Control and Cybernetics

vol. **36** (2007) No. 3

Local cone approximations in optimization

by

M. Castellani[1] and M. Pappalardo[2]

[1] Department of "Sistemi ed Istituzioni per l'Economia"
University of L'Aquila
[2] Department of "Matematica Applicata"
University of Pisa

Abstract: We show how to use intensively local cone approximations to obtain results in some fields of optimization theory such as optimality conditions, constraint qualifications, mean value theorems and error bound.

Keywords: nonsmooth optimization, cone approximations, generalized directional epiderivatives, optimality conditions, constraint qualifications, mean value theorem, error bound.

1. Introduction

Many approaches have been developed in literature for obtaining results in optimization theory. In this paper we propose a systematic approach to many problems of optimization theory (i.e. necessary and sufficient optimality conditions, constraint qualifications, mean value theorems and error bound) which uses local cone approximations of sets. This concept, introduced by Elster and Thierfelder (1988) and studied also in Castellani and Pappalardo (1995), appears to be very useful when used together with separation theorems and generalized derivatives. When intensively applied, we show that it is very flexible in our scope of interest.

The first step of this scheme consists in approximating a set with a cone; in particular, when the set is the epigraph of a function, the cone represents the epigraph of a positively homogeneous function, which will be called generalized derivative. The subsequent step consists in considering the case in which the generalized derivative is the difference between two sublinear functions; finally, the most general situation, is the one in which the generalized derivative is a minimum of sublinear functions.

The paper is organized as follows. Section 2 is devoted to the introduction of the concepts of local cone approximation and directional K–epiderivative. In Section 3 we obtain necessary optimality conditions in two different ways: the former, called "primal", where a necessary optimality condition is classically obtained as nonnegativity of some suitable generalized derivative over some approximation of the feasible set; the latter, called "dual", where a necessary optimality condition is obtained in the form of a zero belonging to some suitable subdifferential of the Lagrangian function. In any case, primal conditions are obtained with weaker assumptions; while dual conditions, more useful for algorithmic schemes, are obtained via alternative theorems and therefore they need stronger assumptions, in other words they need "constraint qualifications". Section 4 develops the analysis of constraint qualifications; Section 5 is devoted to sufficient optimality conditions; Section 6 contains a scheme for mean value theorems and Section 7 shows some applications to establishing error bound.

In the sequel $(\mathbb{X}, \|\cdot\|)$ is a real Banach space, $\mathbb{X}^*$ is its topological dual space endowed with the weak* topology and $\langle\cdot,\cdot\rangle$ is the canonical pairing between $\mathbb{X}^*$ and $\mathbb{X}$. The open ball with center x and radius r is denoted by $B(x,r)$. Given a set A, we indicate by A^c, $\operatorname{cl} A$, $\operatorname{int} A$ and $\operatorname{conv} A$ the complementary, the closure, the interior and the convex hull of A respectively. An extended–value function $f : \mathbb{X} \longrightarrow (-\infty, +\infty]$ is said proper if $\operatorname{dom} f \neq \emptyset$, where

$$\operatorname{dom} f = \{x \in \mathbb{X} : f(x) < +\infty\}$$

is the domain of f. We denote by $\mathcal{E}(\mathbb{X})$ the class of the proper extended–value functions and by $\mathcal{F}(\mathbb{X})$ the subclass of the lower semicontinuous functions. If A is a closed subset of $\mathbb{X}^*$, the support function associated to A is

$$\sigma(x, A) = \sup\{\langle x^*, x\rangle : x^* \in A\};$$

the domain of $\sigma(\cdot, A)$ is called barrier cone of A and it is denoted $\operatorname{barr} A$.

2. Cone approximations and K–epiderivatives

The concepts of directional derivative and subdifferential of a convex function were used for treating convex optimization problems. Since more than thirty years ago a lot of effort was made to establish similar concept in the nonconvex nonsmooth case with the introduction of modifications of the directional derivative. In accordance with such investigations, in Elster and Thierfelder (1988) an axiomatic approach has been proposed for constructing generalized directional derivatives of arbitrary functions: the basic idea is the fact that the epigraphs of the different directional derivatives of a function f can be considered as cone approximations of the epigraph of f.

Definition 1 *A set–valued map $K : 2^{\mathbb{X}} \times \mathbb{X} \rightrightarrows \mathbb{X}$ is said to be a local cone approximation (in short l.c.a.) if to each set $A \subseteq \mathbb{X}$ and each point $x \in \mathbb{X}$ a cone $K(A, x)$ is associated such that the following properties hold:*

(i) $K(A, x) = K(A - x, 0)$,

(ii) $K(A \cap B(x, r), x) = K(A, x)$ *for each* $r > 0$,

(iii) $K(A, x) = \emptyset$ *for each* $x \notin \operatorname{cl} A$,

(iv) $K(A, x) = \mathbb{X}$ *for each* $x \in \operatorname{int} A$,

(v) $K(\varphi(A), \varphi(x)) = \varphi(K(A, x))$ *with* $\varphi : \mathbb{X} \longrightarrow \mathbb{X}$ *linear homeomorphism,*
(vi) $0^+A \subseteq 0^+K(A, x)$ *for each* $x \in \operatorname{cl} A$,
where $0^+A = \{x \in \mathbb{X} : \ A + x \subseteq A\}$ *is the recession cone of* A.

Conditions (i) and (v) require the invariance of the approximation with respect to translations and linear homeomorphisms; conditions (ii), (iii) and (iv) show the local character of K. The last condition is quite technical and it is needed to derive properties concerning the directional K–epiderivatives of functions. The set of all l.c.a. is not empty and the independence of the axioms (i),...,(vi), that is no axiom can be expressed by the others, was proved in Elster and Thierfelder (1988).

Now we cite the main properties of the l.c.a., which will be useful in what follows: most of them can be found in Elster and Thierfelder (1988).

Theorem 1 *Let* K *and* K_i, *with* $i \in I$ *arbitrary index set, be l.c.a.; then*

$$\operatorname{int} K, \quad \operatorname{cl} K, \quad \operatorname{conv} K, \quad \bigcup_{i\in I} K_i, \quad \bigcap_{i\in I} K_i, \quad \sum_{i\in I} K_i,$$

are l.c.a.

