Special Topics in Operations Research 16:711:611

Convex Analysis and Optimization

Spring 2009 Rutgers University Prof. Eckstein

Solutions to Homework 2

1. We start by writing

$$\left(\frac{\|z^k - x\|}{\|z^k - y\|}\right)^2 = \frac{\|z^k - x\|^2}{\|z^k - y\|^2} = \frac{\|z^k\|^2 - \langle z^k, x\rangle + \|x\|^2}{\|z^k\|^2 - \langle z^k, y\rangle + \|y\|^2}.$$

For all large enough k, we should have $\|z^k\| > \|y\|$ and hence $\|z^k\|^2 > \|z^k\|\|y\|$. Applying the Cauchy-Schwarz inequalities and using the equation above, we have that for all such k,

$$\frac{\|z^k\|^2 - \|z^k\|\|x\| + \|x\|^2}{\|z^k\|^2 + \|z^k\|\|y\| + \|y\|^2} \leq \left(\frac{\|z^k - x\|}{\|z^k - y\|}\right)^2 \leq \frac{\|z^k\|^2 + \|z^k\|\|x\| + \|x\|^2}{\|z^k\|^2 - \|z^k\|\|y\| + \|y\|^2}. \tag{1}$$

Taking $k \to \infty$ in the left-hand expression in (1),

$$\lim_{k\to\infty} \frac{\|z^k\|^2 - \|z^k\|\|x\| + \|x\|^2}{\|z^k\|^2 + \|z^k\|\|y\| + \|y\|^2} = \lim_{\alpha\to\infty} \frac{\alpha^2 - \|x\|\alpha + \|x\|^2}{\alpha^2 + \|y\|\alpha + \|y\|^2} = 1.$$

Similarly, taking $k \to \infty$ in the right-hand expression in (1),

$$\lim_{k\to\infty} \frac{\|z^k\|^2 + \|z^k\|\|x\| + \|x\|^2}{\|z^k\|^2 - \|z^k\|\|y\| + \|y\|^2} = \lim_{\alpha\to\infty} \frac{\alpha^2 + \|x\|\alpha + \|x\|^2}{\alpha^2 - \|y\|\alpha + \|y\|^2} = 1.$$

Now, the middle expression in (1) lies between two sequence that are converging to 1, so it must also converge to 1 as $k \to \infty$. The desired result follows by the continuity of the square root function at 1.

2. Suppose $0 < \alpha_1 < \alpha_2$. Then note that we can express $x + \alpha_1 d$ as a convex combination of x and $x + \alpha_2 d$ as follows:

$$\left(1 - \frac{\alpha_1}{\alpha_2}\right) x + \frac{\alpha_1}{\alpha_2}(x + \alpha_2 d) = x - \frac{\alpha_1}{\alpha_2}x + \frac{\alpha_1}{\alpha_2}x + \alpha_1 d = x + \alpha_1 d.$$

Therefore, the convexity of f yields

$$f(x + \alpha_1 d) \leq \left(1 - \frac{\alpha_1}{\alpha_2}\right) f(x) + \frac{\alpha_1}{\alpha_2} f(x + \alpha_2 d).$$

Subtracting $f(x)$ from both sides, we have

$$f(x + \alpha_1 d) - f(x) \leq \frac{\alpha_1}{\alpha_2}\left(f(x + \alpha_2 d) - f(x)\right).$$

Dividing through by $\alpha_1 > 0$, we have

$$\frac{f(x + \alpha_1 d) - f(x)}{\alpha_1} \leq \frac{f(x + \alpha_2 d) - f(x)}{\alpha_2}.$$

Since we chose $0 < \alpha_1 < \alpha_2$ arbitrarily, we must have a nondecreasing function.

3. We assume that C is nonempty; otherwise the problem does not make much sense.

(a) The proof here is essentially identical to the first part of the projection theorem; note that convexity was not used there until we showed that the solution had to be unique. In brief, you pick some fixed $\bar{w} \in C$, and consider the function

$$\bar{g}(w) = \begin{cases} \|w - x\|, & w \in C \text{ and } \|w - x\| \leq \|\bar{w} - x\| \\ +\infty, & \text{otherwise.} \end{cases}$$

This function has bounded domain and may be shown to be lower continuous, so it has a nonempty compact set of minima by the Weierstrass theorem. It is then easily shown that this set of minima is the same as the minima of g over C.

(b) Fix any $\hat{w} \in C$. For any $x \in \mathbb{R}^n$, it is clear that $0 \leq \inf_{w \in C} \|w - x\| \leq \|\hat{w} - x\|$, so we have $0 \leq d_C(x) \leq \|\hat{w} - x\|$ and d_C must be finite-valued.

