Special Topics in Operations Research 16:711:611

Convex Analysis and Optimization

Spring 2009 Rutgers University Prof. Eckstein

Homework 4

1. Consider a proper convex function $f : \mathbb{R}^n \to (\infty, +\infty]$, and its subgradient mapping ∂f from $\mathbb{R}^n$ to subsets of $\mathbb{R}^n$ (sometimes denoted $2^{\mathbb{R}^n}$). That is,

$$\partial f : x \in \mathbb{R}^n \mapsto \{ d \in \mathbb{R}^n \mid f(y) \geq f(x) + \langle d, y - x \rangle \; \forall \, y \in \mathbb{R}^n \} .$$

Show that the point-to-set mapping ∂f has the following property:

$$\left. \begin{array}{r} y \in \partial f(x) \\ y' \in \partial f(x') \end{array} \right\} \quad \Rightarrow \quad \langle x - x', y - y' \rangle \geq 0.$$

Note: point-to-set mappings with this property are called *monotone*.

2. (a) For a nonempty convex set C, define the point-to-set map

$$N_C(x) = \begin{cases} \{ d \in \mathbb{R}^n \mid \langle d, y - x \rangle \leq 0 \; \forall \, y \in C \} , & \text{if } x \in C \\ \emptyset, & \text{if } x \notin C. \end{cases}$$

Show that $N_C = \partial \delta_C$, where δ_C is the indicator function defined in homework assignment 2, that is,

$$\delta_C(x) = \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{if } x \notin C. \end{cases}$$

(b) Suppose that $f : \mathbb{R}^n \to (-\infty, +\infty]$ is a convex function, $x \in \mathbb{R}^n$, $w \in \partial f(x)$, and $d \in N_{\mathrm{dom}\, f}(x)$, where the $N_{(\cdot)}$ operation is as defined in part (a). Show that we also have $w + d \in \partial f(x)$.

(c) Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper convex function, and let $x \in \mathbb{R}^n$ be a point such that $\partial f(x) \neq \emptyset$. Show that $\partial f(x)$ is compact if and only if $x \in \mathrm{int}\,\mathrm{dom}\, f$. (Hint: for the "if" part, adapt the proof given in class for real-valued convex functions.)

(d) Show that if U is a linear subspace of $\mathbb{R}^n$, then $N_U(x) = U^\perp$ for all $x \in U$, where $U^\perp$ denotes the subspace orthogonal to U (by definition, $N_U(x) = \emptyset$ if $x \notin U$).

3. Give an example of how, when $\mathrm{ri}\,\mathrm{dom}\, f_1 \cap \mathrm{ri}\,\mathrm{dom}\, f_2 = \emptyset$, one may have

$$\partial(f_1 + f_2)(x) \supsetneq \partial f_1(x) + \partial f_2(x)$$

for two closed proper convex functions $f_1, f_2 : \mathbb{R}^n \to (\infty, +\infty]$. That is, at some point $x \in \mathbb{R}^n$, there exist subgradients of $f_1 + f_2$ that cannot be expressed as the sum of a subgradient of f_1 at x and a subgradient of f_2 and x.

4. Using the same definition of the $N_{(\cdot)}$ operation as in problem 2(a), prove the following generalization of Proposition 4.7.2 in the textbook: let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper convex function, and $C \subseteq \mathbb{R}^n$ a nonempty convex set. If $\operatorname{ri}\operatorname{dom} f \cap \operatorname{ri} C \neq \emptyset$, then a point $x^* \in C$ minimizes f over C if and only if

$$0 \in \partial f(x^*) + N_C(x^*),$$

and this condition is equivalent to the existence of some $d \in \partial f(x^*)$ such that

$$\langle d, y - x \rangle \geq 0 \quad \forall\, y \in C.$$

5. In this problem, we will prove the following "almost industrial strength" generalization of Proposition 4.2.5(a): let $f : \mathbb{R}^m \to (-\infty, +\infty]$ be a proper convex function and let A be an $m \times n$ matrix. Define $g(x) = f(Ax)$, which is also a convex function. Then $\partial g(x) \supseteq A^\top \partial f(Ax)$ for all $x \in \mathbb{R}^n$. Suppose further that $\operatorname{ri}\operatorname{dom} f \cap \operatorname{im} A \neq \emptyset$, that is, there exists some point in $z \in \operatorname{ri}\operatorname{dom} f$ that may be expressed as $z = Ax$ for some $x \in \mathbb{R}^n$. Then, for any $x \in \mathbb{R}^n$, we have $\partial g(x) = A^\top \partial f(Ax)$.

   (a) Show (without assuming $\operatorname{ri}\operatorname{dom} f \cap \operatorname{im} A \neq \emptyset$) that $\partial g(x) \supseteq A^\top \partial f(Ax)$.

   (b) Define $U = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m \mid z = Ax\}$, which is a linear subspace of $\mathbb{R}^n \times \mathbb{R}^m$, along with the following functions $\mathbb{R}^n \times \mathbb{R}^m \to (-\infty, +\infty]$:

$$F_1(x, z) = f(z)$$

$$F_2(x, z) = \delta_U(x, z) = \begin{cases} 0, & z = Ax \\ +\infty, & z \neq Ax \end{cases}$$

$$F(x, z) = F_1(x, z) + F_2(x, z) = \begin{cases} f(z), & z = Ax \\ +\infty, & z \neq Ax \end{cases}$$

   Show that these functions are convex and that $d \in \partial g(x)$ implies $(d, 0) \in \partial F(x, Ax)$.

   (c) Find expressions for $\partial F_1(x, z)$ and $\partial F_2(x, z)$.

   (d) For the remainder of this problem, assume $\operatorname{ri}\operatorname{dom} f \cap \operatorname{im} A \neq \emptyset$. Show that $\operatorname{ri}\operatorname{dom} F_1$ and $\operatorname{ri}\operatorname{dom} F_2$ must intersect.

   (e) Find an expression for $\partial F(x, z) = \partial(F_1 + F_2)(x, z)$. (Hint: you may use the Moreau-Rockafellar theorem proved in class.)

   (f) Combine the above results to show that $\partial g(x) = A^\top \partial f(Ax)$.