We pay more attention to a particular operation, which is fundamental in order to derive necessary optimality conditions for extremum problems. Starting from the set–valued map $K : 2^{\mathbb{X}} \times \mathbb{X} \rightrightarrows \mathbb{X}$, we consider the new map $K_c : 2^{\mathbb{X}} \times \mathbb{X} \rightrightarrows \mathbb{X}$ defined by:

$$K_c(A, x) = (K(A^c, x))^c, \qquad \forall A \subseteq \mathbb{X}, \ \forall x \in \mathbb{X}. \tag{1}$$

It is trivial to observe that $K(A, x) = K_{cc}(A, x)$ and the following result holds:

Theorem 2 *If* K *is a l.c.a. then* K_c *is a l.c.a. too.*

The notion of l.c.a. is useful for describing generalized directional derivatives of an extended–value function.

Definition 2 *Let* K *be a l.c.a.,* $f \in \mathcal{E}(X)$ *and* $x \in \operatorname{dom} f$*; the directional* K*–epiderivative of* f *at* x *is the positively homogeneous function* $f^K(x, \cdot) : \mathbb{X} \longrightarrow [-\infty, +\infty]$ *defined by*

$$f^K(x, v) = \inf \{y \in \mathbb{R} : (v, y) \in K(\operatorname{epi} f, (x, f(x)))\}$$

where $\operatorname{epi} f = \{(x, y) \in \mathbb{X} \times \mathbb{R} : y \geq f(x)\}$ *is the epigraph of* f*. We assume* $\inf \emptyset = +\infty$.

In this way, we obtain a large family of generalized derivatives. It has been shown that many of the classical generalized derivatives (Dini, Dini-Hadamard, Clarke and so on) can be obtained following this scheme. Using the class of pointwise minimum of sublinear function, we give the following definition:

Definition 3 *Let K be a l.c.a., $f \in \mathcal{E}(X)$ and $x \in \operatorname{dom} f$; the function f is said to be K–MSL–differentiable at x if there exist an index set T and a family $\{\partial_t^K f(x)\}_{t \in T}$ of closed and convex sets in $\mathbb{X}^*$ such that*

$$f^K(x,v) = \min\{\sigma(v, \partial_t^K f(x)) : t \in T\}, \qquad \forall v \in \mathbb{X}.$$

In particular, when T is a singleton, the function is said to be K–subdifferentiable at x and the unique closed and convex set $\partial^K f(x)$ is called K–subdifferential.

When the directional K–epiderivative belongs to the class of difference of sublinear and continuous functions we say that the function f is K–quasidifferentiable (see Demyanov and Rubinov, 1995) at x and the two compact and convex sets $\underline{\partial}^K f(x), \overline{\partial}^K f(x) \subseteq \mathbb{X}^*$ such that

$$f^K(x,v) = \sigma(v, \underline{\partial}^K f(x)) - \sigma(v, \overline{\partial}^K f(x)), \qquad \forall v \in \mathbb{X}$$

are called K–subdifferential and K–superdifferential, respectively. Notice that a K–quasidifferentiable function is K–MSL–differentiable since

$$\begin{aligned} f^K(x,v) &= \sigma(v, \underline{\partial}^K f(x)) - \sigma(v, \overline{\partial}^K f(x)) \\ &= \min\{\sigma(v, \underline{\partial}^K f(x) - \overline{x}^*) \ \overline{x}^* \in \overline{\partial}^K f(x)\}; \end{aligned}$$

actually, the class of K–MSL–differentiable functions is quite wide as the following result, proved in Castellani (2000), shows.

Theorem 3 *Let K be a l.c.a., $f \in \mathcal{E}(\mathbb{X})$ and $x \in \operatorname{dom} f$.*

- *f is K–MSL-differentiable if and only if $f^K(x,\cdot)$ is proper and $f^K(x,0) = 0$;*
- *if $f^K(x,\cdot)$ is Lipschitz continuous then all the sets $\partial_t^K f(x)$ may be chosen compact.*

Since the directional K–epiderivative is strictly related to the epigraph of a function, it is not a surprise that the following result holds when f is K–subdifferentiable.

Theorem 4 *Let K be a l.c.a., $f \in \mathcal{E}(X)$ and $x \in \operatorname{dom} f$; then*

$$\partial^K f(x) = \{x^* \in \mathbb{X}^* : (x^*, -1) \in K^\circ(\operatorname{epi} f, (x, f(x)))\}$$

where K° is the (negative) polar cone of K.

It is possible to show that the topological properties of the l.c.a. K affect the directional K–epiderivative.

Theorem 5 *Let K be a l.c.a., $f \in \mathcal{E}(X)$ and $x \in \operatorname{dom} f$; then*

- *the epigraph of $f^K(x,\cdot)$ is the “vertical” closure of $K(\operatorname{epi} f, (x, f(x)))$ i.e.*

  $$\operatorname{epi} f^K(x,\cdot) = \{(y,\beta) \in \mathbb{X} \times \mathbb{R} : \forall \varepsilon > 0,\ (y, \beta + \varepsilon) \in K(\operatorname{epi} f, (x, f(x)))\};$$

  in particular if K is closed $\operatorname{epi} f^K(x,\cdot) = K(\operatorname{epi} f, (x, f(x)))$ and $f^K(x,\cdot)$ is lower semicontinuous.

- *the strict epigraph of* $f^K(x,\cdot)$*, defined*

$$\text{s-epi}\, f^K(x,\cdot) = \{(v,y) \in \mathbb{X} \times \mathbb{R} : y > f^K(x,v)\}$$

is the "vertical" interior of $K(\text{epi}\, f, (x, f(x)))$ *i.e.*

$$\text{s-epi}\, f^K(x,\cdot) = \{(y,\beta) \in \mathbb{X}\times\mathbb{R} : \forall\varepsilon > 0,\ (y, \beta-\varepsilon) \in K(\text{epi}\, f, (x, f(x)))\};$$

in particular if K *is open* $\text{s-epi}\, f^K(x,\cdot) = K(\text{epi}\, f, (x, f(x)))$ *and* $f^K(x,\cdot)$ *is upper semicontinuous.*

With respect to the l.c.a. K_c we have the following statement:

THEOREM 6 *Let* K *be a l.c.a.,* $f \in \mathcal{E}(X)$ *and* $x \in \text{dom}\, f$*; if, for all* $v \in \mathbb{X}$*,*

$$\inf\{y \in \mathbb{R} : (v,y) \in K_c(\text{epi}\, f, (x, f(x)))\} = \\ = \inf\{y \in \mathbb{R} : (v,y) \in K_c(\text{s-epi}\, f, (x, f(x)))\}$$

then

$$f^{K_c}(x,v) = -(-f)^K(x,v), \qquad \forall v \in \mathbb{X}.$$

In optimization, it is often useful to calculate the directional K–epiderivative of the pointwise maximum of a family of functions. Given $f_i \in \mathcal{E}(\mathbb{X})$ with $i \in I$ a finite index set, define

$$f_{\max}(x) = \max\{f_i(x) : i \in I\}$$

and

$$I_{\max}(x) = \{i \in I : f_{\max}(x) = f_i(x)\}.$$

We prove the following result:

THEOREM 7 *Let* K *be a l.c.a.,* $x \in \cap_{i\in I}\text{dom}\, f_i$ *and suppose* f_i *upper semicontinuous for each* $i \notin I_{\max}(x)$*. If*

$$K(A \cap B, x) \subseteq K(A,x) \cap K(B,x), \quad \forall A, B \subseteq \mathbb{X}, \qquad \forall x \in \mathbb{X}$$

then

$$f^K_{\max}(x,v) \geq \max\{f^K_i(x,v) : i \in I_{\max}(x)\}, \qquad \forall v \in \mathbb{X}.$$

Proof. Notice that $\text{epi}\, f_{\max} = \bigcap_{i\in I} \text{epi}\, f_i$ and therefore, by the assumption on the l.c.a. of the intersection of sets,

$$\begin{aligned} f^K_{\max}(x,v) &= \inf\left\{y \in \mathbb{R} : (v,y) \in K\left(\bigcap_{i\in I} \text{epi}\, f_i, (x, f_{\max}(x))\right)\right\} \\ &\geq \inf\left\{y \in \mathbb{R} : (v,y) \in \bigcap_{i\in I} K\left(\text{epi}\, f_i, (x, f_{\max}(x))\right)\right\}. \end{aligned}$$

The upper semicontinuity implies that $(x, f_{\max}(x)) \in \operatorname{int}\operatorname{epi} f_i$ for each $i \notin I_{\max}(x)$ and then

$$K(\operatorname{epi} f_i, (x, f_{\max}(x))) = \mathbb{X} \times \mathbb{R}, \qquad \forall i \notin I_{\max}(x).$$

Hence

$$\begin{aligned} f^K_{\max}(x, v) &\geq \inf \left\{ y \in \mathbb{R} : (v, y) \in \bigcap_{i \in I_{\max}(x)} K(\operatorname{epi} f_i, (x, f_{\max}(x))) \right\} \\ &= \max_{i \in I_{\max}(x)} \inf \{ y \in \mathbb{R} : (v, y) \in K(\operatorname{epi} f_i, (x, f_i(x))) \} \\ &= \max\{ f^K_i(x, v) : i \in I_{\max}(x) \} \end{aligned}$$

that proves the result. ■

An analogous result holds for the pointwise minimum of lower semicontinuous functions changing the intersection with the union and $\subseteq$ with $\supseteq$ in the assumption and substituting max with min and $\geq$ with $\leq$ in the result.

3. Necessary optimality conditions

Let us consider the following problem

$$\min\{f_0(x) : x \in S\}, \tag{2}$$

where $f_0 \in \mathcal{E}(\mathbb{X})$ and $S \subseteq \mathbb{X}$ is the feasible region. It is immediate to observe that $\overline{x} \in \mathbb{X}$ is a local solution for (2) if and only if $\overline{x} \in S$ and there exists $r > 0$ such that

$$\operatorname{epi} f_0 \cap [S \times (-\infty, f_0(\overline{x}))] \cap B((\overline{x}, f_0(\overline{x})), r) = \emptyset. \tag{3}$$

Even if this expression is easy and quite elegant from the formal viewpoint, in general it is an arduous task to verify it. For this reason it is suitable to replace the sets in (3) with approximations having a simpler structure. After, we study separation between these suitable approximations. In particular take two l.c.a. $K, H : 2^{\mathbb{X}} \times \mathbb{X} \rightrightarrows \mathbb{X}$ such that

$$A \cap B = \emptyset \quad \Longrightarrow \quad K(A, x) \cap H(B, x) = \emptyset, \qquad \forall x \in \mathbb{X}.$$

Such a pair of l.c.a. will be called admissible and the separation will be a necessary condition for the disjunction of the sets A and B. It is obvious that if (K, H) is an admissible pair of l.c.a. and $K' \subseteq K$ and $H' \subseteq H$ are l.c.a. then the pair (K', H') is admissible too. The existence of not trivial pairs of admissible l.c.a., that is to say – different from the identity map, is ensured by the next theorem (see Vlach, 1970, and Castellani and Pappalardo, 1995). We recall that an l.c.a. K is called

- isotone if
  $K(A,x) \subseteq K(B,x), \qquad \forall A \subseteq B \subseteq \mathbb{X},\ \forall x \in \mathbb{X};$
- isotone–dominated if there exists an isotone l.c.a. H such that
  $K(A,x) \subseteq H(A,x), \qquad \forall A \subseteq \mathbb{X},\ \forall x \in \mathbb{X}.$

Theorem 8 *Let K be an isotone l.c.a. and K_c defined in (1). Then K_c is isotone and the pair (K, K_c) is admissible.*

We start with the unconstrained problem $S = \mathbb{X}$ (see Castellani, D'Ottavio and Giuli, 2003).

Theorem 9 *Let K_0 be an isotone–dominated l.c.a. and $\overline{x} \in \mathbb{X}$ be a local solution for (2) with $S = \mathbb{X}$; then*

$$f_0^{K_0}(\overline{x}, v) \geq 0, \qquad \forall v \in \mathbb{X}. \tag{4}$$

Moreover, if f_0 is K_0–MSL–differentiable with respect to the family $\{\partial_t^{K_0} f(\overline{x})\}_{t\in T}$, condition (4) can be equivalently written in dual form

$$0 \in \bigcap_{t\in T} \partial_t^{K_0} f_0(\overline{x}).$$

The points $\overline{x} \in \mathbb{X}$ satisfying (4) are called K_0–stationary points for f_0. Theorem 5.1, proved in Elster and Thierfelder (1988), required $K_0 \subseteq T$; since T is an isotone l.c.a., Theorem 9 extends this result.

We are now interested in the constrained case. Nevertheless, if we replace the space $\mathbb{X}$ with a subset S, we are not able to prove an analogous result to Theorem 9 using only the abstract properties of the l.c.a., but we need a further assumption.

