$ ¥varepsilon $-DIRECTIONAL DERIVATIVE OF A MARGINAL FUNCTION IN PARAMETRIZED CONVEX PROGRAMMING

Shiraishi, Shunsuke
Faculty of Economics, Toyama University

https://doi.org/10.5109/13417

出版情報：Bulletin of informatics and cybernetics. 24 (3/4), pp.177-183, 1991-03. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics, Vol. 24, No. 3~4, 1991

ε-DIRECTIONAL DERIVATIVE OF A MARGINAL FUNCTION IN PARAMETRIZED CONVEX PROGRAMMING

By

Shunsuke SHIRAISHI*

Abstract

A formula for calculating an ε-directional derivative of a marginal function for a parametrized convex programming is given. Earlier result of the author is covered and improved.

1. Introduction

Parametrized mathematical programming problems are stated as follows.

$$(P_1) \qquad \begin{array}{l} \text{minimize} \quad f(y) \\ \text{subject to } y \in \Gamma(x), \end{array}$$

where f is a real-valued function on $\mathbf{R}^n$ and the constraint set Γ is a set-valued mapping from $\mathbf{R}^m$ to $\mathbf{R}^n$. The vector $x \in \mathbf{R}^m$ is viewed as a parameter. The problem (P_1) is rewritten as:

$$(P_2) \qquad \begin{array}{l} \text{minimize} \quad f(y) \\ \text{subject to } (x, y) \in C, \end{array}$$

where C denotes the graph of Γ, i.e.,

$$C = \{(x, y) \in \mathbf{R}^m \times \mathbf{R}^n \mid y \in \Gamma(x)\}.$$

A problem which we intend to consider here is a convex programming problem, that is, the objective function f is convex and the set C is a nonempty closed convex set in $\mathbf{R}^m \times \mathbf{R}^n$. For example, this is precisely the case when

$$\Gamma(x) = \{y \in \mathbf{R}^n | g_i(y) \leq x_i,\ i = 1, \dots, m\},$$

where g_i are convex for $i = 1, \dots, m$. In this case a marginal function φ, which is defined by

* Faculty of Economics, Toyama University, 3190 Gofuku, Toyama 930, Japan

$$\varphi(x) := \inf_{y \in \Gamma(x)} f(y) = \inf_{y \in \mathbf{R}^n} (f(y) + \psi_C(x, y)),$$

is a convex function on $\mathbf{R}^m$, where ψ_c denotes the indicator function of C. The main purpose of this note is to derive a formula of an ε-directional derivative of the marginal function φ in terms of the ε-directional derivative of f.

The ε-directional derivative of convex functions is recognized as a powerful tool in minimization algolismics of convex functions. A lot of minimization methods have been developed, see [7] and references therein. On the other hand, the ε-directional derivative is useful in "convex situations" from the theoretical viewpoint. One can describe almost optimality conditions of convex programming problems by using it. The main principle is stated as follows. A candidate point x_0 minimizes a convex function up to $\varepsilon > 0$ if and only if the ε-directional derivative of the convex function at x_0 is non-negative in any directions. Such optimality conditions are fully described by Strodiot et al [10]. Surprisingly enough, however, the ε-directional derivative is also useful in "nonconvex situations". Hiriart-Urruty showed that a global optimality condition which is not only necessary but also sufficient for a problem of minimizing a difference of two convex functions (d.c. optimization problem) can be described in terms of ε-directional derivatives [5]. To describe the ε-directional derivative of complicated functions in terms of simple one is worth to study. Some of them, that is, the ε-directional derivatives of a sum of convex functions, a maximum of convex functions and composite of linear transformation and convex function etc. were studied [4]. The ε-directional derivative of the marginal function were also investigated for the unconstrained and linear constrained case [3, 4, 9]. All results which we present here are regarded as generalizations of them to some extent.

This note is organized as follows. In section 2, we prepare basic concepts and properties of ε-directional derivative. As we pointed out in our previous paper [9], some duality relation plays an important role to investigate the ε-directional derivative of the marginal function. This interesting fact will be shown again in somewhat special case where the graph of the constraint set C is a closed convex cone with nonempty interior. Section 3 is devoted to this problem. The main results are presented in section 4. Earlier result of the author will be covered and improved.

