LE MATEMATICHE
Vol. XLIX (1994) – Fasc. II, pp. 189-203

VARIATIONAL SETS AND NECESSARY OPTIMALITY CONDITIONS IN NONSMOOTH OPTIMIZATION

MARCO CASTELLANI

In this paper we underline the strict connection between the second–order necessary optimality conditions in nonsmooth optimization and the separation of two particular sets. Introducing two approximating second–order variational sets, we give an alternative proof of a result in [15] (with presence of inequality constraints only) and we build a new result without the presence of regularity conditions.

1. Introduction.

In the last years a lot of first–order necessary conditions in nonsmooth optimization have been given and they have been built using various kinds of directional derivatives (see for example [7], [10], [14], [16], [17] and references therein). Moreover the strict connection between these necessary conditions and the separation of particular local cone approximations has been analysed ([3], [5], [8] and [12]), drawing a general abstract scheme to build such conditions. In this paper we shall adapt this scheme for proving a second–order necessary optimality condition due to Studniarski [15]; moreover it is possible to extend this result either omitting the regularity condition or weakening the assumption about the constraints [4]. The key tools of our proof will be two particular second–order variational sets that represent a generalization of the concept of local cone approximation and they have been investigated also in [1], [9] and [11].

We introduce the notations that will be used in the sequel. $\mathbb{X}$ will be a real Banach space with norm $\|\cdot\|$ and $\mathbb{X}^*$ will denote the dual space of continuous linear functionals on $\mathbb{X}$. For any subset Q of $\mathbb{X}$ we shall denote with $\operatorname{cl} Q$ and $\operatorname{int} Q$ respectively the topological closure and interior of Q; Q^c is the algebraical complement. For any $x_0 \in \mathbb{X}$ and for any $\rho > 0$ we shall denote with $B(x_0, \rho)$ the open ball with radius ρ and centre x_0. Given the function $f : \mathbb{X} \to \mathbb{R}$ the epigraphic of f is the set

$$\operatorname{epif} = \big\{(x, y) \in \mathbb{X} \times \mathbb{R} \mid f(x) \leq y\big\}.$$

If f is lipschitzian at x_0 the *Clarke's directional derivative* of f at x_0 in the direction v is defined by

$$f^\circ(x_0; v) = \limsup_{\substack{x \to x_0 \\ t \downarrow 0}} \frac{f(x + tv) - f(x)}{t}$$

and $f^\circ(x_0; \cdot)$ is a real–valued sublinear function. The *Clarke's subdifferential* of f at x_0 is defined by

$$\partial^\circ f(x_0) = \big\{x^* \in \mathbb{X}^* \mid f^\circ(x_0; v) \geq \langle x^*, v\rangle, \quad \forall v \in \mathbb{X}\big\}$$

and it is a nonempty convex and weak*–compact subset of $\mathbb{X}^*$. In conclusion for any $x_0 \in \mathbb{X}$ we call *subdifferential* of f at x_0 the set

$$\partial f(x_0) = \big\{x^* \in \mathbb{X}^* \mid f(x) \geq f(x_0) + \langle x^*, x - x_0\rangle, \quad \forall x \in \mathbb{X}\big\}$$

and we observe that if f is lipschitzian at x_0 then $\partial^\circ f(x_0) = \partial f^\circ(x_0; \cdot)(0)$.

The mathematical problem that will be studied in this paper is

$$\text{(P)} \qquad \begin{cases} \min f(x) \\ x \in Q = \bigcap\limits_{i \in I} Q_i \end{cases}$$

where $Q_i = \{x \in \mathbb{X} \mid g_i(x) \leq 0\}$ with $f, g_i : \mathbb{X} \to \mathbb{R}$ locally lipschitzian functions and $I = \{1, \ldots, m\}$. Besides, for every $x_0 \in Q$, we shall put

$$I(x_0) = \big\{i \in I \mid g_i(x_0) = 0\big\}.$$

Now we introduce the first and second–order directional derivatives that have been used in the Studniarski's theorem.

