ISSN: 2281-1346

UNIVERSITÀ DI PAVIA
Department of Economics and Management

DEM Working Paper Series

Some Classical Directional Derivatives and Their Use in Optimization

Giorgio Giorgi
(Università di Pavia)

 # 204 (10-21)

Via San Felice, 5
I-27100 Pavia

economiaweb.unipv.it

Some Classical Directional Derivatives and Their Use in Optimization

Giorgio Giorgi*

Abstract. We give an overview, in finite-dimensioonal Euclidean spaces, of some classical directional derivatives (usual directional derivatives, Dini directional derivatives, Hadamard directional derivatives) and of some classical differentiability notions. We take into consideration some applications of the said concepts to convex and generalized convex functions, to nonsmooth unconstrained optimization problems and to nonsmooth constrained scalar and vector optimization pronlems. Also the axiomatic approach to nonsmooth analysis and nonsmooth optimization problems, proposed by K.-H. Elster and J. Thierfelder, is briefly considered.

Key Words: Directional derivatives, generalized directional derivatives, nonsmooth analysis, nonsmooth optimality conditions, nonsmooth vector optimization problems.

AMS 2000 Mathematics Subject Classification: 49K27, 90C20, 90C29, 90C30, 90C46.

1. Introduction

In this paper we give an overview on some classical directional derivatives (usual directional derivatives, Dini directional derivatives, Hadamard directional derivatives) and on some basic notions about differentiability (Gâteaux, Fréchet, Hadamard). For simplicity the treatment will be performed in the n-dimensional Euclidean space $\mathbb{R}^n$. We point out some applications of the said concepts to optimization problems; indeed, in several practical optimization problems the involved functions are not everywhere differentiable. Starting from the seventies of the last century, the necessity of studying nonsmooth (i. e. non differentiable) functions, within optimization theory, gave rise to a new mathematical theory, called Nonsmooth Analysis (this term was introduced by the Canadian mathematician F. H. Clarke).

However, we shall not be concerned with "modern" directional derivatives, due to Clarke, Rockafellar, Michel and Penot, Demyanov and Rubinov, etc., for which there is an abundant literature. Only in Section 6 we give some definitions concerning these more recently proposed directional derivatives.

The work is organized as follows. Section 2 gives an overview on some classical directional derivatives and classical differentiability notions.

* Department of Economics and Management, Via S. Felice, 5 - 27100 Pavia, (Italy). E-mail: giorgio.giorgi@unipv.it

Section 3 is concerned with the use of directional derivatives in convex and generalized convex functions. Section 4 gives some applications of directional derivatives to unconstrained optimization problems and constrained optimization problems with a set constraint, whereas Section 5 takes into consideration directional derivatives in constrained optimization problems with functional constraints. Section 6 gives an overview on the axiomatic approach to non-smooth analysis, proposed by K.-H. Elster and J. Thierfelder. The final Section 7 gives some insights on applications of directional derivatives to vector optimization problems.

2. An Overview on Some Classical Directional Derivatives and Classical Differentiability Introduction Notions

In this section we want to recall progressively various notions of "classical" directional derivatives and differentiability for real-valued functions of several real variables. Many authors consider the possibility of "extended-valued functions", i. e. functions which may assume also infinite values; moreover, in Convex Analysis it is customary to consider functions defined on the whole space $\mathbb{R}^n$. In the present basic overview we consider real-valued functions defined on an open set $X \subset \mathbb{R}^n$. More generally, it is possible to consider a set $X \subset \mathbb{R}^n$ and a related point $x^0 \in int(X)$.

Recall that a derivative is some kind of limit of line segments joining points on the graph of a function. The simplest way to take such a limit is along a line segment containing a point $x^0 \in X$. This leads to the basic definition of *one-sided* (or *radial*) *directional derivative.* Let us define a *direction* in $\mathbb{R}^n$ as a vector $v \in \mathbb{R}^n$, $v \neq 0$ (in some cases it is useful to consider a *normalized direction,* i. e. $\|v\| = 1$).

Definition 1. Let $X \subset \mathbb{R}^n$ be an open set, let $x^0 \in X$ and let $f : X \longrightarrow \mathbb{R}$. We say that f has the *right-sided directional derivative* at x^0 *in the direction* v, if the limit

$$\lim_{t \longrightarrow 0^+} \frac{f(x^0 + tv) - f(x^0)}{t}$$

exists (finite or not).

The result of the above limit is denoted by $D^+f(x^0; v)$. Other notations are used in the literature, such as $f^+(x^0; v)$, $D_v^+ f(x^0)$, etc. For $v = 0$, $D^+f(x^0; v)$ is assumed to be zero.

In order that this definition to make sense we implicitly require that there is some $\varepsilon > 0$ such that $0 \leqq t \leqq \varepsilon$ implies that $x^0 + tv \in X$, so that $f(x^0 + tv)$ is defined. This will be implicitly assumed also in the other definitions of directional derivatives given further. Obviously, if X is open or, more generally, if $x^0 \in int(X)$, this is always possible. The next result shows that the set of the directions v for which $D^+f(x^0; v)$ exists is a *cone,* and that $D^+f(x^0; v)$ is *positively homogeneous* on this cone.

Theorem 1. The right-sided directional derivative $D^+f(x^0; v)$ is positively homogeneous of degree one in v. That is, if $D^+f(x^0;v)$ exists, then

$$D^+f(x^0;\alpha v) = \alpha D^+f(x^0;v),\ \forall\alpha \geqq 0.$$

Proof. This follows at once from

$$\frac{f(x^0+t\alpha v)-f(x^0)}{t} = \alpha\frac{f(x^0+\beta v)-f(x^0)}{\beta},$$

where $\beta = t\alpha$, and letting $t,\beta \longrightarrow 0^+$. □

Similarly, we have the following definition.

Definition 2. Let $X \subset \mathbb{R}^n$ be an open set, let $x^0 \in X$ and let $f : X \longrightarrow \mathbb{R}$. We say that f has the *left-sided directional derivative* at x^0 *in the direction* v, if the limit

$$\lim_{t\longrightarrow 0^-}\frac{f(x^0+tv)-f(x^0)}{t} = D^-f(x^0;v)$$

exists (finite or not). For $v = 0$, $D^-f(x^0;v)$ is assumed to be zero.

Finally, we give the following definition.

Definition 3. Under the same assumptions as before, we say that f has a *(bilateral) directional derivative* at $x^0 \in X$ *in the direction* v or that f is *directionally differentiable* at $x^0 \in X$ *in the direction* v, if the limit

$$\lim_{t\longrightarrow 0}\frac{f(x^0+tv)-f(x^0)}{t} = Df(x^0;v)$$

exists.

It is quite immediate to note that $f : X \longrightarrow \mathbb{R}$ has a left-sided directional derivative at x^0 in the dirrection v *if and only if* f has a right-sided directional derivative at x^0 in the direction $(-v)$. In this case it holds

$$D^-f(x^0;v) = -D^+f(x^0;-v).$$

This explains the fact that in applications (mainly in Convex Analysis and in Optimization) only $D^+f(x^0;v)$ is considered. Furthermore, it is immediate that $f : X \longrightarrow \mathbb{R}$ is directionally differentiable at $x^0 \in X$ in the direction v, if and only if

$$D^-f(x^0;v) = D^+f(x^0;v),$$

i. e.

$$-D^+f(x^0;-v) = D^+f(x^0;v),$$

i. e.

$$D^+f(x^0;v) + D^+f(x^0;-v) = 0.$$

It follows also that if $Df(x^0; v)$ exists, then $Df(x^0; \alpha v) = \alpha Df(x^0; v), \ \forall \alpha \in \mathbb{R}$.

Example 1. Let be $y \in \mathbb{R}^n$ and let be $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be defined by $f(x) = \|x - y\|, \forall x \in \mathbb{R}^n$. This function is right-sided directionally differentiable at every point $x^0 \in \mathbb{R}^n$, in every direction v. The reader is invited to verify that

$$D^+ f(x^0; v) = \begin{cases} \|v\|, & \text{if } x^0 = y; \\ \frac{(x^0 - t)^\top v}{\|x^0 - y\|}, & \text{if } x^0 \neq y. \end{cases}$$

Remark 1. Definition 3 is essentially based on a one-dimensional concept: if we put $\varphi(t) = f(x^0 + tv)$, the functioon φ is the restriction of f to the straight line passing through x^0 and with direction v. It is immediate to note that the difference quotient

$$\frac{\varphi(t) - \varphi(0)}{t}$$

coincides with the difference quotient appearing in Definition 3. Hence, the quantity $Df(x^0; v)$, if it exists, coincides with the usual derivative $\varphi'(0)$.

Example 2. Compute the directional derivative of the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by $f(x, y) = x^2 + y^2 + xy + x$ at $x^0 = (1; -1)^\top$, in the direction $v = (1/\sqrt{2}; 1/\sqrt{2})^\top$.

We have $f(x^0 + tv) = \varphi(t) = 2 + \frac{t}{\sqrt{2}} + \frac{3}{2}t^2$. Hence $Df(x^0; v) = \varphi'(0) = \frac{1}{\sqrt{2}} + 3t\,|_{t=0} = \frac{1}{\sqrt{2}}$.

Remark 2. The existence of the directional derivative at x^0 in a certain direction, gives no information on the existence of the directional derivatives at x^0 in other directions.

Example 3. Consider the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} \frac{x^2}{x^2 + y^2}, & \text{if } (x, y) \neq 0; \\ 0, & \text{if } (x, y) = 0. \end{cases}$$

The reader is invited to verify that $Df(0; e^2) = 0$, but $Df(0; e^1)$ does not exist. Here e^1 and e^2 are the two unit coordinate vectors of $\mathbb{R}^2 : e^1 = (1, 0)^\top; e^2 = (0, 1)^\top$.

If $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is directionally differentiable at $x^0 \in int(X)$ in the direction e^i, being e^i the i-th unit coordinate vector of $\mathbb{R}^n$, i. e.

$$e^i = [0, 0, ..., 1, 0, ..., 0]^\top,$$

with 1 as the i-th element, we say that f is *partially differentiable* at x^0 with respect to the i-th variable x_i, and the quantity

$$Df(x^0; e^i)$$

is the *partial derivative* of f at x^0, with respect to the i-th variable x_i, and denoted as

$$\frac{\partial f}{\partial x_i}(x^0) \text{ or also } f_{x_i}(x^0).$$

If $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ admits at x^0 all n partial derivatives, then the vector

$$\left[\frac{\partial f}{\partial x_1}(x^0), ..., \frac{\partial f}{\partial x_n}(x^0)\right]^\top$$

is called the *gradient* of f at x^0 and denoted as $\nabla f(x^0)$.

Note that f may have directional derivatives in *all* nonzero directions at x^0, yet not be continuous at x^0. Note, moreover, that we may not be able to express the directional derivatives of a given function at a point x^0 as a linear function of the components of the direction $v \in \mathbb{R}^n$.

Example 4. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined as

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y}, & y \neq -x^2; \\ 0, & y = -x^2. \end{cases}$$

Observe that f has directional derivatives at $(0, 0)$ in every direction:

$$\frac{f(tx, ty) - f(0, 0)}{t} = \frac{\left(\frac{t^2xy}{t^2x^2+ty}\right)}{t} = \frac{xy}{tx^2 + y}.$$

If $y \neq 0$, the the limit of this expression is x, as $t \longrightarrow 0$, and if $y = 0$, the limit is 0. Thus the directional derivative exists for every direction (x, y), but the function is *not* continuos at $x^0 = (0, 0)$.

Some authors call the quantity $Df(x^0; v)$ the "first variation" in the sense of Lagrange of f at x^0.

Definition 4. Let be $f : X \longrightarrow \mathbb{R}$, with X open subset of $\mathbb{R}^n$, and let $x^0 \in X$. If $D^+f(x^0; v)$ exists for *all* $v \in \mathbb{R}^n$, then f is said to be *weakly Gâteaux differentiable* at x^0 or also *Gâteaux semidifferentiable* at x^0 or also, less frequently, *hemi-differentiable* at x^0 or also *Dini differentiable* at x^0 or, simply, *directionally differentiable* at x^0.

If f is weakly Gâteaux differentiable at x^0 and the function $v \longrightarrow D^+f(x^0; v)$ is *linear,* then f is said to be *Gâteaux differentiable* at x^0. The quantity $D^+f(x^0; v)$ is also called "Gâteaux differential" of f at x^0 or "Gâteaux derivative" of f at x^0.

If f is Gâteaux differentiable at x^0, the Gâteaux differential of f at x^0 is given by $\nabla f(x^0)^\top v$ and we have therefore

$$D^+f(x^0; v) = \nabla f(x^0)^\top v, \quad \forall v \in \mathbb{R}^n.$$

In other words, it holds

$$f(x^0 + tv) = f(x^0) + t\nabla f(x^0)^\top v + o(t),$$

for $t \longrightarrow 0^+$.

Some authors refer to $Df(x^0; v)$ the definitions of weak Gâteaux differentiability and Gâteaux differentiability. This is, for example, the classical approach of Kolmogorov and Fomin (1980), Kantorovič and Akilov (1980), Ortega and Rheinboldt (1970), Vainberg (1956), etc. For what concerns Gâteaux differentiability, the two definitions are equivalent: indeed linearity implies that $D^+f(x^0; v) = -D^+f(x^0; -v) = D^-f(x^0; v) = Df(x^0; v)$. We point out that there is not uniformity of notations and definitions for what concerns the subjects of the present section.

Note that Gâteaux differentiability of f at x^0 *does not imply continuity of* f at x^0.

Example 5. Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$f(x, y) = \begin{cases} \frac{y}{x}(x^2 + y^2), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Then, for $x \neq 0$,

$$\frac{f(tx, ty) - f(0, 0)}{t} = \frac{\left(\frac{ty}{tx} t^2(x^2 + y^2)\right)}{t} = \frac{ty}{x}(x^2 + y^2) \longrightarrow 0.$$

Thus $D^+f(0, v) = 0$ for every v, so f has a Gâteaux derivative at the origin, namely the zero linear map. However, f is not continuous at the origin. Indeed, consider, e. g., $v(\varepsilon) = (\varepsilon^4, \varepsilon)$. Then $v(\varepsilon) \longrightarrow 0$ as $\varepsilon \longrightarrow 0$, but

$$f(v(\varepsilon)) = \frac{\varepsilon}{\varepsilon^4}(\varepsilon^8 + \varepsilon^2) = \varepsilon^5 + \frac{1}{\varepsilon}.$$

Thus $f(v(\varepsilon)) \longrightarrow \infty$ as $\varepsilon \longrightarrow 0^+$ and $f(v(\varepsilon)) \longrightarrow -\infty$ as $\varepsilon \longrightarrow 0^-$, so $\lim_{\varepsilon \longrightarrow 0} f(v(\varepsilon))$ does not exist and in any case f is not continuous at $(0, 0)$.

Note, moreover, that the existence of all partial derivatives do not assure the existence of $D^+f(x^0; v)$ for all directions $v \in \mathbb{R}^n$, nor the Gâteaux differentiability of f at x^0.

Example 6. Consider the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{(x_1)^2 + (x_2)^2}, & \text{if } (x_1, x_2) \neq (0, 0); \\ 0, & \text{if } (x_1, x_2) = (0, 0). \end{cases}$$

This function possesses at $x^0 = (0, 0)$ both partial derivatives (i. e. $Df(x^0; e^1)$ and $Df(x^0; e^2)$), but $D^+f(x^0; e^1 + e^2)$ does not exist.

We now give the classical definition of *Fréchet differentiability.*

Definition 5. Let $f : X \longrightarrow \mathbb{R}$, with $X \subset \mathbb{R}^n$ open set (or, more generally, $X \subset \mathbb{R}^n$ arbitrary and $x^0 \in int(X)$). We say that f is *Fréchet differentiable* at x^0 (or, simply, *differentiable* at x^0), if there exists a vector $a \in \mathbb{R}^n$, depending only from the point x^0, such that

$$\lim_{v \longrightarrow 0} \frac{f(x^0 + v) - f(x^0) - a^\top v}{\|v\|} = 0,$$

i. e.
$$\lim_{x \longrightarrow x^0} \frac{f(x) - f(x^0) - a^\top (x - x^0)}{\|x - x^0\|} = 0.$$

It is well-known that the vector a is unique and that $a = \nabla f(x^0)$. In this case the quantity $\nabla f(x^0)^\top v$ is also called the *Fréchet differential* of f at x^0. The previous definition is also equivalent to the following condition:
$$f(x^0 + v) = f(x^0) + \nabla f(x^0)^\top v + o(\|v\|),$$
for $v \longrightarrow 0 \in \mathbb{R}^n$. Moreover, it can be shown that Fréchet differentiability at x^0 is equivalent to:

For every $\varepsilon > 0$, there exists $\delta > 0$ such that for every v satisfying $\|v\| < \delta$, it holds
$$\left| f(x^0 + v) - f(x^0) - \nabla f(x^0)^\top v \right| \leqq \varepsilon \|v\| .$$

This relation will be useful to introduce (see further) the stronger notion of *strict differentiability.*

Another equivalent condition of Fréchet differentiability at x^0 is the following one (see, e. g., Nashed (1971)).

Theorem 2. Let $f : X \longrightarrow \mathbb{R}$, with $X \subset \mathbb{R}^n$ open set and $x^0 \in X$. Then f is Fréchet differentiable at x^0 if and only if
$$\lim_{t \longrightarrow 0^+} \frac{f(x^0 + tv) - f(x^0)}{t} = \nabla f(x^0)^\top v$$
and the convergence is uniform, with respect to v, for v varying in a bounded set of $\mathbb{R}^n$ (for example on $B = \{v : \|v\| = 1\}$).

The following results are well-known.

Theorem 3. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open, and let $x^0 \in X$. If f is Fréchet differentiable at x^0, then:

i) f is continuous at x^0 (this is an immediate consequence of the definition of Fréchet differentiability).

ii) f is directionally differentiable at x^0 in every direction $v \in \mathbb{R}^n$ and it holds
$$Df(x^0; v) = \nabla f(x^0)^\top v, \ \forall v \in \mathbb{R}^n.$$

Consequently, f is Gâteaux differentiable at x^0 (the vice-versa does not hold).

iii) If $\{t_j\}$, $j \in \mathbb{N}$, is a sequence on the interval $(0, +\infty)$ converging to zero, and $\{v^j\}$, $j \in \mathbb{N}$, a sequence on $\mathbb{R}^n \diagdown \{0\}$, converging to $v \in \mathbb{R}^n$, then
$$\lim_{j \longrightarrow +\infty} \frac{1}{t_j} \left[f(x^0 + t_j v^j) - f(x^0) \right] = \nabla f(x^0)^\top v.$$

iv) There exist two numbers $r > 0$ and $c > 0$ such that $N(x^0, r) \subset X$ and

$$\left| f(x^0 + v) - f(x^0) \right| \leqq c \, \|v\| \, , \ \forall v \in N(0, r).$$

We give now the definition of *strictly differentiable functions,* a property stronger than Fréchet differentiability.

Definition 6. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open, and let $x^0 \in X$. Then f is *strictly differentiable* at x^0 if there exists a vector $a \in \mathbb{R}^n$, which will be the gradient $\nabla f(x^0)$, such that

$$\lim_{\substack{\bar{x} \longrightarrow x^0 \\ \bar{v} \longrightarrow v \\ t \longrightarrow 0^+}} \frac{f(\bar{x} + t\bar{v}) - f(\bar{x})}{t} = \nabla f(x^0)^\top v, \quad \forall v \in \mathbb{R}^n,$$

or. equivalently,

$$\lim_{\substack{\bar{x} \longrightarrow x^0 \\ x \longrightarrow x^0}} \frac{f(\bar{x}) - f(x) - \nabla f(x^0)^\top (\bar{x} - x)}{\|\bar{x} - x\|} = 0, \ \bar{x} \neq x,$$

i. e.

$$f(\bar{x}) = f(x) + \nabla f(x^0)^\top (\bar{x} - x) + o(\|\bar{x} - x\|).$$

Some authors (e. g. Pourciau (1980), Nijenhuis (1974), Ortega and Rheinboldt (1970)) use the term "strongly differentiable". The above conditions are in turn equivalent to the following one (see, e. g., Alexéev, Tikhomirov and Fomine (1982)):

- For every $\varepsilon > 0$ there exists $\delta > 0$ such that for all x^1 and x^2 verifying the inequalities

$$\left\| x^1 - x^0 \right\| < \delta, \quad \left\| x^2 - x^0 \right\| < \delta$$

we have the following inequality:

$$\left| f(x^1) - f(x^2) - \nabla f(x^0)^\top (x^1 - x^2) \right| \leqq \varepsilon \left\| x^1 - x^2 \right\| .$$

The following implications are well known.

$$\begin{aligned} \left\{ \text{Strict differentiability at } x^0 \right\} \ &\Longrightarrow \ \left\{ \text{Fréchet differentiability at } x^0 \right\} \Longrightarrow \\ &\Longrightarrow \ \left\{ \text{Gâteaux differentiability at } x^0 \right\} . \end{aligned}$$

Moreover,

$$\left\{ \text{Strict differentiability at } x^0 \right\} \Longrightarrow \left\{ \text{Continuity of } f \text{ on a neighborhood of } x^0 \right\} .$$

$$\{\text{Fréchet differentiability at } x^0\} \Longrightarrow \{\text{Continuity of } f \text{ at the point } x^0\}.$$

The above implications cannot be reversed. See Example 5 and the following two examples, taken from Alexéev, Tikhomirov and Fomine (1982).