2. Preliminary Results

Let f be a real convex function on $\mathbf{R}^n$. Given $\varepsilon > 0$, the ε-directional derivative of f at $y_0 \in \mathbf{R}^n$ in the direction $z \in$ is defined by

$$f'_\varepsilon(y_0; z) := \inf_{\lambda > 0} \frac{f(y_0 + \lambda z) - f(y_0) + \varepsilon}{\lambda}$$

The function $f'_\varepsilon(y_0; z)$ is concave with respect to ε and sublinear with respect to d. There is a fundamental way of characterizing $f'_\varepsilon(y_0; z)$ as a support function of some closed convex set. We recall the definition of an ε-subdifferential $\partial_\varepsilon f(y_0)$ of f at y_0;

$$\partial_\varepsilon f(y_0) := \{y^* \in \mathbf{R}^n | f(y\ ) \geq f(y_0) + < y^*, y - y_0 > - \varepsilon \ \forall y \in \mathbf{R}^n\}.$$

The set $\partial_\varepsilon f(y_0)$ is known to be nonempty and compact and reduces to the subdifferential $\partial f(y_0)$ when $\varepsilon = 0$. The following characterization of f'_ε holds and will be used throughout.

PROPOSITION 2.1. [8] *It holds that*

$$f'_\varepsilon(y_0; z) = \psi^*_{\partial_\varepsilon f(y_0)}(z),$$

where ψ^* denotes the support function of a convex set.

In order to ensure that the marginal function φ does not take value $-\infty$, we assume the following hypothesis.

$$\text{(H)} \qquad f_\infty(z) > 0, \ \forall z \neq 0 \text{ satisfying } (0, z) \in 0^+C,$$

where f_∞ and 0^+C denote the recession function of f and the recession cone of C, respectively (See [8]).

PROPOSITION 2.2. *Under the assumption* (**H**), *the marginal function $\varphi(x)$ is finite and the solution set $Y(x) := \{y \in \mathbf{R}^n | \varphi(x) = f(y) + \psi_C(x, y)\}$ is nonempty for all $x \in \mathbf{R}^m$.*

PROOF. For fixed x, set

$$h(y) := f(y) + \psi_C(x, y).$$

Then h is a lower semi-continuous proper convex function and its recession function is given by

$$h_\infty(z) = \begin{cases} f_\infty(z), \text{ if } (0, z) \in 0^+C, \\ +\infty, \text{ otherwise.} \end{cases}$$

It follows from (**H**) that $h_\infty(z) > 0$ for all nonzero z. Thus $h(y)$ attains its minimum (See [8] p. 70). □

Moreover we can prove that $Y(x_0)$ is compact convex. Throughout this note, we assume (**H**). For given $x_0 \in \mathbf{R}^m$, we take an arbitrary element y_0 of $Y(x_0)$. The set $N_\varepsilon(C; x_0, y_0)$ of the ε-normals to C at $(x_0, y_0) \in C$ is defined as the ε-subdifferential of the indicator function of C. It is easy to see that:

$$N_\varepsilon(C; x_0, y_0) = \{(x^*, y^*) | < x^*, x - x_0 > + < y^*, y - y_0 > - \varepsilon \leq 0, \forall (x, y) \in C\}.$$

Several examples of the set of ε-normals can be found in [5, 6].

3. Duality Relation

In this section, we restrict our attention to the case where the set C is a closed convex cone with nonempty interior. In this case, it is well known that the set $N_\varepsilon(C; x_0, y_0)$ reduces to the following [5].

$$N_\varepsilon(C; x_0, y_0) = \{(x^*, y^*) \in C^* | < x^*, x_0 > + < y^*, y_0 > + \varepsilon \geq 0\},$$

where $C^* := \{(x^*, y^*) | < x^*, x > + < y^*, y > \leq 0, \forall (x, y) \in C\}$ (polar cone to C). With the aid of this fact, we can obtain a formula for the ε-subdifferential of φ. The proof is the same with that of Hiriart-Urruty or ours [3, 9]. Hence we omit it.

