DIFFERENTIABLILITY OF A SUPPORT FUNCTION OF AN E-SUBGRADIENT

Claude Lemarechal
Evgeni Nurminski

June 1980
WP-80-101

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Differentiability of a support function of an ϵ - subgradient mapping

Claude Lemarechal

INRIA

Evgeni Nurminski

IIASA

ABSTRACT

Directional differentiability of a support function of an ϵ-subgradient set-valued mapping is proved and formula for a directional derivative is given.

The notion of an ϵ-subgradient proposed by R.T.Rockafellar (Rockafellar1972a). appears to be rather useful for convex nondifferentiable optimization. This is mainly due to their practical advantages from a computational point of view, but it is also important that this mapping has richer analytical properties than the subgradient mapping traditionally studied in convex analysis.

During the last few years it was observed that this mapping has strong continuity properties (Nurminski(1978)a, Hiriart-Urruty1979a). both in finite dimensional and abstract spaces. Here we state a certain kind of differentiability properties of the support function of ϵ-subgradient mapping which may lead to a definition of second order differentiability of convex functions. Throughout the paper we stay within the framework of convex analysis, so $f(x)$ will be a convex locally Lipschitzian function, $f^*(g)$ its conjugate and $\partial_\epsilon f(x)$ will be the set of ϵ-subgradients. For a number of technical reasons we need an additional assumption that the epigraph of function $f(x)$ does not contain nonvertical straight lines. It can be assured for instance by a coercivity assumption with respect to function $f(x)$.

Definition. The function $V_p(x)$ defined as

$$V_p(x) = \max_{g \in \partial_\epsilon f(x)} pg \tag{1}$$

is called the support function of an ϵ-subdifferential $\partial_\epsilon f(x)$.

The support function $V_p(x)$ has an equivalent representation:

$$V_p(x) = \inf_{t > 0} \frac{f(x+tp) - f(x) + \epsilon}{t} \tag{2}$$

which is easy to obtain from dual consideration of problem (1). Notice that due to coercivity problem (2) always has a bounded solution.

Duality can also be used for obtaining some auxiliary results.

Lemma 1. For fixed x and p let g_x and t_x be any solutions of the problems (1) and (2) respectively. Then

$$g_x \in \partial f(x + t_x p)$$

Proof. Notice that the result is obvious for $\epsilon = 0$. Also if $\epsilon > 0$ then $t_x > 0$. Consider the Lagrange function of problem (1):

$$L(g,t) = pg - \frac{1}{t}(f^*(g) - xg)$$

in which the terms not depending on g are omitted. It follows from the optimality of g_x and t_x that

$$0 \in \partial_g L(g_x, t_x)$$

or

$$0 \in p - \frac{1}{t_x}(\partial f^*(g_x) - x)$$

For g_x one has the inclusion

$$x + t_x p \in \partial f^*(g_x)$$

which is equivalent to

$$g_x \in \partial f(x + t_x p)$$

Q.E.D.

Lemma 2. Let $y = x + \tau d$. Then for any solution t_x of problem (2) , any $g \in \partial f(x)$ and $g_y \in \partial_\epsilon f(y)$ such that

$$V_p(y) = pg_y$$

one has

$$V_p(y) - V_p(x) \leq \frac{\tau}{t_x}(g_y - g)d$$

Proof. Let

$$V_p(y) = \inf_{t>0} \frac{1}{t}(f(y+tp)-f(y)+\epsilon) = \frac{1}{t_y}(f(y+t_y p)-f(y)+\epsilon)$$

$$V_p(x) = \frac{1}{t_x}(f(x+t_x p)-f(x)+\epsilon)$$

Then for any $g \in \partial f(x)$

$$t_y V_p(y) = f(y+t_y p)-f(y)+\epsilon \leq f(y+t_y p)-f(x)+\epsilon - g(y-x)$$

and for any $g_y \in \partial f(y+t_y p)$

$$t_x V_p(x) = f(x+t_x p)-f(x)+\epsilon \geq f(y+t_y p)-f(x)+\epsilon + g_y(x + t_x p - y - t_y p)$$

because of Lemma 1 $g_y \in \partial f(y+t_y p)$.

Subtracting these two inequalities we obtain now, taking $g_y \in \partial f(y+t_y p)$ such that

$$V_p(y) = pg_y \ ,$$

$$t_y V_p(y) - t_x V_p(x) \leq g_y(y-x) + (t_y - t_x)V_p(y) - g(y-x)$$

and consequently

$$t_x(V_p(y) - V_p(x)) \leq \tau(g_y - g)d$$

Q.E.D.

These results make it possible to study the differentiability properties of $V_p(x)$.

