A generalized differential of real functionals

V. NOVO, L. RODRÍGUEZ MARÍN*

Recibido: 22 de Junio de 1.994

Presentado por el Académico Numerario D. Pedro Jiménez Guerra

Abstract

With the objective of studying the criteria to optimize nonsmooth functionals, a generalized differential of real functionals is introduced over a Banach space. The main properties of this generalized differential are established, along with their relationships to other theories of generalized differentiation, and the necessary and sufficient conditions for local extremum.

1. Introduction and Preliminary Testing

Recent interest in problems related to the optimization of nonsmooth functions has led to new theories, over the last two decades about the generalized differentiation applicable to this type of optimization problems. Rockafellar [13] studies the case of convex functions in an open set in $\mathbb{R}^n$ and then introduces the subdifferential $(\partial_R f)$. In the same way, Clarke [4] broadens the class of functions considered by Rockafellar by extending the theory of real functions that are locally Lipschitzian in $\mathbb{R}^n$ and by giving the generalized gradient $(\partial_C f)$. Later, Clarke himself generalizes his differential to functions of $\mathbb{R}^n$ in $\mathbb{R}^m$ [5], and to real functional over a Banach space [6]. The development of these theories can be seen in [1],[2],[3],[7],[8],[9] and [12].

In [10] and [11], a new generalized differential is defined between finite dimensional spaces. This differential, different from Clarke's is reduced to the differential in the Fréchet sense for differentiable functions. In this work, we are extending this definition to include the case of real functionals over a Banach space and we are demonstrating that the convex nature of the functionals is connected to a monotonous property of the differential which allows us to apply extremum conditions.

* Dpto. de Matemática Aplicada. Univ. Nacional de Educación a Distancia.Madrid. Spain

G-differentiability of f at a is equivalent to convexity of next generalized directional derivative:

$$d^{+}f(a,v)=\limsup_{t\to 0}\frac{f(a+tv)-f(a)}{t}$$

Summarized below, are some results that we will need later. Supposing that f is a real function defined in an interval $I\subset\mathbb{R}$ and $co(M)$ the convex hull of M. If a, x and x_n are elements of I, and $(x_n)\to a$ when $n\to\infty$ we determine that:

$$F(a,x)=\frac{f(x)-f(a)}{x-a};\, l(f,a,x_n)=\lim_{n\to\infty}F(a,x_n)$$

It is said that f is stable at a if there exists a reduced neighbourhood V of a and a constant $k>0$ such that, $|F(a,x)|\leq k$ for all $x\in V$. $(x_n)\to a$ is said to be a G-derivability sequence of f at a if $l(f,a,x_n)$ exists and is finite. The set of G-derivability sequences of f at a is called $S(f,a)$.

f is G-derivable at a if all sequence $(x_n)\to a$ contains at least one subsequence $(x_k)\in S(f,a)$. The G-derivative of f at a is:

$$\partial f(a)=co\{l(f,a,x_n):(x_n)\in S(f,a)\}$$

If f is stable at a, then $\partial f(a)$ is a convex compact set. Additionally, the stability of f in a is equivalent to the G-derivability. These definitions and the calculus rules of G-derivatives can be seen in [10] and [11], as well as the following mean value theorem for G-derivatives and the necessary extremum condition.

If f is stable at each point of $[a,b]$, then there exist $c\in(a,b)$ and $\xi\in\partial f(c)$ such that $f(b)-f(a)=\xi(b-a)$.

If a is a local extremum of f, then $0\in\partial f(a)$.

In everything that follows, A is an open set of the E real Banach space, E' is the topological dual, f is a real functional defined in A and $a\in A$. If $(t_n)\to 0$ is a sequence of real numbers and $v\in E, v\neq 0$ we denote:

$$F(a,v,t_n)=[f(a+t_n v)-f(a)]/t_n\,;\; l(f,a,v,t_n)=\lim_{n\to\infty}F(a,v,t_n)$$

For each $v \in E, v \neq 0$ the real function g_v is defined in the open set $D = \{t \in \mathbb{R} : a + tv \in A\}$ to be $g_v(t) = f(a + tv)$.

2. G-Differential of Real Functionals

f is stable at $a \in A$ if there exists a V neighbourhood of a and a constant $k > 0$, such that for all x in V, $|f(x) - f(a)| \leq k\|x - a\|$. $S(A)$ denotes the linear space of stable real functionals at each point of A. It is clear that if f is stable at a, then g_v is stable at $t = 0$ for each v in E.

