Journal of Global Optimization (2006) 35: 475–482
DOI 10.1007/s10898-005-6016-5

A Chain Rule for Upper Semicontinuous (CF)-Mappings

TAE-HWAN YOON
Department of Mathematics, University of Seoul, Seoul 130-743, South Korea
(e-mail: yoonthsh@yahoo.co.kr)

(Received 5 November 2004; accepted in revised form 16 December 2005)

Abstract. A chain rule for calculating convexificators of composite functions of the type $f = h \circ g$, with the inner factor g being a transformation of $\mathbb{R}^n$, is proposed. The proof is based on a double application of a mean value theorem for (CF)-mappings due to V.F. Demyanov and V. Jeyakumar (see [4]), along with a stability property for the support of a certain ε-perturbation of (CF)-mappings.

Mathematics Subject Classification (2000). 49J52

Key words: Convexificator, (CF)-mapping, Chain rule

1. Introduction and Preliminaries

Let the function $f: \Omega \to \mathbb{R}$ be locally Lipschitz on an open set $\Omega \subset \mathbb{R}^n$. Then its *upper and lower Dini directional derivatives*

$$f_D^{\uparrow}(x, d) = \limsup_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}$$

and

$$f_D^{\downarrow}(x, d) = \liminf_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}$$

are also Lipschitz as functions of the direction d.

The *upper and lower Clarke directional derivatives* are defined as follows (see [2, 5]):

$$f_{cl}^{\uparrow}(x, d) = \limsup_{\substack{x' \to x \\ t \downarrow 0}} \frac{f(x' + td) - f(x')}{t} \tag{1}$$

and

$$f_{cl}^{\downarrow}(x, d) = \liminf_{\substack{x' \to x \\ t \downarrow 0}} \frac{f(x' + td) - f(x')}{t}. \tag{2}$$

Since f is locally Lipschitz, the limits in (1) and (2) exist and are finite, and the following properties hold:

$$f_{cl}^{\uparrow}(x,d)=\max_{v\in\partial_{cl}f(x)}\langle v,d\rangle$$

and

$$f_{cl}^{\downarrow}(x,d)=\min_{w\in\partial_{cl}f(x)}\langle w,d\rangle,$$

where

$$\partial_{cl}f(x)=co\left\{v\in\mathbb{R}^n \mid \exists\{x_k\}\colon\ x_k\to x,\ x_k\in T(f),\ f^{'}(x_k)\to v\right\}$$

and $T(f)$ is the set of points of Ω where f is differentiable. The set $\partial_{cl}f(x)$, called *Clarke subdifferential* of f at x, is a nonempty, convex and compact set. We note that the relation that the Clarke subdifferential satisfies is usually known as the *strong convexificator condition*.

Michel and Penot proposed the following generalized derivatives (see [6])

$$f_{mp}^{\uparrow}(x,d)=\sup_{q\in\mathbb{R}^n}\left\{\limsup_{t\downarrow 0}\frac{f(x+t(d+q))-f(x+tq)}{t}\right\} \tag{3}$$

and

$$f_{mp}^{\downarrow}(x,d)=\inf_{q\in\mathbb{R}^n}\left\{\liminf_{t\downarrow 0}\frac{f(x+t(d+q))-f(x+tq)}{t}\right\}. \tag{4}$$

We call (3) and (4) the *upper and lower Michel–Penot directional derivative* of f at x in the direction d, respectively. Since f is locally Lipschitz, there exists a convex compact set $\partial_{mp}f(x)$, called *Michel–Penot subdifferential* of f at x, and the following properties hold:

$$f_{mp}^{\uparrow}(x,d)=\max_{v\in\partial_{mp}f(x)}\langle v,d\rangle$$

and

$$f_{mp}^{\downarrow}(x,d)=\min_{w\in\partial_{mp}f(x)}\langle w,d\rangle.$$

Let $h\colon\mathbb{R}^n\to\mathbb{R}$ be a positively homogeneous function of degree 1. Demyanov introduced the concept of convexificator (see [3]). A convex compact set $C\subset\mathbb{R}^n$ is a *convexificator (CF)* of h if

$$\min_{w\in C}\langle w,d\rangle\leqslant h(d)\leqslant\max_{v\in C}\langle v,d\rangle \qquad \forall d\in\mathbb{R}^n.$$

We get

$$\min_{w\in\partial_{cl} f(x)}\langle w,d\rangle \leqslant f_D^{\downarrow}(x,d) \leqslant f_D^{\uparrow}(x,d) \leqslant \max_{v\in\partial_{cl} f(x)}\langle v,d\rangle.$$