Example 7. Consider the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by

$$f(x_1, x_2) = \begin{cases} 1, & \text{if } x_1 = (x_2)^2,\ x_2 > 0; \\ 0, & \text{at all other points.} \end{cases}$$

This function is Gâteaux differentiable at the origin $(0, 0)$, but it is not continuous at the same point.

Example 8. Consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f(x) = \begin{cases} x^2, & \text{if } x \text{ is rational;} \\ 0, & \text{if } x \text{ is not rational.} \end{cases}$$

At $x^0 = 0$ this function is Fréchet differentiable, but it is not strictly differentiable at the same point, as it is not continuous for $x \neq 0$.

Definition 7. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open, and let $x^0 \in X$.

i) The function f is said to be of $\mathcal{C}^1$*-class at* x^0, and denoted by $f \in \mathcal{C}^1(x^0)$, if its gradient $\nabla f(x)$ exists in a neighborhood of x^0 and is continuous at x^0.

ii) The function f is said to be of $\mathcal{C}^1$*-class on* X, and denoted by $f \in \mathcal{C}^1(X)$, if $\nabla f(x)$ is continuous for all $x \in X$. In this case f is also said to be *continuously differentiable on* X.

The following sufficient condition for Fréchet differentiability is well known.

Theorem 4. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open, and let $x^0 \in X$. If $f \in \mathcal{C}^1(x^0)$, then f is Fréchet differentiable at x^0 and also strictly differentiable at x^0. If $f \in \mathcal{C}^1(X)$, then f is Fréchet differentiable on X.

The second part of the previous theorem can be made more precise, on the ground of the following result (see, e. g., Rockafellar and Wets (2009)).

Theorem 5. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open. Then $f \in \mathcal{C}^1(X)$ if and only if f is strictly differentiable on X.

Another classical notion, useful for further considerations, is the definition of *Lipschitz continuous functions* and *locally Lipschitz continuous functions.*

Definition 8. Let $X \subset \mathbb{R}^n$ be a nonempty set and $f : X \longrightarrow \mathbb{R}$. The function f is said to be *Lipschitz* over X (or *Lipschitz continuous* over X) if there exists a real number $k \geqq 0$ such that, for every $x^1, x^2 \in X$ we have

$$\left| f(x^1) - f(x^2) \right| \leqq k \left\| x^1 - x^2 \right\|. \tag{1}$$

The smallest constant k for which the previous relation holds is called "the Lipschitz constant" or "the Lipschitz rank". If $k = 1$, then f is said to be *non-expansive* and if $k < 1$, then f is said to be a *contraction.*

Note that if f is Lipschitz on X, then it is (uniformly) continuous on X, but the converse is not true: take, e. g. the continuous function $f(x) = \sqrt[3]{x}$, $x \in \mathbb{R}$; with $x_2 = 0$ we see that there is no constant $k \geqq 0$ satisfying (1). To understant the meaning of (1), rewrite it as follows

$$\frac{|f(x^1) - f(x^2)|}{\|x^1 - x^2\|} \leqq k, \quad \forall x^1 \neq x^2 \in X.$$

Hence. a function is Lipschitz on the set $X \subset \mathbb{R}^n$ if and only if all its difference quotients are bounded.

Example 9.

i) The function $f(x) = \|x\|$, $x \in \mathbb{R}^n$, is Lipschitz on $\mathbb{R}^n$, with $k = 1$.

ii) The function $f(x) = \|x\|^2$ is not Lipschitz on the whole space $\mathbb{R}^n$. Indeed, by choosing $x^2 = 0$, we have

$$\left\|x^1\right\|^2 \leqq k \left\|x^1\right\|$$

which holds only if $\|x^1\| \leqq k$.

A sufficient condition for f to be Lipschitz on a set contained in its domain, is given by the following proposition.

Theorem 6. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, with X convex set. If f is differentiable on X and if all its partial derivatives are bounded on X, then f is Lipschitz on X. Moreover, for every $M \geqq 0$ such that

$$\left|\frac{\partial f}{\partial x_i}(x)\right| \leqq M, \quad \forall x \in X, \ \forall i = 1, ..., n,$$

then relation (1) holds with $k = \sqrt{n}M$.

Definition 9. Let $X \subset \mathbb{R}^n$ be a nonempty open set and $f : X \longrightarrow \mathbb{R}$. For a point $x^0 \in X$, if there exist a neighborhood $N(x^0)$ of x^0 a nonnegative number k such that

$$\left|f(x^1) - f(x^2)\right| \leqq k \left\|x^1 - x^2\right\|, \quad \forall x^1, x^2 \in N(x^0),$$

then f is said to be *locally Lipschitz at* x^0 or *Lipschitz near* x^0 or *Lipschitz around* x^0, with constant k.

We say that f is locally Lipschitz on X is f is locally Lipschitz at each $x \in X$.

Thus a function which is locally Lipschitz at a point means that the function satisfies the Lipschitz condition in a neighborhood of that point. However, it is important to note that the value of the Lipschitz constant k in general could change as we change the point. Obviously we have the implication

$$f \text{ Lipschitz on } X \subset \mathbb{R}^n \ (X \text{ open}) \Longrightarrow$$

$$\Longrightarrow f \text{ locally Lipschitz at each point of } X,$$

but the converse is not in general true. If however a locally Lipschitz function has a uniform Lipschitz constant k at every point $x^0 \in X$, then f is Lipschitz on X in the sense of Definition 8. A sufficient condition for f to be locally Lipschitz at a point x^0 of its domain is given by the following proposition.

Theorem 7. If a function $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuously differentiable (i. e. of $\mathcal{C}^1$ class) in a neighborhood of $x^0 \in int(X)$, then f is locally Lipschitz at x^0.

Proof. Continuous differentiability around x^0 means that all n partial derivatives of f are continuous on a neighborhood of x^0. It follows that there exist constants $\varepsilon > 0$ and $k \geqq 0$ such that

$$\|\nabla f(x)\| \leqq k, \text{ for all } x \in N(x^0, \varepsilon).$$

Suppose that $x^1, x^2 \in N(x^0, \varepsilon)$. Then, by the classical Mean-Value Theorem, there is $z \in (x^1, x^2) \subset N(x^0, \varepsilon)$ such that

$$f(x^1) - f(x^2) = \nabla f(z)^\top (x^1 - x^2).$$

We now have

$$\left|f(x^1) - f(x^2)\right| \leqq \|\nabla f(z)\| \left\|x^1 - x^2\right\| \leqq k \left\|x^1 - x^2\right\|,$$

i. e. f is Lipschitz continuous at x^0. □

Theorem 7. can be weakened: it can be proved (see, e. g., Nijenhuis (1974)) that if f is *strictly differentiable at* $x^0 \in int(X)$, then f is locally Lipschitz at x^0.

Other classical directional derivatives used in optimization are the *Dini directional derivatives,* introduced by U. Dini (1878) for real functions of one variable. For various considerations and applications of Dini directional derivatives, see, e. g. Bector, Bhatia and Jain (1993), Crouzeix (1981, 1988), Diewert (1981), Giorgi and Komlosi (1993a, b, 1995), Glover (1984), Komlosi (1983, 1995, 2005).

Definition 10. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and $x^0 \in X$. The quantities

$$f^{D^+}(x^0; v) = \limsup_{t \longrightarrow 0^+} \frac{f(x^0 + tv) - f(x^0)}{t}; \tag{2}$$

$$f_{D^+}(x^0; v) = \liminf_{t \longrightarrow 0^+} \frac{f(x^0 + tv) - f(x^0)}{t}, \tag{3}$$

are, respectively, called the *(right-sided) upper Dini directional derivative* of f at x^0 in the direction $v \in \mathbb{R}^n$ and the *(right-sided) lower Dini directional derivative* of f at x^0 in the direction $v \in \mathbb{R}^n$.

Obviously, it is possible to define also the left-sided Dini directional derivatives $f^{D_-}(x^0; v)$ and $f_{D^-}(x^0; v)$.

If the limit

$$\lim_{t \longrightarrow 0^+} \frac{f(x^0 + tv) - f(x^0)}{t} = D^+ f(x^0; v) \tag{4}$$

exists, then

$$f^{D^+}(x^0; v) = f_{D^+}(x^0; v) = D^+ f(x^0; v).$$

Note that the limits in (2), (3) always exist (finite or not), whereas the limit in (4) not always exists. It is easy to see that, for any $v \in \mathbb{R}^n$,

$$f^{D^+}(x^0; v) \geqq f_{D^+}(x^0; v).$$

We must also note that Dini derivatives are positively homogeneous (of degree one) in their second argument, i. e. with respect to the direction v and for all $\lambda > 0$ we have

$$f^{D^+}(x^0; \lambda v) = \lambda f^{D^+}(x^0; v)$$

and

$$f_{D^+}(x^0; \lambda v) = \lambda f_{D^+}(x^0; v).$$

Other "classical" directional derivatives were introduced by the French mathematician Jacques Hadamard.

Definition 11. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and $x^0 \in X$. The quantities

$$f^{H^+}(x^0; v) = \limsup_{t \longrightarrow 0^+,\ \bar{v} \longrightarrow v} \frac{f(x^0 + t\bar{v}) - f(x^0)}{t};$$

$$f_{H^+}(x^0; v) = \liminf_{t \longrightarrow 0^+,\ \bar{v} \longrightarrow v} \frac{f(x^0 + t\bar{v}) - f(x^0)}{t}$$

are called, respectively, the *Hadamard (right-sided) upper directional derivative* of f at the point x^0 in the direction $v \in \mathbb{R}^n$ and the *Hadamard (right-sided) lower directional derivative* of f at the point x^0 in the direction $v \in \mathbb{R}^n$.

Some authors call the above quantities "Dini-Hadamard directional derivatives"; Aubin and Cellina (1984) speak of "contingent derivatives", whereas Penot (1978) uses the term "semiderivatives". See also further. We are not sure that there are not other denominations in the literature. Also in this case it is obviously possible to define Hadamard left-sided directional derivatives $f^{H^-}(x^0; v)$ and $f^{H^-}(x^0; v)$. Note that the limits of Definition 11 always exist but are not nwecessarily finite. Note also that in the one-dimensional case, i. e. $\mathbb{R}^n = \mathbb{R}$, the

Hadamard directional derivatives coincide with the corresponding Dini directional derivatives. If we have that the limit

$$\lim_{t \longrightarrow 0^+,\, \bar{v} \longrightarrow v} \frac{f(x^0 + t\bar{v}) - f(x^0)}{t} \equiv D^H f(x^0; v)$$

exists, then f is called *Hadamard directionally differentiable at* x^0 *in the direction* $v \in \mathbb{R}^n$ (many authors require that the above limit must exist finite). Some authors (e. g. Delfour (2020), Rockafellar and Wets (2009)) speak in this case of the *semiderivative* of f *at* x^0 *in the direction* $v \in \mathbb{R}^n$ and say that f is *semidifferentiable* at x^0 *in the direction* $v \in \mathbb{R}^n$. If this holds for *every* $v \in \mathbb{R}^n$, we say that f is *Hadamard directionally differentiable at* x^0 or that f is *semidifferentiable at* x^0. If $D^H f(x^0; 0)$ exists, then it holds $D^H f(x^0; 0) = 0$.

It is possible to show (see, e. g., Shapiro (1990)) that the above definition of Hadamard directional differentiability is equivalent to the following proposition:

- For any mapping $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{R}^n$ such that

$$\varphi(0) = x^0 \text{ and such that } \frac{\varphi(t) - \varphi(0)}{t} \longrightarrow v \text{ for } t \longrightarrow 0^+,$$

the limit

$$\lim_{t \longrightarrow 0^+} \frac{f(\varphi(t)) - f(x^0)}{t}$$

does exist. Then it holds

$$D^H f(x^0; v) = \lim_{t \longrightarrow 0^+} \frac{f(\varphi(t)) - f(x^0)}{t}.$$

The above characterization gives the Hadamard directional derivative $D^H f(x^0; v)$ along a curve tangential to v. Indeed, some authors speak also, for the case under examination, of *tangential directional derivative.* Other authors (Craven (1986), Craven and Mond (1979)) speak of *arcwise directionally differentiable functions.* Obviously the Hadamard directional derivative $D^H f(x^0; v)$ can also be expressed in terms of sequences:

$$D^H f(x^0; v) = \lim_{n \longrightarrow \infty} \frac{f(x^0 + t_n v^n) - f(x^0)}{t_n},$$

where $\{v^n\} \subset \mathbb{R}^n$ and $\{t_n\} \subset \mathbb{R}_+$ are any sequences such that $v^n \longrightarrow v$ and $t_n \longrightarrow 0^+$.

Robinson (1987) introduced the concept of *Bouligand differentiability* (he called this property "B-differentiability"). However, for real-valued functions on $\mathbb{R}^n$, this property is equivalent to Hadamard directional differentiability, as defined before (see Rockafellar and Wets (2009), page 294).

Note that if $D^H f(x^0; v)$ exists, then also the directional derivative $D^+ f(x^0; v)$ exists and

$$D^H f(x^0; v) = D^+ f(x^0; v).$$

The converse is not necessarily true.

Example 10. Consider the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

$$f(x_1, x_2) = \begin{cases} 0, & \text{if either } x_2 \geqq (x_1)^2 \text{ or } x_2 \leqq 0; \\ 1, & \text{in all other cases.} \end{cases}$$

Let be $x^0 = (0, 0)^\top$. We have $D^+ f(x^0; v) = 0$, $\forall v \in \mathbb{R}^2$. Moreover, we have $D^H f(x^0; v) = 0$, $\forall v = (v_1, v_2)^\top \in \mathbb{R}^2$, with $v_2 \neq 0$. $D^H f(x^0; v)$ does not exist for $v = (v_1, 0)^\top$, $v_1 \in \mathbb{R}$.

An important property of locally Lipschitz functions, which guarantees the converse of the above result, is contained in the following proposition.

Theorem 8. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and $x^0 \in X$; if f is locally Lipschitz at x^0 and $D^+ f(x^0; v)$ exists, then f is also Hadamard directionally differentiable at x^0 in the direction v and it holds

$$D^+ f(x^0; v) = D^H f(x^0; v).$$

Also the Hadamard directional derivatives are positively homogeneous of degree one with respect to the direction $v \in \mathbb{R}^n$.

Definition 12. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and $x^0 \in X$. We say that f is *Hadamard differentiable at* x^0 if $D^H f(x^0; v)$ exists (finite) for all $v \in \mathbb{R}^n$ and this quantity depends *linearly* on $v \in \mathbb{R}^n$.

In the above case we have

$$D^H f(x^0; v) = \nabla f(x^0)^\top v, \quad \forall v \in \mathbb{R}^n,$$

and as a consequence, in finite-dimensional spaces, such as $\mathbb{R}^n$, we have that Hadamard differentiability coincides with Fréchet differentiability.

Theorem 9. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and $x^0 \in X$. Then f is Fréchet differentiable at x^0 if and only if, for all $v \in \mathbb{R}^n$, it holds

$$\lim_{t \longrightarrow 0^+,\ \bar{v} \longrightarrow v} \frac{f(x^0 + t\bar{v}) - f(x^0)}{t} = \nabla f(x^0)^\top v.$$

We have to note that some authors call "Hadamard differentiable" what we have called "Hadamard directionally differentiable", perhaps because in finite-dimensional spaces Hadamard differentiability coincides with Fréchet differentiability. Also for Hadamard directional derivatives there is not uniformity of notations and definitions. Often, in the literature, instead of the limits written in the form

$$t \longrightarrow 0^+,\ \bar{v} \longrightarrow v$$

(for the Hadamard directional derivatives), the same limits are taken in the form

$$(t, \bar{v}) \longrightarrow (0^+, v).$$

The two forms not necessarily coincide. See the paper of F. Giannessi (1995).

Theorem 10. If $D^H f(x^0; \cdot)$ exists (finite) in a neighborhood of $\bar{v} \in \mathbb{R}^n$, then $D^H f(x^0; \cdot)$ is continuous at x^0.

Proof. By assumption there exists $\rho_0 > 0$ such that $D^H f(x^0; v)$ exists for all $v \in B(\bar{v}; \rho_0)$. Let be given $\varepsilon > 0$. Then there exists $\rho \in (0, \rho_0)$ such that for all $t \in (0, \rho)$ and all $v \in B(\bar{v}, \rho)$ it holds

$$\left| \frac{f(x^0 + tv) - f(x^0)}{t} - D^H f(x^0; \bar{v}) \right| \leqq \varepsilon.$$

For $t \longrightarrow 0^+$ it follows, for all $v \in B(\bar{v}, \rho)$,

$$\left| D^+ f(x^0; v) - D^H f(x^0; v) \right| \leqq \varepsilon.$$

Since $D^+ f(x^0; v) = D^H f(x^0; v)$, the thesis follows. □

We have seen (Theorem 8) that if f is locally Lipschitz at x^0, then the existence of $D^+ f(x^0; v)$ implies the existence of $D^H f(x^0; v)$ and the equality $D^+ f(x^0; v) = D^H f(x^0; v)$. As a consequence we have the following important result.

Theorem 11. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open, $x^0 \in X$ and f locally Lipschitz at x^0. Then the following properties are equivalent.

(a) f is Gâteaux differentiable at x^0;

(b) f is Hadamard differentiable at x^0.

We recall that, being X finite-dimensional, under the assumption of Theorem 11 we have also the equivalence between Gâteaux differentiability and Fréchet differentiability. Obviously, for $f : \mathbb{R} \longrightarrow \mathbb{R}$ Fréchet, Hadamard and Gâteaux differentiability at x^0 are equivalent concepts and coincide with the usual classical derivative of f at x^0. It is worth noting, furthermore, that if $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and $x^0 \in X$, has a (finite) directional Hadamard derivative at x^0 in all directions $v \in \mathbb{R}^n$, i. e. $D^H f(x^0; v)$ exists (finite) for all $v \in \mathbb{R}^n$, then f is continuous at x^0, but not necessarily locally Lipschitz continuous at x^0 (see, e. g., Demyanov and Rubinov (1995), Delfour (2020)).

We recall that a Gâteaux differentiable function is not necessarily continuous (Example 5). Obviously, if $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is Hadamard differentiable at $x^0 \in X$, X open, i. e. it is Fréchet differentiable at x^0, then it is continuous at x^0.

For directionally differentiable functions and Hadamard directionally differentiable functions there exist calculus rules (sum, difference, product and quotient): if f_1 and f_2 are, for example, Hadamard directionally differentiable at a point x, then their sum, difference, product and quotient (if $f_2(x) \neq 0$) are also Hadamard directionally differentiable at x and the following formulas hold.

$$D^H(f_1 \pm f_2)(x; v) = D^H f_1(x; v) \pm D^H f_2(x; v);$$

$$D^H(f_1 f_2)(x; v) = f_1(x) D^H f_2(x; v) + f_2(x) D^H f_1(x; v);$$

$$D^H\left(\frac{f_1}{f_2}\right)(x; v) = -\frac{1}{(f_2(x))^2}\left[f_1(x) D^H f_2(x; v) - f_2(x) D^H f_1(x; v)\right].$$

Unfortunately, formulas similar to these ones are no longer valid for Dini and Hadamard upper and lower directional derivatives.

3. Directional Derivatives in Convex and Generalized Convex Functions

Directional derivatives play an important role in Convex Analysis and Optimization Theory. In the present section we give an overview of the main properties of convex and generalized convex functions with regard to directional derivatives. We begin with convex functions; for the related proofs, see, e. g., the fundamental book of Rockafellar (1970) and the books of Bagirov, Karmitsa and Mäkelä (2014), Bertsekas (2009), Bertsekas, Nedic and Ozdaglar (2003), Borwein and Lewis (2000), Dhara and Dutta (2012), Durea and Strugariu (2014), Giorgi, Guerraggio and Thierfelder (2004), Hiriart-Urruty and Lemarechal (1993), Holmes (1972, 1975), Roberts and Varberg (1973), Shimitzu, Ishizuka and Bard (1997).

Also in the present section we consider real-valued functions defined on X, a subset of $\mathbb{R}^n$, even if it is customary in modern Convex Analysis to consider extended-valued functions, defined on the whole space $\mathbb{R}^n$ and assuming also infinite values.