THEOREM 3.1. *It holds that*

$$\partial_\varepsilon \varphi(x_0) = \{x^* | 0 \leq \exists \eta \leq \varepsilon, \exists y^* \in \partial_\eta f(y_0) \ s.t. \ (x^*, -y^*) \in C^*,$$
$$< x^*, x_0 > - < y^*, y_0 > + (\varepsilon - \eta) \geq 0\}.$$

According to Proposition 2.1 and Theorem 3.1, the calculation of the ε-directional derivative of φ at x_0 in the direction d is reduced to calculating an optimal value of a maximization problem (P^*) below. We shall present another minimization problem (D_η^*) and give a duality relation between them.

$$(P^*) \quad \begin{array}{c} \text{maximize} < x^*, d > \\ \text{over all } x^* \text{ satisfying} \\ 0 \leq \exists \eta \leq \varepsilon, \exists y^* \in \partial_\eta f(y_0), \\ (x^*, -y^*) \in C^*, \\ < x^*, x_0 > - < y^*, y_0 > + (\varepsilon - \eta) \geq 0. \end{array}$$

$$(D_\eta^*) \quad \begin{array}{c} \text{minimize } f'_\eta(y_0; z_1) + (\varepsilon - \eta) z_2 \\ \text{subject to } (d, z_1) + z_2(x_0, y_0) \in C, \ z_2 \geq 0 \end{array}$$

Let u^* and $v^*(\eta)$ be optimal solutions of (P^*) and (D_η^*) respectively. Then the following inequality, which is often called weak duality, can be obtained immediately from the definitions of the problem (P^*) and (D_η^*).

PROPOSITION 3.1. *It holds that*

$$\sup_{0 \leq \eta \leq \varepsilon} v^*(\eta) \geq u^*.$$

In order to prove the equality, we assume the following regularity condition.

$$(\mathbf{R}) \qquad \exists \hat{y} \in \mathbf{R}^n \text{ satisfying } (x_0, \hat{y}) \in intC.$$

The condition (**R**) may be referred as usual Slater condition.

THEOREM 3.2. *Under the assumption* (**R**), *the strong duality*

$$\sup_{0 \leq \eta \leq \varepsilon} v^*(\eta) = u^*$$

holds.

PROOF. Let $0 \leq \eta \leq \varepsilon$ be such that for some x^* and y^* the triple (x^*, y^*, η) is feasible for (P^*). Weak duality asserts that $v^*(\eta)$ is finite for this η. Set

$$B: = \{(u, v, w) | u \geq f'_\eta(y_0; z_1) + (\varepsilon - \eta), v = (d + z_2 x_0, z_1 + z_2 y_0)\},$$

$$D: = \{(u, v, w) | u < v^*(\eta), (v, w) \in intC\}.$$

Obviously B and D are convex and disjoint. The separation procedure of two convex sets B and D yields that there exist $\alpha \geq 0$, $(\beta, \gamma) \in C^*$, not all zero, such that

$$\alpha(f'_\eta(y_0; z_1) + (\varepsilon - \eta)z_2) + < \beta, d + z_2x_0 > + < \gamma, z_1 + z_2y_0 > \geq \alpha v^*(\eta), \qquad \forall z_1, \forall z_2 \geq 0. \tag{1}$$

If we suppose $\alpha = 0$, then, by taking $z_1 = z_2(\hat{y} - y_0)$, (1) turns out to be

$$< \beta, d + z_2x_0 > + < \gamma, z_2\hat{y} > \geq 0, \qquad \forall z_2 \geq 0$$

Hence we get $< \beta, x_0 > + < \gamma, \hat{y} > = 0$. Since $(x_0, \hat{y}) \in intC$ and $(\beta, \gamma) \in C^*$, we reach the contradiction $(\beta, \gamma) = 0$. Without loss of generality, we may assume $\alpha = 1$. Then we have:

$$(f'_\eta(y_0; z_1) + (\varepsilon - \eta)z_2) + < \beta, d + z_2x_0 > + < \gamma, z_1 + z_2y_0 > \geq v^*(\eta), \qquad \forall z_1, \forall z_2 \geq 0. \tag{2}$$

By taking the infimum of (2) with respect to z_1 and $z_2 \geq 0$, we have

$$< \beta, d > \geq v^*(\eta). \tag{3}$$

If we divide by z_2 and let z_2 approach infinitely in (2), then we have:

$$< \beta, x_0 > + < \gamma, y_0 > + (\varepsilon - \eta) \geq 0$$

With the same argument, we also have:

$$f'_\eta(y_0; z_1) \geq < - \gamma, z_1 >, \qquad \forall z_1,$$

which implies $-\gamma \in \partial_\eta f(y_0)$. These facts means that $(\beta, -\gamma, \eta)$ is feasible. Thus from (3), we get the required result. □

Rewriting the relation sup $v^*(\eta) = u^*$, we have the following theorem.

