Davide La Torre e Matteo Rocca

$C^{1,1}$ functions and optimality conditions

2002/15

UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

http://eco.uninsubria.it

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

$C^{1,1}$ functions and optimality conditions*

Davide La Torre† Matteo Rocca‡

22nd April 2002

Abstract

In this work we provide a characterization of $C^{1,1}$ functions on $\mathbb{R}^n$ (that is, differentiable with locally Lipschitz partial derivatives) by means of second directional divided differences. In particular, we prove that the class of $C^{1,1}$ functions is equivalent to the class of functions with bounded second directional divided differences. From this result we deduce a Taylor's formula for this class of functions and some optimality conditions. The characterizations and the optimality conditions proved by Riemann derivatives can be useful to write minimization algorithms; in fact, only the values of the function are required to compute second order conditions.

Keywords: Divided differences, Riemann derivatives, $C^{1,1}$ functions, nonlinear optimization, generalized derivatives

1 Introduction

The study of the class of $C^{1,1}$ functions has been renewed since the work of Hiriart-Urruty in his doctoral thesis [7]. The need for investigating these functions, as pointed out in [8], [10], [23], [24] and [25], comes from the fact that several problems of applied mathematics including variational inequalities, semi-infinite programming, iterated local minimization, etc. involve differentiable functions with no hope to be twice differentiable. In [8] the authors introduced the concept of generalized Hessian matrices and derived second order optimality conditions for nonlinear constrained problems. Further applications can be found in [10], [15], [19], [20], [22].

In this section we recall some concepts which are fundamental for understanding the proof of the results.

*This work has been supportted by the F.A.R. 2001, of the University of Insubria.

†Contact author: University of Milan, Department of Economics, Faculty of Political Sciences, via Conservatorio,7, 20122, Milano, Italy. Tel.: +39276074462. Fax: +39276023198 E-mail: davide.latorre@unimi.it

‡University of Insubria, Department of Economics, Faculty of Economics, via Ravasi, 2, 21100, Varese, Italy. Tel: +39258365693. Fax: +39258365617. E-mail: mrocca@eco.uninsubria.it

1.1 Riemann derivatives

In the following we will consider a function $f : \Omega \to \mathbb{R}$, with Ω an open subset of $\mathbb{R}^n$. For such a function we define:

$$\delta_2^d f(x; h) = f(x + 2hd) - 2f(x + hd) + f(x).$$

with $x \in \Omega, h \in \mathbb{R}$ and $d \in \mathbb{R}^n$.

Definition 1.1. *The second Riemann derivative of f at a point $x \in \Omega$ in the direction $d \in \mathbb{R}^n$ is defined as:*

$$f_r''(x; d) = \lim_{h \to 0} \frac{\delta_2^d f(x; h)}{h^2},$$

if this limit exists.

Definition 1.2. *The second upper and lower Riemann derivatives of f at $x \in \Omega$ in the direction $d \in \mathbb{R}^n$ are defined, respectively, as:*

$$\overline{f}_r''(x; d) = \limsup_{h \to 0} \frac{\delta_2^d f(x; h)}{h^2},$$

$$\underline{f}_r''(x; d) = \liminf_{h \to 0} \frac{\delta_2^d f(x; h)}{h^2}.$$

Similarly we can define differences:

$$\Delta_2^d f(x; h) = f(x + hd) - 2f(x) + f(x - hd),$$

and then the corresponding second Riemann-type derivatives $f_R''(x; d)$, $\overline{f}_R''(x; d)$ and $\underline{f}_R''(x; d)$.

For properties of Riemann derivatives one can see [1], [2], [6] and [16].

Lemma 1.1. *Assume that f is bounded in a neighborhood of the point $x_0 \in \Omega$. If, for a fixed $d \in \mathbb{R}^n$, there exist neighborhoods U of the point x_0 and V of $0 \in \mathbb{R}$ such that $\frac{\delta_2^d f(x;h)}{h^2}$ is bounded on $U \times V \backslash \{0\}$, then also $\frac{f(x+hd)-f(x)}{h}$ is bounded on $U \times V \backslash \{0\}$.*