We recall that $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and convex, f differentiable on X, is convex on X if and only if, for every $x, x^0 \in X$

$$f(x) - f(x^0) \geqq \nabla f(x^0)^\top (x - x^0).$$

If f is not differentiable on the open convex set X, then f is convex on X if and only if there exists $u^0 \in \mathbb{R}^n$ such that, for every $x, x^0 \in X$

$$f(x) - f(x^0) \geqq (u^0)^\top (x - x^0).$$

These last considerations allow us to give the following basic definition.

Definition 13. Let $X \subset \mathbb{R}^n$ be a convex set and $f : X \longrightarrow \mathbb{R}$ a convex function on X. A vector $s \in \mathbb{R}^n$ is called a *subgradient* of f at $x^0 \in X$, if, for every $x \in X$ it holds

$$f(x) - f(x^0) \geqq s^\top (x - x^0).$$

The set of all subgradients of f at x^0 is called the *subdifferential* of f at x^0 and denoted by $\partial f(x^0)$.

Example 11. From the previous definition it follows at once that, for $f(x) = |x|$, $x \in \mathbb{R}$, $\partial f(0) = [-1, 1]$.

The subdifferential may be also an empty set: non-subdifferentiability can however occur on the boundary of the domain of f; consider, e. g., $f : [0,1] \longrightarrow [0,1]$, with $f(x) = -\sqrt{x}$. Then f is clearly convex, but $\partial f(0) = \varnothing$.

If the set $\partial f(x^0)$ is nonempty, we say that f is *subdifferentiable* at x^0. We see now the main properties of convex functions with regard to directional derivatives and subdifferentials. We consider $X \subset \mathbb{R}^n$ as an *open* convex set, but all results hold also for $X \subset \mathbb{R}^n$ convex set and $x^0 \in int(X)$.

Theorem 12. If $X \subset \mathbb{R}^n$ is an open convex set, $f : X \longrightarrow \mathbb{R}$ is convex and $x^0 \in X$, then f admits a finite right-sided directional derivative at x^0, $D^+ f(x^0; v)$, and a finite left-sided directional derivative $D^- f(x^0; v)$, for any direction $v \in \mathbb{R}^n$ and it holds

$$D^- f(x^0; v) \leqq D^+ f(x^0; v), \quad \forall v \in \mathbb{R}^n.$$

Theorem 13. If $X \subset \mathbb{R}^n$ is an open convex set, $f : X \longrightarrow \mathbb{R}$ is convex and $x^0 \in X$, then $D^+ f(x^0; \cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a *sublinear function,* i. e.

$$D^+ f(x^0; \lambda v) = \lambda D^+ f(x^0; v), \ \forall v \in \mathbb{R}^n, \ \forall \lambda > 0;$$

$$D^+ f(x^0; v^1 + v^2) \leqq D^+ f(x^0; v^1) + D^+ f(x^0; v^2), \quad \forall v^1, v^2 \in \mathbb{R}^n.$$

Sublinear functions are a class of convex functions: sublinear functions are convex and a convex function which is positively homogeneous (of degree 1) is a sublinear function.

Theorem 14. If $X \subset \mathbb{R}^n$ is an open convex set, $f : X \longrightarrow \mathbb{R}$ is convex and $x^0 \in X$, then

$$D^+ f(x^0; x - x^0) \leqq f(x) - f(x^0), \ \forall x \in X.$$

Remark 3. The above result is indeed a necessary and sufficient condition for a convexity of a directionally differentiable function: see, e. g., Fenchel (1953), page 81, Roberts and Varberg (1973), page 12, Diewert, Avriel and Zang (1981), page 410:

- Let $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a (right-sided) directionally differentiable function on the open convex set X. Then f is convex on X if and only if, for every $x, x^0 \in X$,

$$D^+ f(x^0; x - x^0) \leqq f(x) - f(x^0).$$

Theorem 15. If $X \subset \mathbb{R}^n$ is an open convex set, $f : X \longrightarrow \mathbb{R}$ is convex and $x^0 \in X$, then $s \in \mathbb{R}^n$ is a subgradient of f at x^0 if and only if

$$D^+ f(x^0; v) \geqq s^\top v, \ \forall v \in \mathbb{R}^n.$$

Theorem 16. If $X \subset \mathbb{R}^n$ is an open convex set, $f : X \longrightarrow \mathbb{R}$ is convex and $x^0 \in X$, then it results

$$D^+ f(x^0; v) = \max \left\{ s^\top v : s \in \partial f(x^0) \right\}, \quad \forall v \in \mathbb{R}^n.$$

The next reult has already been stated at the beginning of the present section.

Theorem 17. Let $X \subset \mathbb{R}^n$ be an open convex set; then $f : X \longrightarrow \mathbb{R}$ is convex on X if and only if for every $x^0 \in X$ there exists $s \in \mathbb{R}^n$ such that

$$s^\top (x - x^0) \leqq f(x) - f(x^0), \quad \forall x \in X.$$

In other words: $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is convex on the open convex set X if and only if it is subdifferentiable at every point $x^0 \in X$.

The following result treats the differentiability properties of a convex function on an open convex set $X \subset \mathbb{R}^n$ (or, more generally, on an arbitrary convex set $X \subset \mathbb{R}^n$, with $x^0 \in int(X)$).

Theorem 18. Let $X \subset \mathbb{R}^n$ be an open convex set, $x^0 \in X$ and $f : X \longrightarrow \mathbb{R}$ be convex Then the following assertions are equivalent.

i) f is Fréchet differentiable at x^0;

ii) f is directionally differentiable at x^0 with respect to every direction $v \in \mathbb{R}^n$;

iii) f is Gâteaux differentiable at x^0;

iv) f admits all partial derivatives at x^0, with respect to the variables $x_1, ..., x_n$.

v) f is $\mathcal{C}^1(x^0)$.

Finally, we point out an important property of convex functions.

Theorem 19. If $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is convex on the open convex set X, then f is locally Lipschitz on X.

A remarkable result due to H. Rademacher (1919) says that a function which is locally Lipschitz on an open set $X \subset \mathbb{R}^n$ is differentiable almost everywhere on X. On the grounds of Theorem 19 it holds that a convex function over an open convex set $X \subset \mathbb{R}^n$ is differentiable on X excepts for points of zero (Lebesgue) measure.

Remark 4. We have seen (Theorem 18) that for convex functions defined on an open convex set $X \subset \mathbb{R}^n$, Gâteaux differentiability implies Fréchet differentiability (the reverse implication is obviously always true). This holds (Theorem 11) also for locally Lipschitz functions on an open set $X \subset \mathbb{R}^n$. Another condition which assures that Gâteaux differentiability implies Fréchet differentiability has been established by Chabrillac and Crouzeix (1987). Given $I_1, I_2, ..., I_n$, n open intervals in $\mathbb{R}$ and given a real-valued function f defined on $D = I_1 \times I_2 \times ... \times I_n$, we say that f is *nondecreasing* on D if $f(x) \leqq f(y)$ whenever $x, y \in D$ and $x_i \leqq y_i$, $i = 1, ..., n$. The said authors prove the following interesting result.

• Let $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a nondecreasing function on D. If f is Gâteaux differentiable at $x^0 \in D$, then f is actually Fréchet differentiable at x^0.

Directional derivatives have been used also in studying generalized (nonsmooth) convex functions. We recall that Mangasarian (1965, 1969) has introduced, for differentiable functions, the notion of *pseudoconvex functions,* a class of generalized convex functions lying between convex functions and quasiconvex functions.

Definition 14. Let $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be differentiable on the open convex set X; then f is *pseudoconvex on* X if, for all $x^0, x \in X$,

$$f(x) < f(x^0) \Longrightarrow \nabla f(x^0)^\top (x - x^0) < 0,$$

or equivalently,

$$\nabla f(x^0)^\top (x - x^0) \geqq 0 \Longrightarrow f(x) \geqq f(x^0).$$

Subsequently, Ortega and Rheinboldt (1970) and Thompson and Parke (1973) have given a definition pf pseudoconvex functions without differentiability assumptions.

Definition 15. Let be given $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, with X nonempty convex set. Then f is *pseudoconvex on* X if for every $x, y \in X$ such that $f(y) < f(x)$ there exist a number $c > 0$ and a number $\alpha \in (0, 1]$ such that

$$f((1-a)x + ay) \leqq f(x) - ac, \ \forall a \in (0, \alpha).$$

We now prove that, under Fréchet differentiability assumptions, the characterization of Definition 15 coincides with the one of Definition 14.

Theorem 20. Let $X \subset \mathbb{R}^n$ be an open convex set and let $f : X \longrightarrow \mathbb{R}$ be differentiable on X. Then if f is pseudoconvex on X, following the characterization of Definition 15, then f is pseudoconvex on X, following the characterization of Definition 14, and vice-versa.

Proof.

a) Let be $x^0, x \in X$ such that $f(x) < f(x^0)$. Then, there exists a number $c > 0$ and a number $\alpha \in (0, 1]$ such that

$$f((1-t)x^0 + tx) \leqq f(x^0) - tc, \ \forall t \in (0, \alpha).$$

Consequently, we have

$$\frac{1}{t}\left[f(x^0 + t(x - x^0)) - f(x^0)\right] \leqq -c, \ \forall t \in (0, \alpha).$$

Taking in the above inequality, the limit for $t \longrightarrow 0^+$, we obtain $\nabla f(x^0)^\top (x - x^0) \leqq -c$, i. e. the characterization of Definition 14.

b) Let be $x, y \in X$ such that $f(y) < f(x)$. On the ground of the assumptions we have $\nabla f(x)^\top (y-x) < 0$. Let us choose $c = -\nabla f(x)^\top (y-x)/2$. From the equality

$$\lim_{a \longrightarrow 0^+} \frac{1}{a} [f(x + a(y-x)) - f(x)] = \nabla f(x)^\top (y-x),$$

we have that there exists a number $\alpha \in (0, 1]$ such that

$$\left| \frac{1}{a} [f(x + a(y-x)) - f(x)] - \nabla f(x)^\top (y-x) \right| < 0$$

for every $a \in (0, \alpha)$. Hence we have

$$\frac{1}{a} [f(x + a(y-x)) - f(x)] < -c,$$

for every $a \in (0, \alpha)$ and therefore f is pseudoconvex following the characterization of Definition 15. □

We invite the reader to prove that if $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is convex on the nonempty convex set X, then f is pseudoconvex, following the characterization of Definition 15.

The above notions have been studied also by means of *Dini directional derivatives.* The pioneering paper is the one of Diewert (1981), followd by the paper of Komlosi (1983). See also Giorgi and Komlosi (1993a, b, 1995), Komlosi (1995, 2005). It is possible to use *lower* Dini directional derivatives and *upper* Dini directional derivatives.

Definition 16. Let be given $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, where X is open and convex. Then f is a *lower Dini-pseudoconvex function on* X (LDPC) if, for all $x, x^0 \in X$, we have

$$f(x) < f(x^0) \Longrightarrow f_{D^+}(x^0; x - x^0) < 0,$$

or equivalently

$$f_{D^+}(x^0; x - x^0) \geqq 0 \Longrightarrow f(x) \geqq f(x^0).$$

Definition 17. Let be given $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, where X is open and convex. Then f is an *upper Dini-pseudoconvex function on* X (UDPC) if, for all $x, x^0 \in X$, we have

$$f(x) < f(x^0) \Longrightarrow f^{D^+}(x^0; x - x^0) < 0,$$

or equivalently

$$f^{D^+}(x^0; x - x^0) \geqq 0 \Longrightarrow f(x) \geqq f(x^0).$$

It is obvious that

$$(UDPC) \Longrightarrow (LDPC).$$

Diewert (1981) has proved that the reverse implication does not hold.

Also *quasiconvex functions* have been studied in terms of Dini directional derivatives. Also for this types of studies, the pioneering paper is the one by Diewert (1981). See also Crouzeix (1981, 1998, 2005), Komlosi (2005). We recall that $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, with X nonempty convex set, is *quasiconvex on* X if

$$f(x) \leqq f(y) \Longrightarrow f(\lambda x + (1-\lambda)y) \leqq f(y), \quad \forall x, y \in X, \ \forall \lambda \in [0,1] .$$

The following classical result, essentially due to Arrow and Enthoven (1961), is well known.

Theorem 21. Let $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be differentiable on the open convex set X. Then f is quasiconvex on X if and only if for all $x, y \in X$

$$f(x) \leqq f(y) \Longrightarrow \nabla f(y)^\top (x-y) \leqq 0,$$

or equivalently,

$$\nabla f(y)^\top (x-y) > 0 \Longrightarrow f(x) > f(y).$$

It is quite easy to prove that if $f(x)$ is quasiconvex on the open convex set $X \subset \mathbb{R}^n$, then for all $x, x^0 \in X$ the following implications hold

$$f(x) \leqq f(x^0) \Longrightarrow f^{D^+}(x^0; x - x^0) \leqq 0; \tag{5}$$

$$f(x) \leqq f(x^0) \Longrightarrow f_{D^+}(x^0; x - x^0) \leqq 0. \tag{6}$$

We call *upper Dini-quasiconvex* (UDQC) a function satisfying (5) and *lower Dini-quasiconvex* (LDQC) a function satisfying (6). For a function f defined on an open convex set $X \subset \mathbb{R}^n$ we have

$$\{f \text{ convex}\} \Longrightarrow \{f \text{ quasiconvex}\} \Longrightarrow \{f \text{ (UDQC)}\} \Longrightarrow \{f \text{ (LDQC)}\} .$$

In general the above implications cannot be reversed. It is well known that for differentiable functions pseudoconvexity implies quasiconvexity. This is no longer true for Dini-pseudoconvex functions.

Example 12. Consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f(t) = \begin{cases} 0, & \text{if } 0 < |t| \leqq 1; \\ 1, & \text{if } t = 0. \end{cases}$$

This function is (UDQC) and (LDQC) and also (LDPC) but fails to be quasiconvex.

Other interesting results on directional derivatives applied to quasiconvex functions are due to Crouzeix. See Crouzeix (2005) and the bibliographical references there quoted.

Theorem 22. (Crouzeix). Assume that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is quasiconvex; then $f^{D^+}(x^0; \cdot)$ is quasiconvex.

The lower Dini directional derivative $f_{D^+}(x^0;\cdot)$ is not necessarily quasiconvex, even if f is quasiconvex.

Theorem 23. (Crouzeix). Assume that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is quasiconvex and *directionally differentiable* at x^0, i. e. $Df(x^0; v)$ exists for all $v \in \mathbb{R}^n$. Then f is also *Gâteaux differentiable* at x^0, i. e.

$$Df(x^0; v) = \nabla f(x^0)^\top v, \quad \forall v \in \mathbb{R}^n.$$

We have seen that Gâteaux differentiability and Fréchet differentiability coincide for locally Lipschitz functions $f : X \longrightarrow \mathbb{R}$, X open set of $\mathbb{R}^n$, and for convex functions on the open convex set X. A nice result of Crouzeix states that the same property holds for quasiconvex functions.

Theorem 24. (Crouzeix). Assume that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is quasiconvex on the open convex set $X \subset \mathbb{R}^n$. Then, if f is Gâteaux differentiable at $x^0 \in X$, then f is Fréchet differentiable at x^0.

On the ground of Theorems 23 and 24, we can summarize as follows the related results.

- Let $f : X \longrightarrow \mathbb{R}$ be quasiconvex on the open convex set $X \subset \mathbb{R}^n$. Then it holds:

$$\left\{ f \text{ Fréchet differentiable at } x^0 \in X \right\} \iff$$

$$\iff \left\{ f \text{ Gâteaux differentiable at } x^0 \in X \right\} \iff$$

$$\iff \left\{ f \text{ directionally differentiable (in the bilateral sense) at } x^0 \in X \right\}.$$

Another approach in applying directional derivatives in the study of convex and generalized convex functions is suggested by Ewing (1977) and Kaul and Kaur (1982a).

Definition 18. A set $X \subset \mathbb{R}^n$ is said to be *locally star-shaped at* $x^0 \in X$ if corresponding to x^0 and each $x \in X$, there exists a maximum positive number $a(x^0, x) \leqq 1$ such that

$$(1 - \lambda)x^0 + \lambda x \in X, \; 0 < \lambda < a(x^0, x).$$

If $a(x^0, x) = 1$ for each $x \in X$, then X is said to be *star-shaped at* x^0 and if X is star-shaped at each $x^0 \in X$, then X is a convex set. It may be noted that there exist sets which are locally star-shaped at each of their points, but which are not convex. For example, the set $\Gamma = \{x \in \mathbb{R} : x^3 \leqq 1, \; x \neq 0\}$ is locally star-shaped at each of its points but it is not convex.

Definition 19. A function $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be *semilocally convex at* $x^0 \in X$ if X is locally star-shaped at x^0 and if, corresponding to x^0 and each $x \in X$, there exists a positive number $d(x^0, x) \leqq a(x^0, x) \leqq 1$ (refer Definition 18) such that

$$f((1 - \lambda)x^0 + \lambda x) \leqq (1 - \lambda)f(x^0) + \lambda f(x), \quad 0 < \lambda < d(x^0, x).$$

If $d(x^0, x) = a(x^0, x) = 1$, for each $x \in X$, then f is said to be *convex at* x^0 (see also Mangasarian (1969)). If f is semilocally convex at each $x^0 \in X$, then f is said to be *semilocally convex on* X. Obviously, every convex function is semilocally convex, but the converse does not hold.

Example 13. Consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2, & \text{for } x < 0 \\ 0, & \text{for } x > 0. \end{cases}$$

This function is semilocally convex on $\mathbb{R} \diagdown \{0\}$, but it is not convex on the same set (which is not convex!).

Theorem 25. (Ewing). Let $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be semilocally convex at $x^0 \in X$. Then $D^+f(x^0; x - x^0)$ exists and

$$f(x) - f(x^0) \geqq D^+f(x^0; x - x^0), \quad \forall x \in X.$$

Kaul and Kaur (1982a) introduce also the notion of *semilocally quasiconvex functions* and *semilocally pseudoconvex functions* and investigate some properties of these classes of functions.

Definition 20. Let be $f : X \longrightarrow \mathbb{R}$, $X \subset \mathbb{R}^n$; then f is said to be *semilocally quasiconvex at* $x^0 \in X$ if X is locally star-shaped at x^0 and corresponding to x^0 and each $x \in X$, there exists a positive number $d(x^0, x) \leqq a(x^0, x) \leqq 1$ (refer Definition 18) such that

$$\left\{ f(x) - f(x^0),\ 0 < \lambda < d(x^0, x) \right\} \Longrightarrow f((1 - \lambda)x^0 + \lambda x) \leqq f(x^0).$$

If $d(x^0, x) = a(x^0, x) = 1$ for each $x \in X$, then f is said to be *quasiconvex at* $x^0 \in X$ (see Mangasarian (1969)) and if f is semilocally quasiconvex at every $x^0 \in X$, then f is semilocally quasiconvex on X.

Theorem 26. (Kaul and Kaur). Let be $f : X \longrightarrow \mathbb{R}$, with $X \subset \mathbb{R}^n$ and $x^0 \in X$. If $D^+f(x^0; x - x^0)$ exists for all $x \in X$ and if f is semilocally quasiconvex at x^0, then

$$f(x) - f(x^0) \leqq 0 \Longrightarrow D^+f(x^0; x - x^0) \leqq 0.$$

It is useful to remark that every quasiconvex function is semilocally quasiconvex, but the vice-versa does not hold.

Example 14. The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1, & \text{for } x = 0 \\ 0, & \text{for } x \neq 0 \end{cases}$$

is semilocally quasiconvex on $\mathbb{R}$, but not quasiconvex on $\mathbb{R}$.

Definition 21. Let be $f : X \longrightarrow \mathbb{R}$, with $X \subset \mathbb{R}^n$ and $x^0 \in X$; then f is said to be *semilocally pseudoconvex at* x^0 if for every $x \in X$, the right-sided directional derivative $D^+f(x^0; x - x^0)$ exists for all $x \in X$ and

$$D^+f(x^0; x - x^0) \geqq 0 \Longrightarrow f(x) - f(x^0) \geqq 0.$$

If f is pseudoconvex on X (in the usual sense of Mangasarian), then f is semilocally pseudoconvex on X, but the vice-versa does not hold.

Example 15. The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $f(x) = |x|$ is semilocally pseudoconvex on $\mathbb{R}$, but not pseudoconvex on $\mathbb{R}$.

For other properties of the classes of generalized convex functions introduced above, the reader is referred to Kaul and Kaur (1982a) and for applications to optimality conditions for a nonlinear programming problem, to Kaul and Kaur (1982b, 1984), Kaul and Lyall (1988, 1990), and to Preda, Stancu-Minasian and Batatorescu (1996), Yang (1994) and to Weir (1982). For applications to multiple objective optimization problems (i. e. Pareto optimization problems), see, e. g., Preda (1996), Mukherjee and Mishra (1996), Mukhherjee and Singh (1990).

