Applied Mathematics E-Notes, 3(2003), 130-137 © ISSN 1607-2510
Available free at mirror sites of http://www.math.nthu.edu.tw/∼amen/

Remarks On Second Order Generalized Derivatives For Differentiable Functions With Lipschitzian Jacobian *

Davide La Torre†, Matteo Rocca‡

Received 2 October 2002

Abstract

Many definitions of second-order generalized derivatives have been introduced to obtain optimality conditions for optimization problems with $C^{1,1}$ data. The aim of this note is to show some relations among these definitions.

The class of $C^{1,1}$ functions, that is the class of differentiable functions with Lipschitzian Jacobian, was first brought to attention by Hiriart-Urruty, Strodiot and Hien Nguyen [8]. The need for investigating such functions, as pointed out in [8] and [9], comes from the fact that several problems of applied mathematics including variational inequalities, semi-infinite programming, penalty functions, augmented Lagrangian, proximal point methods, iterated local minimization by decomposition etc. involve differentiable functions with no hope of being twice differentiable. Many second-order generalized derivatives have been introduced to obtain optimality conditions for optimization problems with $C^{1,1}$ data and for this class of functions numerical methods and minimization algorithms have been proposed too [20, 21]. In this paper we will give some relations among several definitions that one can find in literature and we will compare the necessary second-order optimality conditions expressed by means of these derivatives. We will focus our attention on the definitions due to Hiriart-Urruty, Strodiot and Hien Nguyen [8], Liu [14], Yang and Jeyakumar [23], Peano [18] and Riemann [22]. Some of these definitions do not require the hypothesis of $C^{1,1}$ regularity; however, under this assumption, each derivative in the previous list is finite. The definitions introduced by Hiriart-Urruty and Yang-Jeyakumar extend to the second-order, respectively, the notions due to Clarke and Michel-Penot for the first-order case. Peano and Riemann definitions are classical ones. Peano introduced the homonymous definition while he was studying Taylor's expansion formula for real functions. Peano derivatives were studied and generalized in recent years by Ben-Tal and Zowe [1] and Liu [14], who also obtained optimality conditions. Riemann higher-order derivatives were introduced in the theory of trigonometric series. Furthermore they were developed by several authors.

*Mathematics Subject Classifications: 26A24, 26A16.

†Department of Economics, University of Milan, via Conservatorio,7, 20122, Milano, Italy.
‡Department of Economics, University of Insubria, via Ravasi, 2, 21100, Varese, Italy.

In the following Ω will denote an open subset of R^n.

DEFINITION 1. A function $f : \Omega \to R$ is locally Lipschitz at x_0 when there exist a constant K and a neighbourhood U of x_0 such that

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

DEFINITION 2. A function $f : \Omega \to R$ is of class $C^{1,1}$ at x_0 when its first order partial derivatives exist in a neighbourhood 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 1. Let $g : \Omega \subset R^n \to R$ be twice continuously differentiable on Ω and consider $f(x) = [g^+(x)]^2$ where $g^+(x) = \max\{g(x), 0\}$. Then f is $C^{1,1}$ on Ω.

EXAMPLE 2. Let $f_i : R^n \to R$, $i = 0, ..., m$, and consider the following minimization problem:

$$\min f_0(x)$$

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

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

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. By stating $y = 0$ 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 $R^n \times R^m$ with

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

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

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

$$\max g_r(y)$$

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

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

In [11] the authors have proved the following result which gives a characterization of $C^{1,1}$ functions by divided differences.

THEOREM 1. [11] Assume that the function $f : \Omega \to R$ is bounded on a neighbourhood of the point $x_0 \in \Omega$. Then f is of class $C^{1,1}$ at x_0 if and only if there exist neighbourhoods U of x_0 and V of $0 \in R$ such that $\frac{\delta_2^d f(x;h)}{h^2}$ is bounded on $U \times V \backslash \{0\}$, $\forall d \in S^1 = \{d \in R^n : \|d\| = 1\}$ where

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

It is known [23] that if a function f is of class $C^{1,1}$ at x_0 then it can be expressed (in a neighbourhood of x_0) as the difference of two convex functions. The following corollary strenghtens the results in [23].