Definition 1.1. Given the function $f : \mathbb{X} \longrightarrow \mathbb{R}$, the point $x_0 \in \mathbb{X}$ and the directions $v, w \in \mathbb{X}$, we define the following directional derivatives:

(a) the *upper Dini derivative* of f at x_0 in the direction v

$$D_+ f(x_0; v) = \limsup_{t \downarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t};$$

(b) the *lower Dini derivative* of f at x_0 in the direction v

$$D_- f(x_0; v) = \liminf_{t \downarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t};$$

(c) if $D_+ f(x_0; v) \in \mathbb{R}$ the *second–order upper Dini derivative* of f at x_0 in the directions v and w

$$D^2_+ f(x_0; v, w) = \limsup_{t \downarrow 0} \frac{f(x_0 + tv + t^2 w) - f(x_0) - t D_+ f(x_0; v)}{t^2};$$

(d) if $D_- f(x_0; v) \in \mathbb{R}$ the *second–order lower Dini derivative* of f at x_0 in the directions v and w

$$D^2_- f(x_0; v, w) = \liminf_{t \downarrow 0} \frac{f(x_0 + tv + t^2 w) - f(x_0) - t D_- f(x_0; v)}{t^2}.$$

Remark 1.1. If f is lipschitzian at x_0 with constant K, we have that

$$|D_\pm f(x_0; v)| \le K \|v\|;$$

hence the second–order Dini derivatives are well–defined.
Besides $D_\pm f(x_0; 0) = 0$ and it is immediate to observe that

$$D^2_\pm f(x_0; 0, w) = D_\pm f(x_0; w). \qquad \square$$

Remark 1.2. If f is twice Fréchet differentiable at x_0, using the Taylor's formula, we have:

$$D^2_\pm f(x_0; v, w) = \langle \nabla f(x_0), w \rangle + \frac{1}{2} \langle v, \nabla^2 f(x_0) v \rangle. \qquad \square$$

Now, we formulate the main result of [15].

Theorem [Studniarski]. *Let $x_0 \in Q$ be a local optimal solution for* (P) *and suppose that the following regularity condition holds:*

(1.1) *for each set $\{x_i^* \mid i \in I(x_0)\}$ with $x_i^* \in \partial^\circ g_i(x_0)$ and for each $y_i \geq 0$ with $i \in I(x_0)$, the equality $\sum_{i \in I(x_0)} y_i x_i^* = 0$ implies that $y_i = 0$ for all $i \in I(x_0)$.*

Then $D_-^2 f(x_0; v, w) \geq 0$ for all $v, w \in \mathbb{X}$ satisfying the following conditions:

$$D_- f(x_0; v) \leq 0, \tag{1.2}$$

and, for each $i \in I(x_0)$, we have either

$$D_+ g_i(x_0; v) < 0 \tag{1.3}$$

or

$$D_+ g_i(x_0; v) = 0 \qquad \text{and} \qquad D_+^2 g_i(x_0; v, w) \leq 0. \tag{1.4}$$

Remark 1.3. Using the Remark 1.1, we observe that for $v = 0$ the theorem affirms that for each $w \in \mathbb{X}$ such that $D_+ g_i(x_0; w) \leq 0$ for every $i \in I(x_0)$ we have $D_- f(x_0; w) \geq 0$. This result is quite similar to a result in [10]. □

First of all we give an equivalent form of the regularity condition (1.1) that we shall use in our proof.

Theorem 1.1. *Let g_i be lipschitzian at x_0; the following assertions are equivalent:*

(**a**) *for each set $\{x_i^* \mid i \in I\}$ with $x_i^* \in \partial^\circ g_i(x_0)$ and for each $y_i \geq 0$, the equality $\sum_{i \in I} y_i x_i^* = 0$ implies that $y_i = 0$ for all $i \in I$;*

(**b**) *for each $y_i \geq 0$, the inequality*

$$\sum_{i \in I} y_i g_i^\circ(x_0; v) \geq 0, \qquad \forall v \in \mathbb{X}$$

implies that $y_i = 0$ for all $i \in I$;

(**c**) *there exists $v \in \mathbb{X}$ such that $g_i^\circ(x_0; v) < 0$ for every $i \in I$.*

Proof. **(b)** $\Rightarrow$ **(a)** Choosing y_i as in the hypothesis, we have that

$$\sum_{i \in I} y_i g_i^\circ(x_0; v) \geq \sum_{i \in I} y_i \langle x_i^*, v \rangle = \langle 0, v \rangle = 0, \qquad \forall v \in \mathbb{X}.$$

Therefore, by **(b)**, we have $y_i = 0$ for all $i \in I$ and the thesis is proved.
(a) $\Rightarrow$ **(b)** The functions $g_i^\circ(x_0; \cdot)$ are sublinear and $g_i^\circ(x_0; 0) = 0$; therefore, choosing y_i as in the hypothesis, we have that the real–valued convex function $\sum_{i \in I} y_i g_i^\circ(x_0; \cdot)$ has minimum at x_0 and then

$$0 \in \partial\Big(\sum_{i \in I} y_i g_i^\circ(x_0; \cdot)\Big)(0) = \sum_{i \in I} y_i \partial g_i^\circ(x_0; \cdot)(0) = \sum_{i \in I} y_i \partial^\circ g_i(x_0).$$

