Davide La Torre e Matteo Rocca

Mean value theorem for continuous vector functions by smooth approximations

2002/12

UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

http://eco.uninsubria.it

In questi quaderni vengono pubblicati i lavori dei docenti della Facoltà di Economia dell'Università dell'Insubria. La pubblicazione di contributi di altri studiosi, che abbiano un rapporto didattico o scientifico stabile con la Facoltà, può essere proposta da un professore della Facoltà, dopo che il contributo sia stato discusso pubblicamente. Il nome del proponente è riportato in nota all'articolo. I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

These Working papers collect the work of the Faculty of Economics of the University of Insubria. The publication of work by other Authors can be proposed by a member of the Faculty, provided that the paper has been presented in public. The name of the proposer is reported in a footnote. The views expressed in the Working papers reflect the opinions of the Authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

Printed in Italy in April 2002
Università degli Studi dell'Insubria
Via Ravasi 2, 21100 Varese, Italy

Mean value theorem for continuous vector functions by smooth approximations

Davide La Torre* Matteo Rocca†

Abstract

In this note a mean value theorem for continuous vector functions is introduced by mollified derivatives and smooth approximations.

1 Preliminary definitions

In this paper, a generalized mean value theorem for continuous vector functions is proved. This result involves generalized derivatives, defined by smooth approximations, following the approach introduced by Craven and Ermoliev, Norkin, Wets ([3, 4]). In particular, when local lipschitzianity is assumed, our mean value theorem reduces to the well known mean value theorem expressed by means of Clarke's generalized Jacobian [2].

We will make use of the following classical definitions and results of Functional Analysis.

Definition 1.1. A sequence of *mollifiers* is any sequence of functions $\{\phi_\epsilon\} : \mathbb{R}^n \to \mathbb{R}_+$, $\epsilon \downarrow 0$, such that:

- $supp\phi_\epsilon := \{x \in \mathbb{R}^n, |\ \phi_\epsilon(x) > 0\} \subseteq \rho_\epsilon clB,\ \rho_\epsilon \downarrow 0$,
- $\int_{\mathbb{R}^n} \phi_\epsilon(x)dx = 1$,

*University of Milan, Department of Economics, Faculty of Political Sciences, via Conservatorio,7, 20122, Milano, Italy. e-mail: davide.latorre@unimi.it

†University of Insubria, Department of Economics, Faculty of Economics, via Ravasi, 2, 21100, Varese, Italy. e-mail: mrocca@eco.uninsubria.it

where B is the unit ball in $\mathbb{R}^n$, clX means the closure of the set X and dx denotes Lebesgue measure.

Example 1.1. [4] Let ϵ be a positive number.

(i) The functions:

$$\phi_\epsilon(x) = \begin{cases} \frac{1}{\epsilon^n}, & \max_{1,\dots,n} |x_i| \leq \frac{\epsilon}{2} \\ \\ 0, & \text{otherwise} \end{cases}$$

are called Steklov mollifiers.

(ii) The functions:

$$\phi_\epsilon(x) = \begin{cases} \frac{C}{\epsilon^n} \exp\left(\frac{\epsilon^2}{\|x\|^2-\epsilon^2}\right), & \|x\| < \epsilon \\ \\ 0, & \|x\| \geq \epsilon \end{cases}$$

with $C \in \mathbb{R}$ such that $\int_{\mathbb{R}^n} \phi_\epsilon(x)dx = 1$, are called standard mollifiers.

It is easy to check that the second family of functions is smooth.

Definition 1.2. [4] Given a locally integrable function $f : \mathbb{R}^n \to \mathbb{R}^m$ and a sequence of bounded mollifiers, define the functions $f_\epsilon(x)$ through the convolution:

$$f_\epsilon(x) := \int_{\mathbb{R}^n} f(x-z)\phi_\epsilon(z)dz = \int_{\mathbb{R}^n} f(z)\phi_\epsilon(x-z)dz.$$

The sequence $f_\epsilon(x)$ is said a sequence of *mollified functions.*

Remark 1.1. There is no loss of generality in considering $f : \mathbb{R}^n \to \mathbb{R}^m$. The results in this paper remain true also if f is defined on an open subset of $\mathbb{R}^n$.

