x',u') \in \mathbb{R}^n \times \mathbb{R}^n$, and $\alpha \in (0,1)$. If $x-u \notin K$ or $x'-u' \notin K$, then we have either $F(x,u)=+\infty$ or $F(x',u')=+\infty$. In these cases we trivially have

$$F(\alpha x+(1-\alpha)x', \alpha u+(1-\alpha)u') \le +\infty = \alpha F(x,u)+(1-\alpha)F(x',u'),$$

so convexity holds. Consider now the only other possible case, that $x-u$ and $x'-u'$ are both in K. In this case, we have by convexity of K that

$$[\alpha x+(1-\alpha)x'] - [\alpha u-(1-\alpha)u'] = \alpha(x-u)+(1-\alpha)(x'-u') \in K,$$

and so by the definition of F and the convexity of f that

$$\begin{aligned} F(\alpha x+(1-\alpha)x', \alpha u+(1-\alpha)u') &= f(\alpha x+(1-\alpha)x') \\ &\le \alpha f(x)+(1-\alpha)f(x') \\ &= \alpha F(x,u)+(1-\alpha)F(x',u'), \end{aligned}$$

and the convex function inequality holds in all cases. Thus, F is convex.

Note that the assumption on $\overline{\lambda}$ immediately guarantees that the problem is dual feasible. So, it remains only to show that ϕ is closed.

Next we seek to establish that the function $\phi(u)=\inf_{x\in\mathbb{R}^n}\{F(x,u)\}$ is closed by establishing the sufficient condition obtained in class, namely that

$$\begin{aligned} \Pi(\operatorname{epi} F) &= \{(u,z) \mid (x,u,z) \in \operatorname{epi} F\} \\ &= \{(u,z) \mid \exists\, x \in \mathbb{R}^n : (x,u) \in \operatorname{epi} f,\ x-u \in K\} \end{aligned}$$

is closed. Consider now any convergent sequence $(u^k, z_k) \in \Pi(\operatorname{epi} F)$, with limit (u^∞, z_∞). For each k, we see from the above expression for $\Pi(\operatorname{epi} F)$ that for each k there must exist $x^k \in \mathbb{R}^n$ such that $(x^k, z_k) \in \operatorname{epi} f$ and $x^k - u^k \in K$. We then have $f(x^k) \le z^k \to z_\infty < \infty$, and so $\{f(x^k)\}$ is bounded above. Using the assumption on f, we conclude that $\{x^k\}$ is bounded: if it were not, then $\{f(x^k)\}$ would have to have a subsequence diverging to $+\infty$, a contradiction. So $\{x^k\}$ must have a convergent subsequence, say $\{x^k\}_{\mathcal{K}}$, with limit x^∞. We then obtain $x^k - u^k \to_{\mathcal{K}} x^\infty - u^\infty$, which must be in K since K is closed. Since f is closed, we have that

$$\lim_{\substack{k\to\infty \\ k\in\mathcal{K}}} (x^k, z_k) = (x^\infty, z_\infty) \in \operatorname{epi} f$$

Thus, we have established that there exists x^∞ such that $(x^\infty, z_\infty) \in \operatorname{epi} f$ and $x^\infty - u^\infty \in K$, which means that $(u^\infty, z_\infty) \in \Pi(\operatorname{epi} F)$, and so $\Pi(\operatorname{epi} F)$ must be closed.

As proved in class, we then obtain that ϕ is closed. Since F is closed proper convex, we have assumed the existence of a dual feasible point, and ϕ is closed, asymptotic strong duality holds, as also proved in class.

2. (a) Consider any convergent sequence $\{(x^k, u^k, z_k)\} \subset \operatorname{epi} F$, with limit $(x^\infty, u^\infty, z_\infty)$. From the form of $\operatorname{epi} f$, we have $f(x^k) + g(Ax^k + u^k) \le z_k$ for all k. Therefore,

$\liminf_{k\to\infty}\big[f(x^k)+g(Ax^k+u^k)\big] \le z_\infty$. Using that f and g are closed and that $Ax^k+u^k \to Ax^\infty+u^\infty$, we then have

$$\begin{aligned} f(x^\infty)+g(Ax^\infty+u^\infty) &\le \liminf_{k\to\infty} f(x^k)+\liminf_{k\to\infty} g(Ax^k+u^k) \\ &\le \liminf_{k\to\infty}\big[f(x^k)+g(Ax^k+u^k)\big] \\ &\le z_\infty. \end{aligned}$$

This establishes that $(x^\infty,u^\infty,z_\infty)\in \operatorname{epi} F$, and hence that $\operatorname{epi} F$ is closed.

To establish that F is convex:

- The function $(x,u)\mapsto f(x)$ is convex, since f is convex and the map $(x,u)\mapsto x$ is linear.
- The function $(x,u)\mapsto g(Ax+u)$ is convex, since g is convex and the map $(x,u)\mapsto Ax+u$ is linear.
- F must be convex because it is the sum of the two convex functions $(x,u)\mapsto f(x)$ and $(x,u)\mapsto g(Ax+u)$.

Using the definition $L(x,\lambda)=\inf_{u\in\mathbb{R}^m}\{F(x,u)-\langle\lambda,u\rangle\}$ and the given form for F, we have

$$\begin{aligned} L(x,\lambda) &= \inf_{u\in\mathbb{R}^m}\{f(x)+g(Ax+u)-\langle u,\lambda\rangle\} \\ &= f(x)+\inf_{u\in\mathbb{R}^m}\{g(Ax+u)-\langle u,\lambda\rangle\} \\ &= f(x)+\inf_{v\in\mathbb{R}^m}\{g(v)-\langle v-Ax,\lambda\rangle\} \\ &= f(x)+\langle Ax,\lambda\rangle+\inf_{v\in\mathbb{R}^m}\{g(v)-\langle v,\lambda\rangle\} \\ &= f(x)+\langle Ax,\lambda\rangle-\widehat{g}(\lambda), \end{aligned}$$

where the third equality follows by making the substitution $v=Ax+u$ and thus $u=v-Ax$ (since u can range throughout $\mathbb{R}^m$, $Ax+u$ can also take any value in $\mathbb{R}^m$). Note that if $f(x)=+\infty$, then the first infimand above is $+\infty$ for all u, and $L(x,u)=+\infty$ regardless of the value of $\widehat{g}(\lambda)$.

(b) The dual objective is

$$\begin{aligned} F^*(0,\lambda) &= \inf_{x\in\mathbb{R}^n}\{L(x,\lambda)\} \\ &= \inf_{x\in\mathbb{R}^n}\{f(x)+\langle Ax,\lambda\rangle-\widehat{g}(\lambda)\} \\ &= \inf_{x\in\mathbb{R}^n}\{f(x)+\langle Ax,\lambda\rangle\}-\widehat{g}(\lambda) \\ &= \inf_{x\in\mathbb{R}^n}\{f(x)-\langle x,-A^\top\lambda\rangle\}-\widehat{g}(\lambda) \\ &= -\widehat{f}(-A^\top\lambda)-\widehat{g}(\lambda), \end{aligned}$$

precisely as obtained with the Fenchel approach.