Consider any convergent sequence $x^k \to \bar{x}$, and let $\bar{w} \in C$ be some point achieving the minimum of $\|w - \bar{x}\|$ over C. Since $\bar{w} \in C$, we have

$$d_C(x^k) = \inf_{w \in C} \|w - x^k\| \leq \|\bar{w} - x^k\|,$$

in which taking $k \to \infty$ yields

$$\limsup_{k \to \infty} d_C(x^k) \leq \|\bar{w} - \bar{x}\| = d_C(\bar{x}).$$

Thus, $d_C(\bar{x}) \geq \limsup_{k \to \infty} d_C(x^k)$ and d_C is upper semicontinuous.

Next, for each k, let $w^k \in C$ be some point attaining the minimum of $\|w - x^k\|$ over C. The triangle inequality gives that

$$\|w^k - \bar{x}\| \leq \|w^k - x^k\| + \|x^k - \bar{x}\|.$$

Substituting $d_C(\bar{x}) = \inf_{w \in C} \|w - \bar{x}\| \leq \|w^k - \bar{x}\|$ and $d_C(x^k) = \|w^k - x^k\|$ into this inequality produces

$$d_C(\bar{x}) \leq d_C(x^k) + \|x^k - \bar{x}\|$$

for all k. Rearranging,

$$d_C(\bar{x}) - \|x^k - \bar{x}\| \leq d_C(x^k).$$

Taking $k \to \infty$ in this inequality produces

$$d_C(\bar{x}) \leq \liminf_{k \to \infty} d_C(x^k),$$

and so d_C is also lower semicontinuous. Since it is also upper semicontinuous, d_C has to be continuous.

(c) In $\mathbb{R}^1$, consider $C = \{0, 1\}$. Then

$$d_C(x) = \begin{cases} -x, & x \le 0, \\ x, & 0 \le x \le 1/2 \\ 1 - x, & 1/2 \le x \le 1 \\ x - x, & x \ge 1. \end{cases}$$

The graph of this function looks like a "W" and is clearly not convex. In particular, $d_C(0) = 0$, $d_C(1/2) = 1/2$, and $d_C(1) = 0$, so we have $d_C(1/2) > (1/2)d_C(0) + (1/2)d_C(1) = 0$, violating the convexity inequality.

4. (a) The epigraph of δ_X is $X \times [0, \infty)$. If X is closed, this is a Cartesian product of closed sets, and hence closed.

   (b) Again, $\operatorname{epi} f = X \times [0, \infty)$. If X is convex, $\operatorname{epi} f$ is the Cartesian product of convex sets and hence convex.

5. (a) Plugging $w = x$ into the "inf", we have $\hat{f}_\lambda(x) \le f(x) + (1/2\lambda)\|x - x\|^2 = f(x)$.

   (b) From part (a), we now know that $\hat{f}_\lambda(x^*) \le f(x^*)$, so we need only eliminate the possibility that $\hat{f}_\lambda(x^*) < f(x^*)$. If that were the case, then there would have to exist $w \in \mathbb{R}^n$ such that $f(x) + (1/2\lambda)\|w - x\|^2 < f(x^*)$. Rearranging and using that $\|w - x\|^2 \ge 0$, we have

$$f(x) < f(x^*) - (1/2\lambda)\|w - x\|^2 \le f(x^*),$$

yielding $f(x) < f(x^*)$ and contradicting the supposed global optimality of x^*. Therefore, $\hat{f}_\lambda(x^*) = f(x^*)$.

   (c) First, if f is proper, there exists $\bar{x} \in \mathbb{R}^n$ such that $f(\bar{x}) < \infty$. Therefore, for any $x \in \mathbb{R}^n$,

$$\hat{f}_\lambda(x) = \inf_{w \in \mathbb{R}^n} \left\{ f(w) + \frac{1}{2\lambda}\|w - x\|^2 \right\} \le f(\bar{x}) + \frac{1}{2\lambda}\|\bar{x} - x\|^2 < \infty,$$

so f can never take the value $+\infty$. If f is also convex, we now want to show that f cannot take the value $-\infty$. Take an $x \in \mathbb{R}^n$, and consider the compact set $B = \{y \in \mathbb{R}^n \mid \|y - x\|\} \le 1$. Since f is closed and proper, it attains its minimum value $L = \inf\{f(x) \mid x \in B\} > -\infty$ on B by the Weierstrass theorem, and we also know from the immediately preceding result that $L < \infty$. Now consider any point $w \in \mathbb{R}^n \backslash B$, that is, with $\|w - x\| > 1$. Consider the point $z = x + (1/\|w - x\|)(w - x)$, which has $\|z - x\| = \|(1/\|w - x\|)(w - x)\| = 1$ and is hence in B. We can write z as a convex combination of x and w as follows:

$$\left(1 - \frac{1}{\|w - x\|}\right) x + \frac{1}{\|w - x\|} w = x + \frac{1}{\|w - x\|}(w - x) = z.$$