Proof. From the hypotheses we obtain that there exists a number $\delta > 0$ such that $\forall x \in U$ and $\forall h$ with $|h| \leq \delta, h \neq 0$, the following inequalities hold:

$$\left| f(x + hd) - f(x) - 2\left(f\left(x + \frac{h}{2}d \right) - f(x) \right) \right| \leq M \left| \frac{h}{2} \right|^2,$$

$$\left| f\left(x + \frac{h}{2}d \right) - f(x) - 2\left(f\left(x + \frac{h}{4}d \right) - f(x) \right) \right| \leq M \left| \frac{h}{4} \right|^2, \ldots$$

$$\left| f\left(x + \frac{h}{2^{n-1}}d \right) - f(x) - 2\left(f\left(x + \frac{h}{2^n}d \right) - f(x) \right) \right| \leq M \left| \frac{h}{2^n} \right|^2.$$

Multiplying these inequalities by $1, 2, 2^2, \ldots, 2^{(n-1)}$ respectively, we obtain by addition:

$$\left| f(x+hd) - f(x) - 2^n \left(f\left(x + \frac{h}{2^n} d \right) - f(x) \right) \right| \leq 2M \left| \frac{h}{2} \right|^2 ,$$

and hence:

$$\left| 2^n \frac{f(x + \frac{h}{2^n} d) - f(x)}{h} \right| \leq M'$$

for $\frac{1}{2}\delta \leq |h| \leq \delta$, by using the boundedness of f. Hence, writing $\xi = \frac{h}{2^n}$, we have:

$$\left| \frac{f(x+\xi d) - f(x)}{\xi} \right| \leq M' \quad for \ \frac{\delta}{2^{n+1}} \leq |\xi| \leq \frac{\delta}{2^n}, \ n = 0, 1, \ldots ,$$

and the lemma is established, since n can be arbitrarily chosen.

□

In the following we set:

$$f'(x; d) = \lim_{h \to 0} \frac{f(x+hd) - f(x)}{h}, \quad f''(x; d) = \lim_{h \to 0} \frac{f'(x+hd; d) - f'(x; d)}{h},$$

if these limits exist.

1.2 Standard mollifiers

The function:

$$\phi(x) = \begin{cases} C \exp(\frac{1}{\|x\|^2 - 1}), & \text{if } \|x\| < 1 \\ 0, & \text{if } \|x\| \geq 1 \end{cases}$$

is $C^\infty(\mathbb{R}^n)$ and we can choose the constant $C \in \mathbb{R}$ such that:

$$\int_{\mathbb{R}^n} \phi(x) dx = 1.$$

Definition 1.3. *Let $\varepsilon > 0$. The following functions:*

$$\phi_\varepsilon(x) = \frac{\phi(\frac{x}{\varepsilon})}{\varepsilon^n}$$

are called standard mollifiers.

Definition 1.4. *Let $f : \Omega \to \mathbb{R}$. We say that $f \in C_0^k(\Omega)$ if $f \in C^k(\Omega)$ and*

$$spt_f = \overline{\{x \in \Omega : f(x) \neq 0\}} \subset \Omega.$$

Theorem 1.1. *[3] The functions ϕ_ε are $C^\infty(\mathbb{R}^n)$ and satisfy:*

- $\int_{\mathbb{R}^n} \phi_\varepsilon(x) dx = 1$

- $spt_{\phi_\varepsilon} \subset B(0,\varepsilon) = \{x \in \mathbb{R}^n : \|x\| < \varepsilon\}$.

For a bounded function $f : \Omega \to \mathbb{R}$, and $\varepsilon > 0$, we define functions $f_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ by the convolution $f_\varepsilon(x) = \int_\Omega \phi_\varepsilon(y-x)f(y)dy$. Observe that $f_\varepsilon(x) = 0$ if $x \notin \Omega + B(0,\varepsilon)$ and that $f_\varepsilon \in C^\infty(\mathbb{R}^n)$.

Theorem 1.2. *[3] Suppose that $f \in L^1_{loc}(\Omega)$. Then $f_\varepsilon(x) \to f(x)$ a.e. $x \in \Omega$, when $\varepsilon \to 0$. If $f \in C(\Omega)$ then the convergence is uniform on compact subsets of Ω.*

Theorem 1.3. *[3] Let K be a compact subset of Ω. Then $\exists \varepsilon_0 > 0$ such that $\forall \varepsilon \leq \varepsilon_0$ and $\forall x \in K$, the following function:*

$$y \to \phi_\varepsilon(y-x)$$

is $C_0^\infty(\Omega)$.