Theorem 10 *Let (K_0, H) be an admissible pair of l.c.a., H_0 be an l.c.a. and $\overline{x} \in \mathbb{X}$ be a local solution for (2). Suppose that for each $(x,y) \in \mathbb{X} \times \mathbb{R}$*

$$H(S \times (-\infty, y), (x,y)) \supseteq H_0(S,x) \times (-\infty, 0); \tag{5}$$

then

$$f_0^{K_0}(\overline{x}, v) \geq 0, \qquad \forall v \in H_0(S, \overline{x}). \tag{6}$$

Moreover, if $\mathbb{X}$ is finite dimensional and f_0 is K_0–MSL–differentiable with respect to the family $\{\partial_t^{K_0} f_0(\overline{x})\}_{t\in T}$, the cone $H_0(S,\overline{x})$ is convex and the following condition holds

$$-H_0^\circ(S,\overline{x}) \cap \left(\operatorname{barr} \partial_t^{K_0} f_0(\overline{x})\right)^\circ = \{0\}, \qquad \forall t \in T \tag{7}$$

then condition (6) can be equivalently written in dual form

$$0 \in \bigcap_{t\in T} \left(\partial_t^{K_0} f_0(\overline{x}) + H_0^\circ(S,\overline{x})\right).$$

Proof. We prove only the latter part of the theorem (for the first part see Castellani and Pappalardo, 1995). For fixed $t \in T$, we have

$$\operatorname{epi} \sigma(\cdot, \partial_t^{K_0} f_0(\overline{x})) \cap H_0(S, \overline{x}) \times (-\infty, 0) = \emptyset.$$

Since $\operatorname{epi} \sigma(\cdot, \partial_t^{K_0} f_0(\overline{x}))$ is convex, by separation there exists a nonzero vector $(x^*, a) \in \mathbb{X}^* \times \mathbb{R}$ such that

$$\langle x^*, v\rangle + a\alpha \leq 0, \qquad \forall (v, \alpha) \in H_0(S, \overline{x}) \times (-\infty, 0), \tag{8}$$

$$\langle x^*, v\rangle + a\alpha \geq 0, \qquad \forall (v, \alpha) \in \operatorname{epi} \sigma(\cdot, \partial_t^{K_0} f(\overline{x})). \tag{9}$$

From (8), we deduce $a \geq 0$ and $x^* \in H_0^\circ(Q, x_0)$; from (9), for any $\varepsilon > 0$ and for any $v \in \operatorname{dom} \sigma(\cdot, \partial_t^{K_0} f_0(\overline{x}))$ we have

$$\langle x^*, v\rangle + a\sigma(v, \partial_t^{K_0} f_0(\overline{x})) + a\varepsilon \geq 0.$$

The arbitrariness of ε and the nonnegativity of a imply

$$\langle x^*, v\rangle + a\sigma(v, \partial_t^{K_0} f_0(\overline{x})) \geq 0, \qquad \forall v \in \operatorname{dom} \sigma(\cdot, \partial_t^{K_0} f_0(\overline{x})).$$

The constant a is different from zero, otherwise we have

$$-x^* \in (\operatorname{dom} \sigma(\cdot, \partial_t^{K_0} f_0(\overline{x})))^\circ$$

which contradicts the assumption (7); therefore

$$\sigma(v, \partial_t^{K_0} f_0(\overline{x})) \geq \langle -a^{-1} x^*, v\rangle, \qquad \forall v \in \operatorname{dom} \sigma(\cdot, \partial_t^{K_0} f_0(\overline{x})),$$

and $-a^{-1}x^* \in \partial_t^{K_0} f_0(\overline{x})$. Since $H_0^\circ(S, \overline{x})$ is a cone we obtain $a^{-1}x^* \in H_0^\circ(S, \overline{x})$ and the proof is completed. ■

Assumption (7) can be expressed in different ways (see, for instance, Ward and Borwein, 1987, or Ward, 1991); indeed, since if K_0 and H_0 are two cones then $(K_0 + H_0)^\circ = K_0^\circ \cap H_0^\circ$, condition (7) is equivalent to

$$\Big(\operatorname{barr} \partial_t^{K_0} f(\overline{x}) - H_0(S, \overline{x})\Big)^\circ = \{0\}, \qquad \forall t \in T$$

and a sufficient condition for this is expressed by

$$\operatorname{barr} \partial_t^{K_0} f(\overline{x}) - H_0(S, \overline{x}) = \mathbb{X}.$$

Moreover, since

$$0^+ \partial_t^{K_0} f(\overline{x}) = \Big(\operatorname{barr} \partial_t^{K_0} f(\overline{x})\Big)^\circ,$$

many authors, in the K–subdifferential case, use the recession cone of the K–subdifferential instead of $(\operatorname{dom} f_0^{K_0}(\overline{x}, \cdot))^\circ$ in (7).

Now we study the following extremum problem

$$\min\{f_0(x) : f_i(x) \le 0,\ i \in I\}, \tag{10}$$

where $f_i \in \mathcal{E}(\mathbb{X})$ and I is a finite index set. For every feasible point x we denote $I(x) = \{i \in I : f_i(x) = 0\}$ and $I_0(x) = I(x) \cup \{0\}$; moreover we put

$$S_i = \{x \in \mathbb{X} : f_i(x) \le 0\}$$

for each $i \in I$. In the last decades many generalizations of the KKT necessary optimality condition for the problem (10) have been stated without assuming the differentiability of the functions f_i. It has been proved that these necessary optimality conditions are equivalent to the impossibility of systems of suitable directional K–epiderivatives. We show our approach.

Theorem 11 *Let (K_0, H) be an admissible pair of l.c.a., $\overline{x} \in \mathbb{X}$ be a local solution for (10) and f_i be upper semicontinuous for each $i \in I \setminus I(\overline{x})$. Suppose there exist a family of l.c.a. $\{K_i\}_{i \in I(\overline{x})}$ and an l.c.a. H_0 satisfying assumption (5),*

$$\bigcap_{i \in I(\overline{x})} K_i(S_i, \overline{x}) \subseteq H_0(S, \overline{x}). \tag{11}$$

and

$$\{v \in \mathbb{X} : f_i^{K_i}(\overline{x}, v) < 0\} \subseteq K_i(S_i, \overline{x}), \qquad \forall i \in I(\overline{x}). \tag{12}$$

Then the system

$$\begin{cases} f_0^{K_0}(\overline{x}, v) < 0, \\ f_i^{K_i}(\overline{x}, v) < 0, \quad i \in I(\overline{x}), \end{cases} \tag{13}$$

is impossible.