COROLLARY 1. If f is of class $C^{1,1}$ at x_0, then $f = \tilde{f} + p$ where $\tilde{f}$ is convex and p is a polynomial of degree at most two.

PROOF. From the previous theorem we know that a function f is of class $C^{1,1}$ at x_0 if and only if the following difference:

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

is bounded by a constant M for each x in a neighbourhood U of x_0, for each h in a neighbourhood V of 0 and $d \in S^1$. If $p(x) = p(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^n c_{ij} x_i x_j$ is a polynomial of degree 2 in the variables x_1, $x_2, \ldots, x_n$, it is well known that

$$\frac{\delta_2^d p(x;d)}{h^2} = \sum_{i,j=1}^n c_{ij} d_i d_j$$

where $d = (d_1, d_2, \ldots d_n)$. So we can choose the polynomial p such that

$$\sup_{d \in S^1} \frac{\delta_2^d p(x;d)}{h^2} \leq -M$$

and then for the function $\tilde{f}(x) = f(x) - p(x)$ the following inequality holds

$$\frac{\delta_2^d \tilde{f}(x;d)}{h^2} \geq 0, \ \forall x \in U, h \in V, d \in S^1$$

that is $\tilde{f}$ is locally convex.

Now we remember the notions of second-order generalized derivative on which we will focus our attention. The following definitions can be introduced without the hypothesis that f is of class $C^{1,1}$ but were investigated mostly under this hypothesis.

DEFINITION 3. Let us consider a function $f : \Omega \to R$ of class $C^{1,1}$ at x_0.

i) Peano's second order derivative of f at x_0 in the direction $d \in R^n$ is defined as

$$\overline{f}''_P(x_0; d) = 2 \limsup_{t \to 0^+} \frac{f(x_0 + td) - f(x_0) - t\nabla f(x_0)d}{t^2}.$$

ii) Riemann's second upper derivative of f at x_0 in the direction $d \in R^n$ is defined as

$$\overline{f}''_R(x_0; d) = \limsup_{t\to 0^+} \frac{f(x_0 + 2td) - 2f(x_0 + td) + f(x_0)}{t^2}.$$

iii) Yang-Jeyakumar's second upper derivative of f at x_0 in the directions $u, v \in R^n$ is defined as

$$\overline{f}''_Y(x_0; u, v) = \sup_{z\in R^n} \limsup_{t\to 0^+} \frac{\nabla f(x_0 + tu + tz)v - \nabla f(x_0 + tz)v}{t}.$$

iv) Hiriart-Urruty's second upper derivative of f at x_0 in the direction $u, v \in R^n$ is defined as

$$\overline{f}''_H(x_0; u, v) = \limsup_{x'\to x,\ t\to 0^+} \frac{\nabla f(x' + tu)v - \nabla f(x')v}{t}.$$

REMARK 1. Analogously one can define lower derivatives. In particular Peano's second lower derivative is defined by

$$\underline{f}''_P(x_0; d) = 2 \liminf_{t\to 0^+} \frac{f(x_0 + td) - f(x_0) - t\nabla f(x_0)d}{t^2}.$$

REMARK 2. We will denote by $\overline{f}''_D(x_0; d) = \limsup_{t\to 0^+} \frac{\nabla f(x_0+td)d-\nabla f(x_0)d}{t}$ the classical second upper derivative of f at x_0 in the direction $d \in R^n$.

The following result is trivial.

THEOREM 2. If $f : \Omega \to R$ is of class $C^{1,1}$ at x_0, then each derivative in the previous list is finite (the same holds for lower derivatives).

It is known that the following necessary optimality conditions for an unconstrained minimization problem hold.