Hence there exists a set $\{x_i^* \mid i \in I\}$ with $x_i^* \in \partial^\circ g_i(x_0)$ such that $\sum_{i \in I} y_i x_i^* = 0$; by **(a)** the thesis follows.
(c) $\Rightarrow$ **(b)** It is trivial.
(b) $\Rightarrow$ **(c)** Assume, ab absurdo, that **(c)** does not hold and let

$$C = \{\Theta = (\theta_1, \ldots, \theta_m) \in \mathbb{R}^m \mid \exists v \in \mathbb{X} \text{ such that } g_i^\circ(x_0; v) < \theta_i, \ \forall i \in I\}$$

be a nonempty convex subset in $\mathbb{R}^m$. By assumption we have that C is disjointed from the nonpositive orthant

$$\mathbb{R}_-^m = \{\Xi = (\xi_1, \ldots, \xi_m) \in \mathbb{R}^m \mid \xi_i \leq 0, \ \forall i \in I\}.$$

Therefore there exist a nonzero vector $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ and a constant $\alpha \in \mathbb{R}$ such that

$$\sum_{i \in I} y_i \xi_i \leq \alpha \leq \sum_{i \in I} y_i \theta_i, \qquad \forall \Theta \in C, \ \forall \Xi \in \mathbb{R}_-^m.$$

From the first inequality we obtain that $\alpha \geq 0$ and all $y_i \geq 0$; therefore $\sum_{i \in I} y_i \theta_i \geq 0$. For every $\epsilon > 0$ and for every $v \in \mathbb{X}$, we get

$$0 \leq \sum_{i \in I} y_i \big(g_i^\circ(x_0; v) + \epsilon\big).$$

Since ϵ is arbitrary we get

$$\sum_{i \in I} y_i g_i^\circ(x_0; v) \geq 0, \qquad \forall v \in \mathbb{X}$$

and y_i are not all zero; this is absurdo. $\square$

2. Second–order variational sets.

The study of the local properties of sets and functions is often confided to the analysis of their approximations: the most classic example is the Taylor's formula for Fréchet differentiable functions. As regards the sets, the immediate generalization of the linearization through the tangent hyperplane, is given by the use of local cone approximations [8]. In the case that the set is the epigraphic of a function, the local cone approximation is thought as the epigraphic of a generalized directional derivative. Following this scheme, in this section we introduce two kinds of second–order approximating sets that extend the concept of two particular local cone approximations and we investigate some properties.

Definition 2.1. Given a set $Q \subseteq \mathbb{X}$, a point $x_0 \in \operatorname{cl} Q$ and a direction $v \in \mathbb{X}$, we define the following sets:

(a) the *cone of the feasible directions* of Q at x_0

$$D_+(Q;x_0) = \{y \in \mathbb{X} \mid \exists \lambda > 0 \text{ such that } \forall t \in]0,\lambda[,\ x_0 + ty \in Q\};$$

(b) the *cone of the weak feasible directions* of Q at x_0

$$D_-(Q;x_0) = \{y \in \mathbb{X} \mid \forall \lambda > 0,\ \exists t \in]0,\lambda[\text{ such that } x_0 + ty \in Q\};$$

(c) the *second–order variational set of the feasible directions* of Q at x_0 in the direction v

$$\begin{aligned} D_+^2(Q;x_0,v) &= \{y \in \mathbb{X} \mid \exists \lambda > 0 \text{ such that } \forall t \in]0,\lambda[,\ x_0 + tv + t^2 y \in Q\} \\ &= \{y \in \mathbb{X} \mid \forall \{t_n\} \downarrow 0 \text{ definitively } x_0 + t_n v + t_n^2 y \in Q\}; \end{aligned}$$

(d) the *second–order variational set of the weak feasible directions* of Q at x_0 in the direction v

$$\begin{aligned} D_-^2(Q;x_0,v) &= \{y \in \mathbb{X} \mid \forall \lambda > 0,\ \exists t \in]0,\lambda[\text{ such that } x_0 + tv + t^2 y \in Q\} \\ &= \{y \in \mathbb{X} \mid \exists \{t_n\} \downarrow 0 \text{ such that } x_0 + t_n v + t_n^2 y \in Q\}. \end{aligned}$$

Remark 2.1. It is immediate to observe that

$$D_\pm^2(Q;x_0,0) = D_\pm(Q;x_0). \qquad \square$$

Other easy properties of the above defined variational sets are described in the following theorem.