Proof. From the upper semicontinuity of f_i and from assumptions (11) and (12) we deduce

$$\bigcap_{i \in I(\overline{x})} \{v \in \mathbb{X} : f_i^{K_i}(\overline{x}, v) < 0\} \subseteq H_0(S, \overline{x}).$$

The impossibility of system (13) descends from (6). ■

If the system (13) is impossible then the feasible points $\overline{x}$ is said to be a weakly stationary point for (10) with respect to the family $\{K_i\}_{i \in I_0(\overline{x})}$. If $f_i^{K_i}(\overline{x}, \cdot)$, with $i \in I_0(\overline{x})$, are pointwise minimum of sublinear functions it is possible to prove, through a theorem of the alternative for MSL–functions stated in Castellani (2000) that the impossibility of the system (13) is equivalent to a generalized John necessary optimality condition.

THEOREM 12 *In the same assumptions of Theorem 11 and supposing that each f_i is K_i–MSL–differentiable with respect to the family $\{\partial_{t_i}^{K_i} f_i(\overline{x})\}_{t_i\in T_i}$ then, for each $t_i \in T_i$ with $i \in I_0(\overline{x})$, we have*

$$0 \in \operatorname{cl}\operatorname{conv} \bigcup_{i\in I_0(\overline{x})} \partial_{t_i}^{K_i} f_i(\overline{x}). \tag{14}$$

Moreover, if $\operatorname{barr} \partial_{t_i}^{K_i} f_i(\overline{x})$ *are closed, condition (14) is equivalent to affirm that, for each $t_i \in T_i$ with $i \in I_0(\overline{x})$, there exist $\lambda_i \geq 0$ with $i \in I_0(\overline{x})$, not all zero, such that*

$$0 \in \operatorname{cl}\left(\sum_{i\in I_0(\overline{x}),\lambda_i>0} \lambda_i \partial_{t_i}^{K_i} f_i(\overline{x}) + \sum_{i\in I_0(\overline{x}),\lambda_i=0} 0^+ \partial_{t_i}^{K_i} f_i(\overline{x})\right). \tag{15}$$

In particular, if all the sets $\partial_{t_i}^{K_i} f_i(\overline{x})$ are compact (and it happens, for instance, when $f_i^{K_i}(\overline{x},\cdot)$ are Lipschitz), then $\operatorname{barr} \partial_{t_i}^{K_i} f_i(\overline{x}) = \mathbb{X}$ and hence (15) assumes the following simpler form

$$0 \in \sum_{i\in I_0(\overline{x})} \lambda_i \partial_{t_i}^{K_i} f_i(\overline{x}).$$

4. Constraint qualifications

A crucial point in optimization theory is to establish the weakest conditions guaranteeing the first KKT multiplier to be different from zero. We show in this section that, in our approach, this becomes to establish conditions guaranteeing the impossibility of the system

$$\begin{cases} p_0(x) < 0, & \\ p_i(x) \leq 0, & i \in I, \end{cases} \tag{16}$$

starting from the impossibility of system

$$\begin{cases} p_0(x) < 0, & \\ p_i(x) < 0, & i \in I, \end{cases} \tag{17}$$

where $p_i \in \mathcal{E}(\mathbb{X})$ are positively homogeneous functions, with $i \in I \cup \{0\}$ a finite index set and $\cap_{i\in I} \operatorname{dom} p_i \neq \emptyset$. No assumption of convexity or its generalizations will be required. All the results in this section are proved in Castellani (2006).

We recall that the recession function of a positively homogeneous function $p : \mathbb{X} \longrightarrow (-\infty, +\infty]$ is defined as

$$p^\infty(x) = \sup\{p(x+y) - p(y) \ : y \in \operatorname{dom} p\}.$$

Notice that p^∞ is sublinear and if p is sublinear with $p(0)=0$ then $p^\infty(x)=p(x)$ for all $x \in \mathbb{X}$. Denote $P(x)=\max\{p_i(x) : i \in I\}$ and consider the following three conditions:

$$\{x \in \mathbb{X} : P^\infty(x) < 0\} \neq \emptyset, \tag{18}$$

$$\forall x \in \mathbb{X} : P(x)=0,\ \exists v \in \mathbb{X}\ \liminf_{t\downarrow 0} \frac{P(x+tv)-P(x)}{t} < 0 \tag{19}$$

$$\operatorname{cl}\{x \in \mathbb{X} : P(x) < 0\} = \{x \in \mathbb{X}\ P(x) \le 0\}. \tag{20}$$

The relations between these three conditions are stated in the following result:

Theorem 13 *Suppose that the above mentioned assumptions about p_0 and p_i are in force, then (18) implies (19). Moreover, if P is lower semicontinuous, then (19) implies (20).*

The converse implications do not hold in general but only if p_i are sublinear with $p_i(0)=0$. Now we state the main result of this section.

Theorem 14 *Let the system (17) be impossible;*
- (i) *if $\operatorname{dom} p_0^\infty \cap \{x \in X : P^\infty(x) < 0\} \neq \emptyset$, then the system (16) is impossible,*
- (ii) *if (19) holds and p_0 is upper semicontinuous, then the system (16) is impossible,*
- (iii) *if (20) holds and p_0 is upper semicontinuous, then the system (16) is impossible.*

Let us apply Theorem 14 to the optimality condition expressed by the impossibility of the positively homogeneous system (13).

Theorem 15 *Let (K_0,H) be an admissible pair of l.c.a., $\overline{x} \in \mathbb{X}$ be a local solution for (10) and f_i be upper semicontinuous for each $i \in I \setminus I(\overline{x})$. Suppose that H satisfies assumptions (5) and (11). Let $\{K_i\}_{i\in I(\overline{x})}$ be a family of l.c.a. satisfying (12) and such that $f_0^{K_0}(\overline{x},\cdot)$ and $f_i^{K_i}(\overline{x},\cdot)$ are proper. Define*

$$F(\overline{x},v)=\max\{f_i^{K_i}(\overline{x},v) : i \in I(\overline{x})\}$$

and suppose that one of the following assumptions is satisfied:
- (i) $\operatorname{dom} f_0^{0^+K_0}(\overline{x},\cdot) \cap \{v \in \mathbb{X}\ F^\infty(\overline{x},v) < 0\} \neq \emptyset$;
- (ii) $f_0^{K_0}(\overline{x},\cdot)$ *upper semicontinuous and*
  $$\forall v \in \mathbb{X} : F^\infty(\overline{x},v)=0,\ \exists w \in \mathbb{X}\ \liminf_{t\downarrow 0} \frac{F(\overline{x},v+tw)-F(\overline{x},v)}{t} < 0;$$
- (iii) $f_0^{K_0}(\overline{x},\cdot)$ *upper semicontinuous and* $\operatorname{cl}\{v \in \mathbb{X} : F(\overline{x},v) < 0\} = \{v \in \mathbb{X} : F(\overline{x},v) \le 0\}$.

