On subdifferential calculus – highlights 2

September 2012

Remark: (a) Let $x_0 \in S$, with $S \subset \mathbb{R}^n$ convex. Then the subgradient $\partial \chi_S(x_0)$, used in the above proof, coincides with the following convex cone (see Appendix B.3):

$$N_S(x_0) := \{\xi \in \mathbb{R}^n : \xi^t(x - x_0) \leq 0 \ \forall \, x \in S\}.$$

Name: the *normal cone to S at x_0.* Hence, one has $-\bar{\xi} \in N_S(x_0)$ in Theorem 2.10.

(b) If $x_0 \in$ **int** S, then $N_S(x_0) = \{0\}$. So Theorem 2.10 states $0 \in \partial f(\bar{x})$ if $\bar{x} \in \text{int } S$.

Remark: If in Theorem 2.10 f is additionally differentiable, then Theorem 2.10 states:

$$\bar{x} \in S \text{ optimal for } (P) \ \Leftrightarrow -\nabla f(\bar{x}) \in N_S(\bar{x}). \quad (1)$$

Moreover, if $\bar{x} \in \text{int } S$, then it just says:

$$\bar{x} \in S \text{ optimal for } (P) \ \Leftrightarrow \nabla f(\bar{x}) = 0.$$

Exercise: Given m points $x_1, \ldots, x_m$ in $\mathbb{R}^n$, consider

$$(P) \quad \inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} |x - x_i|^2.$$

Use Theorem 2.10 to determine the optimal solution.

Exercise: Let $S \subset \mathbb{R}^n$ be convex. If $f : S \to \mathbb{R}$ is differentiable but perhaps non-convex, then $\Rightarrow$ in (1) continues to hold. Prove this. Show also that $\Leftarrow$ may then fail.

Directional derivatives and the DM-theorem

Definition 2.13: The *directional derivative* of a convex function $f : \mathbb{R}^n \to (-\infty, +\infty]$ at the point $x_0 \in \mathrm{dom} f$ in the direction $d \in \mathbb{R}^n$ is defined as

$$f'(x_0; d) := \lim_{\lambda \downarrow 0} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}.$$

Proposition 2.14: Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a convex function and let x_0 be a point in $\mathrm{dom} f$. Then for every direction $d \in \mathbb{R}^n$ and every $\lambda_1, \lambda_2 \in \mathbb{R}$ such that $\lambda_2 > \lambda_1 > 0$ we have

$$\frac{f(x_0 + \lambda_1 d) - f(x_0)}{\lambda_1} \leq \frac{f(x_0 + \lambda_2 d) - f(x_0)}{\lambda_2}$$

Consequence:

$$f'(x_0; d) = \mathbf{inf}_{\lambda > 0} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}.$$

Hence $f'(x_0, d)$ well-defined (in $[-\infty, +\infty]$)!

Example (continues Exercise 2.1c) Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) := 1 - \sqrt{1 - x^2}$ if $x \in [-1, +1]$ and by $f(x) := +\infty$ if $x < -1$ or $x > 1$. Then for $d = 3$

$$f'(x_0; 3) = \begin{cases} 3f'(x_0) & \text{if } |x_0| < 1 \\ +\infty & \text{if } x_0 = 1 \text{ (by } f = +\infty \text{ on } (1, \infty)) \\ -\infty & \text{if } x_0 = -1 \text{ (by a "real" limit)} \end{cases}$$

Theorem 2.15: Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a convex function and let x_0 be a point in int dom f. Then

$$f'(x_0; d) = \sup_{\xi \in \partial f(x_0)} \xi^t d \text{ for every } d \in \mathbb{R}^n.$$

Proof on p. 11 uses Appendix B, but independent proof also possible.

Theorem 2.17 (Dubovitskii-Milyutin) Let $f_1, \cdots, f_m : \mathbb{R}^n \to (-\infty, +\infty]$ be convex functions and let x_0 be a point in $\cap_{i=1}^m$int dom f_i. Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be given by

$$f(x) := \max_{1 \le i \le m} f_i(x)$$

and let $I(x_0)$ be the (nonempty) set of all $i \in \{1, \cdots, m\}$ for which $f_i(x_0) = f(x_0)$. Then

$$\partial f(x_0) = \text{co } \cup_{i \in I(x_0)} \partial f_i(x_0).$$

Proof of D-M theorem: Write $I := I(x_0)$. If $\xi \in \partial f_i(x_0)$, $i \in I$, then

$$\forall_x f(x) \geq f_i(x) \geq f_i(x_0) + \xi^t(x - x_0)$$

with $f_i(x_0) = f(x_0)$ by $i \in I$. So $\xi \in \partial f(x_0)$. By convexity of $\partial f(x_0)$ this gives

$$K := \text{co } \cup_{i \in I} \partial f_i(x_0) \subset \partial f(x_0).$$

Next, we prove $\xi \notin K \Rightarrow \xi \notin \partial f(x_0)$. By Lemma 2.16 and Exercise 2.18 K is compact, hence closed. By separation Thm. A.2:

$$\exists_{d\in\mathbb{R}^n,\alpha\in\mathbb{R}} \xi^t d > \alpha \geq \max_{i\in I} \sup_{\xi'\in\partial f_i(x_0)} \xi'^t d = \max_{i\in I} f_i'(x_0; d)$$

(= holds by Thm. 2.15). Now

$$f'(x_0; d) \stackrel{!}{=} \lim_{\lambda\downarrow 0} \max_{i\in I} \frac{f_i(x_0+\lambda d) - f_i(x_0)}{\lambda}.$$

So $f'(x_0; d)$ equals

$$\max_{i\in I} \lim_{\lambda\downarrow 0} \frac{f_i(x_0+\lambda d) - f_i(x_0)}{\lambda} = \max_{i\in I} f_i'(x_0; d).$$

Conclusion: $\xi^t d > f'(x_0; d)$. Hence $\xi \notin \partial f(x_0)$. QED

Example: Let $m = 2$, $n = 1$, $f_1(x) = x$, $f_2(x) = -x$ and $x_0 = 0$. Then $f(x) = |x|$, $I(0) = \{1, 2\}$ and the D-M theorem says:

$$\partial f(x_0) = \text{co}\left(\{1\} \cup \{-1\}\right) = [-1, 1],$$

known already by different reasoning.