Theorem 2.1. *The following properties hold:*
(a) *For each $A \subseteq \mathbb{X}$ and for each $x_0 \in \mathbb{X}$ we have*

$$D^2_-(A; x_0, v) = \left(D^2_+(A^c; x_0, v)\right)^c, \qquad \forall v \in \mathbb{X}. \tag{2.1}$$

(b) *$D^2_\pm$ are isotone, that is if $A \subseteq B \subseteq \mathbb{X}$ and $x_0 \in \mathbb{X}$ then*

$$D^2_\pm(A; x_0, v) \subseteq D^2_\pm(B; x_0, v), \qquad \forall v \in \mathbb{X}. \tag{2.2}$$

(c) *For each $A, B \subseteq \mathbb{X}$ and for each $x_0 \in \mathbb{X}$ we have*

$$D^2_+(A \cap B; x_0, v) = D^2_+(A; x_0, v) \cap D^2_+(B; x_0, v), \qquad \forall v \in \mathbb{X}. \tag{2.3}$$

(d) *For each $A \subseteq \mathbb{X}$ and for each $x_0 \in \operatorname{int} A$ we have*

$$D^2_\pm(A; x_0, v) = \mathbb{X}, \qquad \forall v \in \mathbb{X}. \tag{2.4}$$

Proof. (a) It is an immediate consequence of this chain of equivalences:

$$\begin{aligned} w \in \left(D^2_+(A^c; x_0, v)\right)^c &\Leftrightarrow w \notin D^2_+(A^c; x_0, v) \\ &\Leftrightarrow \exists \{t_n\} \downarrow 0 \text{ such that } x_0 + t_n v + t_n^2 w \notin A^c \\ &\Leftrightarrow \exists \{t_n\} \downarrow 0 \text{ such that } x_0 + t_n v + t_n^2 w \in A. \end{aligned}$$

(b) It is trivial.
(c) Let $w \in D^2_+(A; x_0, v) \cap D^2_+(B; x_0, v)$; therefore there exist $\lambda_A, \lambda_B > 0$ such that for every $t_A \in]0, \lambda_A[$ and $t_B \in]0, \lambda_B[$ we have

$$x_0 + t_A v + t_A^2 w \in A \qquad \text{and} \qquad x_0 + t_B v + t_B^2 w \in B.$$

Choosing $\lambda = \min\{\lambda_A, \lambda_B\}$ we get

$$D^2_+(A; x_0, v) \cap D^2_+(B; x_0, v) \subseteq D^2_+(A \cap B; x_0, v).$$

The other inclusion follows immediately from the isotonicity.
(d) It is trivial. □

It is well–known the following connection between the Dini derivatives and the previous local cone approximations

$$D_\pm f(x_0; v) = \inf\left\{\beta \in \mathbb{R} \mid (v, \beta) \in D_\pm\left(\operatorname{epi} f; \left(x_0, f(x_0)\right)\right)\right\}.$$

A similar relation exists between the second–order directional Dini derivatives and the above mentioned variational sets.

Theorem 2.2. *Let* $f : \mathbb{X} \to \mathbb{R}$ *be a function and* $x_0, v \in \mathbb{X}$ *such that* $D_\pm f(x_0; v) \in \mathbb{R}$*; then*

$$D^2_\pm f(x_0; v, w) = \\ = \inf\left\{\beta \in \mathbb{R} \mid (w, \beta) \in D^2_\pm\left(\mathrm{epi}f; \left(x_0, f(x_0)\right), \left(v, D_\pm f(x_0; v)\right)\right)\right\}.$$

where $\inf \emptyset = +\infty$.

Proof. We state the proof for D^2_+ only because the other one can be proved similarly. First of all we observe that

$$(w, \beta) \in D^2_+\Big(\mathrm{epi}f; \left(x_0, f(x_0)\right), \left(v, D_+ f(x_0; v)\right)\Big)$$

if and only if for every $\{t_n\} \downarrow 0$ we have

$$\frac{f(x_0 + t_n v + t_n^2 w) - f(x_0) - t_n D_+ f(x_0; v)}{t_n^2} \leq \beta.$$

Let $D^2_+ f(x_0; v, w) = \ell$; we have three cases:
$\ell = +\infty$. Then there exists $\{t_n\} \downarrow 0$ such that, for every $\beta \in \mathbb{R}$, we have

$$\frac{f(x_0 + t_n v + t_n^2 w) - f(x_0) - t_n D_+ f(x_0; v)}{t_n^2} > \beta;$$

hence $(w, \beta) \notin D^2_+\Big(\mathrm{epi}f; \left(x_0, f(x_0)\right), \left(v, D_+ f(x_0; v)\right)\Big)$ for every β and therefore the thesis.
$\ell = -\infty$. Then for every $\{t_n\} \downarrow 0$ and for every $\beta \in \mathbb{R}$, we have