Consequently, the convexity of f yields

$$f(z) \le \left(1 - \frac{1}{\|w - x\|}\right) f(x) + \frac{1}{\|w - x\|} f(w).$$

Rearranging this inequality, we obtain

$$\frac{1}{\|w-x\|}f(w) \geq f(z) - \left(1 - \frac{1}{\|w-x\|}\right) f(x),$$

which we multiply throught by $\|w-x\| > 1$ to obtain

$$\begin{aligned} f(w) &\geq \|w-x\| f(z) + (\|w-x\| - 1) f(x) \\ &= f(x) + \|w-x\| (f(z) - f(x)) \\ &= f(x) - \|w-x\| (f(x) - f(z)) \\ &\geq f(x) - \|w-x\| (f(x) - L), \end{aligned}$$

the second inequality following because $z \in B$ and thus $f(z) \geq L$. Adding $1/2\lambda\|w-x\|^2$ to both sides,

$$\begin{aligned} f(x) + \frac{1}{2\lambda}\|w-x\|^2 &\geq f(x) - \|w-x\| (f(x) - L) + \frac{1}{2\lambda}\|w-x\|^2 \\ &\geq \inf_{\alpha > 1} \left\{ f(x) - (f(x) - L)\alpha + \frac{1}{2\lambda}\alpha^2 \right\} \\ &\geq \inf_{\alpha \in \mathbb{R}} \left\{ f(x) - (f(x) - L)\alpha + \frac{1}{2\lambda}\alpha^2 \right\} \end{aligned}$$

The infimand in this last expression is a differentiable convex function of α, so we may obtain a global minimum by setting the derivative to 0. This operation gives

$$-(f(x) - L) + \alpha/\lambda = 0 \qquad \Leftrightarrow \qquad \alpha = \lambda(f(x) - L).$$

Plugging in this value of α, we obtain a minimum value of

$$f(x) - \lambda(f(x) - L)^2 + \frac{\lambda^2}{2\lambda}\lambda(f(x) - L)^2 = f(x) - \frac{\lambda}{2}(f(x) - L)^2$$

In conclusion, for all w such that $\|w-x\| > 1$, we have

$$f(x) + \frac{1}{2\lambda}\|w-x\|^2 \geq f(x) - \frac{\lambda}{2}(f(x) - L)^2$$

On the other hand, for all w with $\|w-x\| \leq 1$, we have

$$f(x) + \frac{1}{2\lambda}\|w-x\|^2 \geq f(x) \geq L.$$

Combining these two results,

$$\begin{aligned} \hat{f}_\lambda(x) &= \inf_{w \in \mathbb{R}^n} \left\{ f(w) + \frac{1}{2\lambda}\|w-x\|^2 \right\} \\ &= \min \left\{ \inf_{\|w-x\| \leq 1} \left\{ f(w) + \frac{1}{2\lambda}\|w-x\|^2 \right\}, \inf_{\|w-x\| > 1} \left\{ f(w) + \frac{1}{2\lambda}\|w-x\|^2 \right\} \right\} \\ &\geq \min \left\{ L, f(x) - \frac{\lambda}{2}(f(x) - L)^2 \right\} \\ &> -\infty, \end{aligned}$$

the last inequality following because we showed that L must be finite.

(d) From the definitions,

$$\begin{aligned}
\widehat{\delta_{C\lambda}}(x) &= \inf_{w\in\mathbb{R}^n} \left\{ \delta_C(w) + \frac{1}{2\lambda} \|w-x\|^2 \right\} \\
&= \inf_{w\in C} \left\{ \frac{1}{2\lambda} \|w-x\|^2 \right\} \\
&= \frac{1}{2\lambda} \inf_{w\in C} \left\{ \|w-x\|^2 \right\} \\
&= \frac{1}{2\lambda} \left(\inf_{w\in C} \{ \|w-x\| \} \right)^2 \\
&= \frac{1}{2\lambda} d_C(x)^2.
\end{aligned}$$

Solving $\widehat{\delta_{C\lambda}}(x) = (1/2\lambda) d_C(x)^2$ for $d_C(x)$ (and using that $d_C(x)$ cannot be negative) yields the desired result.