Then the system

$$\begin{cases} f_0^{K_0}(\overline{x},v) < 0, \\ f_i^{K_i}(\overline{x},v) \le 0, \qquad i \in I(\overline{x}). \end{cases} \tag{21}$$

is impossible.

The feasible point $\overline{x}$ is said to be a strongly stationary point for (10) with respect to the family $\{K_i\}_{i\in I_0(\overline{x})}$ if the system (21) is impossible. In order to establish a generalization of the KKT necessary optimality condition, we apply a generalized Farkas lemma for MSL system proved in Glover, Ishizuka, Jeyakumar and Tuan (1996) and we obtain the following result.

THEOREM 16 *Given the assumptions of Theorem 15 and supposing that each f_i is K_i–MSL–differentiable with respect to the family $\{\partial_{t_i}^{K_i} f_i(\overline{x})\}_{t_i\in T_i}$ then, for each $t_i \in T_i$ with $i \in I_0(\overline{x})$, we have*

$$0 \in \mathrm{cl}\left(\partial_{t_0}^{K_0} f_0(\overline{x}) + \mathrm{cone\,conv} \bigcup_{i\in I(\overline{x})} \partial_{t_i}^{K_i} f_i(\overline{x})\right).$$

5. Sufficient optimality conditions

The concept of invexity was introduced by Hanson (1981) as a generalization of differentiable convex functions: the function $f \in \mathcal{E}(\mathbb{X})$ is said to be invex if there exists a function $\eta : \mathbb{X} \times \mathbb{X} \longrightarrow \mathbb{X}$ such that

$$f(x_1) - f(x_2) \geq \langle \nabla f(x_2), \eta(x_1, x_2)\rangle, \qquad \forall x_1 \in \mathbb{X},\ \forall x_2 \in \mathrm{dom}\, f.$$

The name invex descends from a contraction of "invariant convex" and it was proposed by Craven (1981) since he observed that invexity is preserved by bijective coordinate transformations. The concept of invexity was adapted to various nonsmooth classes of functions: Reiland (1990) considered the class of locally Lipschitz functions and he used the generalized gradient of Clarke, Jeyakumar (1987) introduced the notion of approximately quasidifferential functions, Ye (1991) studied the directional differentiable functions.

By exploiting the concept of directional K–epiderivative it is possible to give a unifying definition of invexity for nonsmooth functions.

DEFINITION 4 *Let K be a l.c.a.; the function $f \in \mathcal{E}(X)$ is said K–invex if there exists a function $\eta : \mathbb{X} \times \mathbb{X} \longrightarrow \mathbb{X}$ such that*

$$f(x_1) - f(x_2) \geq f^K(x_2, \eta(x_1, x_2)), \qquad \forall x_1 \in \mathbb{X},\ \forall x_2 \in \mathrm{dom}\, f.$$

The function η is said the kernel of the K–invexity.

We observe that if $f^{K_1}(x,\cdot) \geq f^{K_2}(x,\cdot)$ and f is K_1–invex then f is K_2–invex with respect to the same kernel. The following result is a trivial characterization of K–invexity for K–MSL–differentiable functions.

THEOREM 17 *Let K be an l.c.a. and $f \in \mathcal{E}(X)$ be a K–MSL–differentiable function with respect to the family $\{\partial_t^K f\}_{t\in T}$; then f is K–invex with kernel η if and only if there exists a function $\eta : \mathbb{X} \times \mathbb{X} \longrightarrow \mathbb{X}$ such that for each $x_1 \in \mathbb{X}$ and $x_2 \in \mathrm{dom}\, f$ there exists $t \in T$ with*

$$f(x_1) - f(x_2) \geq \langle x^*, \eta(x_1, x_2)\rangle, \qquad \forall x^* \in \partial_t^K f(x_2).$$

The following result (see Castellani, 2001) is fundamental to understand the structure of this class of functions (for differentiable functions it was proved by Ben–Israel and Mond, 1986).

THEOREM 18 *Let K be an l.c.a.; the function $f \in \mathcal{E}(X)$ is K–invex if and only if every K–stationary point is a global minimum point.*

Adapting the generalizations proposed in Kaul and Kaur (1985) for differentiable functions, we can extend the notion of K–invexity in the following way:

DEFINITION 5 *Let K be an l.c.a.; the function $f \in \mathcal{E}(X)$ is called*

- *K–quasiinvex if there exists a function $\eta : \mathbb{X} \times \mathbb{X} \longrightarrow \mathbb{X}$ such that $f(x_1) \le f(x_2)$ implies $f^K(x_2, \eta(x_1, x_2)) \le 0$ for all $x_1 \in \mathbb{X}$ and $x_2 \in \operatorname{dom} f$;*
- *strictly K–pseudoinvex if there exists a functional $\eta : \mathbb{X} \times \mathbb{X} \longrightarrow \mathbb{X}$ such that $f(x_1) \le f(x_2)$ implies $f^K(x_2, \eta(x_1, x_2)) < 0$ for all $x_1 \in \mathbb{X} \setminus \{x_2\}$ and $x_2 \in \operatorname{dom} f$.*

We observe that every K–invex function is K–quasiinvex with respect to the same kernel.

Now, we use the concept of K–invexity and its generalizations in order to deduce sufficient optimality conditions directly from the impossibility of the systems (13) or (21). Both results were proved in Castellani (2001).

THEOREM 19 *Let $\overline{x} \in \mathbb{X}$ be a strongly stationary point for (10) with respect to the family $\{K_i\}_{i \in I_0(\overline{x})}$. If f_0 is K_0–invex and f_i are K_i–quasiinvex with respect to the same kernel η then $\overline{x}$ is a global solution for (10).*

We have noted that the impossibility of (21) descends from the impossibility of (13) in presence of a regularity condition. Nevertheless, even if we have not regularity but we strengthen the hypothesis of invexity of the constraint functions, the impossibility of the system (13) implies the optimality of $\overline{x}$.

THEOREM 20 *Let $\overline{x} \in \mathbb{X}$ be a weakly stationary point for (10) with respect to the family $\{K_i\}_{i \in I_0(\overline{x})}$. If f_0 is K_0–invex and f_i are strictly K_i–pseudoinvex with respect to the same kernel η then $\overline{x}$ is a global solution for (10).*

6. Mean value theorem

In the last years many authors have introduced different axiomatic approaches in order to derive generalizations of mean value theorems (see Thibault and Zagrodny, 1985, or Aussel, Corvellec and Lassonde, 1995, and references therein). Such an effort has been devoted to avoid redoubling of different results whose, proofs follow the same principles. Nevertheless, the core of these approaches is to construct an axiomatic class of abstract subdifferentials containing as special case all the well-known subdifferentials. The goal of this section is to show that an abstract form of the approximate mean value theorem can be obtained also by means of the concept of directional K–epiderivative. This approach allows us to avoid the analysis of the smoothness regularity of the norm of the Banach space (see references in Aussel, Corvellec and Lassonde, 1995).