2 The main results

Definition 2.1. *A function $f : \Omega \to \mathbb{R}$ is locally Lipschitz at x_0 when there exist a constant K and a neighborhood U of x_0 such that:*

$$|f(x) - f(y)| \leq K\|x - y\|, \quad \forall x, y \in U.$$

Definition 2.2. *A function $f : \Omega \to \mathbb{R}$ is of class $C^{1,1}$ at x_0 when its first order partial derivatives exist in a neighborhood of x_0 and are locally Lipschitz at x_0.*

Some possible applications of $C^{1,1}$ functions are shown in the following examples.

Example 2.1. *Let $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable on Ω and consider[1] $f(x) = [g^+(x)]^2$ where $g^+(x) = \max\{g(x), 0\}$. Then f is $C^{1,1}$ on Ω.*

Example 2.2. *In many problems in engineering applications and control theory ([23], [24] and the references therein) one has to study nonsmooth semi-infinite optimization problems such as the following:*

$$\textit{minimize } f(x)$$

$$\textit{subject to } \max_{t\in[a,b]} \phi_j(x,t) \leq 0, \ j = 1 \dots l$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $\phi_j : \mathbb{R}^n \to \mathbb{R}$ are C^2, $j = 1 \dots l$, $-\infty < a < x < b < +\infty$. One approach for solving this problem is to convert the functional constraints into equality constraints of the form:

$$h_j(x) = \int_a^b [\max\{\phi_j(x,y), 0\}]^2 \, dt = 0, j = 1 \dots l$$

[1] This type of functions arises in some penalty methods.

and apply the methods of nonlinear programming. Hence the problem becomes:

$$\textit{minimize } f(x)$$

$$\textit{subject to } h_j(x) = 0,\ j = 1 \dots l$$

Since ϕ_j is C^2, it is easy see that the function h_j is $C^{1,1}$ with the gradient:

$$\nabla h_j(x) = 2 \int_a^b \max\{\phi_j(x,t), 0\} \nabla \phi_j(x,t) dt,\ j = 1 \dots l.$$

Example 2.3. *Consider the following minimization problem:*

$$\min f_0(x)$$

over all $x \in \mathbb{R}^n$ such that $f_1(x) \leq 0, \dots f_m(x) \leq 0$. Letting r denote a positive parameter, the augmented Lagrangian L_r [21] is defined on $\mathbb{R}^n \times \mathbb{R}^m$ as

$$L_r(x,y) = f_0(x) + \frac{1}{4r} \sum_{i=1}^{m} \{[y_i + 2r f_i(x)]^+\}^2 - y_i^2.$$

From the general theory of duality which yields L_r as a particular Lagrangian, we know that $L_r(x,\cdot)$ is concave and also that $L_r(\cdot,y)$ is convex whenever the minimization problem is a convex minimization problem. Upon setting $y = 0$ in the previous expression, we observe that:

$$L_r(x,0) = f_0(x) + r \sum_{i=1}^{m} [f_i^+(x)]^2$$

is the ordinary penalized version of the minimization problem. L_r is differentiable everywhere on $\mathbb{R}^n \times \mathbb{R}^m$ with:

$$\nabla_x L_r(x,y) = \nabla f_0(x) + \sum_{j=1}^{m} [y_j + 2r f_j(x)]^+ \nabla f_j(x)$$

$$\frac{\partial L_r}{\partial y_i}(x,y) = \max\{f_i(x), \frac{-y_i}{2r}\}, i = 1 \dots m.$$

When the f_i are C^2 on $\mathbb{R}^n$, L_r is $C^{1,1}$ on $\mathbb{R}^{n+m}$. The dual problem corresponding to L_r is by definition:

$$\max g_r(y)$$

over $y \in \mathbb{R}^m$, where $g_r(y) = \inf_{x\in\mathbb{R}^n} L_r(x,y)$. In the convex case with $r > 0$, g_r is again $C^{1,1}$ concave function with the following uniform Lipschitz property on ∇g,

$$|\nabla g_r(y) - \nabla g_r(x)| \leq \frac{1}{2r}|y - y'|,\ \forall y, y' \in \mathbb{R}^m.$$

The following result characterizes a function of class $C^{1,1}$ by the boundness of second-order divided differences.

