Special Topics in Operations Research 16:711:611

Convex Analysis and Optimization

Spring 2009 Rutgers University Prof. Eckstein

Solutions to Homework 7

1. Let G be the duality-generating function of the convexified problem (3), that is,

$$G(x,u) = \begin{cases} c^\top x, & \text{if } x \in \operatorname{conv} X,\ Ax - b + u = 0 \\ +\infty, & \text{otherwise.} \end{cases}$$

We wish to show that $\operatorname{cl}\operatorname{conv} F = G$, that is, $\operatorname{cl}\operatorname{conv}\operatorname{epi} F = \operatorname{epi} G$. Denoting the members of the finite set X by $x^1, \ldots, x^N$, we have

$$\begin{aligned} \operatorname{epi} F &= \{(x,u,z) \mid x \in X,\ Ax - b + u = 0,\ z \geq c^\top x\} \\ &= \{(x, b - Ax, c^\top x + r) \mid x \in X,\ r \geq 0\} \\ &= \left\{(x^i, b - Ax^i, c^\top x^i + r) \mid i = 1, \ldots, N,\ r \geq 0\right\} \end{aligned}$$

$$\begin{aligned} \operatorname{epi} G &= \{(x,u,z) \mid x \in \operatorname{conv} X,\ Ax - b + u = 0,\ z \geq c^\top x\} \\ &= \{(x, b - Ax, c^\top x + r) \mid x \in \operatorname{conv} X,\ r \geq 0\} \end{aligned}$$

Now, any $x \in \operatorname{conv} X$ may be written $\sum_{i=1}^N \alpha_i x^i$, where we have $\alpha_1, \ldots, \alpha_N \geq 0$ and $\sum_{i=1}^N \alpha_i = 1$. So we may express any $(x,u,z) \in \operatorname{epi} G$ as

$$\begin{aligned} (x, b - Ax, c^\top x + r) &= \left(\sum_{i=1}^N \alpha_i x^i, b - A\left[\sum_{i=1}^N \alpha_i x^i\right], c^\top\left[\sum_{i=1}^N \alpha_i x^i\right] + r\right) \\ &= \left(\sum_{i=1}^N \alpha_i x^i, \sum_{i=1}^N \alpha_i (b - Ax^i), \sum_{i=1}^N \alpha_i (c^\top x^i + r)\right) \\ &= \sum_{i=1}^N \alpha_i (x^i, b - Ax^i, c^\top x^i + r), \end{aligned}$$

establishing that it is a convex combination of points from $\operatorname{epi} F$, and hence a member of $\operatorname{conv}\operatorname{epi} F$ and thus of $\operatorname{cl}\operatorname{conv}\operatorname{epi} F$. So, $\operatorname{epi} G \subseteq \operatorname{cl}\operatorname{conv}\operatorname{epi} F$.

Now consider any member of $\operatorname{conv}\operatorname{epi} F$, which must be expressible as a convex combination of points in $\operatorname{epi} F$, which in turn are expressible as $(x^i, b - Ax^i, c^\top x^i + r)$, where $1 \leq i \leq N$ and $r \geq 0$. Thus, any member of $\operatorname{conv}\operatorname{epi} F$ has the form

$$(x,u,z) = \sum_{k=1}^K \beta_k (x^{i(k)}, b - Ax^{i(k)}, c^\top x^{i(k)} + r_k),$$

where $i(k) \in \{1, \ldots, N\}$ for each k, $\beta_1, \ldots, \beta_K \geq 0$, and $\sum_{k=1,\ldots}^{K} \beta_k = 1$. We then have

$$\begin{aligned}
(x, u, z) &= \sum_{k=1}^{K} \beta_k (x^{i(k)}, b - Ax^{i(k)}, c^\top x^{i(k)} + r_k) \\
&= \left(\sum_{k=1}^{K} \beta_k x^{i(k)}, b - A\left[\sum_{k=1}^{K} \beta_k x^{i(k)}\right], c^\top \left[\sum_{k=1}^{K} \beta_k x^{i(k)}\right] + \sum_{k=1}^{K} \beta_k r_k \right) \\
&= (x, b - Ax, c^\top x + r),
\end{aligned}$$

where $x = \sum_{k=1}^{K} \beta_k x^{i(k)} \in \operatorname{conv} X$ and we let $r = \sum_{k=1}^{K} \beta_k r_k \geq 0$. Thus, the point is a member of $\operatorname{epi} G$, and we have $\operatorname{conv} \operatorname{epi} F \subseteq \operatorname{epi} G$. Since $\operatorname{conv} X$ is a polyhedral set (described by a finite number of linear constraints), we see from the form of $\operatorname{epi} G$ that it is also polyhedral, and hence closed. This means that $\operatorname{cl} \operatorname{conv} \operatorname{epi} F \subseteq \operatorname{epi} G$ as well. Therefore, $\operatorname{cl} \operatorname{conv} \operatorname{epi} F = \operatorname{epi} G$, and so $\operatorname{cl} \operatorname{conv} F = G$.

