BULL. AUSTRAL. MATH. SOC. 49A52, 90C30
VOL. 45 (1992) [37–41]

A FORMULA ON THE APPROXIMATE SUBDIFFERENTIAL OF THE DIFFERENCE OF CONVEX FUNCTIONS

J.E. MARTÍNEZ-LEGAZ AND A. SEEGER

We give a formula on the ε-subdifferential of the difference of two convex functions. As a by-product of this formula, one recovers a recent result of Hiriart-Urruty, namely, a necessary and sufficient condition for global optimality in nonconvex optimisation.

1. THE ε-SUBDIFFERENTIAL OF A DC-FUNCTION

Whether the extended-real-valued function $f : X \longrightarrow \overline{\mathbb{R}}$ is convex or not, we use the standard expression

$$(1) \qquad \partial_\varepsilon f(x_0) := \{u \in X^* : f(x) \geqslant f(x_0) + \langle u, x - x_0 \rangle - \varepsilon \quad \text{for all} \quad x \in X\}$$

as definition for the ε-subdifferential of f at $x_0 \in X$. Here X and X^* are locally convex (real) topological linear spaces paired in duality by a bilinear form $\langle \cdot, \cdot \rangle : X^* \times X \longrightarrow \mathbb{R}$. As it is customary, we assume that ε is a nonnegative real number and that x_0 is a point at which f is finite. The particular instance $\varepsilon = 0$ corresponds, of course, to the usual subdifferential

$$(2) \qquad \partial f(x_0) := \{u \in X^* : f(x) \geqslant f(x_0) + \langle u, x - x_0 \rangle \quad \text{for all} \quad x \in X\} .$$

It is important to note that in the nonconvex case, the ε-subdifferential mapping $\partial_\varepsilon f : X \rightrightarrows X^*$ may be empty-valued at some points.

Formulas for evaluating the ε-subdifferential of a convex function can be found, for instance, in Kutateladze [4] and Hiriart-Urruty [1]. These authors established calculus rules for most of the operations preserving convexity (like addition, inf-convolution, upper envelope, *et cetera*). They did not consider, however, the case of the subtraction, an operation which does not preserve the convexity in general.

The purpose of this note is to write a formula on the ε-subdifferential of a DC-function, that is, of a function f which can be represented as the difference

$$x \in X \longrightarrow f(x) := g(x) - h(x)$$

Received 8th January 1991
The first author acknowledges financial support from the Dirección General de Investigación Científica y Técnica (DGICYT), under project PS89-0058.

of two convex functions $g, h : X \longrightarrow \mathbb{R} \cup \{+\infty\}$. Since g and h may take the value $+\infty$ at the same time, we adopt here the rule $(+\infty) - (+\infty) = +\infty$. The class of DC–functions has received a great deal of attention in recent time. For a survey on this topic, one may consult Hiriart–Urruty [2]. In particular, one can find there a formula on the Fenchel conjugate of the difference of two convex functions. To the best of our knowledge, a formula on the ε–subdifferential of such type of difference has not been established yet.

In next theorem,

$$A \overset{*}{-} B = \{u \in X^* : u + B \subset A\}$$

stands for the "star–difference" between two sets A and B in X^* (see [**2**, p.56]).

Theorem 1. *Let $g, h : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be two lower–semicontinuous proper convex functions, finite at $x_0 \in X$. Then, for every $\varepsilon \geqslant 0$, one has*

$$\partial_\varepsilon (g - h)(x_0) = \bigcap_{\lambda \geqslant 0} \left\{\partial_{\varepsilon+\lambda} g(x_0) \overset{*}{-} \partial_\lambda h(x_0)\right\}. \tag{3}$$

Setting $\varepsilon = 0$, one gets in particular

$$\partial (g - h)(x_0) = \bigcap_{\lambda \geqslant 0} \left\{\partial_\lambda g(x_0) \overset{*}{-} \partial_\lambda h(x_0)\right\}. \tag{4}$$

Proof: By definition, $u \in \partial_\varepsilon (g - h)(x_0)$ if and only if

$$(g - h)(x) \geqslant (g - h)(x_0) + \langle u, x - x_0 \rangle - \varepsilon \quad \text{for all} \quad x \in X$$

or, what is equivalent,

$$g(x) - h_u(x) \geqslant g(x_0) - h_u(x_0) - \varepsilon \quad \text{for all} \quad x \in X, \tag{5}$$

where $h_u := h + \langle u, \cdot \rangle$. But, according to the Toland–Singer duality theorem [**6**, **7**], one can write

$$\inf_{x \in X} \{g(x) - h_u(x)\} = \inf_{y \in X^*} \{(h_u)^*(y) - g^*(y)\}, \tag{6}$$

where the convention $(+\infty) - (+\infty) = \infty$ applies on both sides of this equality, and $\varphi^* : X^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ stands for the Fenchel conjugate of $\varphi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$. Consequently, (5) is equivalent to

$$(h_u)^*(y) - g^*(y) \geqslant g(x_0) - h_u(x_0) - \varepsilon \quad \text{for all} \quad y \in X^*. \tag{7}$$

Now, introduce the notation

$$p(y) := (h_u)^*(y) + h_u(x_0) - \langle y, x_0 \rangle$$
$$q(y) := g^*(y) + g(x_0) - \langle y, x_0 \rangle$$

and write (7) in the form

$$(8) \qquad p(y) \geqslant q(y) - \varepsilon \quad \text{for all} \quad y \in X^* .$$