$$\frac{f(x_0 + t_n v + t_n^2 w) - f(x_0) - t_n D_+ f(x_0; v)}{t_n^2} \leq \beta;$$

hence $(w, \beta) \in D^2_+\Big(\mathrm{epi}f; \left(x_0, f(x_0)\right), \left(v, D_+ f(x_0; v)\right)\Big)$ for all β and therefore the thesis.
$\ell \in \mathbb{R}$. Then for each $\epsilon > 0$ we have, for every $\{t_n\} \downarrow 0$,

$$\frac{f(x_0 + t_n v + t_n^2 w) - f(x_0) - t_n D_+ f(x_0; v)}{t_n^2} \leq \ell + \epsilon$$

and there exists $\{t_n\} \downarrow 0$ such that

$$\frac{f(x_0 + t_n v + t_n^2 w) - f(x_0) - t_n D_+ f(x_0; v)}{t_n^2} \geq \ell - \epsilon.$$

From the first inequality we have

$$(w, \ell + \epsilon) \in D_+^2\Big(\text{epi} f;\ \big(x_0, f(x_0)\big), \big(v, D_+ f(x_0; v)\big)\Big),$$

from the second inequality we have

$$(w, \ell - \epsilon) \notin D_+^2\Big(\text{epi} f;\ \big(x_0, f(x_0)\big), \big(v, D_+ f(x_0; v)\big)\Big);$$

therefore the thesis is proved. □

The strict lower level set of the function f

$$L_f(x_0) = \big\{x \in \mathbb{X} \mid f(x) < f(x_0)\big\},$$

plays a key role in optimization; in fact we have that $x_0 \in Q$ is a local optimal solution for (P) if and only if there exists $\rho > 0$ such that

$$L_f(x_0) \cap Q \cap B(x_0, \rho) = \emptyset. \tag{2.5}$$

The verification of this condition is quite arduous; therefore, in order to get more tractable optimality conditions, the sets $L_f(x_0)$ and Q are usually approximated by other sets with a simpler structure, for instance the local cone approximations. In the sequel we shall show that the previous second–order variational sets are an effective pair of approximating sets.

Theorem 2.3. *Let $x_0 \in Q$ be a local optimal solution for* (P)*; then:*

$$D_\pm^2\big(L_f(x_0); x_0, v\big) \cap D_\mp^2(Q; x_0, v) = \emptyset, \qquad \forall v \in \mathbb{X}. \tag{2.6}$$

Proof. It is an immediate application of the properties (2.1) and (2.2); in fact, using the characterization of optimality (2.5), we have

$$\begin{aligned} L_f(x_0) \cap B(x_0, \rho) \subseteq Q^c &\Rightarrow D_+^2\big(L_f(x_0); x_0, v\big) \subseteq D_+^2(Q^c; x_0, v), \ \ \forall v \in \mathbb{X} \\ &\Rightarrow D_+^2\big(L_f(x_0); x_0, v\big) \subseteq \big(D_-^2(Q; x_0, v)\big)^c, \ \ \forall v \in \mathbb{X} \\ &\Rightarrow D_+^2\big(L_f(x_0); x_0, v\big) \cap D_-^2(Q; x_0, v) = \emptyset, \ \forall v \in \mathbb{X}. \end{aligned}$$

Similar proof for the other disjunction. □

The crucial point of this paper is to be able to identify the directions belonging to the second–order variational sets $D_\pm^2\big(L_f(x_0); x_0, v\big)$ by means of the first and second–order Dini derivatives.

Theorem 2.4. *Let f be a lipschitzian function with constant K;*
(a) *if $D_{\pm} f(x_0; v) < 0$ then*

$$D^2_{\pm}\big(L_f(x_0); x_0, v\big) = \mathbb{X}; \tag{2.7}$$

(b) *if $D_{\pm} f(x_0; v) = 0$ and $D^2_{\pm} f(x_0; v, w) < 0$ then*

$$w \in \operatorname{int} D^2_{\pm}\big(L_f(x_0); x_0, v\big); \tag{2.8}$$

(c) *if there exists $y \in \mathbb{X}$ such that $f^{\circ}(x_0; y) = -L < 0$ then for every $v, w \in \mathbb{X}$ such that $D_{\pm} f(x_0; v) = 0$ and $D^2_{\pm} f(x_0; v, w) = 0$ we have*

$$w \in \operatorname{cl} D^2_{\pm}\big(L_f(x_0); x_0, v\big). \tag{2.9}$$

Proof. We state the proofs for D^2_{+} only because the other cases can be proved similarly.
(a) By assumption there exists $L > 0$ such that

$$-L = D_{+} f(x_0; v) = \limsup_{t \downarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}.$$