Proposition 1.1. *[4] Let $f \in C(\mathbb{R}^n)$. Then f_ϵ converges continuously to f, i.e. $f_\epsilon(x_\epsilon) \to f(x)$ for all $x_\epsilon \to x$. In fact f_ϵ converges uniformly to f on every compact subset of $\mathbb{R}^n$ as $\epsilon \downarrow 0$.*

Mollified functions have also some differentiability properties, under suitable regularity assumptions on f and the associated mollifiers, as stated in the following:

Proposition 1.2. *[5] Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be locally integrable. Whenever the mollifiers ϕ_ϵ are of class C^k, so are the associated mollified functions. Furthermore if ϕ_ϵ are of class $C^{k,1}$, that is k-times differentiable with locally lipschitzian Jacobians, then so are the associated mollified functions.*

By means of mollified functions it is possible to define generalized directional derivatives for a nonsmooth function f. Such an approach has been deepened by several authors (see e.g. [3, 4]) in the scalar case.

Definition 1.3. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a locally integrable function, let $\epsilon_n \downarrow 0$ as $n \to +\infty$ and consider the sequence $f_n := f_{\epsilon_n}$ of mollified functions with associated mollifiers $\phi_{\epsilon_n} \in C^1$. Given $x, d \in \mathbb{R}^n$ we define the following sets:

$$\partial f(x_0; d) = \{l = \lim_{n \to +\infty} \nabla f_n(x_n)d, x_n \to x_0\}$$

$$\partial_\infty f(x_0; d) = \{l = \lim_{n \to +\infty} t_n \nabla f_n(x_n)d, x_n \to x_0, t_n \downarrow 0^+\} \backslash \{0\}.$$

Proposition 1.3.

- *$\partial f(x_0; d)$ is a closed subset of $\mathbb{R}^m$.*
- *$\partial_\infty f(x_0; d)$ is a closed cone of $\mathbb{R}^m$.*
- *$\xi \partial f(x_0; d) \subseteq \partial(\xi f)(x_0; d)$, $\forall \xi \in \mathbb{R}^m$. If f is locally lipschitzian then the equality holds.*

Proof. Omitted since trivial. □

Proposition 1.4. *If f is locally lipschitzian then $\partial f(x_0; d) \subseteq \partial_C f(x_0)d$, where $\partial_C f(x_0)$ is Clarke's generalized Jacobian of f at x_0 [2].*

Proof. In fact, $\forall \xi \in \mathbb{R}^m$, the following inclusion holds [4]:

$$\partial(\xi f)(x_0; d) \subseteq \partial_C(\xi f)(x_0)d.$$

Hence:

$$\xi \partial f(x_0; d) \subseteq \xi \partial_C f(x_0)d$$

and then the thesis follows by a standard separation argument. □

Corollary 1.1. *If f is C^1 then $\partial f(x_0; d) = \nabla f(x_0)d$.*

Proof. If f is C^1, then $\partial_C f(x_0)d = \nabla f(x_0)d$ [2] and then the thesis follows from the previous proposition. □

2 Generalized mean value theorem

Theorem 2.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a given continuous function. Then the following mean value theorem holds:*

$$f(x) - f(y) \in \left\{ \operatorname{conv}_{\delta \in [x,y]} \partial f(\delta; x - y) + \operatorname{conv}_{\delta \in [x,y]} \partial_\infty f(\delta; x - y) \cup \{0\} \right\} \bigcup$$
$$\operatorname{conv}_{\delta \in [x,y]} \left\{ \partial_\infty f(\delta; x - y) + f(x) - f(y) \right\}.$$

where $\operatorname{conv}_{\delta \in [x,y]} A(\delta)$ *denotes the convex hull of the sets* $A(\delta)$*,* $\delta \in [x, y]$.