Theorem 2.1. *Assume that the function $f : \Omega \to \mathbb{R}$ is bounded on a neighborhood of the point $x_0 \in \Omega$. Then f is of class $C^{1,1}$ at x_0 if and only if there exist neighborhoods U of x_0 and V of $0 \in \mathbb{R}$ such that $\frac{\delta_2^d f(x;h)}{h^2}$ is bounded on $U \times V \backslash \{0\}$, $\forall d \in S^1 = \{d \in \mathbb{R}^n : \|d\| = 1\}$.*

Proof. i) Sufficiency. From lemma 1.1, since $\frac{\delta_2^d f(x;h)}{h^2}$ is bounded on $U \times V\backslash\{0\}$, $\forall d \in S^1$, the same holds for $\frac{f(x+hd)-f(x)}{h}$. Observe that this last fact implies that f is locally Lipschitz at x_0 and hence continuous in a neighborhood of x_0. For every x in a neighborhood of x_0 and for ε "sufficiently small", we have, for $d \in S^1$:

$$f'_\varepsilon(x; d) = \lim_{h\to 0} \frac{f_\varepsilon(x + hd) - f_\varepsilon(x)}{h} =$$

$$\lim_{h\to 0} \frac{1}{h}\left[\int_\Omega \phi_\varepsilon(y - x - hd)f(y)dy - \int_\Omega \phi_\varepsilon(y - x)f(y)dy\right].$$

Putting $z = y - hd$, we obtain:

$$\int_\Omega \phi_\varepsilon(y - x - hd)f(y)dy = \int_{\Omega-\{hd\}} \phi_\varepsilon(z - x)f(z + hd)dz.$$

From theorem 1.3, we know that, for ε "sufficiently small", the functions $z \to \phi_\epsilon(z - x)$ are $C_0^\infty(\Omega)$ and hence, if also $|h|$ is "small enough", we get:

$$\int_{\Omega-\{hd\}} \phi_\varepsilon(z - x)f(z + hd)dz = \int_\Omega \phi_\varepsilon(z - x)f(z + hd)dz.$$

It follows that:

$$f'_\varepsilon(x; d) = \lim_{h\to 0} \int_\Omega \frac{f(z + hd) - f(z)}{h} \phi_\varepsilon(z - x)dz.$$

Furthermore one can easily see that:

$$f''_\varepsilon(x; d) = \lim_{h\to 0} \frac{f_\varepsilon(x + 2hd) - 2f_\varepsilon(x + hd) + f_\varepsilon(x)}{h^2}$$

and similarly deduce:

$$f''_\varepsilon(x; d) = \lim_{h\to 0} \int_\Omega \frac{\delta_2^d f(z; h)}{h^2} \phi_\varepsilon(z - x)dz.$$

From the boundedness of $\frac{\delta_2^d f(x;h)}{h^2}$ and of $\frac{f(x+hd)-f(x)}{h}$, we obtain the existence of a constant M such that $|f'_\varepsilon(x; d)| \le M$ and $|f''_\varepsilon(x; d)| \le M$, for every $d \in S^1$, x in a neighborhood $\tilde{U}$ of x_0 and ε "sufficiently small". Hence, for every $x \in \tilde{U}$ and $d \in S^1$, there exists a sequence ε_n converging to 0 such that $f'_{\varepsilon_n}(x; d)$ converges to a limit which we denote by $\alpha(x; d)$. Observe that $\alpha(x; d)$ is bounded on $\tilde{U}$ whenever $d \in S^1$. For every $x \in \tilde{U}$, $d \in S^1$ and h with $|h|$ "small enough", we can write:

$$f_{\varepsilon_n}(x + hd) = f_{\varepsilon_n}(x) + hf'_{\varepsilon_n}(x; d) + \frac{1}{2}h^2 f''_{\varepsilon_n}(\xi_n; d),$$

where $\xi_n \in (x, x+hd)$.
Recalling theorem 1.2, taking the limit for $n \to +\infty$, it follows that $f''_{\varepsilon_n}(\xi_n; d)$ converges to a limit which we denote by $\beta(x, h, d)$. Moreover:

$$f(x+hd) = f(x) + h\alpha(x; d) + \frac{1}{2}h^2\beta(x, h, d).$$

Observing that $\beta(x, h, d)$ is bounded for $x \in \tilde{U}$, $|h|$ "sufficiently small" and $d \in S^1$, it follows that $\alpha(x; d) = f'(x; d)$.
Furthermore, $\forall d \in S^1$ the functions $f''_{\varepsilon_n}(x; d)$ are bounded on $\tilde{U}$ uniformly with respect to ε and thus the functions $f'_{\varepsilon_n}(x; d)$ satisfy the following uniform Lipschitz condition:

$$\left|f'_{\varepsilon_n}(y; d) - f'_{\varepsilon_n}(x; d)\right| \leq B\|y - x\|, \forall x, y \in \tilde{U}.$$

Since $f'_{\varepsilon_n}(y; d)$ and $f'_{\varepsilon_n}(x; d)$ converge to $f'(y; d)$ and $f'(x; d)$ respectively, we see that $f'(x; d)$ is Lipschitz on $\tilde{U}$, $\forall d \in \mathbb{R}^n$. Taking $d = e^i, i = 1, \ldots, n$ (where e^i is the i-th fundamental vector of $\mathbb{R}^n$), we obtain the thesis.
ii) Necessity. Assume that f is of class $C^{1,1}$ at x_0. Set:

$$\overline{\Delta}_2^d f(x; s, t) = f(x + sd + td) - f(x + td) - f(x + sd) + f(x),$$

where $d \in S^1$, $x \in \Omega$, $s, t \in \mathbb{R}$ and $|s|$ and $|t|$ are "sufficiently small". Applying the mean value theorem, we obtain:

$$\frac{\overline{\Delta}_2^d f(x; s, t)}{st} = \frac{< \nabla f(x + \theta td + sd) - \nabla f(x + \theta td), d >}{s},$$

where $\theta \in (0, 1)$. Since f is of class $C^{1,1}$ at x_0 it follows easily that there exist a constant M, a neighborhood $\tilde{U}$ of x_0 and a number $\delta > 0$ such that, $\forall d \in S^1$ we have:

$$\left|\frac{\overline{\Delta}_2^d f(x; s, t)}{st}\right| \leq M, \quad \forall x \in \tilde{U}, \quad |s| < \delta, \ |t| < \delta.$$

Now the thesis follows observing that if $s = t = h$, then $\overline{\Delta}_2^d f(x; s, t) = \delta_2^d f(x; h)$.

□

Corollary 2.1. *Assume that the function f is bounded on a neighborhood of $x_0 \in \Omega$. Then f is of class $C^{1,1}$ at x_0 if and only if there exist neighborhoods U of x_0 and V of $0 \in \mathbb{R}$ such that $\frac{\Delta_2^d f(x;h)}{h^2}$ is bounded on $U \times V \backslash \{0\}$, $\forall d \in S^1$.*

Proof. The proof is straightforward remembering that:

$$\delta_2^d f(x; h) = \Delta_2^d f(x + hd; h).$$

□

Corollary 2.2. *If f is of class $C^{1,1}$ at x_0, there exist sequences ε_n converging to 0 and $\xi_n \in (x_0, x_0 + hd)$ such that $f''_{\varepsilon_n}(\xi_n; d)$ converges to a limit $\beta(x_0, h, d)$ and it holds:*

$$f(x_0 + hd) = f(x_0) + f'(x_0; d)h + \frac{\beta(x_0, h, d)}{2}h^2.$$

Proof. It is enclosed in the proof of the previous theorem. □

Theorem 2.2. *(Taylor's formula) Let f be a function of class $C^{1,1}$ at x_0.*

(i) If the function $x \to \overline{f}''_r(x;d)$ is upper semicontinuous in a neighborhood of x_0, for a fixed $d \in S^1$, then there exists $\xi \in [x_0, x_0 + hd]$ such that, for h "small enough" we have:

$$f(x_0 + hd) \le f(x_0) + hf'(x_0;d) + \frac{h^2}{2!}\overline{f}''_r(\xi;d)$$

(ii) If the function $x \to \underline{f}''_r(x;d)$ is lower semicontinuous in a neighborhood of x_0, for a fixed $d \in S^1$, then there exists $\xi \in [x_0, x_0 + hd]$ such that for h "small enough" we have:

$$f(x_0 + hd) \ge f(x_0) + hf'(x_0;d) + \frac{h^2}{2!}\underline{f}''_r(\xi;d)$$

Proof. i) Without loss of generality, the term $\beta(x_0;h;d)$ in the previous corollary 2.2 can be expressed as:

$$\beta(x_0;h;d) = \lim_{n\to+\infty} f''_{\epsilon_n}(\xi_n;d)$$

for some sequences $\xi_n \to \xi \in [x_0, x_0 + hd]$ and $\epsilon_n \to 0$. Similarly to the proof of theorem 2.1, one can write that[2]:

$$f''_{\varepsilon_n}(\xi_n, d) = \int_\Omega \phi''_{\varepsilon_n}(y - \xi_n; d) f(y)dy =$$

$$\int_{B(0,1)} \lim_{h\to 0} \frac{\delta_2^d \phi_{\varepsilon_n}(y - \xi_n; h)}{h^2} f(y)dy =$$

$$\lim_{h\to 0} \int_{B(0,1)} \frac{\delta_2^d \phi_{\varepsilon_n}(y - \xi_n; h)}{h^2} f(y)dy =$$

$$\lim_{h\to 0} \int_{B(0,1)} \frac{\delta_2^d f(\xi_n + \varepsilon_n y; h)}{h^2} \phi_{\varepsilon_n}(y)dy \le$$

$$\int_{B(0,1)} \limsup_{h\to 0} \frac{\delta_2^d f(\xi_n + \epsilon_n y; h)}{h^2} \phi(y)dy = \int_{B(0,1)} \overline{f}''_r(\xi_n + \epsilon_n y; d)\phi(y)dy.$$

Now using the upper semicontinuity of $\overline{f}''_r(\cdot;d)$ we have:

$$\beta(x;h;d) \le \int_{B(0,1)} \limsup_{n\to+\infty} \overline{f}''_r(\xi_n + \epsilon_n y; d)\phi(y)dy \le$$

$$\int_{B(0,1)} \overline{f}''_r(\xi;d)\phi(y)dy = \overline{f}''_r(\xi;d)$$

and the proof is complete.
ii) It is similar to the previous proof and we omit it. □

[2] In the proof of this theorem we will use the following generalized version of Fatou's lemma: *if f_n is a sequence of measurable functions, $f_n \le M$ and $E \subset \mathbb{R}^n$ is a subset of finite measure, then* $\limsup_{n\to+\infty} \int_E f_n \le \int_E \limsup_{n\to+\infty} f_n$

Theorem 2.3. *Assume that f is continuous and $f_r''(x;d)$ exists on a neighborhood of the point x_0 , $\forall d \in S^1$. Then f is of class $C^{1,1}$ at x_0 if and only if there exist a neighborhood U of x_0 and a function $g \in L^1(U)$ such that the following assumptions hold:*

(i) $\exists M \geq 0$ such that $|f_r''(x;d)| \leq M$, $\forall x \in U$, $\forall d \in S^1$,

(ii) $\left|\frac{\delta_2^d f(x;h)}{h^2}\right| \leq g(x)$, for $|h|$ "small enough" $(h \neq 0)$, $d \in S^1$ and a.e. $x \in U$.

Proof. i) Sufficiency. Arguing in a fashion similar to that of theorem 2.1 and using Lebesgue theorem, we obtain for ε "sufficiently small", for every x in a neighborhood of x_0 and $d \in S^1$:

$$f_\varepsilon''(x;d) = \lim_{h\to 0} \int_\Omega \frac{\delta_2^d f(z;h)}{h^2} \phi_\varepsilon(z-x)dz =$$

$$\int_\Omega \lim_{h\to 0} \frac{\delta_2^d f(z;h)}{h^2} \phi_\varepsilon(z-x)dz =$$

$$\int_\Omega f_r''(z;d)\phi_\varepsilon(z-x)dz.$$

It follows that $\forall d \in S^1$ $f_\varepsilon''(x,d)$ is bounded on U (uniformly with respect to ε). Using the integral representation of divided differences (see for instance [9], ch. 6, th. 2), we have:

$$\frac{\delta_2^d f_\varepsilon(x;h)}{h^2} = \int_0^1 dt_1 \int_0^{t_1} f_\varepsilon''(x + t_2hd + t_1hd; d)dt_2.$$