4. Directional Derivatives in Unconstrained Optimization Problems and in Set-Constrained Optimization Problems

We give first some definitions and properties concerning some *local cone approximations* frequently used in optimization theory, of a set $S \subset \mathbb{R}^n$ at a point $x^0 \in S$. Other local cone approximations will be considered in Section 6, within a more general context. For more notions on this subject the reader may consult Aubin and Frankowska (1990), Bazaraa and Shetty (1976), Giorgi and Guerraggio (1992a,b, 2002), Giorgi, Guerraggio and Thierfelder (2004), Palata (1989).

Definition 22. Let be $S \subset \mathbb{R}^n$, S non empty, and $x^0 \in S$.

a) The *Bouligand tangent cone* to S at x^0 or *contingent cone* to S at x^0 or *cone of tangent directions* to S at x^0 is given by

$$T(S, x^0) = \left\{ y \in \mathbb{R}^n : \exists t_n \longrightarrow 0^+,\ \exists y^n \longrightarrow y \text{ such that } x^0 + t_n y^n \in S,\ \forall n \right\},$$

or, in terms of neighborhoods,

$$T(S, x^0) = \left\{ \begin{array}{c} y \in \mathbb{R}^n : \forall N(y),\ \forall \lambda > 0,\ \exists t \in (0, \lambda),\ \exists \bar{y} \in N(y) \\ \text{such that } x^0 + t\bar{y} \in S \end{array} \right\}.$$

b) The *radial tangent cone* to S at x^0 or *cone of radial directions* to S at x^0 or *cone of weakly feasible directions* to S at x^0 is given by

$$WF(S, x^0) = \left\{ y \in \mathbb{R}^n : \exists t_n \longrightarrow 0^+ \text{ such that } x^0 + t_n y \in S,\ \forall n \right\}$$

or, in terms of neighborhoods,

$$WF(S,x^0) = \left\{ y \in \mathbb{R}^n : \forall \lambda > 0,\ \exists t \in (0,\lambda) \text{ such that } x^0 + ty \in S \right\}.$$

c) The *cone of feasible directions* to S at x^0 or *cone of radial interior directions* to S at x^0 is given by

$$F(S,x^0) = \left\{ y \in \mathbb{R}^n : \exists \varepsilon > 0,\ \forall t \in (0,\varepsilon) \text{ it holds } x^0 + ty \in S \right\}.$$

d) The *cone of interior displacements* to S at x^0 or *cone of interior directions* to S at x^0 or *inner tangent cone* to S at x^0 is given by

$$I(S,x^0) = \left\{ y \in \mathbb{R}^n : \exists \varepsilon > 0,\ \forall t \in (0,\varepsilon),\ \forall z \in N(y,\varepsilon) \text{ it holds } x^0 + tz \in S \right\}.$$

It is easy to note that the above defined sets are indeed cones and that the following inclusion relations hold:

$$I(S,x^0) \subset F(S,x^0) \subset WF(S,x^0) \subset T(S,x^0).$$

We note also that $0 \in F(S,x^0)$ and hence $0 \in WF(S,x^0)$ and $0 \in T(S,x^0)$, whereas $I(S,x^0)$ may be empty. If $x^0 \in int(S)$, all these cones coincide with the whole space $\mathbb{R}^n$. Moreover, if $N(x^0)$ is a neighborhood of x^0, then $T(S,x^0) = T(S \cap N(x^0), x^0)$. Similarly for the other cones of Definition 22. In general these cones are *not* convex.

Example 16. Let be

$$S_1 = \left\{ (x_1,x_2) \in \mathbb{R}^2 : x_1 \geqq 0, -(x_1)^2 \leqq x_2 < \sqrt{x_1} \right\}.$$

Let be $x^0 = (0,0)^\top$. We have

$$F(S_1,x^0) = WF(S_1,x^0) = \left\{ (y_1,y_2) \in \mathbb{R}^2 : y_1 > 0,\ y_2 \geqq 0 \right\} \cup \{(0,0)\}.$$

$$T(S_1,x^0) = \mathbb{R}^2_+.$$

$$I(S_1,x^0) = \left\{ (y_1,y_2) \in \mathbb{R}^2 : y_1 > 0,\ y_2 > 0 \right\}.$$

Example 17. Let be

$$S_2 = \left\{ (x_1,x_2) \in \mathbb{R}^2 : x_1 \geqq 0,\ x_2 = -(x_1)^2 \text{ or } x_2 = \sqrt{x_1} \right\}.$$

Let be $x^0 = (0,0)$. We have $I(S_2,x^0) = \varnothing$.

Theorem 27. $T(S,x^0)$ is a closed cone, $I(S,x^0)$ is an open cone, whereas $WF(S,x^0)$ and $F(S,x^0)$ have not such topological properties.

Theorem 28. Let $S \subset \mathbb{R}^n$ be a nonempty convex set. Then it holds:

(i)

$$I(S,x^0) = cone\left[int(S) - x^0\right];$$

(ii)

$$WF(S,x^0) = F(S,x^0) = cone(S - x^0);$$

(iii)

$$T(S,x^0) = cl\left[cone(S - x^0)\right].$$

If, moreover, $int(S) \neq \varnothing$, then

(iv)

$$I(S,x^0) = int(T(S,x^0));$$
$$T(S,x^0) = cl(I(S,x^0)).$$

Theorem 29. Let be given $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^n$, with $A \cap B \neq \varnothing$; let be $x^0 \in A \cap B$. It holds

$$\begin{array}{rcl} F(A,x^0) \cap F(B,x^0) & = & F(A \cap B, x^0); \\ I(A,x^0) \cap I(B,x^0) & = & I(A \cap B, x^0); \\ WF(A,x^0) \cap F(B,x^0) & \subset & WF(A \cap B, x^0); \\ T(S,x^0) \cap I(B,x^0) & \subset & T(A \cap B, x^0). \end{array}$$

Finally, we note that it is possible to characterize the Bouligand tangent cone $T(S,x^0)$ by means of the lower Dini directional derivative of the *distance function.* We recall that, given a set $S \subset \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$, the *distance function* $x \longrightarrow d_S(x)$ is defined by

$$d_S(x) = \inf_{y \in S} \|x - y\|, \; x \in \mathbb{R}^n,$$

being $\inf(\varnothing) = +\infty$.

Theorem 30. It holds

$$T(S,x^0) = \left\{ y \in \mathbb{R}^n : \liminf_{t \longrightarrow 0^+} \frac{d_S(x^0 + ty)}{t} = 0 \right\}.$$

Now we consider an *unconstrained minimization problem,* i. e. the problem

$$(P): \qquad \min f(x), \quad x \in X \subset \mathbb{R}^n,$$

where either X is an open set or X is arbitrary and $x^0 \in int(X)$, with x^0 optimal point for (P). The following result is now quite classical. See, e. g., Ben-Tal and Zowe (1985), Demyanov and Rubinov (1995).

Theorem 31. Let x^0 be a local solution of problem (P) and let f admit the (finite) right-sided directional derivative $D^+f(x^0;v)$ for all directions $v \in \mathbb{R}^n$. Then it holds

$$D^+f(x^0;v) \geqq 0, \; \forall v \in \mathbb{R}^n. \tag{7}$$

This result is an immediate consequence of the definition of $D^+f(x^0;v)$. Some authors call a point x^0 satisfying (7) an "inf-stationary point of f on X". A "sup-stationary point" is defined in a similar way. Obviously, if the (bilateral) directional derivative $Df(x^0;v)$ exists for all $v \in \mathbb{R}^n$, from the fact that x^0 is a local solution of (P) it follows $Df(x^0;v) = 0$, $\forall v \in \mathbb{R}^n$. The same holds if f is Gâteaux differentiable at x^0 : we have $\nabla f(x^0)^\top v = 0$, $\forall v \in \mathbb{R}^n$, i. e. $\nabla f(x^0 = 0 \in \mathbb{R}^n$.

It is possible, following Demyanov and Pevnyi (1974) and Hiriart-Urruty (1982), to define also a *second-order* one-sided directional derivative:

$$D^{2+}f(x^0;v) = \lim_{t \longrightarrow 0^+} \frac{f(x^0+tv) - f(x^0) - tD^+f(x^0;v)}{t^2},$$

provided that this limit exists (finite or not). Then, we have the following refinement of Theorem 31: if x^0 is a local solution of (P) and $D^+f(x^0;v)$ and $D^{2+}f(x^0;v)$ exist (finite) for all $v \in \mathbb{R}^n$, then we have

$$D^+f(x^0;v) \geqq 0, \ \forall v \in \mathbb{R}^n$$

and

$$D^+f(x^0;v) = 0 \Longrightarrow D^{2+}f(x^0;v) \geqq 0.$$

For other questions and applications of second-order directional derivatives to nonsmooth optimization problems, see, e. g., Huang and Ng (1994), Huang (2005), Studniarski (1991), Yang (1996, 1999). Also for this topic the literature bis abundant.

A famous error of J. L. Lagrange is contained in his claim that if a smooth function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ has all its directional derivatives at a point x^0, that are nonnegative for all directions, then x^0 is a local minimum point for f. Only at the end of the 18th century Giuseppe Peano gave the following counterexample. Take $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, with

$$f(x_1,x_2) = (x_2 - 2(x_1)^2)(x_2 - (x_1)^2),$$

and consider $x^0 = (0,0)$. This function satisfies at x^0 the assumptions of Lagrange, but it changes its sign on *every* neighborhood of x^0, hence x^0 cannot be a local minimizer for f. In order to obtain sufficient (first-order) optimality conditions for (P), we have to make further assumptions.

Theorem 32. Let $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be locally Lipschitz at $x^0 \in X$. If

$$D^+f(x^0;v) > 0, \ \forall v \in \mathbb{R}^n, \ v \neq 0,$$

then x^0 is a strict local minimizer for (P).

We shall give the proof under a more general assumption (Theorem 35).

We now revert to *convex optimization,* i. e. in problem (P) the function f isa *convex* on the open convex set $X \subset \mathbb{R}^n$. We recall that in this case $D^+f(x^0;v)$ exists finite for all $v \in \mathbb{R}^n$.

Theorem 33. Let $X \subset \mathbb{R}^n$ be an open convex set and $f : X \longrightarrow \mathbb{R}$ a convex function; let $x^0 \in X$. Then the following assertions are equivalent.

(a) f admits a local minimum point at x^0;

(b) f admits a global minimum point at x^0;

(c) It holds $0 \in \partial f(x^0)$;

(d) It holds $D^+ f(x^0; v) \geqq 0,\ \forall v \in \mathbb{R}^n$.

Proof. The relations (a) $\Longleftrightarrow$ $(b) \Longrightarrow (d)$ are evident. The implication (d)$\Longrightarrow$(c) follows from Theorem 15. The implication (c)$\Longrightarrow$(b) follows at once from the definition of subdifferential of f at x^0 : the inequality $f(x) - f(x^0) \geqq s^\top (x - x^0)$, $\forall x \in X$, when $s = 0$ gives condition (b). □

Theorem 31 can be generalized by using other types of (more general) directional derivatives. As a consequence of a general result of Elster and Thierfelder (1988b) we have the following necessary optimality conditions for (P). See also Section 6 of the present paper.

Theorem 34. Let x^0 be a local solution for problem (P). Then it holds

$$f^{D^+}(x^0; v) \geqq 0,\ \forall v \in \mathbb{R}^n,$$

or also the sharper condition

$$f_{D^+}(x^0; v) \geqq 0,\ \forall v \in \mathbb{R}^n.$$

Or also, in terms of Hadamard directional derivatives,

$$f^{H^+}(x^0; v) \geqq 0,\ \forall v \in \mathbb{R}^n,$$

or also the sharper condition

$$f_{H^+}(x^0; v) \geqq 0,\ \forall v \in \mathbb{R}^n.$$

We recall that if f is locally Lipschitz, then the Dini and the correspondent Hadamard directional derivatives are finite and coincide (see Aubin and Cellina (1984)).

Now we turn to sufficient first-order optimality conditions for (P) in terms of Dini directional derivatives. See, e. g., Qi (2001), Demyanov and Rubinov (1995).

Theorem 35. Suppose that in problem (P) $f : X \longrightarrow \mathbb{R}$ is a locally Lipschitz function. If it holds

$$f_{D^+}(x^0; v) > 0, \quad \forall v \in \mathbb{R}^n,\ v \neq 0,$$

then x^0 is a strict local minimizer for (P).

Proof. By assumption, there exist a neighborhood $N(x^0)$ and a constant $L \geqq 0$ such that, for any $x, y \in N(x^0)$

$$|f(x) - f(y)| \leqq L \|x - y\|.$$

Ab absurdo, assume that x^0 is not a strict local minimum point of f. Then, there exists a sequence $\{x^k\} \subset N(x^0)$ such that $x^k \longrightarrow x^0$, but $x^k \neq x^0$ and $f(x^k) \leqq f(x^0)$ for all k. Without loss of generality, assume that

$$\left\{(x^k - x^0)\big/ \left\|x^k - x^0\right\|\right\} \longrightarrow v.$$

Then $\|v\| = 1$. Let be

$$y^k \equiv x^0 + v \left\|x^k - x^0\right\|.$$

Then,

$$\left\|y^k - x^0\right\| = \left\|x^k - x^0\right\|$$

and

$$\left|f(y^k) - f(x^k)\right| \leqq L \left\|y^k - x^k\right\| = L \left\|x^k - x^0\right\| \left\|\frac{x^k - x^0}{\|x^k - x^0\|} - v\right\|.$$

This implies that

$$\lim_{k \longrightarrow \infty} \frac{f(y^k) - f(x^0)}{\|x^k - x^0\|} \leqq \lim_{k \longrightarrow \infty} \frac{f(y^k) - f(x^k)}{\|x^k - x^0\|} \leqq \lim_{k \longrightarrow \infty} L \left\|\frac{x^k - x^0}{\|x^k - x^0\|} - v\right\| = 0.$$

This shows that the Dini lower directional derivative of f at x^0 in the direction v is nonpositive, contradicting the assumption. □

Analogously the upper Dini directional derivative can be used to obtain sufficient conditions for a maximum. We remark finally that if

$$f_{D^+}(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n$$

and if f is *lower Dini-pseudoconvex* (Definition 16) on the open convex set $X \subset \mathbb{R}^n$, then x^0 is a *global* minimizer of f over X.

If in (P) the objective function f is not locally Lipschitz, then in order to obtain first-order sufficient local optimality conditions in terms of directional derivatives, it is possible to use the *Hadamard directional derivative.* See Demyanov and Rubinov (1995, page 236), Studniarski (1986).

Theorem 36. Let be given $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open and $x^0 \in X$; if

$$f_{H^+}(x^0; v) > 0, \quad \forall v \in \mathbb{R}^n, \; v \neq 0,$$

then x^0 is a strict local minimum point for (P).

Note that in the above theorem the function is not assumed to be locally Lipschitz, nor continuous. If f is locally Lipschitz on X, then Theorem 36 collapses to Theorem 35, being, under the said assumption, $f_{D^+}(x^0; v) = f_{H^+}(x^0; v)$. Note that all the first-order sufficient conditions for the strict optimality of x^0 in problem (P) have an inevitable nonsmooth nature. In the smooth case they would make no sense.

Now we consider a constrained minimization problem with a *set constraint* (or *abstract constraint*), i. e.

$$(P_0): \qquad \min f(x), \quad x \in S \subset X \subset \mathbb{R}^n,$$

where S is a closed subset of the open set X. When f is differentiable on X and x^0 is a local minimizer of f over S, then a well known necessary optimality condition for (P_0) is

$$\nabla f(x^0)^\top y \geqq 0, \quad \forall y \in T(S, x^0),$$

or, equivalently,

$$-\nabla f(x^0) \in (T(S, x^0))^*,$$

where $(T(S, x^0))^*$ is the (negative) polar cone of $T(S, x^0)$. The following necessary optimality conditions for (P_0) are due to D. E. Ward (1987) and are special cases of a more general result.

Theorem 37. Let $x^0 \in S$ be a local solution for (P_0). Then we have

i)

$$f^{H^+}(x^0; y) \geqq 0, \quad \forall y \in T(S, x^0);$$

ii)

$$f_{H^+}(x^0; y) \geqq 0, \quad \forall y \in I(S, x^0);$$

iii)

$$f^{D^+}(x^0; y) \geqq 0, \quad \forall y \in WF(S, x^0);$$

iv)

$$f_{D^+}(x^0; y) \geqq 0, \quad \forall y \in F(S, x^0).$$

We recall that if f is locally Lipschitz on the open set $X \subset \mathbb{R}^n$, then

$$f^{H^+}(x^0; y) = f^{D^+}(x^0; y)$$

and

$$f_{H^+}(x^0; y) = f_{D^+}(x^0; y).$$

(See, e. g., Aubin and Cellina (1984)).

We now assume that f is locally Lipschitz on X. We have also the following result (see Castellani and Pappalardo (1995)).

Theorem 38. Assume that in (P_0) the objective function f is locally Lipschitz on the open set $X \subset \mathbb{R}^n$ and that $x^0 \in S$ is a local solution for (P_0). Then we have

$$f^{D^+}(x^0; y) \geqq 0, \quad \forall y \in T(S, x^0).$$

Moreover, being $f^{D^+}(x^0;y) \geqq f_{D^+}(x^0;y)$, we have also the more accurate necessary optimality condition

$$f_{D^+}(x^0;y) \geqq 0, \quad \forall y \in T(S,x^0).$$

Obviously, under the assumptions of Theorem 38, if f is right-sided directionally differentiable at x^0, then we have

$$D^+f(x^0;y) \geqq 0, \quad \forall y \in T(S,x^0).$$

Under the assumptions of Theorem 37, if $D^+f(x^0;y)$ exists (finite) for all $y \in WF(S,x^0)$, then we have

$$D^+f(x^0;y) \geqq 0, \quad \forall y \in WF(S,x^0)$$

and if f is Gâteaux differentiable at x^0, then we have

$$\nabla f(x^0)^\top y \geqq 0, \ \forall y \in (WF(S,x^0))^{**},$$

where $(WF(S,x^0))^{**}$ is the *bipolar cone* of $WF(S,x^0)$. Indeed, in this case we have $\nabla f(x^0)^\top y \geqq 0$, $\forall y \in WF(S,x^0)$. But from the linearity we obtain also the last written inequality with the bipolar cone, a sharper result, as, for any set $S \subset \mathbb{R}^n$, it always holds $S \subset S^{**}$ (see, e. g., Ben-Israel (1969), Uzawa (1958a)).

For what concerns first-order sufficient optimality conditions for (P_0), we have some nice results, due to Correa and Hiriart-Urruty (1989), Penot (1984) and Studniarski (1986).

Theorem 39. Let $x^0 \in S$ in problem (P_0). If

$$f_{H^+}(x^0;y) > 0, \quad \forall y \in T(S,x^0), \ y \neq 0,$$

then x^0 is a strict local minimizer for (P_0).

It can be proved that the above condition assures that x^0 is also an *isolated local minimum point of order 1,* i. e. there exist $\alpha > 0$ and $N(x^0)$ such that

$$f(x) > f(x^0) + \alpha \left\| x - x^0 \right\|, \quad \forall N(x^0) \cap S.$$

If we assume that f is locally Lipschitz on X, then the thesis of Theorem 39 holds under the condition

$$f_{D^+}(x^0;y) > 0, \quad \forall y \in T(S,x^0), \ y \neq 0.$$

Always under the assumption that f is locally LIpschitz on X, then

$$D^+f(x^0;y) > 0, \quad \forall y \in T(S,x^0), \ y \neq 0,$$

is a sufficient condition for the strict local optimality of x^0 for (P_0).

A "converse" of Theorem 39 is provided by Studniarski (1986).

Theorem 40. Let $x^0 \in S$ in problem (P_0). If x^0 is an isolated local minimum point of order 1, then

$$f^{H^+}(x^0;y) > 0, \quad \forall y \in T(S,x^0), \ y \neq 0.$$

Now suppose that in (P_0) f is *convex* on the open convex set $X \subset \mathbb{R}^n$ and that S is convex. Then we have the following result.

Theorem 41. Let be in (P_0) the set $S \subset X$ a nonempty convex set and let $X \subset \mathbb{R}^n$ be an open convex set. Let $f : X \longrightarrow \mathbb{R}$ be convex on X. Then the following properties are equivalent.

(a) x^0 is a global minimum point of f over S.

(b) x^0 is a local minimum point of f over S.

(c) It holds $D^+ f(x^0; x - x^0) \geqq 0, \ \forall x \in S$.

Proof. We recall that, being f convex over X, the right-sided directional derivative $D^+ f(x^0; y)$ exists finite for every $y \in \mathbb{R}^n$. The equivalence (a) $\Longleftrightarrow$ (b) is well known. The implication (b)$\Longrightarrow$(c) stems form Theorem 38 and Theorem 28. The implication (c)$\Longrightarrow$(a) follows from the convexity of f :

$$D^+ f(x^0; x - x^0) \leqq f(x) - f(x^0), \quad \forall x \in S. \qquad \square$$

Theorem 41 is often given in terms of the *subdifferential* of f at x^0 : see, e. g., Rockafellar (1970), Dhara and Dutta (2012), Durea and Strugariu (2014). If we impose another suitable inequality on the Dini directional derivatives of f, we can obtain sufficient *global* optimality conditions for (P_0), without assuming that f is convex or generalized convex. The following result is due to Correa and Hiriart-Urruty (1989).