The directional G-derivative of f at a with respect to a vector v is defined by:

$$\partial_v f(a) = co\{l(f, a, v, t_n) : (t_n) \in S(g_v, 0)\}; \ \partial_0 f(a) = \{0\}$$

Proof of two next propositions are the same as th.2.4 and prop.2.5 [11].

Proposition 2.1 If f is stable at a, then for each v we have:
(i) $\partial_v f(a)$ is a non-empty convex compact subset in $\mathbb{R}$.
(ii) There exists $k > 0$, such that $\partial_v f(a) \subset [-k\|v\|, k\|v\|]$.

Proposition 2.2 If f is stable at a, then the set-valued function $T: E \rightrightarrows \mathbb{R}$ defined by $T(v) = \partial_v f(a)$ is an odd prefan in Ioffe's terminology [8]. (i.e., $T(v)$ is a convex compact set, $0 \in T(0)$ and $T(\lambda v) = \lambda T(v)$ for all $\lambda \in \mathbb{R}$).

If f is stable at a, then the generalized directional derivatives $d^+ f(a,.) : E \to \mathbb{R}$ and $d^- f(a,.) : E \to \mathbb{R}$ are defined by:

$$d^+ f(a, v) = \limsup_{t \to 0} \frac{f(a + tv) - f(a)}{t}$$

$$d^- f(a, v) = \liminf_{t \to 0} \frac{f(a + tv) - f(a)}{t}$$

Proposition 2.3 If f is stable at a, then:
i) $d^+ f(a,.) : E \to \mathbb{R}$ is positively homogeneous.
ii) $d^+ f(a, v) = -d^- f(a, -v)$, for each v.
iii) $\partial_v f(a) = [d^- f(a, v), d^+ f(a, v)]$.

Proof.
i) If $\lambda > 0$ we have

$$d^{+}f(a,\lambda v)=\limsup_{t\to 0}\frac{f(a+t\lambda v)-f(a)}{t}=\lambda\,\limsup_{t\to 0}\frac{f(a+t\lambda v)-f(a)}{\lambda t}=\lambda d^{+}f(a,v);$$

ii) For each v in E, we have

$$d^{+}f(a,v)=\limsup_{t\to 0}\frac{f(a+tv)-f(a)}{t}=\limsup_{t\to 0}-\frac{f(a+(-t)(-v))-f(a)}{-t}=$$

$$=\limsup_{\mu\to 0}-\frac{f(a+\mu(-v))-f(a)}{\mu}=-\liminf_{\mu\to 0}\frac{f(a+\mu(-v))-f(a)}{\mu}=-d^{-}f(a,-v).$$

iii) It is clair that

$$d^{+}f(a,v)=\sup_{(t_n)\in S(g_v,0)}\left\{\lim_{n\to\infty}\frac{f(a+t_n v)-f(a)}{t_n}\right\}=\sup\partial_v f(a)$$

$$d^{-}f(a,v)=\inf_{(t_n)\in S(g_b,0)}\left\{\lim_{n\to\infty}\frac{f(a+t_n v)-f(a)}{t_n}\right\}=\inf\partial_v f(a).$$

Definition 2.4 It is said that f is G-differentiable at a, if for each v and each $l\in\partial_v f(a)$ there exists a linear continuous selection $\xi\in E'$ from the set valued function T, such that $\xi(v)=l$. The set of these selections is the G-differential of f at a and is denoted by $\partial f(a)$.

From this definition it follows that $\partial f(a)(v)=\partial_v f(a)$ and $d^{+}f(a,.)$ is the support function of $\partial f(a)$ (13).

$$d^{+}f(a,v)=\sup\{\xi(v):\xi\in\partial f(a)\}$$

Proposition 2.5 If f is G-differentiable at a, then $\partial f(a)$ is a convex and *-weak compact set in E'.

Proof. $\partial f(a)$ is convex because $\partial_v f(a)$ is for each v in E. We will show that $\partial f(a)$ is bounded. If $\xi\in\partial f(a)$ then

$$\|\xi\|=\sup_{\|v\|=1}|\xi(v)|\le\sup_{\|v\|=1}|T(v)|\le k$$

Suppose E' with the *-weak topology and let $\xi \in \overline{\partial f(a)}$. For each v in E, there exists a sequence $(\xi_n) \subset \partial f(a)$ such that $\lim_{n\to\infty} \xi_n(v) = \xi(v)$. Because $\xi_n \in \partial f(a)$ for each $n \in \mathbb{N}$, we deduce that $\xi_n(v) \in \partial_v f(a)$ and since $\partial_v f(a)$ is closed, we find $\xi(v) \in \partial_v f(a)$ for each v in E and therefore $\xi \in \partial f(a)$.