THEOREM 3. [6, 8, 14, 23] Let $f : \Omega \to R$ be a $C^{1,1}$ function and assume that $x_0 \in \Omega$ is a local minimum point of f. Then $\forall d \in R^n$ the derivatives $\overline{f}''_H(x_0; d, d)$, $\overline{f}''_Y(x_0; d, d)$, $\overline{f}''_P(x_0; d)$, $\overline{f}''_R(x_0; d)$, $\underline{f}''_P(x_0; d)$, are greater or equal than zero.

In the following we will give some relations among the derivatives in Definition 3 that we will allow us to state that the "best" necessary optimality conditions for an unconstrained minimization problem are those expressed by means of Peano's second upper derivative.

REMARK 3. In [23] is given the following chain of inequalities

$$\overline{f}''_D(x_0; d) \le \overline{f}''_Y(x_0; d, d) \le \overline{f}''_H(x_0; d, d).$$

Furthermore $\overline{f}''_Y(x_0; d, d) = \overline{f}''_H(x_0; d, d)$ if and only if the map $\overline{f}''_Y(\cdot; d, d)$ is upper semicontinuous [23].

REMARK 4. In [5] is given the following characterization of Hiriart-Urruty's generalized derivative

$$\overline{f}''_H(x_0; u, v) = \limsup_{y\to x_0, s,t\to 0^+} \frac{\overline{\Delta}_2^{u,v} f(y; s, t)}{st}$$

where

$$\overline{\Delta}_2^{u,v} f(y;s,t) = f(y+su+tv) - f(y+su) - f(y+tv) + f(y).$$

From this characterization one can trivially deduce that $\overline{f}''_R(x_0;d) \leq \overline{f}''_H(x_0;d,d)$.

LEMMA 1. Let $f : \Omega \subset R^n \to R$ be a given $C^{1,1}$ function. Then $\overline{f}''_R(x;d) \leq 2\overline{f}''_P(x;d) - \underline{f}''_P(x;d)$.

PROOF. Choose $t_n \to 0^+$ as $n \to +\infty$, such that

$$L = \lim_{n\to+\infty} \frac{f(x+2t_n d) - 2f(x+t_n d) + f(x)}{t_n^2}.$$

Obviously $L \leq \overline{f}''_R(x;d)$. Let now

$$s_{1,n} = 2\frac{f(x+t_n d) - f(x) - t_n \nabla f(x) d}{t_n^2}$$

and

$$s_{2,n} = 2\frac{f(x+2t_n d) - f(x) - 2t_n \nabla f(x) d}{t_n^2},$$

eventually by extracting subsequences, $s_{1,n} \to s_1$ and $s_{2,n} \to s_2$ with $s_1 \geq \underline{f}''_P(x;d)$ and $s_2 \leq 4\overline{f}''_P(x;d)$. By a simple calculation, we have

$$s_{2,n} - 2s_{1,n} = 2\frac{f(x+2t_n d) - 2f(x+t_n d) + f(x)}{t_n^2} \to 2L$$

and then

$$2L = \lim_{n\to+\infty} s_{2,n} - 2s_{1,n} \leq 4\overline{f}''_P(x;d) - 2\underline{f}''_P(x;d).$$

THEOREM 4. Let f be a function of class $C^{1,1}$ at x_0. Then

i) $\overline{f}''_P(x_0;d) \leq \overline{f}''_D(x_0;d) \leq \overline{f}''_Y(x_0;d,d) \leq \overline{f}''_H(x_0;d,d)$.

ii) $\overline{f}''_P(x_0;d) \leq \overline{f}''_R(x_0;d) \leq \overline{f}''_Y(x_0;d,d) \leq \overline{f}''_H(x_0;d,d)$.