Theorem 42. Assume that in (P_0) the set $S \subset X$ is *convex* and that $f : X \longrightarrow \mathbb{R}$ is *continuos* on X. A sufficient condition for $x^0 \in S$ to be a global minimum point of f on S is that

$$f_{D^+}(x; x^0 - x) \leqq 0, \quad \forall x \in S.$$

If, moreover, the above inequality is strict, whenever $x \in S \diagdown \{x^0\}$, then x^0 is a strict global minimum point of f on S.

Note that in Theorem 42 the Dini lower directional derivative is taken at the (variable) point x, not at the point x^0. Theorem 42 is a nonsmooth generalization of a result of Boisko (1986) for the differentiable case.

5. Directional Derivatives in Optimization Problems with Functional Constraints

Optimality conditions in terms of subdifferentials or directional derivatives for a constrained convex optimization problem (i. e. a constrained optimization problem with convex objective function and convex constraints) have been obtained since the beginning of modern Convex Analysis and modern Nonsmooth Calculus. The reader is referred to the works quoted at the beginning of the previous section. Let us consider, for example, the following mathematical programming problem.

$$(P_1) : \qquad \begin{cases} \min f(x) \\ \text{subject to: } g_i(x) \leqq 0, \ i = 1, ..., m, \\ x \in X \subset \mathbb{R}^n, \end{cases}$$

where $f : X \longrightarrow \mathbb{R}$ and every $g_i : X \longrightarrow \mathbb{R}$, $i = 1, ..., m$, are convex on the open convex set $X \subset \mathbb{R}^n$. The *feasible set* of (P_1) is denoted by K :

$$K = \{x \in X : g_i(x) \leqq 0, \ i = 1, ..., m\}.$$

If $x^0 \in K$, the set of the indices of the *active constraints* at x^0 is denoted by $I(x^0)$:

$$I(x^0) = \left\{i : g_i(x^0) = 0\right\}.$$

In terms of (unilateral) directional derivatives, under the assumptions made above, we have the following Fritz John-type necessary optimality conditions.

Theorem 43. If $x^0 \in K$ is a local solution of the convex problem (P_1), then there exist multipliers $u_0 \geqq 0, u_1 \geqq 0, ..., u_m \geqq 0$, *not all zero,* such that

$$u_0 D^+ f(x^0; v) + \sum_{i=1}^{m} u_i D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n;$$

$$u_i g_i(x^0) = 0, \quad i = 1, ..., m.$$

We recall that, under our assumptions, the right-sided directional derivatives which appear in the above inequality, exist finite at every point of the open convex set $X \subset \mathbb{R}^n$. In order to obtain that in Theorem 43 the multiplier u_0 is positive, i. e. $u_0 = 1$, we must impose some *constraint qualification.* Under our assumptions, the most natural and simple constraint qualification is the *Slater c. q.:*

- There exists a point $\bar{x} \in K$ such that $g_i(\bar{x}) < 0$, $i \in I(x^0)$.

Theorem 44. Let $x^0 \in K$ be a local solution of the convex problem (P_1) and let the Slater c. q. be verified. Then there exist multipliers $u_1 \geqq 0, ..., u_m \geqq 0$ such that the following Karush-Kuhn-Tucker-type necessary optimality conditions hold:

$$D^+ f(x^0; v) + \sum_{i=1}^{m} u_i D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n;$$

$$u_i g_i(x^0) = 0, \quad i = 1, ..., m.$$

Theorems 43 and 44 will be proved further, under more general assumptions.

Remark 5. If in the convex problem (P_1), besides the functional constraints, we have also a *set constraint,* i. e. $x \in C \subset X$, where C is a closed convex set, the Fritz John conditions of Theorem 43 are expressed as follows: there exist multipliers $u_0 \geqq 0, u_1 \geqq 0, ..., u_m \geqq 0$, *not all zero,* such that

$$u_0 D^+ f(x^0; x - x^0) + \sum_{i=1}^{m} u_i D^+ g_i(x^0; x - x^0) \geqq 0, \quad \forall x \in C;$$

$$u_i g_i(x^0) = 0, \quad i = 1, ..., m.$$

Similarly for Theorem 44, under the Slater constraint qualification:

- There exists $\bar{x} \in C$ such that $g_i(\bar{x}) < 0$, $i \in I(x^0)$.

One of the first authors to consider explicitly nonconvex nonsmooth constrained optimization problems of the type (P_1), has been B. N. Pshenichnyi (1971). If we consider, as almost always usual in the present paper, a real-valued function $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, where X is an open set, then f is said to be *quasi-differentiable at $x^0 \in X$ in the sense of Pshenichnyi,* if $D^+ f(x^0; v)$ exists (finite) for all $v \in \mathbb{R}^n$ and there exists a nonempty closed convex subset $Mf(x^0) \subset \mathbb{R}^n$ such that

$$D^+ f(x^0; v) = \max_{y \in Mf(x^0)} \left\{ y^\top v \right\}, \quad \forall v \in \mathbb{R}^n.$$

The set $Mf(x^0)$ is also called the *quasi-differential* of f at x^0. Hence the quasi-differentiability condition of Pshenichnyi imposes that the function $v \longrightarrow D^+ f(x^0; v)$ is *convex* for every $v \in \mathbb{R}^n$. Obviously, a convex function on the open convex set $X \subset \mathbb{R}^n$ is quasi-differentiable on X, with $Mf(x^0) = \partial f(x^0)$, however a function may be quasi-differentiable in the sense of Pshenichnyi, i. e. $D^+ f(x^0; v)$ is convex, $\forall v \in \mathbb{R}^n$, without being itself convex.

Example 18. The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} x + x^2 \sin \frac{1}{x}, \text{ for } x > 0; \\ 0, \text{ for } x = 0; \\ x^2 \sin \frac{1}{x}, \text{ for } x < 0, \end{cases}$$

is not convex, however at $x^0 = 0$ has a right-sided directional derivative

$$D^+ f(x^0; v) = \begin{cases} v, \text{ for } v > 0, \\ 0, \text{ for } v \leqq 0, \end{cases}$$

which is convex for $v \in \mathbb{R}$.

We consider now (P_1) under the assumptions that f and every g_i, $i = 1, ..., m$, admit at $x^0 \in K$ *finite and convex* right-sided directional derivatives, with respect to every direction $v \in \mathbb{R}^n$, and that the constraints g_i, $i \notin I(x^0)$, are *continuous at* x^0. The weaker assumption that the directional derivatives are finite and convex, in the direction v, on a convex cone $C(x^0) \subset \mathbb{R}^n$, requires only formal variants in what follows.

Let be $x^0 \in K$; the *generalized linearizing cone at* x^0 for the above problem (P_1) is defined as

$$L(x^0) = \left\{ v \in \mathbb{R}^n : D^+ g_i(x^0; v) \leqq 0, \ \forall i \in I(x^0) \right\}.$$

Now we fit to (P_1) the approach adopted by Berge and Ghouila-Houri (1965) for the differentiable case. We define therefore the following problems.

(P_2) (Local minimum problem along directions). Find $x^0 \in K$ such that for every $v \in \mathbb{R}^n$ there exists $\lambda_0 > 0$ such that

$$\left\{x^0 + \lambda v \in K,\ \lambda \in (0, \lambda_0]\right\} \Longrightarrow f(x^0 + \lambda v) \geqq f(x^0).$$

(P_3) (Feasible directions problem). Find a point $x^0 \in K$ such that

$$v \in L(x^0) \Longrightarrow D^+ f(x^0; v) \geqq 0.$$

(P_4) (Generalized Karush-Kuhn-Tucker problem). Find a pair $(x^0, u^0) \in K \times \mathbb{R}^m_+$ such that

$$D^+ f(x^0; v) + \sum_{i \in I(x^0)} u_i^0 D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n,$$

$$u_i^0 g_i(x^0) = 0, \quad i = 1, ..., m.$$

(P_5) (Saddle points problem). Find a pair $(x^0, u^0) \in X \times \mathbb{R}^m_+$ such that, with $\mathcal{L}(x, u) = f(x) + \sum_{i=1}^m u_i g_i(x)$, it holds

$$\mathcal{L}(x^0, u) \leqq \mathcal{L}(x^0, u^0) \leqq \mathcal{L}(x, u^0), \quad \forall (x, u) \in X \times \mathbb{R}^m_+.$$

For the developments of our analysis we need two preliminary lemmas. The first result is due to Fan, Glicksberg and Hoffman (1957) and may be considered a generalization to the nonlinear (convex) case of the theorem of the alternative of Gordan (see Mangasarian (1969)).

Lemma 1. Let be $S \subset \mathbb{R}^n$ a nonempty convex set and let $f : S \longrightarrow \mathbb{R}^m$ be convex. Then, either the system

$$\begin{cases} f_1(x) < 0 \\ \vdots \\ f_m(x) < 0 \end{cases}$$

has a solution $x \in S$, or it holds

$$p_1 f_1(x) + p_2 f_2(x) + ... + p_m f_m(x) \geqq 0,\ \forall x \in S,$$

for some $p_1 \geqq 0, p_2 \geqq 0, ..., p_m \geqq 0$, *not all zero,* but never both.

Lemma 2. If $x^0 \in K$ is a local solution of (P_1), then the system

$$\begin{cases} D^+ f(x^0; v) < 0 \\ D^+ g_i(x^0; v) < 0, \quad i \in I(x^0), \end{cases}$$

admits no solution $v \in \mathbb{R}^n$.

Proof. Ab absurdo let us suppose that the system admits a solution $\bar{v} \in \mathbb{R}^n$. Let us consider the constraints g_i, with $i \notin I(x^0)$. Being these constraints continuous at x^0, there will exist a sufficiently small number, say $\lambda_0 > 0$, such that

$$g_i(x^0 + \lambda\bar{v}) < 0, \quad \forall i \notin I(x^0), \; \lambda \in (0, \lambda_0] \,.$$

For $i \in I(x^0)$, it is possible to find $\lambda_0 > 0$ such that

$$g_i(x^0 + \lambda\bar{v}) < 0, \quad \forall i \in I(x^0), \; \lambda \in (0, \lambda_0] \,,$$

being

$$g_i(x^0 + \lambda\bar{v}) = \lambda D^+ g_i(x^0 + \lambda\bar{v}) + o(\lambda), \quad \text{for } \lambda \longrightarrow 0^+.$$

Similarly, for the objective function we can find $\lambda_0 > 0$ such that

$$f(x^0 + \lambda\bar{v}) - f(x^0) < 0, \; \lambda \in (0, \lambda_0] \,. \tag{8}$$

So we have found a point $x^0 + \lambda\bar{v} \in K$ such that (8) holds, in contradiction with the local optimality of x^0 on K. □

The following result is a generalization of the well known Fritz John Theorem to a nonlinear programming problem of the type (P_1), where the related functions are endowed with (finite) convex right-sided directional derivatives.

Theorem 45. Let in (P_1) the directional derivatives $D^+ f(x^0; v)$ and $D^+ g_i(x^0; v)$, $i \in I(x^0)$, be *convex* in v for all $v \in \mathbb{R}^n$. If x^0 is a local solution of (P_1), then there exist numbers $\bar{p}_0 \geqq 0$, $\bar{p}_i \geqq 0$, $i \in I(x^0)$, *not all zero,* such that

$$\bar{p}_0 D^+ f(x^0; v) + \sum_{i \in I(x^0)} \bar{p}_i D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n;$$

$$\bar{p}_i g_i(x^0) = 0, \; i = 1, ..., m.$$

Proof. It is sufficient to apply Lemma 2 and subsequently Lemma 1. By choosing $\bar{p}_i = 0$, $i \notin I(x^0)$, we obtain the second part of the thesis. □

Note that the convexity assumption of $D^+ g_i(x^0; v)$ with respect to v, is required only for the active constraints at x^0. Obviously Theorem 45 applies also to the convex case: we recall that a convex function on an open convex set $X \subset \mathbb{R}^n$, has finite (right-sided) sublinear directional derivatives on X, for every direction $v \in \mathbb{R}^n$. Note, moreover, that if the functions involved in (P_1) are *Gâteaux differentiable at* x^0, then we have the usual Fritz John conditions

$$\bar{p}_0 \nabla f(x^0) + \sum_{i \in I(x^0)} \bar{p}_i \nabla g_i(x^0) = 0 \in \mathbb{R}^n,$$

$$\bar{p}_i g_i(x^0) = 0, \; i = 1, ..., m.$$

In order to obtain Karush-Kuhn-Tucker-type necessary optimality conditions for (P_1), expressed by (right-sided) directional derivatives, we have to consider some *constraint qualifications.* For example:

- $(CQ)_1$: *Slater constraint qualification.* The constraints g_i, $i = 1, ..., m$, are convex on the open convex set $X \subset \mathbb{R}^n$ and there exists $\bar{x} \in K$, with $g_i(\bar{x}) < 0$, $\forall i = 1, ..., m$.

- $(CQ)_2$: *Karlin constraint qualification.* The functions g_i, $i = 1, ..., m$, are convex on the open convex set $X \subset \mathbb{R}^n$ and for every vector $p \in \mathbb{R}^m_+$, $p \neq 0$, there exists $\bar{x} \in \mathbb{R}^n$ such that

$$\sum_{i=1}^{m} p_i g_i(\bar{x}) < 0.$$

It is well known (see, e. g., Mangasarian (1969)) that the two above constraint qualifications are equivalent: it is sufficient to apply Lemma 1. Other two constraint qualifications which do not require convexity assumptions on the constraints are the following ones.

- $(CQ)_3$. The directional derivatives $D^+g_i(x^0; g)$, $i \in I(x^0)$, are convex in v, for all $v \in \mathbb{R}^n$, and there exists $\bar{v} \in \mathbb{R}^n$ such that

$$D^+g_i(x^0; \bar{v}) < 0, \ \forall i \in I(x^0).$$

- $(CQ)_4$. The directional derivatives $D^+g_i(x^0; g)$, $i \in I(x^0)$, are convex in v, for all $v \in \mathbb{R}^n$ and for all $u_i \geqq 0$, $i \in I(x^0)$, u_i not all zero, there exists $\bar{v} \in L(x^0)$ with

$$\sum_{i \in I(x^0)} u_i D^+g_i(x^0; \bar{v}) < 0.$$

Being $(CQ)_1 \iff (CQ)_2$, we have also $(CQ)_3 \iff (CQ)_4$. Then we have $(CQ)_1 \Longrightarrow (CQ)_3$: being g_i, $i = 1, ..., m$, convex, also $D^+g_i(x^0; v)$ is a convex function of $v \in \mathbb{R}^n$, for $i = 1, ..., m$. Choose $\bar{v} = \bar{x} - x$. For each $i \in I(x^0)$ we have

$$D^+g_i(x^0; \bar{v}) = \lim_{\lambda \longrightarrow 0^+} \frac{g_i(x^0 + \lambda(\bar{x} - x^0) - g_i(x^0)}{\lambda} \leqq$$

$$\leqq \lim_{\lambda \longrightarrow 0^+} \frac{\lambda g_i(\bar{x}) + (1 - \lambda)g_i(x^0) - g_i(x^0)}{\lambda} = g_i(\bar{x}) < 0.$$

It is now possible to obtain a Karush-Kuhn-Tucker-type necessary optimality condition for (P_1).

Theorem 46. Let in (P_1) the directional derivatives $D^+f(x^0; v)$, $D^+g_i(x^0; v)$, $i \in I(x^0)$, be convex in v, for all $v \in \mathbb{R}^n$. If $x^0 \in K$ is a local solution of (P_1) and one of the above constraint qualifications is satisfied, then there exists a vector $u^0 \in \mathbb{R}^m_+$ such that the pair (x^0, u^0) is solution of problem (P_4).

Proof. From Theorem 45 there exist multipliers $\bar{p}_0 \geqq 0$, $\bar{p}_i \geqq 0$, $i \in I(x^0)$, *not all zero,* such that

$$\bar{p}_0 D^+ f(x^0; v) + \sum_{i \in I(x^0)} \bar{p}_i D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n.$$

Let us suppose that $(CQ)_4$ holds and suppose that in the above inequality it holds $\bar{p}_0 = 0$. Then we have

$$\sum_{i \in I(x^0)} \bar{p}_i D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n,$$

being $\bar{p}_i \geqq 0$, $i \in I(x^0)$, not all zero. This is in contradiction with $(CQ)_4$, hence $\bar{p}_0 > 0$ and with the choice

$$u_i^0 = \frac{\bar{p}_i}{\bar{p}_0}, \ i \in I(x^0)$$

we obtain

$$D^+ f(x^0; v) + \sum_{i \in I(x^0)} u_i^0 D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n.$$

Then, by choosing $u_i^0 = 0$ for $i \notin I(x^0)$, we have that (x^0, u^0) is solution of (P_4). □

Theorem 47. If the pair (x^0, u^0) is solution of (P_5), then it is also solution of (P_4).

Proof. By assumption, we have

$$f(x^0) + \textstyle\sum_{i=1}^m u_i g_i(x^0) \leqq f(x^0) + \sum_{i=1}^m u_i^0 g_i(x^0) \leqq f(x) + \sum_{i=1}^m u_i^0 g_i(x),$$
$$\forall (x, u) \in X \times \mathbb{R}_+^m.$$

From these inequalities it follows (see, e. g., Mangasarian (1969))

$$\sum_{i=1}^m u_i^0 g_i(x^0) = 0$$

and hence also $g_i(x^0) \leqq 0$, $i = 1, ..., m$. On the other hand we have

$$\mathcal{L}(x^0 + \lambda v, u^0) \geqq \mathcal{L}(x^0, u^0), \ \forall v \in \mathbb{R}^n, \ \forall \lambda > 0.$$

Hence

$$\frac{\mathcal{L}(x^0 + \lambda v, u^0) - \mathcal{L}(x^0, u^0)}{\lambda} \geqq 0, \ \forall v \in \mathbb{R}^n.$$

For $\lambda \longrightarrow 0^+$ we have that the pair (x^0, u^0) is solution of (P_4). □

Theorem 48. Let in (P_1) the objective function f and every constraint g_i, $i = 1, ..., m$, be convex on the open convex set $X \subset \mathbb{R}^n$. If (x^0, u^0) is a solution of (P_4), then (x^0, u^0) is also a solution of (P_5).

Proof. It is sufficient to choose $v = x - x^0$ and take into account the convexity of f and each g_i, $i = 1, ..., m$:

$$f(x^0 + \lambda v) - f(x^0) \leqq \lambda f(x) + (1-\lambda) f(x^0) - f(x^0) =$$
$$= \lambda \left[f(x) - f(x^0) \right], \quad \lambda \in (0,1), \ x \in X$$

and hence

$$D^+ f(x^0; v) \leqq f(x) - f(x^0), \ \forall x \in X.$$

In correspondence it holds

$$D^+ g_i(x^0; v) \leqq g_i(x) - g_i(x^0), \quad \forall i = 1, ..., m, \ \forall x \in X.$$

Being (x^0, u^0) solution of (P_4), it holds

$$0 \leqq D^+ f(x^0; v) + \sum_{i=1}^{m} u_i^0 D^+ g_i(x^0; v) \leqq f(x) - f(x^0) + \sum_{i=1}^{m} \left[g_i(x) - g_i(x^0) \right],$$

that is

$$\mathcal{L}(x^0, u^0) \leqq \mathcal{L}(x, u^0), \quad \forall x \in X.$$

The inequality

$$\mathcal{L}(x^0, u) \leqq \mathcal{L}(x^0, u^0), \quad \forall u \in \mathbb{R}^m_+,$$

follows immediately form the conditions

$$\sum_{i=1}^{m} u_i^0 g_i(x^0) = 0; \quad \sum_{i=1}^{m} u_i g_i(x^0) \leqq 0. \qquad \square$$

Theorem 49. If (x^0, u^0) is solution of (P_4), then x^0 is also solution of (P_3).

Proof. From the assumptions we have

$$D^+ f(x^0; v) \geqq - \sum_{i \in I(x^0)} u_i^0 D^+ g_i(x^0; v), \quad \forall v \in \mathbb{R}^n.$$

It follows

$$D^+ f(x^0; v) \geqq 0, \ \forall v \in L(x^0). \qquad \square$$

Theorem 50. Let the directional derivatives $D^+ f(x^0; v)$ and $D^+ g_i(x^0; v)$, $i \in I(x^0)$, be convex in v for every $v \in \mathbb{R}^n$. Suppose that the constraint qualification $(CQ)_3$ is satisfied. If $x^0 \in K$ is solution of (P_3), then there exists $u^0 \in \mathbb{R}^m_+$ such that (x^0, u^0) is solution of (P_4).