Lemma 2.6 Let f be stable at a and $\xi \in E'$. If $\xi(v) \le d^+ f(a,v)$ then $d^- f(a,v) \le \xi(v) \le d^+ f(a,v)$, and $\xi(v) \in T(v)$.

Proof. Because $d^+ f(a,v) = -d^- f(a,-v)$, we have

$$-\xi(v) = \xi(-v) \le d^+ f(a,-v) = -d^- f(a,v)$$

then $\xi(v) \ge d^- f(a,v)$ and from 2.3 (iii) it follows that $\xi(v) \in T(v)$.

Theorem 2.7 Let f be stable at a, then the following propositions are equivalent:

(i) T is a set-valued fan.

(ii) f is G-differentiable at a.

(iii) $d^+ f(a,.)$ is a convex functional.

Proof: The proof of (i)$\Leftrightarrow$(ii) is analogous to that of theorem 2.9 [11].

(ii) $\Rightarrow$ (iii). Because $v \to d^+ f(a,v)$ is the supremum of linear functionals.

(iii) $\Rightarrow$ (ii). Since f is a stable functional at a, we have $\partial_u f(a) \ne \varnothing$ for all u in E. Let u in E, $l \in \partial_u f(a)$ and α be the linear functional of $L = \{v = \lambda u : \lambda \in \mathbb{R}\}$ in $\mathbb{R}$ such that $\alpha(u) = l$. For all $v \in L$ from 2.2 it follows that:

$$\alpha(v) = \alpha(\lambda u) = \lambda\alpha(u) \in \lambda\partial_u f(a) = \partial_{\lambda u} f(a) = \partial_v f(a)$$

and consequently $\alpha(v) \le d^+ f(a,v)$ for all $v \in L$.

Because $d^+ f(a,.)$ is a convex and positively homogeneous functional, there exists a linear functional $\xi : E \to \mathbb{R}$ such that:

$$\xi(v) = \alpha(v),\ \forall v \in L \quad and \quad \xi(w) \le d^+ f(a,w)\ \forall w \in E.$$

From 2.6 we have $\xi(w) \in T(w)$ for all w in E and being ξ continuous (prop. 2.1 ii) f is G-differentiable in a.

Corollary 2.8. If f is G-differentiable at a, then

$$\partial f(a)=\partial_R d^+ f(a,0)$$

Proof. From 2.7 we know that $d^+ f(a,.)$ is a convex functional, then there exists $\partial_R d^+ f(a,0)$. Let ξ be an element in $\partial_R d^+ f(a,0)$, we have:

$$d^+ f(a,v)\geq d^+ f(a,0)+\xi(v-0), \forall v\in E.$$

Because $d^+ f(a,0)=0$ we have $\xi(v)\leq d^+ f(a,v)$ for all v in E, and consequently $\xi\in\partial f(a)$.
Proof of the other inclusion is analogous.

It is easy to prove that if f is a G-differentiable functional at a, then:

$$\partial(\lambda f)(a)=\lambda\partial f(a)$$

where $\lambda\in\mathbb{R}$. If f, g are G-differentiable at a, then

$$\partial(f+g)(a)\subset\partial f(a)+\partial g(a)$$

In the following theorem, the relationship with other differential definitions is studied.

Theorem 2.9
(i) $\partial f(a)=\{\xi\}$ iff is Gâteaux differentiable at a.
(ii) If f is convex in A, then $\partial f(a)=\partial_R f(a)$ for each a in A.
(iii) If f is locally Lipschitzian in A, then $\partial f(a)\subset\partial_C f(a)$ for each a in A.
Proof: (i) is a direct consequence of definition 2.4.
(iii) is deduced by taking into account that

$$d^+ f(a,v)\leq \limsup_{h\to 0,\lambda\downarrow 0}\left[f(a+h+\lambda v)-f(a+h)\right]/\lambda$$

for each a in A and all v in E.
(ii) If $t_1,t_2\in D,\alpha\geq 0,\beta\geq 0,\alpha+\beta=1$, then

$$g_v(\alpha t_1+\beta t_2)=f\left[a+(\alpha t_1+\beta t_2)v\right]=f\left[\alpha(a+t_1 v)+\beta(a+t_2 v)\right]\leq$$
$$\leq\alpha f(a+t_1 v)+\beta f(a+t_2 v)=\alpha g_v(t_1)+\beta g_v(t_2)$$

which means that for each v in E, g_v is a convex function and therefore

$$\partial_v f(a) = \partial g_v(0) = [f'_-(a;v), f'_+(a;v)]$$

where f'_-, f'_+ are the one-sided derivatives (see [13]). From lemma 2.6 we deduce that $\xi \in \partial f(a)$ if and only if $\xi(v) \leq f'_+(a;v)$ for each v in E and consequently $\partial f(a) = \partial_R f(a)$.