Proof. In fact, for the scalar function ξf_n, we have:

$$\xi f_n(x) - \xi f_n(y) = \xi \nabla f_n(\delta_n(\xi))(x - y)$$

where $\xi \in \mathbb{R}^m$ and $\delta_n(\xi) \in (x, y)$. So we have:

$$\xi f_n(x) - \xi f_n(y) \in \xi A_n$$

where $A_n = \{\nabla f_n(\delta)(x - y), \delta \in [x, y]\}$ and obviously A_n is compact. So by a standard separation argument, we have:

$$f_n(x) - f_n(y) \in \operatorname{conv} A_n$$

where conv stands for the convex hull of A_n. Let now $l_n = f_n(x) - f_n(y)$. For all $n \in \mathbb{N}$, by Charatheodory theorem, we have:

$$l_n = \sum_{j=1}^{m+1} \lambda_{j,n} a_{j,n},$$

where $\sum_{j=1}^{m+1} \lambda_{j,n} = 1$, $\lambda_{j,n} \geq 0$, $j = 1, \dots, m + 1$, $a_{j,n} \in A_n$. Then:

$$l_n = \sum_{j \in I_1} \lambda_{j,n} a_{j,n} + \sum_{j \in I_2} \lambda_{j,n} a_{j,n} + \sum_{j \in I_3} \lambda_{j,n} a_{j,n}$$

where:

- for all $j \in I_1$ the sequence $a_{j,n}$ is bounded and it converges to $a_{j,0}$. Since $a_{j,n} \in A_n$, $\forall n \in \mathbb{N}$, then $a_{j,n} = \nabla f_n(\delta_{j,n})(x - y)$, $\delta_{j,n} \in [x, y]$. Eventually by extracting a subsequence, we have $\delta_{j,n} \to \delta_j \in [x, y]$ and then:

$$a_{j,0} = \lim_{n \to +\infty} a_{j,n} = \lim_{n \to +\infty} \nabla f_n(\delta_{j,n})(x - y) \in \partial f(\delta_j; d).$$

- for all $j \in I_2$, the sequence $a_{j,n}$ is unbounded but the sequence $\lambda_{j,n}a_{j,n}$ is bounded and it converges to $a_{j,*}$.
- for all $j \in I_3$, the sequence $\lambda_{j,n}a_{j,n}$ is unbounded but there exists $j_0 \in I_3$ such that the sequence $\frac{\lambda_{j,n}a_{j,n}}{\|\lambda_{j_0,n}a_{j_0,n}\|}$ converges to $a_{j,\infty}$, $\forall j = 1, \ldots, m+1$.

We now consider the case in which I_3 is not empty. Then:

$$0 = \lim_{n\to+\infty} \frac{l_n}{\|\lambda_{j_0,n}a_{j_0,n}\|} =$$

$$\lim_{n\to+\infty} \sum_{j\in I_3} \frac{\lambda_{j,n}a_{j,n}}{\|\lambda_{j_0,n}a_{j_0,n}\|} = \sum_{j\in\ I_3} a_{j,\infty},$$

with $a_{j_0,\infty} \neq 0$. Since $a_{j,n} = \nabla f_n(\delta_{j,n})(x-y)$, $\delta_{j,n} \to \delta_j$, $\frac{\lambda_{j,n}}{\|\lambda_{j_0,n}a_{j_0,n}\|} \to 0$ for every $j \in I_3$, we have $a_{j,\infty} \in \partial_\infty f(\delta_j; d) \cup \{0\}$. Furthermore $a_{j_0,\infty} \neq 0$ and then:

$$0 \in \operatorname{conv}_{\delta\in[x,y]} \partial_\infty f(\delta; x-y).$$

We now consider the case in which I_3 is empty. Eventually extracting subsequences, let $\lambda_{j,0} = \lim_{n\to+\infty} \lambda_{j,n}$. Then, we have $\lambda_{j,0} = 0\ \forall j \in I_2$, $\sum_{j\in I_1} \lambda_{j,0} = 1$ and $a_{j,*} \in \partial_\infty f(\delta_j; x-y) \cup \{0\}$. So:

$$l = \lim_{n\to+\infty} l_n = \sum_{j\in I_1} \lambda_{j,0}a_{j,0} + \sum_{j\in I_2} a_{j,*}$$