Proof. By assumption, the system

$$\begin{cases} D^+ f(x^0; v) < 0 \\ D^+ g_i(x^0; v) \leqq 0, \ i \in I(x^0), \end{cases}$$

has no solution $v \in \mathbb{R}^n$. A fortiori the system

$$\begin{cases} D^+ f(x^0; v) < 0 \\ D^+ g_i(x^0; v) < 0, \ i \in I(x^0), \end{cases}$$

has no solution $v \in \mathbb{R}^n$. By Lemma 1, there exist nonnegative multipliers, not all zero, u_0, $u_1^0, ..., u_m^0$, such that

$$u_0 D^+ f(x^0; v) + \sum_{i \in I(x^0)} u_i^0 D^+ g_i(x^0; v) \geqq 0, \quad \forall v \in \mathbb{R}^n.$$

But being $(CQ)_3$ verified, it will be $u_0 = 1$, i. e. (x^0, u^0), with $u_i^0 = 0$ for $i \notin I(x^0)$, is solution of (P_4). □

Theorem 51. If $x^0 \in K$ is solution of (P_3), then it is also solution of (P_2).

Proof. Let us consider any direction $v \in \mathbb{R}^n$. We have two cases.

a) $v \notin L(x^0)$. Therefore it will exist an index $i_* \in I(x^0)$ such that $D^+ g_{i_*}(x^0; v) > 0$. For λ_1 sufficiently small we have then

$$g_{i_*}(x^0 + \lambda v) > g_{i_*}(x^0) = 0, \quad \forall \lambda \in (0, \lambda_1]\,.$$

Therefore in this case $x^0 + \lambda v \in K$ only for $\lambda = 0$, and trivially we have

$$f(x^0 + \lambda v) = f(x^0).$$

b) $v \in L(x^0)$. We have therefore

$$\begin{aligned} &D^+ f(x^0; v) \geqq 0 \\ &D^+ g_i(x^0; v) \leqq 0, \quad \forall i \in I(x^0). \end{aligned}$$

For λ_2 sufficiently small it follows, with $\lambda \sqsubseteq (0, \lambda_2]$,

$$\begin{aligned} &f(x^0 + \lambda v) \geqq f(x^0) \\ &g_i(x^0 + \lambda v) \leqq 0, \quad \forall i \in I(x^0) \end{aligned}$$

and also

$$g_i(x^0 + \lambda v) \leqq 0, \ \forall i \notin I(x^0).$$

Hence there exists $\lambda^* > 0$ such that, for every $v \in \mathbb{R}^n$ it holds

$$\{\lambda \in (0, \lambda^*]\,, \ x^0 + \lambda v \in K\} \Longrightarrow f(x^0 + \lambda v) \geqq f(x^0),$$

i. e. x^0 is solution of (P_2). □

Remark 6. The assumption on the convexity of the (right-sided) directional derivatives involved in (P_1) can be weakened, but not completely skipped, if we wish to obtain a Fritz John or a Karush-Kuhn-Tucker multipliers rule. A proposal of weakening the above assumption is given by the class of *preinvex functions:* see Ben-Israel and Mond (1986) and Weir and Mond (1988).

Definition 23. Let $S \subset \mathbb{R}^n$ be a nonempty set. The function $\varphi : S \longrightarrow \mathbb{R}$ is *preinvex* on S if there exists an n-dimensional vector function $\eta(x, y) : S \times S \longrightarrow \mathbb{R}^n$ such that, for all $x, y \in S$ and all $\lambda \in [0, 1]$, we have

$$\varphi(y + \lambda\eta(x, y)) \leqq \lambda\varphi(x) + (1 - \lambda)\varphi(y).$$

It can be shown (see Weir and Mond (1988)) that convex functions are preinvex, but the converse does not hold. The main tool which allows a generalization of the previous results is the following lemma, which is in turn a generalization of Lemma 1.

Lemma 3. Let $S \subset \mathbb{R}^n$ be a nonempty set and let $f : S \longrightarrow \mathbb{R}^m$ be preinvex on S, with respect to $\eta(\cdot, \cdot)$, i. e. each component of f is preinvex on S, with respect to the *same* vector function $\eta(\cdot, \cdot)$. Then, either the system

$$\begin{cases} f_1(x) < 0 \\ \quad\vdots \\ f_m(x) < 0 \end{cases}$$

has a solution $x \in S$, or it holds

$$\lambda_1 f_1(x) + ... + \lambda_m f_m(x) \geqq 0, \quad \forall x \in S,$$

with $\lambda_1 \geqq 0, ..., \lambda_m \geqq 0$, *not all zero,* but never both.

Without any convexity (or generalized convexity) assumptions on the directional derivatives of the functions involved in (P_1), it is not in general possible to obtain multipliers rules of the Fritz John-type or of the Karush-Kuhn-Tucker-type for (P_1). However, it is possible to obtain other types of multipliers rules. An interesting result has been obtained by Craven (2000), by means of *variable* multipliers, i. e. multipliers which depend on the directions $v \in \mathbb{R}^n$. See also the papers of Dinh, Lee and Tuan (2005) and Dinh and Tuan (2003). More precisely, we have the following necessary optimality conditions for (P_1).

Theorem 52. Let be given problem (P_1), where f and every g_i, $i = 1, ..., m$, admit finite right-sided directional derivatives at $x^0 \in K$; moreover, let every g_i, $i \notin I(x^0)$, be continuous at x^0. If $x^0 \in K$ is a local solution of (P_1), then for each direction $v \in \mathbb{R}^n$, there exist multipliers $\lambda_0(v) \geqq 0$, $\lambda_1(v) \geqq 0, ..., \lambda_m(v) \geqq 0$, *not all zero,* such that

$$\begin{aligned} &\lambda_0(v) D^+ f(x^0; v) + \textstyle\sum_{i=1}^m \lambda_i(v) D^+ g_i(x^0; v) \geqq 0, \\ &\lambda_i(v) g_i(x^0) = 0, \quad i = 1, ..., m. \end{aligned}$$

We again point out that the multiplier vector $\lambda = [\lambda_0, \lambda_1, ..., \lambda_m] \in \mathbb{R}^{m+1}$ depends on the direction $v \in \mathbb{R}^n$, i. e. $\lambda(\cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}_+^{m+1} \diagdown \{0\}$, with $0 \in \mathbb{R}^{m+1}$.

Another approach which avoids convexity assumptions on the functions involved in (P_1) or on their directional derivatives, is offered by Jeyakumar (1987) by means of what he calls "approximately quasidifferentiable functions".

Definition 24. A function $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, X open set, is said to be *approximately quasidifferentiable at* $x^0 \in X$, if f is continuous in a neighborhood of x^0 and if there exists a convex compact subset $Qf(x^0) \subset \mathbb{R}^n$ such that

$$f^{D^+}(x^0; v) \leqq \max_{y \in Qf(x^0)} \left\{ y^\top v \right\}, \quad \forall v \in \mathbb{R}^n.$$

Jeyakumar calls $Qf(x^0)$ an *approximate quasidifferential of f at x^0*. It must be noted that an approximate quasidifferential at a point is not necessarily unique. It can be shown that the class of approximately quasidifferentiable functions strictly contains the class of locally Lipschitz functions and the class of quasidifferentiable functions (in the sense of Pshenichnyi). The author proves the following Fritz John-type theorem for (P_1).

Theorem 53. Consider problem (P_1), with $x^0 \in K$, X open set and with f and every g_i, $i = 1, ..., m$, approximately quasidifferentiable at x^0. Denote by $Qf(x^0)$ and by $Qg_i(x^0)$, $i = 1, ..., m$, the related approximate quasidifferentials. If x^0 is a local solution of (P_1), then there exist multipliers $\lambda_0 \geqq 0, \lambda_1 \geqq 0, ..., \lambda_m \geqq 0$, *not all zero,* such that

$$0 \in \lambda_0 Qf(x^0) + \sum_{i=1}^{m} \lambda_i Qg_i(x^0);$$

$$\lambda_i g_i(x^0) = 0, \quad i = 1, ..., m.$$

Under an appropriate constraint qualification, then Jeyakumar (1987) obtain also for (P_1) a Karush-Kuhn-Tucker-type theorem. In the same paper also suffcient optimality conditions, in terms of approximate quasidifferentials, are obtained. For further comments on the various differentiability assumptions used to obtain multiplier rules for a nonlinear programming problem, see, e. g., Blot (2016), Fernandez (1997), Giorgi and Zuccotti (2016), Halkin (1974), Pourciau (1980).

Directional derivatives are also important in the study of stability and sensitivity of a *parametric* nonlinear programming problem. Also for these topics the literature is abundant. We recommend the classical books of Fiacco (1983a), and Shimitzu, Ishizuka and Bard (1997), and the more advanced book of Bonnans and Shapiro (2000). The interested reader may consult the following papers: Fiacco (1983b), Fiacco and Hutzler (1982), Gauvin and Dubeau (1982, 1984), Gauvin and Tolle (1977), Gauvin and Janin (1990), Gollan (1984), Hogan (1973), Janin (1984), Kaul (1985), Rockafellar (1984), Ralph and Dempe (1995).

We report only the following classical result, which maybe goes back to Uzawa (1958). See also Gauvin (1980) and Horst (1984a,b).

Let us consider the nonlinear parametric programming problem

$$\begin{cases} \max f(x) \\ \text{subject to: } g_i(x) \leqq b_i,\ i = 1, ..., m, \\ x \in X \subset \mathbb{R}^n, \end{cases}$$

where $f : X \longrightarrow \mathbb{R}$ is *concave* on the open convex set X and every g_i, $i = 1, ..., m$, is *convex* on the same set X.

It can be shown that the *optimal value function* (or *marginal function*)

$$\varphi(b) \equiv \max \{f(x) : g_i(x) \leqq b_i,\ i = 1, ..., m\}$$

is concave on its domain

$$\mathfrak{B} \equiv \left\{ b \in \mathbb{R}^m : \exists x(b) \text{ such that } f(x(b)) \equiv \max_x \{f(x) \mid g(x) \leqq b\} \right\}$$

and that, under the said assumptions, the set

$$\mathfrak{B}_0 \equiv \{b \in \mathbb{R}^m : \exists \bar{x} \in \mathbb{R}^n \text{ such that } g(\bar{x}) < b\}$$

is open and convex. Moreover, it holds, for $b \in \mathfrak{B} \cap \mathfrak{B}_0$,

$$\left[\frac{\partial \varphi}{\partial b_i}(b) \right]_+ \leqq y_i(b) \leqq \left[\frac{\partial \varphi}{\partial b_i}(b) \right]_- ,$$

where $\left[\frac{\partial \varphi}{\partial b_i}(b) \right]_+$ and $\left[\frac{\partial \varphi}{\partial b_i}(b) \right]_-$ are, respectively, the right-sided and the left-sided partial derivative of φ with respect to b_i, and $y_i(b)$ is the i-th Lagrangian multiplier associated to the optimal value $x(b)$ which satisfies the saddle point conditions of the Lagrangian function.

6. The Axiomatic Approach of K.-H. Elster and J. Thierfelder to Nonsmooth Optimization

As we have previously mentioned, starting from the 60's and 70's of the last century, several mathematicians have studied the possibility to generalize the classical concepts of differentiability (Gâteaux, Fréchet, Hadamard, etc.) in order to treat problems described by nonsmooth functions. Besides the approaches of Rockafellar (1970) to convex functions and of Pshenichnyi (1971), mentioned previously, it is worth mentioning the approaches of Clarke (1983), Demyanov and Rubinov (1995), Mordukhovich (2006) and Rockafellar (1980, 1981), that will not be treated in the present paper.

The variety of the various approaches, proposed to study nonsmooth functions and nonsmooth optimization problems, has led to define axiomatic constructions which include, as particular cases, several of the said above approaches. We briefly examine the axiomatic approach of K.-H. Elster and J. Thierfelder (1985, 1988a,b, 1989), but we point out also the interesting approaches of Giannessi (1989, 2005) and Komlosi and Pappalardo (1994).

The approach of Elster and Thierfelder is based on an axiomatic definition of *local cone approximation* of a set at a point. In a previous section we have introduced and used various local cone approximations. Elster and Thierfelder give the following general axiomatic definition (they consider a locally convex Hausdorff space, but we continue to consider the Euclidean space $\mathbb{R}^n$). See also the papers of Giorgi and Guerraggio (2002), Ioffe (1986) and Ward (1987, 1988, 1989).

Definition 25. A map $K : 2^{\mathbb{R}^n} \times \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ is a *local cone approximation* if for each set $S \subset \mathbb{R}^n$ and each point $x^0 \in \mathbb{R}^n$ a cone $K(S, x^0)$ is associated such as the following properties are fulfilled:

1. $K(S, x^0) = K(S - x^0, 0)$;
2. $K(S \cap N(x^0, \varepsilon), x^0) = K(S, x^0)$, $\forall \varepsilon > 0$;
3. $K(S, x^0) = \varnothing$, $\forall x^0 \notin cl(S)$;
4. $K(S, x^0) = \mathbb{R}^n$, $\forall x^0 \in int(S)$;
5. $K(\varphi(S), \varphi(x^0)) = \varphi(K(S, x^0))$, for any linear homeomorphism $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$;
6. $0^+S \subset 0^+K(S, x^0)$, $\forall x^0 \in cl(S)$, where

$$0^+S = \{y \in \mathbb{R}^n : a + ty \in S,\ \forall t > 0,\ \forall a \in S\}$$

is the *recession cone* of S (see Rockafellar (1970)). Moreover, we set $0^+\varnothing = \mathbb{R}^n$.

Theorem 54. The axioms 1.-6. are independent, i. e. for each axiom there exists a map $K(\cdot, \cdot)$ whic fails exactly the said axiom and satisfies the remaining axioms.

Almost all local cone approximations used in optimization theory verify the previous axioms. We give below a list of the most used local cone approximations which are a particular case of the axiomatic definition described above (some of these cones have already been presented and used in the present paper). We adopt the various descriptions in terms of neighborhoods.

Definition 26. Let be $S \subset \mathbb{R}^n$ and $x^0 \in \mathbb{R}^n$.

- The cone
$$F(S, x^0) = \big\{y \in \mathbb{R}^n : \exists \delta > 0,\ \forall t \in (0, \delta) : x^0 + ty \in S\big\}$$

is called *cone of feasible directions* to S at x^0.

- The cone
$$WF(S, x^0) = \big\{y \in \mathbb{R}^n : \forall \delta > 0\ \exists t \in (0, \delta) : x^0 + ty \in S\big\}$$

is called *cone of weakly feasible directions* or *radial tangent cone* to S at x^0.

- The cone

$$T(S, x^0) = \big\{y \in \mathbb{R}^n : \forall \delta > 0\ \exists \bar{y} \in N(y, \delta),\ \exists t \in (0, \delta) : x^0 + t\bar{y} \in S\big\}$$

is called *Bouligand tangent cone* or *contingent cone* to S at x^0.

- The cone

$$I(S,x^0)=\left\{y\in\mathbb{R}^n : \exists\delta>0,\ \forall\bar{y}\in N(y,\delta),\ \forall t\in(0,\delta) : x^0+t\bar{y}\in S\right\}$$

is called *cone of interior directions* or *cone of interior displacements* to S at x^0.

- The cone

$$A(S,x^0)=\left\{y\in\mathbb{R}^n : \forall N(y)\ \exists\lambda>0,\ \forall t\in(0,\lambda)\ \exists\bar{y}\in N(y) : x^0+t\bar{y}\in S\right\}$$

is called *cone of attainable directions* or *Kuhn-Tucker tangent cone* or *Ursescu tangent cone* to S at x^0.

- The cone

$$Q(S,x^0)=\left\{y\in\mathbb{R}^n : \exists N(y),\ \forall\lambda>0,\ \exists t\in(0,\lambda),\ \forall\bar{y}\in N(y) : x^0+t\bar{y}\in S\right\}$$

is called *cone of quasi-interior directions* to S at x^0.

- The cone

$$T^o(S,x^0)=\left\{\begin{array}{l} y\in\mathbb{R}^n : \forall N(y),\ \exists N(x^0),\ \exists\lambda>0,\ \forall t\in(0,\lambda), \\ \forall\bar{x}\in N(x^0)\cap S\cup\{x^0\},\ \exists\bar{y}\in N(y) : \bar{x}+t\bar{y}\in S \end{array}\right\}$$

is called *Clarke tangent cone* to S at x^0.

- The cone

$$H(S,x^0)=\left\{\begin{array}{l} y\in\mathbb{R}^n : \exists N(x^0),\ \exists\lambda>0,\ \forall t\in(0,\lambda), \\ \forall\bar{x}\in N(x^0)\cap S\cup\{x^0\},: \bar{x}+ty\in S \end{array}\right\}$$

is called *Rockafellar hypertangent cone* to S at x^0.

- The cone

$$E(S,x^0)=\left\{\begin{array}{l} y\in\mathbb{R}^n : \exists N(y),\ \exists N(x^0),\ \exists\lambda>0,\ \forall t\in(0,\lambda), \\ \forall\bar{x}\in N(x^0)\cap S\cup\{x^0\},\ \forall\bar{y}\in N(y) : \bar{x}+t\bar{y}\in S \end{array}\right\}$$

is called *cone of epi-Lipschitzian directions* to S at x^0

Remark 7. The descriptions of the cones $T^o(S,x^0)$, $H(S,x^0)$ and $E(S,x^0)$ are slightly different from the original definitions (see. e. g., Aubin and Frankowska (1990)), where the point $\bar{x}$ belongs to the set $S\cap N(x^0)$. The present description, taken from Elster and Thierfelder (1985, 1988a, b), allows to verify the third axiom of Definition 25. However, the consideration of the set $S\cap N(x^0)\cup\{x^0\}$ does not involve the original behaviour of the map. More precisely, Giorgi and Guerraggio (1992a) have shown that if $x^0\in cl(S)$, the descriptions given in Definition 26 for $T^o(S,x^0)$, $H(S,x^0)$ and $E(S,x^0)$ coincide with the original definitions.

For a quick overview of the main properties of the cones previously defined, it is useful the following scheme.

$$\begin{array}{ccccc} E(S,x^0) & \subset & I(S,x^0) & \subset & Q(S,x^0) \\ \cap & & \cap & & \cap \\ H(S,x^0) & \subset & F(S,x^0) & \subset & WF(S,x^0) \\ \cap & & \cap & & \cap \\ T^o(S,x^0) & \subset & A(S,x^0) & \subset & T(S,x^0) \end{array}$$

With regard to this scheme the following assertions hold true.

- The cones of the first row are open and it holds

$$x^0 \in int(S) \iff 0 \in K(S,x^0).$$

- The cones of the third row are closed and it holds

$$x^0 \in cl(S) \iff 0 \in K(S,x^0).$$

The cones of the second row verify the property

$$x^0 \in S \iff 0 \in K(S,x^0).$$

- The cones of the first column are *convex;* the cones of the second and third column are *isotone,* i. e.

$$S_1 \subset S_2 \implies K(S_1,x^0) \subset K(S_2,x^0),\ \forall x^0 \in \mathbb{R}^n.$$

By means of the axiomatic characterization of a local cone approximation, always following Elster and Thierfelder (1988a,b), but see also Ward (1987, 1988, 1989), it is possible to give the following definition of *generalized directional derivative.*

Definition 27. Let be $f : \mathbb{R}^n \longrightarrow [-\infty,+\infty]$, $x^0 \in \mathbb{R}^n$ such that $|f(x^0)| < +\infty$ and $K(\cdot,\cdot)$ a local cone approximation, according to Definition 25. Then the function $f^K(x^0;\cdot) : \mathbb{R}^n \longrightarrow [-\infty,+\infty]$ defined by

$$f^K(x^0;y) = \inf\left\{\beta \in \mathbb{R} : (y,\beta) \in (epi\ f,(x^0 f(x^0)))\right\},\ \forall y \in \mathbb{R}^n,$$

is called the K*-directional derivative of* f at x^0. It is assumed $\inf(\varnothing) = +\infty$.

It is worth noting that Bazaraa and Goode (1973) were perhaps the first authors to notice the connection between the *Dini directional derivatives* and an appropriate local cone approximation of the epigraph of f at $(x^0, f(x^0))$. It is quite immediate to remark that $f^K(x^0;\cdot)$ is positively homogeneous. Moreover, it can be proved that the topological properties of the local cone approximation $K(\cdot,\cdot)$ are reflected on the K-directional derivatives, as described in the following theorem, due to Elster and Thierfelder (1988b).

Theorem 55. Let be $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x^0 \in \mathbb{R}^n$ and $K(\cdot,\cdot)$ a local cone approximation. Then:

(i) If $K(epi\ f, (x^0, f(x^0)))$ is convex, then $f^K(x^0;\cdot)$ is sublinear.

(ii) It holds

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

In particular, if $K(epi\ f,\ (x^0, f(x^0)))$ is closed, it holds

$$epi\ f^K(x^0;\cdot) = K(epi\ f, (x^0, f(x^0)))$$

and $f^K(x^0;\cdot)$ is lower semicontinuous.