Hence, the Clarke subdifferential of f at x is a convexificator of both functions

$$h(d)=f_D^{\uparrow}(x,d) \quad \text{and} \quad h(d)=f_D^{\downarrow}(x,d).$$

Also we conclude that

$$\min_{w\in\partial_{mp} f(x)}\langle w,d\rangle \leqslant f_D^{\downarrow}(x,d) \leqslant f_D^{\uparrow}(x,d) \leqslant \max_{v\in\partial_{mp} f(x)}\langle v,d\rangle.$$

Thus, the Michel–Penot subdifferential of f at x is also a convexificator of both functions

$$h(d)=f_D^{\uparrow}(x,d) \quad \text{and} \quad h(d)=f_D^{\downarrow}(x,d).$$

We call a convexificator $C^+(x)$ $(C^-(x))$ of the function $h(d)=f_D^{\uparrow}(x,d)$ $(f_D^{\downarrow}(x,d))$ an *upper (lower) convexificator* of f at x. If $C(x)$ is a convexificator of both functions $f_D^{\uparrow}(x,d)$ and $f_D^{\downarrow}(x,d)$, we say that $C(x)$ is a *convexificator* of f at x.

Now we define a (CF)-mapping for a locally Lipschitz function, introduced by Demyanov and Jeyakumar (see [4]). A mapping $C^+\,(C^-)\colon \Omega\to 2^{\mathbb{R}^n}$ is an *upper (lower) (CF)-mapping* of f on Ω if for every $x\in\Omega$ the convexificator $C^+(x)$ $(C^-(x))$ satisfies the following inequalities

$$\min_{w\in C^+(x)}\langle w,d\rangle \leqslant f_D^{\uparrow}(x,d) \leqslant \max_{v\in C^+(x)}\langle v,d\rangle \qquad \forall d\in\mathbb{R}^n$$

$$\left(\min_{w\in C^-(x)}\langle w,d\rangle \leqslant f_D^{\downarrow}(x,d) \leqslant \max_{v\in C^-(x)}\langle v,d\rangle \qquad \forall d\in\mathbb{R}^n\right).$$

A mapping $C\colon \Omega\to 2^{\mathbb{R}^n}$ is called a *(CF)-mapping* of f if for every $x\in\Omega$ the convexificator $C(x)$ satisfies the inequalities

$$\min_{w\in C(x)}\langle w,d\rangle \leqslant f_D^{\downarrow}(x,d) \leqslant f_D^{\uparrow}(x,d) \leqslant \max_{v\in C(x)}\langle v,d\rangle \qquad \forall d\in\mathbb{R}^n.$$

The following is the Mean-Value Theorem for a (CF)-mapping of f by Demyanov and Jeyakumar (see [4]).

THEOREM 1. (Mean-Value Theorem). *Let* $C\colon \Omega \to 2^{\mathbb{R}^n}$ *be a (CF)-mapping of a locally Lipschitz function* f. *If the interval* $co\{x_1, x_2\} \subset \Omega$, *then there exist a* $\gamma \in (0,1)$ *and a* $v \in C(x_1 + \gamma(x_2 - x_1))$ *such that*

$$f(x_2) - f(x_1) = \langle v, x_2 - x_1 \rangle.$$

Let $\Gamma\colon \Omega \to 2^{\mathbb{R}^n}$ be a set-valued map, where Ω is an open set $\subset \mathbb{R}^n$. We define Γ to be *upper semicontinuous at* $x \in \Omega$ if for any $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$\Gamma(x') \subset \Gamma(x) + \varepsilon B \quad \text{for all } x' \in x + \delta B,$$

where B is the open unit ball in $\mathbb{R}^n$.

2. Main result

Let $f = h \circ g$, where $g\colon \Omega \to \mathbb{R}^n$ and $h\colon \mathbb{R}^n \to \mathbb{R}$ are locally Lipschitz functions, and Ω is an open set $\subset \mathbb{R}^n$. The component functions of g will be denoted g_i $(i = 1, 2, \ldots, n)$ and B denotes the open unit ball in $\mathbb{R}^n$. We assume that each g_i is Lipschitz at x and h is Lipschitz at $g(x)$: this implies that f is Lipschitz at x. Assume that $C_{g_i}\colon \Omega \to 2^{\mathbb{R}^n}$ and $C_h\colon \mathbb{R}^n \to 2^{\mathbb{R}^n}$ are upper semicontinuous (CF)-mappings of g_i and h, respectively.