Therefore there exists $\lambda_0 > 0$ such that for each $t \in]0, \lambda_0[$

$$f(x_0 + tv) - f(x_0) \leq -\frac{L}{2} t.$$

Fix $w \in \mathbb{X}$; then for each $t \in]0, \lambda_0[$ we have

$$\begin{aligned} & f(x_0 + tv + t^2 w) - f(x_0) = \\ & \quad = f(x_0 + tv + t^2 w) - f(x_0 + tv) + f(x_0 + tv) - f(x_0) \leq \\ & \quad \leq K \|(x_0 + tv + t^2 w) - (x_0 + tv)\| - \frac{L}{2} t = t \left(tK\|w\| - \frac{L}{2} \right). \end{aligned}$$

Hence, choosing $\lambda = \min\left\{\lambda_0, \frac{L}{3K\|w\|}\right\}$, we get the thesis.
(b) By assumption there exists $L > 0$ such that

$$\begin{aligned} -L > D^2_{+} f(x_0; v, w) &= \limsup_{t \downarrow 0} \frac{f(x_0 + tv + t^2 w) - f(x_0) - tD_{+} f(x_0; v)}{t^2} \\ &= \limsup_{t \downarrow 0} \frac{f(x_0 + tv + t^2 w) - f(x_0)}{t^2}. \end{aligned}$$

Then, for each $\{t_n\} \downarrow 0$, we have

$$\frac{f(x_0 + t_n v + t_n^2 w) - f(x_0)}{t_n^2} \leq -L + \frac{L}{2} = -\frac{L}{2}$$

and therefore $w \in D_+^2(L_f(x_0); x_0, v)$. Fix $w' \in \mathbb{X}$; then

$$\begin{aligned}
&f(x_0 + t_n v + t_n^2 w') - f(x_0) = \\
&= f(x_0 + t_n v + t_n^2 w') - f(x_0 + t_n v + t_n^2 w) + f(x_0 + t_n v + t_n^2 w) - f(x_0) \leq \\
&\leq K \|(x_0 + t_n v + t_n^2 w') - (x_0 + t_n v + t_n^2 w)\| - \frac{L}{2} t_n^2 = \\
&= \left(K \|w - w'\| - \frac{L}{2} \right) t_n^2 .
\end{aligned}$$

Choosing w' such that $\|w' - w\| \leq \frac{L}{3K}$ we get the thesis.

(c) Fix $\tau > 0$; therefore we have

$$\begin{aligned}
&D_+^2 f(x_0; v, w + \tau y) = \\
&= \limsup_{t \downarrow 0} \frac{f\big(x_0 + tv + t^2(w + \tau y)\big) - f(x_0) - t D_+ f(x_0; v)}{t^2} \leq \\
&\leq \limsup_{t \downarrow 0} \frac{f(x_0 + tv + t^2 w + t^2 \tau y) - f(x_0 + tv + t^2 w)}{t^2} + \\
&+ \limsup_{t \downarrow 0} \frac{f(x_0 + tv + t^2 w) - f(x_0) - t D_+ f(x_0; v)}{t^2} \leq \\
&\leq \tau \limsup_{\substack{x \to x_0 \\ s \downarrow 0}} \frac{f(x + sy) - f(x)}{s} + D_+^2 f(x_0; v, w) = \tau f^\circ(x_0; y),
\end{aligned}$$

where we have placed in the last inequality $x = x_0 + tv + t^2 w$ and $s = \tau t^2$. Then, for each $\tau > 0$ we get

$$D_+^2 f(x_0; v, w + \tau y) \leq \tau f^\circ(x_0; y) = -\tau L.$$

Hence, by (2.8), we have $w + \tau y \in \operatorname{int} D_+^2(L_f(x_0); v, w)$ for every $\tau > 0$ and therefore, for $\tau \downarrow 0$, we achieve the thesis. $\square$

3. Second–order necessary optimality conditions.

This section is devoted to give an alternative proof of the Studniarski's result by means of the above mentioned variational sets. Moreover we extend this result also in absence of a regularity condition.