An l.c.a. K is said to be convex–regular if for each $f \in \mathcal{E}(X)$ and for each continuous convex function g we have

$$(f+g)^K(x,v) \leq f^K(x,v) + g'(x,v), \qquad \forall x \in \operatorname{dom} f \text{ and } v \in \mathbb{X},$$

where

$$g'(x,v) = \lim_{t \to 0^+} \frac{g(x+tv) - g(x)}{t}$$

is the classical directional derivative.

Theorem 21 *Let K be an isotone–dominated and convex–regular l.c.a. and $f \in \mathcal{F}(X)$; then, for each $a, b \in \mathbb{X}$ with $a \in \operatorname{dom} f$, and for each $r \leq f(b)$ there exist $\overline{x} \in [a,b)$ and a sequence $\{x_k\} \subseteq \operatorname{dom} f$ with $x_k \to \overline{x}$ and $f(x_k) \to f(\overline{x})$ such that*

$$\liminf_{k \to +\infty} f^K(x_k, b-a) \geq r - f(a)$$

and

$$\liminf_{k \to +\infty} f^K(x_k, b-x_k) \geq \frac{\|b-\overline{x}\|}{\|b-a\|}(r - f(a)).$$

The proof of Theorem 21, given in Castellani, D'Ottavio and Giuli (2003), follows the line of the proof given by Thibault (1995) for a generalization of the well–known approximate mean value theorem proved by Zagrodny (1988).

Moreover, by means of the previous approximate mean value theorem we may adapt the results in Aussel, Corvellec and Lassonde (1995) using l.c.a.

Corollary 1 *Let K be an isotone–dominated and convex–regular l.c.a. and $f \in \mathcal{F}(\mathbb{X})$. Suppose there exist $L > 0$ such that, for each $x \in \operatorname{dom} f$*

$$f^K(x,v) \leq L\|v\|, \qquad \forall v \in \mathbb{X}$$

then f is a Lipschitz function with constant L.

Proof. For fixed $x_1, x_2 \in \mathbb{X}$ with $x_1 \in \operatorname{dom} f$, $r \leq f(x_2)$ and $\varepsilon > 0$, from Theorem 21 we deduce that there exists $x' \in \operatorname{dom} f$ near to $[x_1, x_2)$ such that

$$r - f(x_1) \leq f^K(x', x_2 - x_1) + \varepsilon \leq L\|x_2 - x_1\| + \varepsilon.$$

Hence $x_2 \in \operatorname{dom} f$ and, since ε is arbitrary, we achieve the inequality

$$f(x_2) - f(x_1) \leq L\|x_2 - x_1\|.$$

Changing the role of x_1 with x_2 we conclude the proof. ■

The next result is related to the monotonicity of a lower semicontinuous function. We recall that a convex and pointed cone C defines the partial ordering relation $\leq_C$ on $\mathbb{X}$ by

$$x_1 \leq_C x_2 \qquad \Longleftrightarrow \qquad x_2 - x_1 \in C.$$

A function f is C–decreasing if $x_1 \leq_C x_2$ implies $f(x_1) \geq f(x_2)$.

Corollary 2 *Let K be an isotone–dominated and convex–regular l.c.a. and $f \in \mathcal{F}(X)$; if*

$$f^K(x,v) \leq 0, \qquad \forall v \in C$$

where C is a convex and ponted cone, then f is C–decreasing.

Proof. Suppose, by contradiction, that there exist $x_1, x_2 \in \mathbb{X}$ with $x_2 - x_1 \in C$ and $x_2 \in \operatorname{dom} f$ such that $f(x_1) > f(x_2)$. From Theorem 21 we deduce that there exists $x' \in \operatorname{dom} f$ near to $[x_1, x_2)$ such that $f^K(x', x_2 - x_1) > 0$, which contradicts the assumption. ■

7. Error bound

Roughly speaking, the solution set of an inequality system is said to have an error bound if the involved functions provide an upper estimate for the distance from any point to the solution set. More precisely, given a function $f \in \mathcal{E}(\mathbb{X})$ and denoting the solution set of the inequality by

$$S = \{x \in \mathbb{X} : f(x) \leq 0\},$$

we say that S has a local error bound if it is nonempty and there exist two constants $\mu > 0$ and $a > 0$ such that

$$d(x,S) \leq \mu f_+(x), \qquad \forall x \in f^{-1}(-\infty, a)$$

where $d(x,S) = \inf_{x' \in S} \|x - x'\|$ and $f_+(x) = \max\{0, f(x)\}$ is the positive part of f.

In the last years, the pioneering result of Hoffman (1952) who gave an error bound on the distance from any point to the solution set of a linear system in $\mathbb{R}^n$, has been extended along many directions (see Azé, 2003, for an interesting survey). More recently, some authors weakened the convexity assumption on the functions using some tools of nonsmooth analysis: generalized subdifferentials and the Ekeland variational principle (see for instance Wu and Ye, 2001 and 2002).

It is possible to adapt the method presented in Wu and Ye (2002) for the directional K–epiderivatives and to obtain the following error bound result:

THEOREM 22 *Let K be an isotone–dominated and convex–regular l.c.a. and $f \in \mathcal{F}(\mathbb{X})$. Suppose that*

(i) there exists $a > 0$ such that $f^{-1}(-\infty, a) \neq \emptyset$,

(ii) there exists $m > 0$ such that, for each $x \in f^{-1}(0, a)$ there is $v = v(x) \in \mathbb{X}$ such that

$$f^K(x, v) < -m\|v\|;$$

then

$$d(x, S) \leq m^{-1} f_+(x), \qquad \forall x \in f^{-1}(-\infty, a).$$

Proof. Suppose, by contradiction, that there exists $\overline{x} \in f^{-1}(0, a)$ such that

$$d(\overline{x}, S) > m^{-1} f_+(\overline{x}) = m^{-1} f(\overline{x}).$$

Let $t \in \mathbb{R}$ with $t > 1$ such that

$$d_S(\overline{x}) > t m^{-1} f_+(\overline{x}) = r;$$

then, applying Ekeland's variational principle to the lower semicontinuous function

$$f_+(x) + \delta(x, \operatorname{cl} B(\overline{x}, r)) = f(x) + \delta(x, \operatorname{cl} B(\overline{x}, r))$$

at the point $\overline{x}$ with $\varepsilon = f(\overline{x})$ and $\lambda = r t^{-1}$, where $\delta(\cdot, \operatorname{cl} B(\overline{x}, r))$ is the indicator function of the closed ball $\operatorname{cl} B(\overline{x}, r)$, we obtain that there exists $z \in \mathbb{X}$ such that