THEOREM. 3.3.

$$\varphi'_\varepsilon(x_0; d) = \sup_{0 \leq \eta \leq \varepsilon} \inf\{f'_\eta(y_0; z_1) + (\varepsilon - \eta)z_2 | (d, z_1) + z_2(x_0, y_0) \in C, z_2 \geq 0\}.$$

4. ε-Directional Derivative of a Marginal Function

In this section, we turn back to the general problem:

$$\begin{array}{ll} \text{minimize} & f(y) \\ \text{subject to} & (x, y) \in C, \end{array}$$

where f is a convex function on $\mathbf{R}^n$ and C is a closed convex set (not necessarily a cone). We direct our attention to the fact that our marginal function has two structures in nature. The structures are sum operation of two functions and inf operation. For convex functions which are constructed by each operation, Hiriart-Urruty gave the ε-directional derivative of them [3, 4]. His results are as follows:

PROPOSITION 4.1. *Given two convex functions f_1 and f_2 from $\mathbf{R}^n$ to $\mathbf{R}$, it holds that for all $\varepsilon > 0$:*

$$(f_1 + f_2)'_\varepsilon(y_0; z) = \max_{\substack{\varepsilon_1 \geq 0,\ \varepsilon_2 \geq 0, \\ \varepsilon_1 + \varepsilon_2 = \varepsilon}} \{f'_{\varepsilon_1}(y_0; z) + f'_{\varepsilon_2}(y_0; z)\}$$

PROPOSITION 4.2. *If the constraint set C of problem (P_1) is the whole space $\mathbf{R}^n \times \mathbf{R}^m$, then the ε-directional derivative of φ is the lower semi-continuous hull of the convex function*

$$d \to \inf_{z \in \mathbf{R}^n} f'_\varepsilon(x_0, y_0; d, z).$$

If $f_\infty(0, z) > 0$ for all nonzero $z \in \mathbf{R}^n$, then

$$\varphi'_\varepsilon(x_0; d) = \min_{z \in \mathbf{R}^n} f'_\varepsilon(x_0, y_0; d, z).$$

Using the indicator function of the convex set C, $\varphi(x)$ can be written as the infimum of the function $f(x, y) = f(y) + \psi_C(x, y)$ with respect to y. Under the assumption (**H**), it is easy to see thaf $f_\infty(0, z) > 0$ for all non zero z (See Prop 2.2). Hence, with the aid of Hiriart-Urruty's results, we have the following:

$$\varphi'_\varepsilon(x_0; d) = \min_{z \in \mathbf{R}^n} \max_{0 \leq \eta \leq \varepsilon} [f'_\eta(y_0; z) + \psi^*_{N_{\varepsilon-\eta}(C; x_0, y_0)}(d, z)].$$

Since the function in the bracket in the above formula is convex with respect to z and concave in η and η moves the compact convex set $[0, \varepsilon]$, it follows from the min-max theorem ([1], Theorem 1) that the operators "min−max" commute. Hence we obtain the following theorem.

THEOREM 4.1. *The ε-directional derivative of φ is given as*

$$\varphi'_\varepsilon(x_0; d) = \max_{0 \leq \eta \leq \varepsilon} \min_{z \in \mathbf{R}^n} [f'_\eta(y_0; z) + \psi^*_{N_{\varepsilon-\eta}(C; x_0, y_0)}(d, z)],$$

for all $d \in \mathbf{R}^m$ and the operators max−min commute.