Proceeding exactly as in the proof of theorem 2.10 [11], we obtain:

Theorem 2.10 If f es G-differentiable at A and $[a, a+h] \subset A$, then there exist $\theta \in (0,1)$ and $\xi \in \partial f(a + \theta h)$, such that

$$f(a+h) - f(a) = \xi(h).$$

3. Extremum Conditions

Throughout this section, we assume that f is G-differentiable at each point of A.

Theorem 3.1 If a is a local extremum of f, then $0 \in \partial f(a)$.
Proof. As f has a local extremum in a, for each v, g_v has a local extremum in 0; it follows that $0 \in \partial g_v(0) = \partial_v f(a)$ for each v. We have $0(v) = 0 \in \partial_v f(a)$ and consequently $0 \in \partial f(a)$.

The following definition tranfers the concept of monotony to the case of set-valued functions.

Definition 3.2 It is said that the set-valued function $M: A \rightrightarrows E'$ is non decreasing (non increasing) monotonous in A, if for all x, y in A, $\xi \in M(x), \eta \in M(y)$ the following is verified

$$(\xi - \eta)(x - y) \geq 0 \quad [(\xi - \eta)(x - y) \leq 0].$$

Theorem 3.3 Let A be an open convex set, and f be G-differentiable in A. The set-valued function $\partial f: A \rightrightarrows E'$ is non decreasing (non increasing) monotonous in A if, and only if, f is a convex (concave) functional in A.
Proof. If f is a convex functional in A, then $\partial f(x) = \partial_R f(x)$ and for fixed x, y in A we have (see [13])

$$\xi \in \partial f(x) \ \ iff \ \ f(y) \geq f(x) + \xi(y - x)$$
$$\eta \in \partial f(y) \ \ iff \ \ f(x) \geq f(y) + \eta(x - y)$$

for all $\xi \in \partial f(x), \eta \in \partial f(y)$ it follows that

$$\xi(x - y) \geq f(x) - f(y) \geq \eta(x - y); (\xi - \eta)(x - y) \geq 0$$

and ∂f is non-decreasing monotonous in A.
Conversely, suppose ∂f is non-decreasing monotonous in A. f is a convex functional in A if, and only if,

$$f[a + \lambda(b - a)] \leq f(a) + \lambda[f(b) - f(a)] \tag{1}$$

for all *a, b* in *A* and $\lambda \in [0,1]$. (1) is equivalent to

$$(1 - \lambda)[f(x) - f(a)] \leq \lambda [f(b) - f(x)] \tag{2}$$

where $x = a + \lambda(b - a)$. From 2.10, there exist c in (a, x) and $\xi \in \partial f(c)$ such that $f(x) - f(a) = \xi(x - a)$ and there exist d in (x, b) and $\eta \in \partial f(d)$, so that $f(b) - f(x) = \eta(b - x)$ which (2) is equivalent to

$$(1 - \lambda)\xi(x - a) \leq \lambda\eta\,(b - x) \tag{3}$$

and since $(1 - \lambda)(x - a) = \lambda(b - x) = \lambda(1 - \lambda)(b - a)$, (3) is equivalent to

$$\xi(b - a) \leq \eta(b - a)$$

It is clear that $d - c = \alpha(b - a)$ for some $0 < \alpha < 1$; since *f* is non-decreasing monotonous in *A*, $(\xi - \eta)(c - d) \geq 0$ must be true and therefore

$$(\xi - \eta)(a - b) = \frac{1}{\alpha}(\xi - \eta)(c - d) \geq 0$$

from where $\xi(b - a) \leq \eta(b - a)$ and consequently *f* is a convex functional in *A*.

The proof for a concave functional is analogous since $-f$ is convex and $\partial(-f)(x) = -\partial f(x)$.

Corollary 3.4 Let f be G-differentiable in A and $0 \in \partial f(a)$. If $\partial > 0$ exists such that ∂f is non-decreasing (non-increasing) monotonous in $B(a,\delta) \subset A$, then f reaches a local minimum (maximum) in a.

The proof is a direct consequence of Theorem 3.3.