For x and h in suitable neighborhoods of x_0 and 0 respectively, the left member in the previous inequality is bounded by a constant M (uniformly with respect to ε). Sending ε to 0 and recalling theorem 1.2, we get the existence of neighborhoods U of x_0 and V of $0 \in \mathbb{R}$ such that $\forall d \in S^1$ $\frac{\delta_2^d f(x;h)}{h^2}$ is bounded on $U \times V \backslash \{0\}$. The thesis now follows recalling theorem 2.1.

ii) Necessity. The proof is similar to that of the necessary condition in theorem 2.1. □

Remark 2.1. *Hypothesis (ii) in the previous theorem cannot be omitted. In fact, as is easily seen, the function $f(x) = |x|$ satisfies hypothesis (i) but not hypothesis (ii) in a neighborhood of $x_0 = 0$ and is not of class $C^{1,1}$ at 0.*

Remark 2.2. *Theorems 2.1 and 2.3 extend the elementary condition which relates the Lipschitz condition on f' and the boundedness of f''. We generalize this relation without requiring any differentiability hypothesis and linking the existence and the Lipschitz behaviour of f' to the boundedness of $\frac{\delta_2^d f(x,h)}{h^2}$ or of the directional Riemann derivatives.*

Remark 2.3. *Conditions similar to those of theorem 2.3, expressed in terms of $f_R''(x;d)$ can be proved in analogous way.*

3 Optimality conditions for unconstrained optimization problems

The aim of this section is to study necessary and sufficient conditions for $C^{1,1}$ unconstrained optimization problems. These conditions are proved by using the generalized Taylor's expansions given in the previous section and Riemann derivatives. These optimality conditions can be used to write minimization algorithms; in fact for the computation of the next results only the values of the function are required. In the following we will suppose that, for any $d \in S^1$, the function $x \to \overline{f}''_r(x, d)$ is upper semicontinuous and that the function $x \to \underline{f}''_r(x; d)$ is lower semicontinuous in a neighborhood of x_0.

So we consider the following unconstrained optimization problem:

$P1)$ $$\min_{x\in\Omega} f(x)$$

where Ω is an open subset of $\mathbb{R}^n$.

Theorem 3.1. *(necessary condition) If f is $C^{1,1}$ and x_0 is a local minimum point then $\nabla f(x_0) = 0$ and $\overline{f}''_r(x_0, d) \geq 0$ $\forall d \in S^1$.*

Proof. From Taylor's formula we obtain, for h "small enough":

$$f(x_0 + hd) \leq f(x_0) + h < \nabla f(x_0), d > + \frac{h^2}{2}\overline{f}''_r(\xi, d)$$

where $\xi \in [x_0, x_0 + hd]$. So

$$0 \leq f(x_0 + hd) - f(x_0) \leq \overline{f}''_r(\xi, d)$$

and taking the limit when $h \to 0$ we obtain the thesis. □

Theorem 3.2. *(sufficient condition) If f is $C^{1,1}$, $\nabla f(x_0) = 0$ and $\underline{f}''_r(x_0+\alpha d, d) > 0$, $\forall \alpha \in (0, 1)$ and $\forall d \in S^1$, then x_0 is a strict local minimum of f on Ω.*

Proof. On the contrary suppose that x_0 is not a strict local minimum; then there exists a sequence x_k such that $x_k \to x_0$, when $k \to +\infty$ and $f(x_k) \leq f(x_0)$ $\forall k \in \mathbb{N}$. So $x_k = x_0 + \delta_k u_k$, where $\|u_k\| = 1$ and $\delta_k \to 0$ when $k \to +\infty$. So we have:

$$f(x_k) \geq f(x_0) + \delta_k^2 \frac{\underline{f}''_r(\xi_k, u_k)}{2}$$

where $x_0 \leq \xi_k \leq x_0 + \delta_k u_k$. This implies that

$$0 \geq f(x_k) - f(x_0) \geq \delta_k^2 \frac{\underline{f}''_r(\xi_k, u_k)}{2}$$

and then, when k is "sufficiently large", we obtain:

$$\underline{f}''_r(\xi_k, u_k) \leq 0$$

which contradicts the hypothesis. □

References

[1] Ash J.M.: *Very generalized Riemann derivatives.* Real Analysis Exhange, 12, 1985, 10-29.