If we apply this theorem to linear constrained case, we can get our previous result as a corollary. The problem to which the theorem is applicable is:

$$\begin{array}{ll} \text{minimize} & f(y) \\ \text{subject to} & Ay \leq x, \end{array}$$

where f is convex on $\mathbf{R}^n$ and A is a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$. In this case, the set C is:

$$C = \{(x, y) | Ay \leq x\}$$

The only thing to do is to calculate $\psi^*_{N_{\varepsilon-\eta}(C; x_0, y_0)}(d, z)$, which is equivalent to calculate an optimal value of the following maximization problem:

$$\begin{array}{ll} \text{maximize} & \langle d, x^* \rangle + \langle z_1, y^* \rangle \\ \text{subject to} & (x^*, y^*) \in N_\varepsilon(C; x_0, y_0) \end{array} \tag{4}$$

We can express the set $N_\varepsilon(C; x_0, y_0)$ in a simple form.

$$N_\varepsilon(C; x_0, y_0) = \{(-\lambda, A^T \lambda) \geq 0, \langle Ay_0 - x_0, \lambda \rangle \geq -\varepsilon\}$$

Hence the problem 4 reduces to the following linear programming problem:

$$\begin{aligned} \text{maximize} \quad & < Az_1 - d, \lambda > \\ \text{subject to} \quad & < Ay_0 - x_0, \lambda > \geq - \varepsilon \\ & \lambda \geq 0. \end{aligned}$$

By the duality theorem of linear programming problem, we have:

$$\psi^*_{N_{\varepsilon-n}(C;x_0,y_0)}(d, z) = \min_{z_2} \{\varepsilon z_2 | Az_1 + z_2(Ay_0 - x_0) \leq d, z_2 \geq 0\}.$$

Hence we get our previous result under weaker assumption.(regularity condition is not needed).

COROLLARY 4.1. ([9])

$$\varphi'_\varepsilon(x_0; d) = \max_{0 \leq \eta \leq \varepsilon} \min \{f'_\eta(y_0; z) + (\varepsilon - \eta)z_2 | Az_1 + z_2(Ay_0 - x_0) \leq d, z_2 \geq 0\}.$$

References

[1] GOL'SHTEIN, E. G.: *Theory of convex programming,* American Mathematical Society, Providence, Rhode Island, (1972).

[2] HIRIART-URRUTY, J.-B.: *The approximate first-order and second-order directional derivatives for a convex function*, In J. P. Cecconi and T. Zolezzi, editors, Mathematical Theories of Optimization, Lecture Notes in Mathematics **979**, Springer-Verlag, Berlin, (1983), 144–177.

[3] HIRIART-URRUTY, J.-B.: *Approximate first-order and second-order directional derivatives of a marginal function in convex optimization*, Journal of Optimization Theory and Applications, **48**, (1986), 127–140.

[4] HIRIART-URRUTY, J.-B.: *Calculus rules on the approximate second-order directional derivative of a convex function*, SIAM Journal on Control and Optimization, **22**, (1984), 381–404.

[5] HIRIART-URRUTY, J.-B.: *From convex optimization to nonconvex optimization,* In F. H. Clarke, V. F. Demyanov, and F. Giannessi, editors, Nonsmooth Optimization and Related Topics, Plenum, (1989), 218–239.

[6] HIRIART-URRUTY, J.-B.: *ε-subdifferential calculus*, In J.-P. Aubin and P. Vinter, editors, Convex Analysis and Optimization, Pitman Publishers, New York, New York, (1982), 43–92.

[7] LEMARÉCHAL, C.: *Constructing bundle methods for convex optimization*, In J.-B. Hiriart-Urruty, editor, *FERMAT Days* **85**: Mathematics for Optimization, North-Holland, Amsterdam, (1986), 201–240.

[8] ROCKAFELLAR, R. T.: *Convex analysis*, Princeton University Press, Princeton, New Jersey, (1970).

[9] SHIRAISHI, S.: *First-order and-second order ε-directional derivatives of a marginal function in convex programming with linear inequality constraints*, to appear in Journal of Optimization Theory and Applications.

[10] STRODIOT, J.-J., NGUYEN, V. H. and HEUKEMES, N.: *ε-optimal solutions in nondifferentiable convex programming and some related questions*, Mathematical Programming, **25**, (1983), 307–328.

Received September, 11, 1990
Communicated by N. Furukawa