Special Topics in Management Science 26:711:685

Convex Analysis and Optimization

Fall 2023 Rutgers University Prof. Eckstein

Solutions to Homework 6

1. (a) $f^*(y) = \sup_{x\in\mathbb{R}}\{xy - \frac{1}{2}x^2\}$. The supremand is a differentiable concave function, so we just set $\frac{\partial}{\partial x}(xy - \frac{1}{2}x^2) = 0$ obtaining $y - x = 0$, that is, $x = y$. Substituing $x = y$ into $xy - \frac{1}{2}x^2$, we have $f^*(y) = y^2 - \frac{1}{2}y^2 = \frac{1}{2}y^2$. Thus, in this case, $f^* = f$. Aside: in $\mathbb{R}^n$, $f(x) = \frac{1}{2}\|x\|^2$ is the unique function for which $f^* = f$.

   (b) In this case, we have

$$\begin{aligned} f^*(y) &= \sup_{x\in[a,b]}\{xy - 0\} \\ &= \begin{cases} ay, & \text{if } y < 0 \\ 0, & \text{if } y = 0 \\ by, & \text{if } y > 0 \end{cases} \\ &= \max\{ay, by\}. \end{aligned}$$

   Note: this example is a special case of problem 2(b) below.

   (c) Here, we have $f^*(y) = \sup_{x\in\mathbb{R}}\{xy - e^x\}$. Note that if $y < 0$, then we have $\lim_{x\to-\infty}(xy - e^x) = +\infty$, and so the supremum is $+\infty$. For $y = 0$, we have $f^*(0) = \sup_{x\in\mathbb{R}}\{-e^x\} = 0$ (although this sup is not attained). For $y > 0$, we note that we are maximizing a differentiable concave function, and so try setting the derivative of the supremand to 0:

$$\frac{\partial}{\partial x}(xy - e^x) = 0 \quad\Leftrightarrow\quad y - e^x = 0 \quad\Leftrightarrow\quad y = e^x \quad\Leftrightarrow\quad x = \log y,$$

   where log denotes the natural logarithm. Substituting $x = \log y$ into $xy - e^x$, we obtain $f^*(y) = (\log y)y - e^{\log y} = y\log y - y$. Summarizing,

$$f^*(y) = \begin{cases} y\log y - y, & y > 0, \\ 0, & y = 0, \\ +\infty, & y < 0. \end{cases}$$

   Note: using L'Hôpital's rule, it is possible to show that $\lim_{y\downarrow 0}(y\log y - y) = 0$, so f^* is right-continuous at 0; however, it turns out that f^* is not differentiable at 0, nor does it have any subgradients there, which is a result of the sup in the conjugate formula not being attained. Essentially, as $y \downarrow 0$, the function $y\log y - y$ approaches 0, but its slope becomes vertical and its derivative approaches $-\infty$. The function $y\log y$ is well known in physics and statistics, and is often known as the *Gibbs entropy* function.

2. (a) Note that epi f is a cone if and only if, for any $\alpha \geq 0$,

$$z \geq f(x) \quad \Rightarrow \quad \alpha z \geq f(\alpha x).$$

If $f(0) = 0$, the above is clearly true whenever $\alpha = 0$. And if $f(\alpha x) = \alpha f(x)$ for $\alpha > 0$, the above is also true whenever $\alpha > 0$ since we have that $z \geq f(x)$ implies $\alpha z \geq \alpha f(x) = f(\alpha x)$. Thus, positive homogeneity of f implies that epi f is a cone.

Conversely, suppose epi f is a cone and f is proper. Since epi f must be nonempty and we define nonempty cones to always contain 0, we have $(0, 0) \in \operatorname{epi} f$ and hence $f(0) \leq 0$. But if $f(0) < 0$, we have from epi f being a cone that all points of the form $(\alpha \cdot 0, \alpha f(0)) = (0, \alpha f(0)) \in \operatorname{epi} f$ for $\alpha > 0$, implying $f(0) = -\infty$ and contradicting that it is supposed to be a proper function mapping $\mathbb{R}^n \to (-\infty, +\infty]$. So $f(0) = 0$.

Next consider any $x \in (\operatorname{dom} f)\backslash\{0\}$ and $\alpha > 0$. We have $(x, f(x)) \in \operatorname{epi} f$, and so epi f being a cone implies $(\alpha x, \alpha f(x)) \in \operatorname{epi} f$ and thus $f(\alpha x) \leq \alpha f(x)$. But we also have $(\alpha x, f(\alpha x)) \in \operatorname{epi} f$ and hence $(\frac{1}{\alpha}\alpha x, \frac{1}{\alpha} f(\alpha x)) = (x, \frac{1}{\alpha} f(\alpha x)) \in \operatorname{epi} f$, implying $f(x) \leq \frac{1}{\alpha} f(\alpha x)$, that is, $\alpha f(x) \leq f(\alpha x)$. It follows that $f(\alpha x) = \alpha f(x)$.

Finally, consider $x \notin \operatorname{dom} f$ and $\alpha > 0$. Then $(x, z) \notin \operatorname{epi} f$ for any $z \in \mathbb{R}$. Then we cannot have $(\alpha x, w) \in \operatorname{epi} f$ for any $w \in \mathbb{R}$, because otherwise from epi f being a cone we would have $\frac{1}{\alpha}(\alpha x, w) = (x, \frac{w}{\alpha}) \in \operatorname{epi} f$, a contradiction. Thus, if $f(x) = \infty$, we also have $f(\alpha x) = \infty$ and $f(\alpha x) = \alpha f(x)$ holds in this case too.

We conclude that f is positively homogeneous.