(iii) It holds

$$epi^o\ f^K(x^0;\cdot) = \left\{(y,\beta) \in \mathbb{R}^n \times \mathbb{R} : \forall \varepsilon > 0,\ (y, \beta - \varepsilon) \in K(epi\ f,\ (x^0, f(x^0)))\right\},$$

where

$$epi^o\ f^K(x^0;y) = \left\{(y,\beta) : f^K(x^0;y) < \beta\right\}$$

is the *strict epigraph* of the K-directional derivative.

In particular, if $K(epi\ f,\ (x^0, f(x^0)))$ is open, it holds

$$epi^o\ f^K(x^0;y) = K(epi\ f, (x^0, f(x^0))$$

and $f^K(x^0;\cdot)$ is upper semicontinuous.

By means of Definition 27 it is possible to get a family of generalized directional derivatives. In particular, if we make use of the local cone approximations previously recalled, we obtain the following results. We use the following notations, taken from Rockafellar (1980, 1981):

$$(\bar{x},\alpha) \downarrow x^0 \iff (\bar{x},\alpha) \longrightarrow (x^0, f(x^0)) \text{ and } \alpha \geqq f(\bar{x});$$

$$(\bar{x},\alpha) \uparrow x^0 \iff (\bar{x},\alpha) \longrightarrow (x^0, f(x^0)) \text{ and } \alpha \leqq f(\bar{x});$$

$$\bar{x} \longrightarrow_f x^0 \iff (\bar{x}, f(\bar{x})) \longrightarrow (x^0, f(x^0)).$$

Also the definitions of "lim sup inf" and "lim inf sup" operations are taken from Rockafellar (1980, 1981). Let $g : \mathbb{R}^n \longrightarrow [-\infty,+\infty]$ and $h : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow [-\infty,+\infty]$ extended real-valued functions. We have

$$\liminf_{\bar{y}\longrightarrow y} g(\bar{y}) = \sup_{U(y),}\inf_{\bar{y}\in U(y)} g(\bar{y});$$

$$\limsup_{\bar{y}\longrightarrow y} = \inf_{U(y),}\sup_{\bar{y}\in U(y)} g(\bar{y});$$

$$\limsup_{\bar{z}\longrightarrow z,}\inf_{\bar{y}\longrightarrow y} h(\bar{y},\bar{z}) = \sup_{U_1(y)}\inf_{U_2(y)}\sup_{\bar{z}\in U_2(z)}\inf_{\bar{y}\in U_1(y)} h(\bar{y},\bar{z});$$

$$\liminf_{\bar z \longrightarrow z,\ \bar y \longrightarrow y} \sup h(\bar y, \bar z) = \inf_{U_1(y)} \sup_{U_2(y)} \inf_{\bar z \in U_2(z)} \sup_{\bar y \in U_1(y)} h(\bar y, \bar z).$$

Let be $f : \mathbb{R}^n \longrightarrow [-\infty, +\infty]$ and $x^0 \in \mathbb{R}^n$. Then:

- The *lower Hadamard directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f_{H^+}(x^0; y) = f^T(x^0; y) = \liminf_{(\bar y, t) \longrightarrow (y, 0^+)} \frac{f(x^0 + t\bar y) - f(x^0)}{t}, \quad \forall y \in \mathbb{R}^n.$$

- The *upper Hadamard directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f^{H^+}(x^0; y) = f^I(x^0; y) = \limsup_{(\bar y, t) \longrightarrow (y, 0^+)} \frac{f(x^0 + t\bar y) - f(x^0)}{t}, \quad \forall y \in \mathbb{R}^n.$$

- The *lower Dini directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f_{D^+}(x^0; y) = f^{WF}(x^0; y) = \liminf_{t \longrightarrow 0^+} \frac{f(x^0 + ty) - f(x^0)}{t}, \quad \forall y \in \mathbb{R}^n.$$

- The *upper Dini directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f^{D^+}(x^0; y) = f^F(x^0; y) = \limsup_{t \longrightarrow 0^+} \frac{f(x^0 + ty) - f(x^0)}{t}, \quad \forall y \in \mathbb{R}^n.$$

- The *lower Ursescu directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f^A(x^0; y) = \limsup_{t \longrightarrow 0^+,\ \bar y \longrightarrow y} \inf \frac{f(x^0 + t\bar y) - f(x^0)}{t}, \quad \forall y \in \mathbb{R}^n.$$

- The *upper Ursescu directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f^Q(x^0; y) = \liminf_{t \longrightarrow 0^+,\ \bar y \longrightarrow y} \sup \frac{f(x^0 + t\bar y) - f(x^0)}{t}, \quad \forall y \in \mathbb{R}^n.$$

- The *Clarke generalized directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f^o(x^0; y) = f^H(x^0; y) = \limsup_{(\bar x, \alpha) \downarrow x^0,\ t \longrightarrow 0^+} \frac{f(\bar x + ty) - \alpha}{t}, \quad \forall y \in \mathbb{R}^n,$$

Here H is the hypertangent cone (do not make confusion with the upper Hadamard directional derivative!).

- The *Clarke-Rockafellar generalized directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f^{\uparrow}(x^0; y) = f^{T^o}(x^0; y) = \limsup_{\substack{(\bar{x}, \alpha) \downarrow x^0 \\ t \longrightarrow 0^+}} \inf_{\bar{y} \longrightarrow y} \frac{f(\bar{x} + t\bar{y}) - \alpha}{y}, \quad \forall y \in \mathbb{R}^n.$$

- The *epi-Lipschitzian directional derivative* at x^0 in the direction $y \in \mathbb{R}^n$ is

$$f^E(x^0; y) = \limsup_{\substack{(\bar{x}, \alpha) \downarrow x^0, \ \bar{y} \longrightarrow y \\ t \longrightarrow 0^+}} \frac{f(\bar{x} + t\bar{y}) - \alpha}{y}, \quad \forall y \in \mathbb{R}^n.$$

Remark 8. When f is lower semicontinuous, the convergence $(\bar{x}, \alpha) \downarrow x^0$ becomes simply $\bar{x} \longrightarrow_f x^0$ and, moreover, if f is continuous, it becomes $\bar{x} \longrightarrow x^0$.

If f is locally Lipschitz, then:

a)

$$f^o(x^0; y) = f^{\uparrow}(x^0; y) = \limsup_{x \longrightarrow x^0, \ t \longrightarrow 0^+} \frac{f(\bar{x} + ty) - f(\bar{x})}{t},$$

i. e. we obtain the usual definition of the Clarke directional derivative (see Clarke (1983)).

b)

$$f^A(x^0; \cdot) = f_D^+(x^0; y) = f_{H^+}(x^0; y).$$

c)

$$f^Q(x^0; y) = f^{D^+}(x^0; y) = f^{H^+}(x^0; y).$$

It follows that f is (right-sided) directionally differentiable at x^0 in the direction $y \in \mathbb{R}^n$ if and only if

$$f^{WF}(x^0; \cdot) = f^F(x^0; \cdot).$$

Moreover, f is *Gâteaux differentiable* at x^0 in the direction $y \in \mathbb{R}^n$ if and only if

$$f^{WF}(x^0; \cdot) = f^F(x^0; \cdot)$$

is linear.

Similarly to the inclusion scheme concerning the various local cone approximations, we obtain the following scheme showing the relationships between the various generalized directional derivatives previously considered.

$$\begin{array}{ccccc} f^E(x^0; \cdot) & \geqq & f^{H^+}(x^0; \cdot) & \geqq & f^Q(x^0; \cdot) \\ \begin{array}{c}\|\!\vee\end{array} & & \begin{array}{c}\|\!\vee\end{array} & & \begin{array}{c}\|\!\vee\end{array} \\ f^o(x^0; \cdot) & \geqq & f^{D^+}(x^0; \cdot) & \geqq & f_{D^+}(x^0; \cdot) \\ \begin{array}{c}\|\!\vee\end{array} & & \begin{array}{c}\|\!\vee\end{array} & & \begin{array}{c}\|\!\vee\end{array} \\ f^{\uparrow}(x^0; \cdot) & \geqq & f^A(x^0; \cdot) & \geqq & f_{H^+}(x^0; \cdot) \end{array}$$

The following assertions hold true.

1) The directional derivatives of the first row of the scheme are upper semicontinuous and it holds

$$f^K(x^0;0) \geqq 0.$$

The directional derivatives of the third row of the scheme are lower semicontinuous and it holds

$$f^K(x^0;0) \leqq 0.$$

For the directional derivatives of the second row of the scheme it holds

$$f^K(x^0;0) = 0.$$

2) The directional derivatives of the first column of the scheme are convex (more precisely: sublinear). The directional derivatives of the second and third column are *isotone,* in the sense that they verify the following property:

$$\left.\begin{array}{l} f_1(\cdot) \leqq f_2(\cdot) \\ f_1(x^0) = f_2(x^0) \end{array}\right\} \Longrightarrow f_1^K(x^0;\cdot) \leqq f_2^K(x^0;\cdot).$$

In a similar way with respect to the definition of K-directional derivative, it is possible to introduce the concept of *K-subdifferential.*

Definition 28. Let be $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x^0 \in \mathbb{R}^n$ and $K(x^0, y)$ be a local cone approximation. The set (possibly empty)

$$\partial^K f(x^0) = \left\{\xi \in \mathbb{R}^n : f^K(x^0;y) \geqq \xi^\top y,\ \forall y \in \mathbb{R}^n\right\}$$

is said the *K-subdifferential* of f at x^0 and the elements $\xi \in \partial^K f(x^0)$ are said the *K-subgradients* of f at x^0.

Note that $0 \in \partial^K f(x^0)$ if and only if $f^K(x^0;y) \geqq 0$, $\forall y \in \mathbb{R}^n$. When $\partial^K f(x^0) \neq \varnothing$, then $\partial^K f(x^0)$ is a closed and convex set. As the K-directional derivative of f is directly related to the local cone approximation $K(\cdot,\cdot)$ of its epigraph, something similar holds true also for the K-subdifferential.

Theorem 56. Let be $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x^0 \in \mathbb{R}^n$ and $K(x^0, \cdot)$ a local cone approximation. Then it holds

$$\partial^K f(x^0) = \left\{\xi \in \mathbb{R}^n : (\xi - 1) \in K^*(epi\ f, (x^0, f(x^0)))\right\},$$

where K^* is the polar cone of K.

Proof. We have the following chain of equivalences:

$$x^0 \in \partial^K f(x^0) \iff \inf\left\{\beta \in \mathbb{R} : (y,\beta) \in K(epi\ f, (x^0, f(x^0)))\right\} \geqq$$

$$\geqq \xi^\top y,\ \forall y \in \mathbb{R}^n \iff \beta \geqq \xi^\top y,\ \forall (y,\beta) \in K(epi\ f, (x^0, f(x^0))) \iff$$

$$\iff (\xi, -1)^\top (y,\beta) \leqq 0,\ \forall (y,\beta) \in K(epi\ f, (x^0, f(x^0))) \iff$$

$$\iff (\xi, -1) \in K^*(epi\ f, (x^0, f(x^0))).$$

□

Now we consider briefly some optimality conditions expressed in terms of K-directional derivatives. We begin with an *unconstrained minimization problem.*

Theorem 57. Let be $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ and let $x^0 \in int(X)$ be a local minimum point of f over X. If $K(\cdot,\cdot)$ is any local cone approximation such that $K(\cdot,\cdot) \subset T(\cdot,\cdot)$, then it holds

i) $f^K(x^0; y) \geqq 0$, $\forall y \in \mathbb{R}^n$;

ii) $0 \in \partial^K f(x^0)$.

Proof.

i) Let us assume $f^K(x^0; y) < 0$ for a vector $y \in \mathbb{R}^n$. Then, because $K(\cdot,\cdot) \subset T(\cdot,\cdot)$, we have

$$f^T(x^0; y) \leqq f^K(x^0; y) < 0$$

and hence

$$\liminf_{\bar{y} \longrightarrow y,\ t \longrightarrow 0^+} \frac{f(x^0 + t\bar{y}) - f(x^0)}{t} < 0,$$

which means $\forall N(y)$, $\forall \lambda > 0$, $\exists t \in (0, \lambda)$, $\exists \bar{y} \in N(y)$:

$$\frac{f(x^0 + t\bar{y}) - f(x^0)}{t} < 0,$$

which contradicts the assumption that x^0 is an unconstrained local minimum point of f.

ii) The assertion follows from Definition 28. □

For example, we have, under the assumptions of Theorem 57,

$$f^F(x^0; y) = f^{D^+}(x^0; y) = \limsup_{t \longrightarrow 0^+} \frac{f(x^0 + ty) - f(x^0)}{t} \geqq 0,\ \forall y \in \mathbb{R}^n,$$

or also the sharper condition

$$f^{WF}(x^0; y) = f_{D^+}(x^0; y) = \liminf_{t \longrightarrow 0^+} \frac{f(x^0 + ty) - f(x^0)}{t} \geqq 0,\ \forall y \in \mathbb{R}^n,$$

or also, in terms of upper Hadamard directional derivatives,

$$f^I(x^0; y) = f^{H^+}(x^0; y) = \limsup_{(\bar{y},t) \longrightarrow (y,0^+)} \frac{f(x^0 + t\bar{y}) - f(x^0)}{t} \geqq 0,\ \forall y \in \mathbb{R}^n,$$

or also the sharper condition

$$f^T(x^0; y) = f_{H^+}(x^0; y) = \liminf_{(\bar{y},t)\longrightarrow(y,0^+)} \frac{f(x^0 + t\bar{y}) - f(x^0)}{t} \geqq 0, \ \forall y \in \mathbb{R}^n.$$

Now we consider a minimization problem of the type (P_0), i. e. with a set constraint, and of type (P_1), i. e. with inequality constraints. First we introduce the following sets.

- $$D_f^K(x^0) = \left\{ y \in \mathbb{R}^n : f^K(x^0; y) < 0 \right\}$$

is the *cone of descent directions of* f *at* x^0.

- $$C_f^K(x^0) = \left\{ y \in \mathbb{R}^n : f^K(x^0; y) \leqq 0 \right\}$$

is the *linearizing cone of* f *at* x^0.

- $$D_M^K(x^0) = \underset{i \in M}{\cap} D_{g_i}^K(x^0);$$

- $$C_M^K(x^0) = \underset{i \in M}{\cap} C_{g_i}^K(x^0),$$

where $M = \{1, ..., m\}$.

Obviously these cones are convex, if $K(x^0; \cdot)$ is convex.

In the following we assume, when it is necessary, that the local cone approximation $K(\cdot, \cdot)$ satisfies the following conditions.

(A_1) K is convex and closed.

(A_2) $z \in S \iff 0 \in K(S, z)$.

(A_3) $K(\cdot, \cdot) \subset T(\cdot, \cdot)$.

(A_4) $int(K(\cdot, \cdot)) \subset I(\cdot, \cdot)$.

Let us consider problem (P_0) :

$$\min f(x), \quad x \in S \subset \mathbb{R}^n.$$

Theorem 58. If $x^0 \in S$ is a local solution of (P_0) and $K(\cdot, \cdot)$ satisfies conditions (A_3) and (A_4), then

i)

$$D_f^{int(K)}(x^0) \cap K(S, x^0) = \varnothing;$$

ii)

$$D_f^K(x^0) \cap int(K(S, x^0)) = \varnothing.$$

Theorem 59. If $x^0 \in S$ is a local solution of (P_0), if $(A_1), (A_3)$ and (A_4) are verified, and if one of the following conditions is verified:

(B_1) $\quad dom\ f^{int(K)}(x^0; \cdot) \cap K(S, x^0) = \varnothing;$

(B_2) $\quad dom\ f^K(x^0; \cdot) \cap int(K(S, x^0)) = \varnothing,$

then it holds

$$0 \in \partial^K f(x^0) + K^*(s, x^0).$$

Now let us consider problem (P_1), i. e.

$$(P_1): \qquad \begin{cases} \min f(x) \\ \text{subject to: } g_i(x) \leqq 0,\ i = 1, ..., m, \\ x \in X \subset \mathbb{R}^n, \end{cases}$$

with X open set of $\mathbb{R}^n$. In order to avoid confusion with the cones $K(\cdot, \cdot)$, we denote by S_1 the feasible set of (P_1). Elster and Thierfelder (1988b) obtain for (P_1) the following Karush-Kuhn-Tucker-type necessary optimality conditions.

Theorem 60. Let $x^0 \in S_1$ be a local solution of (P_1) and let $(A_1), (A_3), (A_4)$, either (B_1) or (B_2) be verified. Moreover, the following constraint qualification is satisfied:

$$(CQ)_1: \qquad K^*(S_4, x^0) \subset B^K_{I(x^0)}(x^0),$$

where

$$B^K_{I(x^0)}(x^0) = \left\{ \xi \in \mathbb{R}^n : \xi = \sum_{i \in I(x^0)} \lambda_i \xi^i, \quad \lambda_i \geqq 0,\ \xi^i \in \partial^K g_i(x^0),\ i \in I(x^0) \right\}$$

is the *cone of K-gradients* of g_i, $i \in I(x^0)$, at x^0.

Then, there exist multipliers $\lambda_i \geqq 0$, $i \in I(x^0)$, such that

i)

$$0 \in \partial^K f(x^0) + \sum_{i \in I(x^0)} \lambda_i \partial^K g_i(x^0);$$

ii)

$$f^K(x^0; y) + \sum_{i \in I(x^0)} \lambda_i g_i^K(x^0; y) \geqq 0,\ \forall y \in \mathbb{R}^n.$$

Let us now consider the following further constraint qualifications ($x^0 \in S_1$).

- $(CQ)_2$. *Generalized Guignard-Gould-Tolle constraint qualification*:

$$(K(S_4, x^0))^* \subset (C^K_{I(x^0)}(x^0))^*,$$
$$\partial^K g_i(x^0) \neq \varnothing,\ \forall i \in I(x^0),$$

$$B^K_{I(x^0)}(x^0) \text{ closed, either } (B_1) \text{ or } (B_2).$$

- $(CQ)_3$. *Generalized Abadie constraint qualification*:

$$\begin{aligned} C^K_{I(x^0)}(x^0) &\subset K(S_4, x^0), \\ \partial^K g_i(x^0) &\neq \varnothing,\ \forall i \in I(x^0), \end{aligned}$$

$$B^K_{I(x^0)}(x^0) \text{ closed, either } (B_1) \text{ or } (B_2).$$

- $(CQ)_4$. *First generalized Slater constraint qualification:*

$$D^{int(K)}_{I(x^0)}(x^0) \neq \varnothing,\quad \partial^K g_i(x^0) \neq \varnothing,\ \forall i \in I(x^0),$$

$$B^K_{I(x^0)}(x^0) \text{ closed, either } (B_1) \text{ or } (B_2).$$

$(CQ)_5$. *Second generalized Slater constraint qualification:*

$$dom\ f^K(x^0, \cdot) \cap D^{int(K)}_{I(x^0)}(x^0) \neq \varnothing,$$

$$\partial^K g_i(x^0) \neq \varnothing,\ \forall i \in I(x^0),\ B^K_{I(x^0)}(x^0) \text{ closed.}$$

We have the following result.

Theorem 61. (Elster and Thierfelder (1988b)). Let be $x^0 \in S_1$ and let conditions $(A_1)-(A_4)$ be verified. Moreover, let be verified the condition

$$(A_5): \qquad D^K_{I(x^0)}(x^0) \subset K(S_4, x^0).$$

Then we have the following implications:

$$(CQ)_5 \Longrightarrow (CQ)_4 \Longrightarrow (CQ)_3 \Longrightarrow (CQ)_2 \Longrightarrow (CQ)_1.$$

The same authors obtain also Fritz John-type optimality conditions for (P_1) in trems of K-directional derivatives.

Theorem 62. Let $x^0 \in S_1$ be a local solution of (P_1) and let the conditions $(A_1)-(A_5)$ be verified. Then:

(i) There exist multipliers $\lambda_i \geqq 0$, $i \in \{0\} \cup I(x^0)$, *not all zero,* such that

$$\lambda_0 f^{int(K)}(x^0; y) + \sum_{i \in I(x^0)} \lambda_i g_i^K(x^0; y) \geqq 0,$$

$$\forall y \in dom\ f^{int(K)}(x^0;\cdot) \bigcap_{i\in I(x^0)} dom\ g_i^K(x^0;\cdot).$$

(ii) There exist multipliers $\lambda'_i \geqq 0$, $i \in \{0\} \cup I(x^0)$, *not all zero,* such that

$$\lambda'_0 f^K(x^0;y) + \sum_{i\in I(x^0)} \lambda'_i g_i^{int(K)}(x^0;y) \geqq 0,$$

$$\forall y \in dom\ f^K(x^0;\cdot) \bigcap_{i\in I(x^0)} dom\ g_i^{int(K)}(x^0;\cdot).$$

Under other appropriate conditions, the same authors obtain the following version of the Fritz John necessary optimality conditions for (P_1) :

- There exist multipliers $u_0 \geqq 0$, $u_i \geqq 0$, *not all zero,* such that

$$u_0 f^K(x^0;y) + \sum_{i\in I(x^0)} u_i g_i^K(x^0;y) \geqq 0, \quad \forall y \in \mathbb{R}^n;$$

$$0 \in u_0 \partial^K f(x^0) + \sum_{i\in I(x^0)} u_i \partial^K g_i(x^0).$$

Under an appropriate constraint qualification, it is possible to obtain $u_0 \neq 0$, i. e. $u_0 = 1$, in the above conditions.