PROOF.

i) From the previous remarks, it is only necessary to prove the inequality $\overline{f}''_P(x_0;d) \leq \overline{f}''_D(x_0;d)$. If we take the function $\phi_1(t) = f(x_0+td) - t\nabla f(x_0)d$ and $\phi_2(t) = t^2$, applying Cauchy's theorem, we obtain

$$2\frac{f(x_0+td) - f(x_0) - t\nabla f(x_0)d}{t^2} = 2\frac{\phi_1(t) - \phi_1(0)}{\phi_2(t) - \phi_2(0)} =$$

$$2\frac{\phi_1'(\xi)}{\phi_2'(\xi)} = \frac{\nabla f(x_0+\xi d)d - \nabla f(x_0)d}{\xi},$$

where $\xi = \xi(t) \in (0,t)$, and then $\overline{f}''_P(x_0;d) \leq \overline{f}''_D(x_0;d)$.

ii) From the previous remarks, it is only necessary to prove the inequalities $\overline{f}''_P(x_0; d) \leq \overline{f}''_R(x_0; d) \leq \overline{f}''_Y(x_0; d, d)$. Concerning the first inequality, from the definition of $\overline{f}''_P(x_0; d)$ we have

$$f(x_0 + td) = f(x_0) + t\nabla f(x_0)d + \frac{t^2}{2}\overline{f}''_P(x_0; d) + g(t)$$

where $\limsup_{t\to 0^+} \frac{g(t)}{t^2} = 0$ and

$$f(x_0 + 2td) = f(x_0) + 2t\nabla f(x_0)d + 2t^2\overline{f}''_P(x_0; d) + g(2t)$$

where $\limsup_{t\to 0^+} \frac{g(2t)}{t^2} = 4\limsup_{t\to 0^+} \frac{g(2t)}{4t^2} = 0$. Then

$$\frac{f(x_0 + 2td) - 2f(x_0 + td) + f(x_0)}{t^2} = \frac{t^2\overline{f}''_P(x_0; d) + g(2t) - g(t)}{t^2} \geq$$

$$\overline{f}''_P(x_0; d) + \limsup_{t\to 0^+} \frac{g(2t)}{t^2} - \limsup_{t\to 0^+} \frac{g(t)}{t^2}.$$

Then $\overline{f}''_R(x_0; d) \geq \overline{f}''_P(x_0; d)$. For the second inequality, we define $\phi_1(t) = f(x_0 + 2td) - 2f(x_0 + td)$ and $\phi_2(t) = t^2$. Then, by Cauchy's theorem, we obtain

$$\frac{f(x_0 + 2td) - 2f(x_0 + td) + f(x_0)}{t^2} = \frac{\phi_1(t) - \phi_1(0)}{\phi_2(t) - \phi_2(0)} =$$

$$\frac{\phi'_1(\xi)}{\phi'_2(\xi)} = \frac{\nabla f(x_0 + 2\xi d)d - \nabla f(x_0 + t\xi)d}{\xi},$$

where $\xi = \xi(t) \in (0, t)$, and then $\overline{f}''_R(x_0; d) \leq \overline{f}''_Y(x_0; d, d)$.

THEOREM 5. Let f be a function of class $C^{1,1}$ at x_0. If x_0 is a local minimum for f, then $\nabla f(x_0) = 0$ and the following chain of inequalities holds

$$0 \leq \underline{f}''_P(x_0; d) \leq \overline{f}''_P(x_0; d)$$

$$\leq \overline{f}''_R(x_0; d) \leq 2\overline{f}''_D(x_0; d) \leq 2\overline{f}''_Y(x_0; d, d) \leq 2\overline{f}''_H(x_0; d, d)$$

Indeed, the proof follows from Theorem 4 and Lemma 1.

REMARK 5. From the previous theorem one can conclude that the "best" optimality conditions for an unconstrained minimization problem with $C^{1,1}$ objective function are those expressed by means of Peano's derivatives.