(b) We have $\delta_X^*(y) = \sup_{x \in \mathbb{R}^n}\{\langle x, y\rangle - \delta_X(x)\}$, but since $\delta_X(x) = \infty$ for $x \notin X$, the supremand is $-\infty$ whenever $x \notin X$, so we can restrict the "sup" to $x \in X$. For $x \in X$, we also have $\delta_X(x) = 0$, reducing the supremand to $\langle x, y\rangle$. In summary, $\delta_X^*(y) = \sup_{x \in X}\{\langle x, y\rangle\}$. Since X is nonempty, we clearly have $\delta_X^*(0) = \sup_{x \in X}\{\langle x, 0\rangle\} = 0$. Finally, we note that for $\alpha > 0$,

$$\delta_X^*(\alpha y) = \sup_{x \in X}\{\langle x, \alpha y\rangle\} = \sup_{x \in X}\{\alpha\langle x, y\rangle\} = \alpha \sup_{x \in X}\{\langle x, y\rangle\} = \alpha \delta_X^*(\alpha y),$$

establishing that δ_X^* is positively homogeneous.

(c) Suppose f is positively homogeneous. Then $f^*(y) = \sup_{x \in \mathbb{R}^n}\{\langle x, y\rangle - f(x)\} \geq \langle x, 0\rangle - f(0) = 0$, and f^* must be nonnegative everywhere. Now, if $f^*(y) > 0$, it means that there exists an $x \in \mathbb{R}^n$ with $\langle x, y\rangle - f(x) > 0$. Applying positive homogeneity, we have for any $\alpha > 0$ that $\langle \alpha x, y\rangle - f(\alpha x) = \alpha\langle x, y\rangle - \alpha f(x) = \alpha\big(\langle x, y\rangle - f(x)\big)$. By taking $\alpha \to \infty$, we obtain $f^*(y) = \sup_{z \in \mathbb{R}^n}\{\langle z, y\rangle - f(z)\} \geq \sup_{\alpha > 0}\{\langle \alpha x, y\rangle - f(\alpha x)\} = \sup_{\alpha > 0}\{\alpha\big(\langle x, y\rangle - f(x)\big)\} = +\infty$.

Thus, f^* can only take the values 0 or $+\infty$, meaning that it is the indicator function of the set $C = \operatorname{dom} f^*$. Being a conjugate function, f^* is convex and closed, and therefore C, being one of f^*'s level sets, must be convex and closed. Note: it is also straightforward to show that $C = \partial f(0)$.

(d) Suppose $y \in K^*$. Then, using the formula from part (b), we know that $\delta_K^*(y) = \sup_{x \in K}\{\langle x, y\rangle\} = 0$ because every member of $\{\langle x, y\rangle \mid x \in K\}$ must be nonpositive, and setting $x = 0 \in K$ yields a value of 0. Conversely, if $y \notin K^*$, there exists

$\bar{x} \in K$ such that $\langle \bar{x}, y \rangle > 0$, and since K is a cone, we have $\alpha\bar{x} \in K$ for all $\alpha \geq 0$. Letting $\alpha \to \infty$ then shows that $\delta_K^*(y) = \sup_{x \in K}\{\langle x, y \rangle\} \geq \sup_{\alpha \geq 0}\{\langle \alpha\bar{x}, y \rangle\} = \sup_{\alpha \geq 0}\{\alpha\langle \bar{x}, y \rangle\} = +\infty$. Thus, $\delta_K^*(y)$ is 0 for $y \in K^*$ and $+\infty$ for $y \notin K^*$, that is, $\delta_K^* = \delta_{K^*}$.

3. The Fenchel dual takes the form $\min_{y \in \mathbb{R}^m} g^*(y) + f^*(-M^\top y)$. So, we need to compute f^* and g^*. For f^*, we have

$$
\begin{aligned}
f^*(y) &= \sup_{x \in \mathbb{R}^n} \big\{ \langle x, y \rangle - f(x) \big\} \\
&= \sup_{x \geq 0} \big\{ \langle x, y \rangle - \langle x, c \rangle \big\} \\
&= \sup_{x \geq 0} \big\{ \langle x, y - c \rangle \big\} \\
&= \begin{cases} 0, & \text{if } y \leq c \\ +\infty, & \text{otherwise.} \end{cases}
\end{aligned}
$$

The last equality is because, if $y_j > c_j$ for any index j, then the sup can be driven to $+\infty$ by taking $x_j \to \infty$ and thus $(y_j - c_j)x_j \to \infty$. If every element of $y - c$ is nonpositive, on the other hand, the largest possible value is zero, attained at $x = 0$.

Finding g^* is a special case of problem 2b with $X = \{b\}$, so we have

$$
g^*(w) = \sup_{x = b} \big\{ \langle w, b \rangle \big\} = \langle w, b \rangle.
$$

So, the dual function $g^*(y) + f^*(-M^\top y) = g^*(y) + f^*(-A^\top y)$, takes the form

$$
g^*(y) + f^*(-A^\top y) = \begin{cases} \langle b, y \rangle + 0, & \text{if } -A^\top y \leq c \\ +\infty, & \text{otherwise,} \end{cases}
$$

which means that minimizing it is equivalent to the problem

$$
\begin{array}{ll} \min_{y \in \mathbb{R}^m} & b^\top y \\ \text{ST} & -A^\top y \leq c, \end{array}
$$

which is turn equivalent to

$$
\begin{array}{ll} \max_{y \in \mathbb{R}^m} & -b^\top y \\ \text{ST} & -A^\top y \leq c. \end{array}
$$

Making the substitution $u = -y$ we obtain the further equivalent problem

$$
\begin{array}{ll} \max_{u \in \mathbb{R}^m} & b^\top u \\ \text{ST} & A^\top u \leq c. \end{array}
$$