Proof of Studniarski's theorem. Placing

$$Q' = \bigcap_{i\in I(x_0)} Q_i \qquad \text{and} \qquad Q'' = \bigcap_{i\in I\setminus I(x_0)} Q_i,$$

we note that, thanks to the continuity of g_i, $x_0 \in \operatorname{int} Q''$. Being $x_0 \in Q$ a local optimal solution for (P), using (2.6), we get

$$D^2_-\big(L_f(x_0); x_0, v\big) \cap D^2_+(Q; x_0, v) = \emptyset, \qquad \forall v \in \mathbb{X}.$$

Moreover, by (2.3) and by (2.4), we obtain

$$D^2_-\big(L_f(x_0); x_0, v\big) \cap D^2_+(Q'; x_0, v) = \emptyset, \qquad \forall v \in \mathbb{X}.$$

We observe that if $A, B \subseteq \mathbb{X}$ and $A \cap B = \emptyset$ then $\operatorname{int} A \cap \operatorname{cl} B = \emptyset$; therefore

$$\operatorname{int} D^2_-\big(L_f(x_0); x_0, v\big) \cap \operatorname{cl} D^2_+(Q'; x_0, v) = \emptyset, \qquad \forall v \in \mathbb{X}. \tag{3.1}$$

Suppose, ab absurdo, that there exist $v, w \in \mathbb{X}$ satisfying (1.2) and either (1.3) or (1.4) and such that $D^2_+ f(x_0; v, w) < 0$; therefore by (1.2) and either (2.7) or (2.8) we get

$$w \in \operatorname{int} D^2_-\big(L_f(x_0); x_0, v\big).$$

Defining $g(x) = \max_{i\in I(x_0)} g_i(x)$ we have $Q' = \{x \in \mathbb{X} \mid g(x) \leq 0\}$; by Theorem 1.1, using the assumption (**c**), and by a well–known result in [7], we have that there exists $y \in \mathbb{X}$ such that $g^\circ(x_0; y) < 0$. Moreover it is immediate to observe that either $D_+g(x_0; v) < 0$ or $D_+g(x_0; v) = 0$ and $D^2_+g(x_0; v, w) \leq 0$. Therefore, using again the Theorem 2.4, we have

$$w \in \operatorname{cl} D^2_+(Q'; x_0, v)$$

that contradicts (3.1). □

Remark 3.1. We observe that, as a matter of fact, the assumption (1.2) is breakable in two cases:

(**a**) if $D_-f(x_0; v) = 0$ we get our statement, that is, $D^2_- f(x_0; v, w) \geq 0$;

(**b**) if $D_-f(x_0; v) < 0$ then $D^2_-\big(L_f(x_0); x_0, v\big) = \mathbb{X}$ and therefore there is no $w \in \mathbb{X}$ satisfying either the hypothesis (1.3) or (1.4). □

Taking advantage of this method of proof, we are able to prove a second–order necessary optimality condition without using regularity conditions but weakening the test directions. In the sequel, for every $x_0 \in Q$, we consider the following sets:

$$D(x_0) = \{v \in \mathbb{X} \mid D_- f(x_0; v) \leq 0, \quad D_+ g_i(x_0; v) \leq 0, \quad \forall i \in I(x_0)\}$$

and, for every $v \in D(x_0)$,

$$\Lambda(x_0, v) = \{i \in I(x_0) \mid D_+ g_i(x_0; v) = 0\}.$$

Theorem 3.1. *Let $x_0 \in Q$ be a local optimal solution for* (P). *For every $v \in D(x_0)$ there is no $w \in \mathbb{X}$ such that*

(i) *if $D_- f(x_0; v) < 0$, solves*

$$D_+^2 g_i(x_0; v, w) < 0, \qquad i \in \Lambda(x_0, v); \tag{3.2}$$

(ii) *if $D_- f(x_0; v) = 0$, solves*

$$\begin{cases} D_-^2 f(x_0; v, w) < 0 \\ D_+^2 g_i(x_0; v, w) < 0, \qquad i \in \Lambda(x_0, v). \end{cases} \tag{3.3}$$

Proof. Fix $v \in D(x_0)$; for every $i \in I(x_0) \setminus \Lambda(x_0, v)$ we have $D_+^2(Q_i; x_0, v) = \mathbb{X}$ and hence, similarly at the previous proof, we get

$$D_-^2\big(L_f(x_0); x_0, v\big) \cap \bigcap_{i \in \Lambda(x_0, v)} D_+^2(Q_i; x_0, v) = \emptyset.$$

(i) Let $D_- f(x_0; v) < 0$ and, ab absurdo, suppose there exists $w \in \mathbb{X}$ satisfying the system (3.2); then we have $D_-^2\big(L_f(x_0); x_0, v\big) = \mathbb{X}$ and $w \in \operatorname{int} D_+^2(Q_i; x_0, v)$ for each $i \in \Lambda(x_0, v)$ against our assumption.

(ii) Let $D_- f(x_0; v) = 0$ and, ab absurdo, suppose there exists $w \in \mathbb{X}$ satisfying the system (3.3); then we have $w \in \operatorname{int} D_-^2\big(L_f(x_0); x_0, v\big)$ and $w \in \operatorname{int} D_+^2(Q_i; x_0, v)$ against our assumption. □

If the functions f and g_i are twice Fréchet differentiable at x_0 the Theorem 3.1 collapses in a result obtained in [2].