Next, we claim that the respective dual functions $F^* = -\widehat{F}$ and $G^* = -\widehat{G}$ are equal, for which it is sufficient to show that $\widehat{F} = \widehat{G}$. Note that both F and G are bounded below by $\min_{i=1,\ldots,N}\{c^\top x^i\}$, so neither can take the value $-\infty$, and both must have an affine minorant. In view of this fact, we note that since $\operatorname{epi} G$ is nonempty and polyhedral, G must be closed, proper and convex. Because G is closed proper convex, both F and G have affine minorants, and we have proved $\operatorname{cl} \operatorname{conv} F = G$, we respectively have

$$\widehat{\widehat{G}} = G \qquad\qquad \widehat{\widehat{F}} = \operatorname{cl} \operatorname{conv} F = G$$

Because they were each obtained by taking a conjugate, we also know that $\widehat{F}$ and $\widehat{G}$ are closed proper convex functions, which means that

$$\widehat{\widehat{\widehat{G}}} = \widehat{G} \qquad\qquad \widehat{\widehat{\widehat{F}}} = \widehat{F}.$$

Since we know $\widehat{\widehat{G}} = \widehat{\widehat{F}} = G$, it then follows that $\widehat{G} = \widehat{F}$, and so $F^* = G^*$.

Note that since $\operatorname{epi} G$ is polyhedral, it is clear that $\Pi(\operatorname{epi} G)$ is also polyhedral and thus closed. We have already shown that G is closed proper convex, and it is clearly also dual feasible, since $\operatorname{conv} X$ must be bounded, and so the corresponding dual function $G^*(0, \lambda) = \inf_{x \in \operatorname{conv} X}\{c^\top x + \langle \lambda, Ax - b\rangle\}$ cannot be $-\infty$. Thus, strong duality holds for G, and the maximum value of $G^*(0, \cdot)$ is exactly the optimum value of (3). Because $F^* = G^*$, the maximum value of $F^*(0, \cdot)$ is also the optimum value of (3). Of course, this is usually lower than the optimum value $\inf_x\{F(x, 0)\}$ of the corresponding discrete problem.

2. In the general Rockafellar duality setup, as covered in class, we can find a subgradient of $-q(\lambda) = -F^*(0, \lambda) = \widehat{F}(0, \lambda)$ as follows:
   - Find some x^{k+1} attaining the infimum over x in $q(\lambda^k) = \inf_x\{L(x, \lambda^k)\}$.
   - Identify some u^{k+1} attaining the minimum over u in

$$L(x^{k+1}, \lambda^k) = \inf_u \{F(x^{k+1}, u) - \langle \lambda^k, u\rangle\}$$

- This u^{k+1} is a subgradient of $-q$ at λ^k.

In this particular case, the Lagrangian has the form

$$L(x, \lambda) = f(x) + \langle \lambda, Ax \rangle - \widehat{g}(\lambda),$$

as proved on the previous homework. When minimizing $L(x, \lambda^k)$ with respect to x in the first step of the procedure above, we may ignore the constant $-\widehat{g}(\lambda^k)$, obtaining

$$x^{k+1} \in \text{Arg min}_{x \in \mathbb{R}^n} \{f(x) + \langle \lambda^k, Ax \rangle\} = \text{Arg min}_{x \in \mathbb{R}^n} \{f(x) + \langle A^\top \lambda^k, x \rangle\}.$$

Now, we must find u^{k+1} attaining the infimum in the formula the Lagrangian, that is, $L(x^{k+1}, \lambda^k) = \inf_u \{F(x^{k+1}, u) - \langle \lambda^k, u \rangle\}$. In this case, we have

$$\inf_u \{f(x^{k+1}) + g(Ax^{k+1} + u) - \langle \lambda^k, u \rangle\},$$

or, ignoring the constant $f(x^{k+1})$,

$$\inf_u \{g(Ax^{k+1} + u) - \langle \lambda^k, u \rangle\}.$$

Making a change of variables $y = Ax^{k+1} + u$, hence $u = y - Ax^{k+1}$, we can equivalently find y^{k+1} attaining the inf in either of

$$\inf_y \{g(y) - \langle \lambda^k, y - Ax^{k+1} \rangle\} = \inf_y \{g(y) - \langle \lambda^k, y \rangle\} + Ax^{k+1},$$

and then set $u^{k+1} = y^{k+1} - Ax^{k+1}$. Once we have identified this subgradient u^{k+1}, we calculate

$$\begin{aligned} \lambda^{k+1} &= \lambda^k - \alpha_k u^{k+1} \\ &= \lambda^k - \alpha_k (y^{k+1} - Ax^{k+1}) \\ &= \lambda^k + \alpha_k (Ax^{k+1} - y^{k+1}). \end{aligned}$$

Thus, the overall subgradient iteration may be written

$$\begin{aligned} x^{k+1} &\in \text{Arg min}_x \{f(x) + \langle A^\top \lambda^k, x \rangle\} \\ y^{k+1} &\in \text{Arg min}_y \{g(y) - \langle \lambda^k, y \rangle\} \\ \lambda^{k+1} &= \lambda^k + \alpha_k (Ax^{k+1} - y^{k+1}). \end{aligned}$$

Note that this calculation acheives *decomposition*: we have two separate minimands, one involving f, and the other involving g, but we do not have to minimize any expression simultaneously involving both f and g.