[2] Ash J.M.: *Generalizations of the Riemann derivative.*, Tran. Amer. Math. Soc., 126, 1997, 181-199.

[3] H. Brezis: *Analyse fonctionelle- Theorie et applications.* Masson Editeur, Paris, 1963.

[4] X. Chen: *Convergence of BFGS method for LC^1 convex constrained optimization.* SIAM J. Control Optim., 34, 1996, 2051-2063.

[5] R. Cominetti, R. Correa: *A generalized second-order derivative in nonsmooth optimization.* SIAM J. Control Optim., 28, 1990, 789-809.

[6] A. Guerraggio, M. Rocca: *Derivate dirette di Riemann e di Peano.* Convessitá e Calcolo Parallelo, Verona, 1997.

[7] J.B. Hiriart-Hurruty: *Contributions a la programmation mathematique: deterministe et stocastique.* (Doctoral thesis), Univ. Clermont-Ferrand, 1977.

[8] J.B. Hiriart-Urruty, J.J. Strodiot , V. Hien Nguyen: *Generalized Hessian matrix and second order optimality conditions for problems with $C^{1,1}$ data.* Appl. Math. Optim., 11, 1984, 43-56.

[9] E. Isaacson, B.H. Keller: *Analysis of numerical methods.* Wiley, New York, 1966.

[10] D. Klatte, K. Tammer: *On the second order sufficient conditions for $C^{1,1}$-optimization problems.*, Optimization, 19, 1988, 169-180.

[11] D. Klatte: *Upper Lipschitz behavior of solutions to perturbed $C^{1,1}$ programs.* Math. Program. (Ser. B), 88, 2000, 285-311.

[12] D. La Torre, M. Rocca: *$C^{k,1}$ functions and Riemann derivatives.* Real Analysis Exchange, 25, 1999-2000,743-752.

[13] D. La Torre, M. Rocca: *Higher order smoothness conditions and differentiability of real functions.* Real Analysis Exchange, 26, 2000-01, 657-667.

[14] D. La Torre: *Characterizations of $C^{1,1}$ functions, Taylor's formulae and optimality conditions.* (Doctoral Thesis), Department of Mathematics, University of Milan, 2001.

[15] D.T. Luc: *Taylor's formula for $C^{k,1}$ functions.* SIAM J. Optimization, 5, 1995, 659-669.

[16] Marcinkiewicz J., Zygmund A.: *On the differentiability of functions and summability of trigonometrical series.* Fund. Math., 26, 1936, 1-43.

[17] P. Michel, J.P. Penot: *Second-order moderate derivatives.* Nonlin. Anal., 22, 1994, 809-824.

[18] L. Qi, W. Sun: *A nonsmooth version of Newton's method.* Math. Program., 58, 1993, 353-367.

[19] L. Qi: *Superlinearly convergent approximate Newton methods for LC^1 optimization problems.* Math. Program., 64, 1994, 277-294.

[20] L. Qi: *LC^1 functions and LC^1 optimization.* Operations Research and its applications (D.Z. Du, X.S. Zhang and K. Cheng eds.), World Publishing, Beijing, 1996, 4-13.

[21] T.D. Rockafellar: *Proximal subgradients, marginal values, and augmented lagrangians in nonconvex optimization.* Mathematics of Operations Research, 6, 1981, 427-437.

[22] W. Sun, R.J.B. de Sampaio, J. Yuan: *Two algorithms for LC^1 unconstrained optimization.* J. Comput. Math., 18, 2000, 621-632.

[23] X.Q. Yang, V. Jeyakumar: *Generalized second-order directional derivatives and optimization with $C^{1,1}$ functions.* Optimization, 26, 1992, 165-185.

[24] X.Q. Yang: *Second-order conditions in $C^{1,1}$ optimization with applications.* Numer. Funct. Anal. and Optimiz., 14, 1993, 621-632.

[25] X.Q. Yang: *Generalized second-order derivatives and optimality conditions.* Nonlin. Anal., 23, 1994, 767-784.