Corollary 3.1. *Let $x_0 \in Q$ be a local optimal solution for* (P). *For every $v \in D(x_0)$ there is no $w \in \mathbb{X}$ such that*

(i) *if $\langle \nabla f(x_0), v \rangle < 0$, solves*

$$\langle \nabla g_i(x_0), w \rangle + \langle v, \nabla^2 g_i(x_0) v \rangle < 0, \qquad i \in \Lambda(x_0, v);$$

(ii) *if $\langle \nabla f(x_0), v \rangle = 0$, solves*

$$\begin{cases} \langle \nabla f(x_0), w \rangle + \langle v, \nabla^2 f(x_0) v \rangle < 0 & \\ \langle \nabla g_i(x_0), w \rangle + \langle v, \nabla^2 g_i(x_0) v \rangle < 0, & i \in \Lambda(x_0, v). \end{cases} \qquad \square$$

4. Conclusion.

It is possible to extend this result considering also non–lipschitzian functions even if we have to change directional derivatives and second–order variational sets [4]; other results have been obtained using metric regularity properties [18]. Besides, following this scheme, it is possible to build second–order necessary optimality condition in presence of abstract constraints [6]. Another important application of the second–order directional derivatives seems to be in the research of sufficient second–order optimality conditions that extend the well-known Penissi's result [13].

REFERENCES

[1] J.P. Aubin - H. Frankowska, *Set-valued analysis,* Birkhauser, Boston, 1990.

[2] A. Ben-Tal, *Second-order and related extremality conditions in nonlinear programming,* J. Optim. Theory Appl., 31 (1980), pp. 143-165.

[3] M. Castellani, *Refinements on solvability in mathematical programming,* Report No. 3.205(772), Dept. of Mathematics, Pisa, Italy. Submitted to SIAM J. Optim..

[4] M. Castellani, *A necessary second-order optimality condition in nonsmooth mathematical programming,* Report No. 3.211 (843), Dept. of Mathematics, Pisa, Italy. Submitted to Oper. Res. Lett..

[5] M. Castellani - M. Pappalardo, *First order cone approximations and necessary optimality conditions,* Report No. 3.204 (771), Dept. of Mathematics, Pisa, Italy. Submitted to Optimization.

[6] M. Castellani - M. Pappalardo, *Local second order approximations and applications in optimization,* to appear.

[7] F.H. Clarke, *Nonsmooth analysis and optimization,* Wiley, New York, 1984.

[8] K.H. Elster - J. Thierfelder, *Abstract cone approximations and generalized differentiability in nonsmooth optimization,* Optimization, 19 (1988), pp. 315-341.

[9] N. Furukawa - Y. Yoshinaga, *Higher-order variational sets, variational derivatives and higher-order necessary conditions in abstract mathematical programming,* Bull. Inform. Cybernet.,23 (1988), pp. 9-40.

[10] J.B. Hiriart - Urruty, *On optimality conditions in nondifferentiable programming,* Math. Programming, 14 (1978), pp. 73-86.

[11] K.H. Hoffmann - H.J. Kornstaedt, *Higher-order necessary conditions in abstract mathematical programming,* J. Optim. Theory Appl., 26 (1978), pp. 531-567.

[12] S. Komlosi - M. Pappalardo, *A general scheme for first order approximations in optimization,* Optimization Methods and Software, 3 (1994), pp. 143-152.

[13] E.J. Penissi, *An indirect sufficiency proof for the problem of Lagrange with differential inequalities as added side conditions,* Trans. Amer. Math. Soc., 74 (1953), pp. 177-198.

[14] J.B. Penot, *Variations on the theme of nonsmooth analysis,* in Lecture Notes in Econom. and Math. Systems, 225, Springer, New York, 1985.

[15] M. Studniarski, *Second-order necessary conditions for optimality in nonsmooth nonlinear programming,* Jour. Math. Anal. Appl., 154 (1991), pp. 303-317.

[16] D.E. Ward, *Convex subcones of the contingent cone in nonsmooth calculus and optimization,* Trans. Amer. Math. Soc., 302 (1987), pp. 661-682.

[17] D.E. Ward, *Directional derivatives calculus and optimality conditions in nonsmooth mathematical programming,* J. Inform. Optim. Sci., 10 (1989), pp. 81-96.

[18] D.E. Ward, *A chain rule for parabolic second-order epiderivatives,* Optimization, 28 (1994), pp. 223-236.

Dipartimento di Matematica,
Università di Pisa,
Via F. Buonarroti 2,
56127 Pisa (ITALY),
e-mail: castella@dm.unipi.it