Obviously $\sum_{j\in I_2} a_{j,*} \in \operatorname{conv}_{\delta\in[x,y]} \partial_\infty f(\delta, x-y) \cup \{0\}$. So we have:

$$f(x) - f(y) \in \left\{\operatorname{conv}_{\delta\in[x,y]} \partial f(\delta, x-y) + \operatorname{conv}_{\delta\in[x,y]} \partial_\infty f(\delta, x-y) \cup \{0\}\right\} \bigcup$$

$$\operatorname{conv}_{\delta\in[x,y]} \left\{\partial_\infty f(\delta, x-y) + f(x) - f(y)\right\}.$$

□

Corollary 2.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$. If we define a generalized upper derivative as:*

$$Df(x; d) = \limsup_{n\to+\infty, x_n\to x_0} \nabla f_n(x_n)d$$

then the following mean value theorem holds:

$$f(x) - f(y) \leq Df(\xi; x-y)$$

where $\xi \in [x, y]$.

Proof. We only consider the case in which $Df(s, x-y) < +\infty$ $\forall s \in [x, y]$ (if $\exists \xi \in [x, y]$ such that $Df(s, x-y) = +\infty$ the thesis is trivial). Then $\partial_\infty f(s, x-y) \subset (-\infty, 0)$, $\forall s \in [x, y]$. If, ab absurdo, $f(x)-f(y) \in \text{conv}_{\delta \in [x,y]} \{\partial_\infty f(\delta, x-y) + f(x) - f(y)\}$ then $\exists \xi \in [x, y]$ such that $f(x)-f(y) \in (-\infty, 0)+f(x)-f(y)$ that is $0 \in (-\infty, 0)$. So $f(x) - f(y) \in \text{conv}_{\delta \in [x,y]} \partial f(\delta, x-y) + \text{conv}_{\delta \in [x,y]} \partial_\infty f(\delta, x-y)$. Then $f(x) - f(y) = a + b$ where $a \in \text{conv}_{\delta \in [x,y]} \partial f(\delta, x-y)$ and $b \in \text{conv}_{\delta \in [x,y]} \partial_\infty f(\delta, x-y)$. Then $\exists \xi \in [x, y]$ such that $a \leq \sup_{l \in \partial f(\xi, x-y)} l$, that is $a \leq Df(\xi, x-y)$, and $b \leq 0$. So the thesis follows. □

Corollary 2.2. *If f is locally Lipschitz, then we have:*

$$f(x) - f(y) \in \text{conv}_{\delta \in [x,y]} \partial_C f(\delta)(x-y).$$

Proof. We know that (proposition 1.4) at any point δ, $\partial f(\delta; x-y) = \partial_C(\delta)(x-y)$. furthermore, form the Lipschitz hypothesis it follows easily that $\partial_\infty f(\delta; x-y) = \emptyset$, whenever δ. So the thesis follows. □

References

[1] Borwein J.M., Fitzpatrick S.P., Giles J.R.: *The differentiability of real functions on normed linear space using generalized subgradients.* J. Math. Anal. Appl., 128, 1987, 512-534.

[2] Clarke F.H.: *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[3] Craven B.D.: *Nonsmooth multiobjective programming*, Numerical Functional Analysis and Optimization, 10 1989, 49-64.

[4] Ermoliev Y.M., Norkin V.I., Wets R.J-B.: *The minimization of semicontinuous functions: mollifier subgradients.* SIAM J. Control and Optimization, 33, 1, 1995, 149-167.

[5] Evans M.J., Weil C.E.: *Peano derivatives: a survey.* Real Analysis Exchange, 7, 1981-82, 5-23.

[6] Jeyakumar V., Luc D.T.: *Approximate Jacobian matrices fro nonsmooth continuous maps and C^1 optimization*, SIAM J. Control and Optim. 36, 5, 1998, 1815-1832.