REMARK 6. A similar chain of inequalities holds $\forall x \in \Omega$ when f is convex and of class $C^{1,1}$. In fact in this case it is not difficult to prove that $\overline{f}''_R(x; d) \leq 4\overline{f}''_D(x; d)$ and then, reasoning as in the proof of Theorem 5, we obtain the following chain of inequalities

$$0 \leq \underline{f}''_P(x; d) \leq \overline{f}''_P(x; d)$$

$$\leq \overline{f}''_R(x; d) \leq 4\overline{f}''_D(x; d) \leq 4\overline{f}''_Y(x; d, d) \leq 4\overline{f}''_H(x; d, d).$$

References

[1] A. Ben-Tal and J. Zowe, Directional derivatives in nonsmooth optimization, Journal of Optimization Theory and Applications, 47(4)(1985), 483–490.

[2] J. M. Borwein J. M., S. P. Fitzpatrick and J. R. Giles, The differentiability of real functions on normed linear space using generalized subgradients, Journal of Mathematical Analysis and Applications, 128(1987), 512–534.

[3] W. L. Chan, L. R. Huang and K. F. Ng, On generalized second-order derivatives and Taylor expansion formula in nonsmooth optimization, SIAM Journal on Control and Optimization, 32(3)(1994), 591–611.

[4] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[5] R. Cominetti and R. Correa, A generalized second-order derivative in nonsmooth optimization, SIAM Journal Control and Optimization, 28(1990), 789–809.

[6] I. Ginchev and A. Guerraggio, Second order optimality conditions in nonsmooth unconstrained optimization, Pliska Studia Mathematica Bulgarica, 12(1998), 39–50.

[7] A. Guerraggio and D.T. Luc, On optimality conditions for $C^{1,1}$ vector optimization problems, Journal of Optimization Theory and Applications, 109(3)(2001), 615–629.

[8] J. B. Hiriart-Urruty, J. J. Strodiot and H. V. Nguyen, Generalized Hessian matrix and second order optimality conditions for problems with $C^{1,1}$ data, Applied Mathematics and Optimization, 11(1984), 43–56.

[9] D. Klatte and K. Tammer, On second-order sufficient optimality conditions for $C^{1,1}$ optimization problems, Optimization, 19(1988), 169–179.

[10] D. Klatte, Upper Lipschitz behavior of solutions to perturbed $C^{1,1}$ programs. Mathematical Programming (Ser. B), 88(2000), 285–311.

[11] D. La Torre and M. Rocca, $C^{1,1}$ functions and optimality conditions, Journal of Computational Analysis and Applications, to appear.

[12] D. La Torre and M. Rocca, On $C^{1,1}$ constrained optimization problems, Journal of Computational Analysis and Applications, to appear.

[13] D. La Torre and M. Rocca, A characterization of $C^{k,1}$ functions, Real Analysis Exchange, 27(2)(2002), 515–534.

[14] L. Liu and M. Krisek, Second order optimality conditions for nondominated solutions of multiobjective programming with $C^{1,1}$ data, Applications of Mathematics, 45(5)(2000), 381–397.

[15] D.T. Luc, Taylor's formula for $C^{k,1}$ functions, SIAM Journal on Optimization, 5(1995), 659–669.

[16] P. Michel and J. P. Penot, Calcul sous-différential pour des fonctions lischitziennes an nonlipschitziennes, Comptes Rendus de l'Academie des Sciences Paris, 298(1984), 269–272.

[17] H. W. Oliver, The exact Peano derivative, Transaction Amererican Mathematical Sociecty, 76(1954), 444–456.

[18] G. Peano, Sulla formula di Taylor, Atti dell'Accademia delle Science di Torino, 27(1891-92), 40–46.

[19] L. Qi and W. Sun, A nonsmooth version of Newton's method, Mathematical Programming, 58(1993), 353–367.

[20] L. Qi, Superlinearly convergent approximate Newton methods for LC^1 optimization problems, Mathematical Programming, 64(1994), 277-294.

[21] 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.

[22] B. Riemann, Uber die darstellbarkeit einen function durch eine trigonometrische reihe, Ges. Werke, 2 Aufl., Leipzig (1982), 227–271.

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