(a) $\|z - \overline{x}\| \leq r t^{-1}$,

(b) $f(z) + \delta(z, \operatorname{cl} B(\overline{x}, r)) \leq f(\overline{x}) + \delta(\overline{x}, \operatorname{cl} B(\overline{x}, r)) = f(\overline{x})$,

(c) the function $\varphi = f_+ + \delta(\cdot, \operatorname{cl} B(\overline{x}, r)) + m\| \cdot - z\|$ assumes global minimum at z.

From *(a)* and *(b)* we deduce that $z \in f^{-1}(0, a)$ and

$$\varphi = f + m\| \cdot - z\|$$

in a suitable neighborhood of z; hence, by means of Theorem 9 for each $v \in \mathbb{X}$, we have

$$0 \leq \varphi^K(z, v) = \left(f + m\| \cdot - z\|\right)^K(z, v) \leq f^K(z, v) + m\|v\|.$$

Therefore, $f^K(z, v) \geq -m\|v\|$ which contradicts the assumption. ■

Since $m\|v\| = \sigma(v, mB^*)$ where B^* is the closed unit ball in the dual space $\mathbb{X}^*$, if the function f is K–MSL–differentiable with respect to the family $\{\partial_t^K f\}_{t \in T}$, we have

$$\begin{aligned} f^K(x, v) + m\|v\| &= \min_{t \in T} \sigma(v, \partial_t^K f(x)) + \sigma(v, mB^*) \\ &= \min_{t \in T} \sigma(v, \partial_t^K f(x) + mB^*); \end{aligned}$$

hence, exploiting the dual characterization of K–stationary points, we may write assumption (ii) in the following form

$$0 \notin \bigcap_{t \in T} \left(\partial_t^K f(x) + mB^* \right).$$

In particular, if f is K–quasidifferentiable, the equivalent dual expression of assumption (ii) is

$$\overline{\partial}^K f(x) \not\subseteq \underline{\partial}^K f(x) + mB^*.$$

References

Aussel, D., Corvellec, J.–N. and Lassonde, M. (1995) Mean value property and subdifferential criteria for lower semicontinuous functions. *Trans. Amer. Math. Soc.* **347**, 4144–4161.

Azé, D. (2003) A survey on error bounds for lower semicontinuous functions. In: *Proceedings of 2003 MODE-SMAI Conference* **13**, 1–17.

Ben–Israel, A. and Mond, B. (1986) What is invexity? *J. Austral. Math. Soc. Ser. B* **28**, 1–9.

Castellani, M. (2000) A dual representations for proper positively homogeneous functions. *J. Global Optim.* **16**, 393–400.

Castellani, M. (2001) Nonsmooth invex functions and sufficient optimality conditions. *J. Math. Anal. Appl.* **255**, 319–332.

Castellani, M. (2006) On constraint qualifications in nonlinear programming. Submitted.

Castellani, M., D'Ottavio, A. and Giuli, M. (2003) A mean value theorem for K–directional epiderivatives. In: G.P. Crespi et al., eds., *Recent Advances in Optimization*, 21–34.

Castellani, M. and Pappalardo, M. (1995) First order cone approximations and necessary optimality conditions. *Optimization* **35**, 113–126.

Craven, B.D. (1981) Invex functions and constrained local minima. *Bull. Austral. Math. Soc.* **24**, 357–366.

Craven, B.D. and Glover, B.M. (1985) Invex functions and duality. *J. Austral. Math. Soc. Ser. A* **39**, 1–20.

Demyanov, V.F. and Rubinov, A.M. (1995) *Constructive Nonsmooth Analysis.* Peter Lang, Frankfurt am Main.

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

Glover, B.M., Ishizuka, Y., Jeyakumar, V. and Tuan, H.D. (1996) Complete characterizations of global optimality for problems involving the pointwise minimum of sublinear functions. *SIAM J. Optim.* **6**, 362–372.

Hanson, M.A. (1981) On sufficiency of the Kuhn–Tucker conditions. *J. Math. Anal. Appl.* **80**, 545–550.

Hoffman, A.J. (1952) On approximate solutions of systems of linear inequalities. *J. Res. Nat. Bur. Standards* **49**, 263–265.

Jeyakumar, V. (1987) On optimality conditions in nonsmooth inequality constrained minimization. *Numer. Funct. Anal. Optim.* **9**, 535–546.

Kaul, R.N. and Kaur, S. (1985) Optimality criteria in nonlinear programming involving nonconvex functions. *J. Math. Anal. Appl.* **105**, 104–112.

Komlosi, S. and Pappalardo, M. (1994) A general scheme for first order approximations in optimization. *Optim. Methods Softw.* **3**, 143–152.

Reiland, T.W. (1990) Nonsmooth invexity. *Bull. Austral. Math. Soc.* **42**, 437–446.

Thibault, L. (1995) A note on the Zagrodny mean value theorem. *Optimization* **35**, 127–130.

Thibault, L. and Zagrodny, D. (1985) Integration of subdifferentials of lower semicontinuous functions on Banach spaces. *J. Math. Anal. Appl.* **189**, 33–58.

Vlach, M. (1970) On necessary conditions of optimality in linear space. *Comment. Math. Univ. Carolin.* **11**, 501–513.

Ward, D.E. (1991) Chain rules for nonsmooth functions. *J. Math. Anal. Appl.* **158**, 519–538.

Ward, D.E. and Borwein, J.M. (1987) Nonsmooth calculus in finite dimensions. *SIAM J. Control Optim.* **25**, 1312–1340.

Wu, Z. and Ye, J.J. (2001) Sufficient conditions for error bounds. *SIAM J. Optim.* **12**, 421–435.

Wu, Z. and Ye, J.J. (2002) On error bounds for lower semicontinuous functions. *Math. Programming* **92**, 301–314.

Ye, Y.L. (1991) d–invexity and optimality conditions. *J. Math. Anal. Appl.* **162**, 242–249.

Zagrodny, D. (1988) Approximate mean value theorem for upper subderivatives. *Nonlinear Anal.* **12**, 1413–1428.