Finally, here are two examples. The first shows a function with G-derivative for which the Clarke's derivative does not exist, and the second shows the application of extremum conditions to a simple case.

Example 3.5

Suppose $f(x) = x sin(1/x)$ *if* $x \in (-1,1)$ *and* $x \neq 0; f(0) = 0$. f is stable in $(-1,1)$ and therefore G-derivable. Nevertheless, it is not locally Lipschitzian. $\partial f(0) = [-1,1]$ while there is not derivative in the Clarke's sense at 0.

Example 3.6 Suppose that $f: \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(x,y) = y$, if $y \geq 0$ and $f(x,y) = -y$ if $y < 0$. Let $v = (v_1, v_2) \in \mathbb{R}^2$, $v \neq 0$, then
$\partial_v f(x,0) = [-|v_2|, |v_2|]$ and $\partial f(x,0) = (0,[-1,1])$ for each $x \in \mathbb{R}$.
As in the remaining points, f is Fréchet differentiable, we find:

$$
\begin{array}{ll}
\partial f(x,y) = (0,1) & \textit{if } y > 0 \\
\partial f(x,y) = (0,[-1,1]) & \textit{if } y = 0 \\
\partial f(x,y) = (0,-1) & \textit{if } y < 0
\end{array}
$$

Later the points that verify the necessary extremum condition are those in $(x,0)$. On the other hand, given $(x_1,y_1),(x_2,y_2) \in \mathbb{R}^2$ the following is found:

$$
< \partial f(x_1,y_1) - \partial f(x_2,y_2), (x_1,y_1) - (x_2,y_2) > =
$$

$$
\begin{array}{lll}
0 \quad \textit{if} \quad y_1 y_2 > 0 & \textit{or} & y_1 = y_2 = 0 \\
2|y_1 - y_2| & \textit{if} & y_1 y_2 < 0 \\
|[0,2]y_2| & \textit{if} & y_1 = 0,\ y_2 \neq 0 \\
|[0,2]y_1| & \textit{if} & y_1 \neq 0,\ y_2 = 0
\end{array}
$$

Then ∂f is non-decreasing monotonous in $\mathbb{R}^2$ and from theorem 3.3 and corollary 3.4 we conclude that each point $(x,0)$ is a local minimum of f.

References

[1] AUBIN, J. P. FRANKOWSKA, H. *Set-valued analysis. System, Control, Foundation and Applications*. Vol. 2. Birkhauser, Boston (1990).

[2] BORWEIN, J.M. The differentiability of real functions on normed linear spaces using generalized subgradients. *Journal of Math. Anal. Appl. 128, num. 2 (1987)*, pp. 512-534.

[3] BUTLER, G.J., TIMOURIAN, J.G., VIGER, C. The rank theorem for locally Lipschitz continuous functions. *Canad. Math. Bull. 31, num. 2 (1988)*, pp. 217-226.

[4] CLARKE, F. Generalized gradients and applications. *Transac. Amer. Math. Soc. 305 (1975)*, pp. 246-262.

[5] CLARKE, F. On the inverse function theorem. *Pacific Journal of Mathematics 64, num. 1 (1976)*, pp. 97-102.

[6] CLARKE, F. Generalized gradients of Lipschitz functionals. *Adv. in Math. 40 (1981)*, pp. 52-67.

[7] HIRIART-URRUTY, J.B. Miscellanies of nonsmooth analysis and optimization. *Lect. Notes in Econ. and Math. Systems 255 (1985)*, pp. 8-24.

[8] IOFFE, A. Nonsmooth analysis: differential calculus of non-differentiable mappings. *Trans. Amer. Math. Society 266, num. 1 (1981)*, pp. 1-56.

[9] LEBOURG, G. Valeur moyenne pour gradient généralisé. *C.R. Acad. Sci. Paris 281, num. 10 (1975)*, pp. 795-797.

[10] NOVO, V. Optimización de funciones no derivables. *Rev. Acad. Cienc. Zaragoza 46 (1991)*, pp. 37-50.

[11] NOVO, V., MARIN, L.R. An extension of the inverse function theorem. *Rev. Real Acad. Cienc. Ex. Fis. Nat. Tomo 84. Cuad. 4 (1990)*, pp. 575-588.

[12] POURCIAU, H.B. Global invertibility of nonsmooth mappings. *Jour. Math. Anall Apple. 131 (1988)*, pp. 170-179.

[13] ROCKAFELLART, T. *Convex Analysis*. Princeton Univ. Press (1970).