The inequality (8) relating the nonnegative functions p and q can be expressed in terms of an inclusion

$$(9) \qquad \{y \in X^* : p(y) \leqslant \lambda\} \subset \{y \in X^* : q(y) \leqslant \varepsilon + \lambda\} \quad \text{for all} \quad \lambda \geqslant 0$$

between their corresponding level sets. But, from the very definition of p and q, one sees that

$$\{y \in X^* : \ p(y) \leqslant \lambda\} = \partial_\lambda h_u(x_0) = u + \partial_\lambda h(x_0),$$

and

$$\{y \in X^* : \ q(y) \leqslant \varepsilon + \lambda\} = \partial_{\varepsilon+\lambda}\, g(x_0).$$

Summarising, one has proved that $u \in \partial_\varepsilon (g - h)(x_0)$ if and only if

$$(10) \qquad u + \partial_\lambda h(x_0) \subset \partial_{\varepsilon+\lambda} g(x_0) \quad \text{for all} \quad \lambda \geqslant 0 .$$

This is precisely what formula (3) says. □

REMARK 2. The lower–semicontinuity and the convexity of g are not essential assumptions in Theorem 1. In fact, these assumptions have been used only for writing equality (6). It is known that formula (6) is still valid if $g : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is arbitrary. On the other hand, formula (6) has been extended to the more general conjugation framework of Moreau. Consequently, a formula similar to (3) can be obtained for the corresponding generalised concept of ε–subdifferential. For the above mentioned extensions of formula (6), see for instance Martínez–Legaz [**5**, Theorem 3.1].

2. AN APPLICATION TO DC–PROGRAMMING

The different consequences and applications of Theorem 1 will not be explored in this short note. We shall mention, however, the application we had in mind when we established formula (3). Recall that a point $x_0 \in X$ is said to be an ε–minimum of the function $f : X \longrightarrow \overline{\mathbb{R}}$, if $f(x_0)$ is finite and

$$f(x_0) - \varepsilon \leqslant f(x) \quad \text{for all} \quad x \in X .$$

As an illustration on the use of Theorem 1, we exhibit a necessary and sufficient condition for ε–minimality due to Hiriart–Urruty [**3**].

https://doi.org/10.1017/S0004972700036984 Published online by Cambridge University Press

COROLLARY 3. (see [**3**, Theorem 4.4]) *Let g and h be as in Theorem 1. A necessary and sufficient condition for $x_0 \in X$ be an ε–minimum of $x \in X \mapsto f(x) := g(x) - h(x)$ is that*

$$\partial_\lambda h(x_0) \subset \partial_{\varepsilon+\lambda} g(x_0) \quad \text{for all} \quad \lambda \geqslant 0. \tag{11}$$

In particular, $x_0 \in X$ is a global minimum of $f = g - h$ if and only if

$$\partial_\lambda h(x_0) \subset \partial_\lambda g(x_0) \quad \text{for all} \quad \lambda \geqslant 0. \tag{12}$$

PROOF: Condition (11) is equivalent to $0 \in \partial_\varepsilon (g - h)(x_0)$. ∎

REMARK 4. Theorem 1 can, in turn, be derived from Corollary 3. Indeed, starting from (5) it suffices to apply the optimality condition (11) to the ε–minimum x_0 of the function $g - h_u$. In this way, one gets

$$\partial_\lambda h_u(x_0) \subset \partial_{\varepsilon+\lambda} g(x_0) \quad \text{for all} \quad \lambda \geqslant 0,$$

which is equivalent to (10). The proof we gave in Section1 was inspired on a proof of (12) due to Pham Dinh Tao, communicated to us by Hiriart–Urruty (personal communication).

REFERENCES

[1] J.-B. Hiriart-Urruty, 'ε–subdifferential calculus', in *Convex analysis and optimization: Research Notes in Mathematics* 57, pp. 43–92 (Pitman Publishers, 1982).

[2] J.-B. Hiriart-Urruty, 'Generalized differentiability, duality and optimization for problems dealing with differences of convex functions', in *Convexity and duality in optimization: Lecture Notes in Econom. and Math. Systems* 256, pp. 37–70 (Springer-Verlag, Berlin, Heidelberg, New York, 1986).

[3] J.-B. Hiriart-Urruty, 'From convex optimization to nonconvex optimization', in *Nonsmooth optimization and related topics*, Editors F.H. Clarke, V.F. Demyanov, F.Giannessi, pp. 219–239 (Plenum Press, 1989).

[4] S. S. Kutateladze, 'Convex ε-programming', *Soviet Math. Dokl.* **20** (1979), 391–393.

[5] J.-E. Martínez-Legaz, 'Generalized conjugation and related topics', in *Generalized convexity and fractional programming with economic applications: Lect. Notes in Econom. and Math. Systems* 345, pp. 198–218 (Springer-Verlag, Berlin, Heidelberg, New York, 1990).

[6] I. Singer, 'A Fenchel–Rockafellar type duality theorem for maximization', *Bull. Austral. Math. Soc.* **29** (1979), 193–198.

[7] J. Toland, 'A duality principle for nonconvex optimization and the calculus of variations', *Arch. Rational Mech. Anal.* **71** (1979), 41–61.

Dept. d'Economia i d'Història Econòmica
Universitat Autònoma de Barcelona
08193-Bellaterra, Barcelona
Spain

and

Dept. de Matemàtica Aplicada i Anàlisi
Universitat de Barcelona
08071 - Barcelona
Spain

Dept. de Matemàtica Aplicada i Anàlisi
Universitat de Barcelona
08071 - Barcelona
Spain