LEMMA 1. *Assume that* $d \in \mathbb{R}^n$ *is fixed, and* v_ε *and* v_0 *are given by*

$$\sup\left\{\sum_i \beta_i \langle \alpha_i, d\rangle\colon \alpha_i \in C_{g_i}(x_i),\, \beta \in C_h(u),\, x_i \in x + \varepsilon B,\, u \in g(x) + \varepsilon B\right\}$$

and

$$\max\left\{\sum_i \beta_i \langle \alpha_i, d\rangle\colon \alpha_i \in C_{g_i}(x),\, \beta \in C_h(g(x))\right\},$$

respectively. Then $\lim_{\varepsilon \to 0^+} v_\varepsilon = v_0$ *holds.*

Proof. Let any $\delta > 0$ be given. We can choose ε so that

$$C_{g_i}(x + \varepsilon B) \subset C_{g_i}(x) + \delta B, \tag{5}$$

$$C_h(g(x) + \varepsilon B) \subset C_h(g(x)) + \delta B \tag{6}$$

because C_{g_i}, C_h are assumed to be upper semicontinuous. Now it follows that

$$
\begin{aligned}
v_\varepsilon &\leqslant \sup_{\alpha_i \in C_{g_i}(x)+\delta B} \sup \left\{ \sum_i \beta_i \langle \alpha_i, d\rangle \colon \beta \in C_h(u),\ u \in g(x)+\varepsilon B \right\} && \text{by (5)} \\
&\leqslant \sup_{\alpha_i \in C_{g_i}(x)+\delta B} \sup_{\beta \in C_h(g(x))+\delta B} \left\{ \sum_i \beta_i \langle \alpha_i, d\rangle \right\} && \text{by (6)} \\
&\leqslant v_0 + \sum_i \delta \langle \delta, |d| \rangle \\
&\leqslant v_0 + n\delta^2 |d|.
\end{aligned}
$$

Therefore v_ε is bounded above by $v_0 + n\delta^2|d|$ for all ε sufficiently small, which yields $\lim_{\varepsilon\to 0^+} v_\varepsilon = v_0$ since $v_\varepsilon \geqslant v_0$. □

LEMMA 2. *Assume that* $d \in \mathbb{R}^n$ *is fixed, and* w_ε *and* w_0 *are given by*

$$
\inf \left\{ \sum_i \beta_i \langle \alpha_i, d\rangle : \alpha_i \in C_{g_i}(x_i), \beta \in C_h(u), x_i \in x+\varepsilon B, u \in g(x)+\varepsilon B \right\}
$$

and

$$
\min \left\{ \sum_i \beta_i \langle \alpha_i, d\rangle : \alpha_i \in C_{g_i}(x), \beta \in C_h(g(x)) \right\},
$$

respectively. Then $\lim_{\varepsilon\to 0^+} w_\varepsilon = w_0$ *holds.*

Proof. The proof is similar to that of Lemma 1 except using the minimum. □

REMARK. We note that the set over which the maximum (minimum) in Lemma 1 (2) is taken is the Cartesian product of compact sets (the underlying space is finite-dimensional) under compact-valued upper semicontinuous maps. This guarantees the existence of the maximum v_0 (minimum w_0). Also, we notice that v_ε, w_ε, v_0, w_0 in Lemmas 1, 2 are functions of d.

The following main theorem states that a convexificator for the composite function can be obtained as a set consisting of all linear combinations of elements from convexificators of each component of g, with the coefficients in the combination being component of vectors from a convexificator for the outer factor h.

THEOREM 2. (Chain Rule). *Let* $f = h \circ g$, *where* $g\colon \Omega \to \mathbb{R}^n$ *and* $h\colon \mathbb{R}^n \to \mathbb{R}$ *are locally Lipschitz functions. Assume that* $C_{g_i}\colon \Omega \to 2^{\mathbb{R}^n}$ *and* $C_h\colon \mathbb{R}^n \to 2^{\mathbb{R}^n}$ *are upper semicontinuous (CF)-mappings of* g_i *and* h, *respectively. Then the set*

$$\left\{\sum_i \beta_i \alpha_i : \alpha_i \in C_{g_i}(x),\ \beta \in C_h(g(x))\right\} \tag{7}$$

is a convexificator of f *at* x *and*

$$x \mapsto \left\{\sum_i \beta_i \alpha_i : \alpha_i \in C_{g_i}(x),\ \beta \in C_h(g(x))\right\}$$

is a (CF)-mapping of f.