Let T_x be the set of solutions of problem (2) and G_x be the solution set of problem (1). The following theorem holds:

Theorem. The support function $V_p(x)$ is directionally differentiable and its directional derivative is given by the formula:

$$V'_p(x,d) = \max_{g_x \in G_x} \min_{\substack{t \in T_x \\ g \in \partial f(x)}} \frac{g_x - g}{t} d$$

Proof. Let $y = x + \tau d$, $\tau > 0$. Using Lemma 2 twice one can obtain two-side bounds of the kind:

$$\frac{V_p(y) - V_p(x)}{\tau} \le \min_{\substack{g \in \partial f(x) \\ t \in T_x}} \frac{g_y - g}{t} d$$

$$\frac{V_p(y) - V_p(x)}{\tau} \ge \max_{\substack{g \in \partial f(y) \\ t \in T_y}} \frac{g_x - g}{t} d$$

for any $g_x \in G_x$ and $g_y \in G_y$. Passing to the limit when τ goes to $+0$ one can assume that correspondent sequences of ϵ-subgradients and subgradients in upper and lower boundaries converge to some limits. From u.s.c of T_x ,see for instance (Hogan 1973a). with related bibliography, it follows that the correspondent sequence of solutions $t^*_y \in T_y$ can be assumed convergent to some limit in T_x as well.

Then

$$\lim_{t \to +0} \frac{V_p(y) - V_p(x)}{\tau} \ge \lim_{t \to +0} \max_{\substack{g \in \partial f(x) \\ t \in T_y}} \frac{g_x - g}{t} d = \frac{g_x - g}{\bar{t}_x} d \ge \min_{\substack{t \in T_x \\ g \in \partial f(x)}} \frac{g_x - g}{t} d$$

for arbitrary $g_x \in G_x$ and some $g \in \partial f(x)$.

Also

$$\overline{\lim_{t \to +0}} \frac{V_p(y) - V_p(x)}{\tau} \le \overline{\lim_{t \to 0}} \min_{\substack{g \in \partial f(x) \\ t \in T_x}} \frac{g_y - g}{t} d = \min_{\substack{t \in T_x \\ g \in \partial f(x)}} \frac{g_x - g}{t} d$$

where $g_x \in \partial_\epsilon f(x)$.

So

$$\min_{\substack{t\in T_x \\ g\in\partial f(x)}} \frac{g_x - g}{t} d \le \min_{\substack{t\in T_x \\ g\in\partial f(x)}} \frac{\bar{g_x} - g}{t} d \le \max_{g\in G_x} \min_{\substack{t\in T_x \\ g\in\partial f(x)}} \frac{g_x - g}{t} d$$

for arbitrary $g_x \in G_x$ and some $\bar{g_x} \in G_x$ which means that

$$\max_{g\in G_x} \min_{\substack{t\in T_x \\ g\in\partial f(x)}} \frac{g_x - g}{t} d \ge \min_{\substack{t\in T_x \\ g\in\partial f(x)}} \max_{g\in G_x} \frac{g_x - g}{t} d$$

So far as strict inequality is impossible in this case then

$$\underline{\lim}_{t\to +0} \frac{V_p(y) - V_p(x)}{t} = \overline{\lim}_{t\to +0} \frac{V_p(y) - V_p(x)}{t} = \max_{g_x\in G_x} \min_{\substack{g\in\partial f(x) \\ t_x\in T_x}} \frac{g_x - g}{t} d \qquad (3)$$

Q.E.D.

A final remark should be made on the order of *min max* operations in expression (3). The proof of the theorem shows that it is irrelevant in which order these operations are performed. An additional argument for that is that the function $f(x,y,\theta) = \theta(x-y)b$ for $0 < \theta_0 \le \theta \le \Theta_0$ has a saddle point in variables (y,θ) and x, where x and y are taken from some compact sets.

References

Rockafellar1972a. R.T. Rockafellar, *Convex Analysis*, Princeton University Press (1972).

Nurminski(1978)a. E. Nurminski, "Continiuity of E-subgradient mappings," *Kibernetika (Cybernetics)* 5, pp. 148-149 (790-791) (1977 (1978)).

Hiriart-Urruty1979a. J.-B. Hiriart-Urruty, "Lipschitz r-continuity of the approximate subdifferential of a convex function," Technical report, Universite de Clermont-Ferrand II Complexe Scientifique des Cezeaux Departement de Mathematiques Appliquees Boite Postale n 45 (D).

Hogan 1973a. W.W. Hogan, "Point-to-set maps in mathematical programming," *SIAM Review* **15**, pp. 591-603 (1973).