Proof. Assume that $d \in \mathbb{R}^n$ is fixed. We can find a positive t near 0 such that

$$\frac{f(x+td) - f(x)}{t} - \varepsilon \leqslant f_D^{\downarrow}(x,d) \leqslant f_D^{\uparrow}(x,d) \leqslant \frac{f(x+td) - f(x)}{t} + \varepsilon.$$

The degree of nearness is chosen to guarantee

$$td \in \varepsilon B,\ g(x+td) \in g(x) + \varepsilon B.$$

By the Mean-Value Theorem in Preliminaries (Theorem 1), we may write

$$\begin{aligned} f(x+td) - f(x) &= h(g(x+td)) - h(g(x)) \\ &= \sum_i \beta_i \{g_i(x+td) - g_i(x)\} \end{aligned}$$

where $\beta \in C_h(u)$ and u is a point in the segment $[g(x+td), g(x)]$ and hence in $g(x) + \varepsilon B$. By another application of the Mean-Value Theorem, the last term above can be expressed as

$$\sum_i \beta_i \langle \alpha_i, td\rangle$$

where $\alpha_i \in C_{g_i}(x_i)$ and x_i is a point in the segment $[x+td, x]$ and hence in $x + \varepsilon B$. Therefore

$$\sum_i \beta_i \langle \alpha_i, d\rangle - \varepsilon \leqslant f_D^{\downarrow}(x,d) \leqslant f_D^{\uparrow}(x,d) \leqslant \sum_i \beta_i \langle \alpha_i, d\rangle + \varepsilon$$

holds. Then we have

$$w_\varepsilon - \varepsilon \leqslant f_D^{\downarrow}(x,d) \leqslant f_D^{\uparrow}(x,d) \leqslant v_\varepsilon + \varepsilon$$

and hence

$$w_0 \leqslant f_D^{\downarrow}(x,d) \leqslant f_D^{\uparrow}(x,d) \leqslant v_0$$

where v_ε, w_ε, v_0, w_0 are defined and $\lim_{\varepsilon\to 0^+} v_\varepsilon = v_0$, $\lim_{\varepsilon\to 0^+} w_\varepsilon = w_0$ are proved in Lemmas 1, 2 above. Therefore the set (7) is a convexificator of f at x and thus

$$x \mapsto \left\{ \sum_i \beta_i \alpha_i : \alpha_i \in C_{g_i}(x),\ \beta \in C_h\,(g(x)) \right\}$$

is a (CF)-mapping of f as required. □

EXAMPLE 1. Let $f = h \circ g$, where $g\colon \Omega \to \mathbb{R}^n$ and $h\colon \mathbb{R}^n \to \mathbb{R}$ are locally Lipschitz functions. Then the set-valued maps $x \mapsto \partial_{cl} g_i(x)$ and $x \mapsto \partial_{cl} h(x)$ are (CF)-mappings which are upper semicontinuous. Therefore the set

$$\left\{ \sum_i \beta_i \alpha_i : \alpha_i \in \partial_{cl} g_i(x),\ \beta \in \partial_{cl} h\,(g(x)) \right\}$$

is a convexificator of f at x and thus

$$x \mapsto \left\{ \sum_i \beta_i \alpha_i : \alpha_i \in \partial_{cl} g_i(x),\ \beta \in \partial_{cl} h\,(g(x)) \right\}$$

is a (CF)-mapping of f by Theorem 2 above. More in general, we can find the class of mappings called CUSCOs, whose (CF)-mappings are particular examples (see [1]).

Acknowledgement

The author is thankful to the referees for their insightful and constructive comments.

References

1. Borwein, J.M. and Zhu, Q.J. (1999), *Multifunctional and functional analytic techniques in nonsmooth analysis*, NATO Sci. Ser. C Math. Phys. Sci., 528, Kluwer Academic Publishers, Dordrecht.
2. Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York.
3. Demyanov, V.F. (1994), Convexification and concavification of a positively homogeneous function by the same family of linear functions, Università di Pisa, Report 3,208,802.

4. Demyanov, V.F. and Jeyakumar, V. (1997), Hunting for a smaller convex subdifferential, *Journal of Global Optimization* 10, 305–326.
5. Demyanov, V.F. and Rubinov, A.M. (1995), *Constructive Nonsmooth Analysis*, Verlag Peter Lang, Frankfurt a/M.
6. Michel, P. and Penot, J.P. (1984), Calcul sous-différential pour les fonctions lipschitziennes et non-lipschitziennes, *C. R. Acad. Sc. Paris, Ser. I*, 298, 269–272.
7. Rockafellar, R.T. (1970), *Convex Analysis*, Princeton University Press, Princeton N.J.