7. Applications to Vector Optimization Problems

The literature on applications of Nonsmooth Analysis and classical directional derivatives to Vector Optimization Problems is huge. We quote only the following papers for what concerns "classical directional derivatives" (usual directional derivatives, Dini and Hadamard directional derivatives): Aggarwal and Bhatia (1989), Bector, Bhatia and Jain (1993), Bhatia and Aggarwal (1992), Giorgi, Jimenez and Novo (2004), Ginchev, Guerraggio and Rocca (2006), Jimenez and Novo (2002a, b, 2003, 2008), Jimenez, Novo and Sama (2009), Ishizuka (1992), Kaul, Lyall and Kaur (1988), Mukherjee and Mishra (1996), Mukherjee and Singh (1990), Novo and Jimenez (2004), Preda and Stancu-Minasian (1997), Preda and Chitescu (1999), Luu and Nguyen (2009), Khan and Tuan (2007), Nguyen and Luu (2007), Taa (1999).

The main reference books on vector optimization are Ehrgott (2005), Jahn (2005), Luc (1989), Miettinen (1999), Sawaragi, Nakayama and Tanino (1985). Here we give only some hints on necessary optimality conditions for a vector optimization problem, expressed by means of "classical" directional derivatives.

We consider the following multiobjective (or Pareto) nonlinear programming problem with equality and inequality constraints.

$$\min f(x), \quad \text{subject to } x \in S, \tag{9}$$

where

$$S = \{x \in \mathbb{R}^n : g(x) \leqq 0,\ h(x) = 0\},$$

and $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}^r$, $r < n$.

We denote by

$$f_i,\ i \in I = \{1, 2, ..., p\};$$
$$g_j,\ j \in J = \{1, 2, ..., m\};$$
$$h_k,\ k \in K = \{1, 2, ..., r\},$$

the components of the functions f, g and h, respectively. The set

$$J_0 = \left\{j \in J : g_j(x^0) = 0\right\}$$

is the set of the active indices of g at $x^0 \in S$.

We denote

$$G = \{x \in \mathbb{R}^n : g(x) \leqq 0\};$$
$$H = \{x \in \mathbb{R}^n : h(x) = 0\};$$
$$S = G \cap H$$

and we consider the following conditions:

• $(H1)$ f and g are *Hadamard directionally differentiable at* $x^0 \in S$, *with convex derivative* (i. e. each component of f and g is Hadamard directionally differentiable at $x^0 \in S$, with convex derivative).

• $(H2)$ h is *Fréchet differentiable at* $x^0 \in S$, with the Jacobian $\nabla h(x^0)$ having maximal row-rank (i. e. $\nabla h_k(x^0)$, $k \in K$, are linearly independent).

We recall that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is Hadamard directionally differentiable at x^0 if its Hadamard directional derivative

$$D^H f(x^0; v) = \lim_{(t,u) \longrightarrow (0^+, v)} \frac{f(x^0 + tu) - f(x^0)}{t}$$

exists for all directions $v \in \mathbb{R}^n$.

We recall that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is (right-sided) directionally differentiable at x^0 if $D^+ f(x^0; v)$ exists for all directions $v \in \mathbb{R}^n$. Moreover (see Section 2):

- If f is Fréchet differentiable at x^0, then $\nabla f(x^0)^\top v = D^H f(x^0; v)$.
- If there exists $D^H f(x^0; v)$, then there exists $D^+ f(x^0; v)$ and both derivatives are equal.
- In particular, if f is locally Lipschitz at x^0 and there exists $D^+ f(x^0; v)$, then there exists $D^H f(x^0; v)$.
- If f is Hadamard directionally differentiable at x^0, then f is continuous at x^0 and $D^H f(x^0; \cdot)$ is continuous over $\mathbb{R}^n$. This property is not true for $D^+ f(x^0; v)$.

Some authors (e. g. Penot (1978)) have introduced the notion of *Dini subdifferential.*

Definition 29. Let $f:\mathbb{R}^n \longrightarrow \mathbb{R}$ be directionally differentiable at x^0. The *Dini subdifferential of f at x^0* is

$$\partial_D f(x^0) = \left\{\xi \in \mathbb{R}^n : \xi^\top v \leqq D^+ f(x^0; v), \quad \forall v \in \mathbb{R}^n\right\}.$$

If $D^+ f(x^0;\cdot)$ is a convex function, then there exists the Dini subdifferential. If $D^+ f(x^0;\cdot)$ is not a convex function, then $\partial_D f(x^0)$ may be the empty set.

We recall that, given problem (9), a point $x^0 \in S$ is said to be a *local weak efficient point* for the same problem (or also a *local weak Pareto solution*), if there exists a neighborhood $B(x^0,\delta)$ of x^0 such that

$$S_f \cap S \cap B(x^0,\delta) = \varnothing,$$

where

$$S_f = \left\{x \in \mathbb{R}^n : f(x) < f(x^0)\right\}.$$

In other words,

$$f(x) \notin f(x^0) - int(\mathbb{R}^p_+), \quad \forall x \in S \cap B(x^0,\delta)$$

or

$$f(x) \in f(x^0) + \mathbb{R}^p \diagdown - int(\mathbb{R}^p_+), \quad \forall x \in S \cap B(x^0,\delta).$$

We define also the *strict critical cone* for the objective function

$$C_0(f,x^0) = \left\{v \in \mathbb{R}^n : D^H f_i(x^0;v) < 0, \quad \forall i \in I\right\}$$

and the strict critical cone for the set of inequality constraints

$$C_0(G,x^0) = \left\{v \in \mathbb{R}^n : D^H g_j(x^0;v) < 0, \quad \forall j \in J_0\right\}.$$

The following two results are essential in obtaining a Fritz John-type multiplier rule for problem (9).

Theorem 63. (Jimenez and Novo (2002b). Under the assumptions $(H1)$ and $(H2)$, we have

$$C_0(G,x^0) \cap \ker(\nabla h(x^0)) \subset T(S,x^0).$$

(Here $T(S,x^0)$ is, as usual, the Bouligand tangent cone to S at $x^0 \in S$).

Theorem 64. (Jimenez and Novo (2002a)). Let us suppose that $\varphi_1,\varphi_2,...,\varphi_q : \mathbb{R}^n \longrightarrow \mathbb{R}$ are sublinear functions and $\psi_1,\psi_2,...,\psi_r : \mathbb{R}^n \longrightarrow \mathbb{R}$ are linear functions given by $\psi_k(u) = c^k u$, $k \in K = \{1,2,...,r\}$. Then, one and only one of the following assertions is true.

(a) There exists $v \in \mathbb{R}^n$ such that

$$\begin{cases} \varphi_i(x^0;v) < 0, \quad \forall i = 1,2,...,q; \\ \psi_k(v) = 0, \quad \forall k = 1,2,...,r. \end{cases}$$

(b) There exists $(\xi, \nu) = (\xi_1, \xi_2, ..., \xi_q, \nu_1, \nu_2, ..., \nu_r) \in \mathbb{R}^{q+r}$, $\xi \neq 0, \xi \geqq 0$, such that

$$0 \in \sum_{i=1}^{q} \xi_i \partial\varphi_i(0) + \sum_{k=1}^{r} \nu_k c^k.$$

(Here $\partial\varphi(\cdot)$ is the usual subdifferential of Convex Analysis).

Finally, we recall a classical first-order necessary optimality condition for a vector problem of the type (9), where the constraint set S is not specified. See, e. g., Taa (1999).

Theorem 65. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ be Hadamard directionally differentiable at $x^0 \in S$. If x^0 is a local weak efficient point for f over S, then

$$T(S, x^0) \cap C_0(f, x^0) = \varnothing.$$

Proof. Let be $v \in T(S, x^0)$; then there exist sequences $\{t_n\} \longrightarrow 0^+$, $\{v^n\} \longrightarrow v$ such that $x^0 + t_n v^n \in S$ for all n. Since x^0 is a local efficient solution of our problem, then there exists an integer n_0 such that for all $n \geqq n_0$,

$$f(x^0 + t_n v^n) \in f(x^0) + \mathbb{R}^p \diagdown - int(\mathbb{R}^p_+).$$

Then by our assumptions it follows that

$$(D^H f_1(x^0; v), ..., D^H f_p(x^0; v)) \in \mathbb{R}^p \diagdown - int(\mathbb{R}^p_+),$$

i. e. $C_0(f, x^0) \cap T(S, x^0) = \varnothing$. □

We are now ready to prove the main result of the present section.

Theorem 66. Let us consider problem (9) and assume that the previous conditions $(H1)$ and $(H2)$ are satisfied. If x^0 is a local weak efficient solution of (9), then there exists $(\lambda, \mu, \nu) \in \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^r$ such that

$$(\lambda, \mu) \geqq 0, \quad (\lambda, \mu) \neq 0, \tag{10}$$

$$0 \in \sum_{i=1}^{p} \lambda_i \partial_D f_i(x^0) + \sum_{j=1}^{m} \mu_j \partial g_j(x^0) + \sum_{k=1}^{r} \nu_k \nabla h_k(x^0), \tag{11}$$

$$\mu_j g_j(x^0) = 0, \quad j = 1, ..., m. \tag{12}$$

If, in addition, $C_0(S, x^0) \neq \varnothing$, then $\lambda \neq 0$.

Proof. As x^0 is a local weak efficient point for (9), we have (Theorem 65)

$$T(S, x^0) \cap C_0(f, x^0) = \varnothing, \tag{13}$$

i. e. there exists no $v \in \mathbb{R}^n$ such that

$$\begin{cases} D^H f_i(x^0; v) < 0, & \forall i \in I \\ v \in T(S, x^0). \end{cases} \tag{14}$$

Now, from Theorem 63 we have

$$C_0(G, x^0) \cap \ker(\nabla h(x^0)) \subset T(S, x^0).$$

So, taking (14) into account, there exists no $v \in \mathbb{R}^n$ such that

$$\begin{cases} D^H f_i(x^0; v) < 0, & \forall i \in I, \\ D^H g_j(x^0; v) < 0, & \forall j \in J_0, \\ \nabla h_k(x^0)^\top v = 0, & \forall k \in K \end{cases} \tag{15}$$

and using Theorem 64 the conclusion follows by choosing $\mu_j = 0$ for $j \notin J_0$.

For the second part, let us suppose that $C_0(S, x^0) \neq \varnothing$, that is there exists $w \in \mathbb{R}^n$ such that

$$D^H g_j(x^0; w) < 0, \ \forall j \in J_0; \quad \nabla h_k(x^0)^\top w = 0, \quad \forall k \in K. \tag{16}$$

Assume that $\lambda = 0$. The conditions (10)-(12) imply that

$$\sum_{j \in J_0} \mu_j D^H g_j(x^0; u) + \sum_{k=1}^{r} \nu_k \nabla h_k(x^0)^\top u \geqq 0, \ \forall u \in \mathbb{R}^n,$$

with $\mu \neq 0$. For $u = w$ we have a contradiction, since from (16) it follows that

$$\sum_{j \in J_0} \mu_j D^H g_j(x^0; w) + \sum_{k=1}^{r} \nu_k \nabla h_k(x^0)^\top w < 0.$$

Consequently $\lambda \neq 0$. □

Remark 9. It is also possible to obtain the thesis of Theorem 66 under the assumptions $(H1)'$ and $(H2)$, where:

- $(H1)'$: each component of f is Hadamard directionally differentiable at $x^0 \in S$ and for each $j \in J$, g_j is either Dini-quasiconvex and continuous on a neighborhood of x^0, with convex derivative or Fréchet differentiable at x^0.

See Novo and Jimenez (2004). Here, a function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ is Dini-quasiconvex at $x^0 \in X \subset \mathbb{R}^n$, X convex set, if for every $x \in X$,

$$\varphi(x) \leqq \varphi(x^0) \Longrightarrow D^+\varphi(x^0; x - x^0) \leqq 0.$$

Finally, we point out that Theorem 3.2 of Preda and Chitescu (1999) is not correct, as shown by Giorgi, Jimenez and Novo (2004). Also the paper of Mukherjee and Mishra (1996) contains some errors, as shown by Yang (1994).

References

S. AGGARWAL and D. BHATIA (1989), *Pseudolinearity and efficiency via Dini derivatives,* Indian J. Pure and Appl. Math., **20,** 1173-1183.

V. ALEXEEV, V. TIKHOMIROV and S. FOMINE (1982), Commande Optimale, Editions MIR, Moscou.

K. J. ARROW and A. C. ENTHOVEN (1961), *Quasi-concave programming,* Econometrica, **29,** 522-552.

J. P. AUBIN and A. CELLINA (1984), Differential Inclusions, Springer Verlag, Berlin.

J. P. AUBIN and F. FRANKOWSKA (1990), Set-valued Analysis, Birkhäuser, Boston.

V. I. AVERBUKH and O. G. SMOLYANOV (1967), *The theory of differentiation in linear topological spaces,* Russian Math. Surveys, **22,** 201-258.

V. I. AVERBUKH and O. G. SMOLYANOV (1968), *The various definitions of the derivative in linear topological spaces,* Russian Math. Surveys, **23,** 67-113.

A. BAGIROV, N. KARMITSA and M. M. MÄKELÄ (2014), Introduction to Nonsmooth Optimization, Springer, Heidelberg.

M. S. BAZARAA and J. J. GOODE (1973), *Extensions of optimality conditions via supporting functions,* Math. Programming, **5,** 267-285.

M. S. BAZARAA and C. M. SHETTY (1976), Foundations of Optimization, Springer-Verlag, Berlin.

C. R. BECTOR, D. BHATIA and P. JAIN (1993), *Generalized concavity and duality in multiobjective nonsmooth programming,* Utilitas Mathematica, **43,** 71-78.

A. BEN-ISRAEL (1969), *Linear equations and inequalities on finite dimensional, real or complex, vector spaces: a unified theory,* J. Math. Anal. Appl.,**27,** 367-389.

A. BEN-ISRAEL and B. MOND (1986), *What is invexity?,* Journal of the Australian Math. Soc., Ser. B, **28,** 1-9.

A. BEN-TAL and J. ZOWE (1985), *Directional derivatives in nonsmooth optimization,* J. Optim. Theory Appl., **47,** 483-490.

C. BERGE and A. GHOUILA-HOURI (1965), Programming, Games and Transportation Networks, John Wiley & Sons, New York.

D. P. BERTSEKAS (2009), Convex Optimization Theory, Athena Scientific, Belmont, Mass.

D. P. BERTSEKAS, A. NEDIC and A. E. OZDAGLAR (2003), Convex Analysis and Optimization, Athena Scientific, Belmont, Mass.

D. BHATIA and S. AGGARWAL (1992), *Optimality and duality for multiobjective nonsmooth programming,* European J. of Operational Research, **57,** 360-367.

J. BLOT (2016), *On the multiplier rules,* Optimization, **65,** 947-955.

M. W. BOISKO (1986), *A first derivative test for functions of several variables,* Amer. Math. Monthly, **93,** 558-561.

J. F. BONNANS and A. SHAPIRO (2000), Perturbation Analysis of Optimization Problems, Springer Verlag, New York..

J. M. BORWEIN and A. S. LEWIS (2000), Convex Analysis and Nonlinear Optimization, Springer, New York..

M. CASTELLANI (1997), *Second-order optimality conditions for nonsmooth mathematical programming,* Atti XXI Convegno Annuale A.M.A.S.E.S., Roma, 10-13 Settembre 1997, Università "La Sapienza", Università "Tor Vergata", "L.U.I.S.S. Guido Carli", Roma, 217-229..

M. CASTELLANI and M. PAPPALARDO (1995), *First-order cone approximations and necessary optimality conditions,* Optimization, **35,** 113-126..

Y. CHABRILLAC and J.-P. CROUZEIX (1987), *Continuity and differentiability properties of monotone real functions of several real variables,* Math. Programming Study, **30,** 1-16..

F. H. CLARKE (1983), Optimization and Nonsmooth Analysis, J. Wiley, New York..

R. CORREA and J.-B. HIRIART-URRUTY (1989), *A first order sufficient conditions for optimality in nonsmooth optimization,* Math. Nachr., **144,** 309-319..

B. D. CRAVEN (1986), *On quasidifferentiable optimization,* J. Australian Math. Soc. (Series A), **41,** 64-78..

B. D. CRAVEN (2000), *Lagrange multipliers for nonconvex optimization;* in X. Yang, A. I. Meese, M. Fischer and L. Jennings (Eds.), Progress in Optimization. Contributions from Australasia, Kluwer Academic Publishers, Dordrecht, Boston, London, 123-128..

B. D. CRAVEN and B. MOND (1979), *Lagrangean conditions for quasidifferentiable optimization;* in A, Prékopa (Ed.), Survey of Mathematical Programming, Vol. 1, North Holland, Amsterdam, 177-192..

J.-P. CROUZEIX (1981), *Some differentiability properties of quasiconvex functions on* $\mathbb{R}^n$; in A. Auslender, W. Oettli and J. Stoer (Eds.), Optimization and Optimal Control, Springer Verlag, Berlin, 9-20..

J.-P. CROUZEIX (1998), *Some properties of Dini derivatives of quasiconvex functions;* in F. Giannessi, S. Komlosi and T. Rapcsàk (Eds.), New Trends in Mathematical Programming, Kluwer Academic Publishers, Dordrecht, 41-57..

J.-P. CROUZEIX (2005), *Continuity and differentiability of quasiconvex functions;* in N. Hadjisavvas, S, Komlosi and S. Schaible (Eds.), Handbook of Generalized Convexity and Generalized Monotonicity, Springer, New York, 121-149..

M. C. DELFOUR (2020), Introduction to Optimization and Hadamard Semidifferential Calculus (second edition), SIAM-Society for Industrial and Applied Mathematics, Philadelphia..

V. F. DEMYANOV and A. B. PEVNYI (1974), *Expansion with respect to a parameter of the extremal values of game problems,* USSR Computational Mathematics and Mathematical Physics, **14,** 33-45..

V. F. DEMYANOV and A. M. RUBINOV (1995), Constructive Nonsmooth Analysis, Verlag Peter Lang, Frankfurt am Main..

A. DHARA and J. DUTTA (2012), Optimality Conditions in Convex Optimization. A Finite-Dimensional View, CRC Press, Boca Raton, London, New York..

W. E. DIEWERT (1981), *Alternative characterizations of six kinds of quasiconcavity in the nondifferentiable case with applications to nonsmooth programming;* in S. Schaible and W. T. Ziemba (Eds.), Generalized Concavity in Optimization and Economics, Academic Press, New York, 51-93.

W. E. DIEWERT, M. AVRIEL and I. ZANG (1981), *Nine kinds of quasiconcavity and concavity,* Journal of Economic Theory, **25,** 397-420.

N. DINH and L. A. TUAN (2003), *Directional Kuhn-Tucker conditions and duality for quasi-differentiable programs,* Acta Math. Vietnamica, **28,** 17-38.

N. DINH, G. M. LEE and L. A. TUAN (2005), *Generalized Lagrange multipliers for nonconvex directionally differentiable programs;* in V. Jeyakumar and A. Rubinov (Eds.), Continuous Optimization. Current Trends and Modern Applications, Springer, New York, 293-319.

U. DINI (1878), Fondamenti per la Teorica delle Funzioni di Variabili Reali, T. Nistri, Pisa. Reprinted on 1990 by the Unione Matematica Italiana, Bologna.

M. DUREA and R. STRUGARIU (2014), An Introduction to Nonlinear Optimization Theory, De Gruyter Open Ltd, Warsaw/Berlin.

M. EHRGOTT (2005), Multicriteria Optimization, Second Edition, Springer, New York.

K.-H. ELSTER and J. THIERFELDER (1985), *The general concept of cone approximations in nondifferentiable optimization;* in V. F. Demyanov and D. Pallaschke (Eds.), Nondifferentiable Optimization. Motivations and Applications, Springer Verlag, Berlin, 170-189.

K.-H. ELSTER and J. THIERFELDER (1988a), *On Cone approximations and generalized directional derivatives;* in V. F. Demyanov and F. Giannessi (Eds.), Nonsmooth Optimization and Related Topics, Plenum Press, New York, 133-154.

K.-H. ELSTER and J. THIERFELDER (1988b), *Abstract cone approximations and generalized differentiability in nonsmooth optimization,* Optimization, **19,** 315-341.

K.-H. ELSTER and J. THIERFELDER (1989), *Generalized notions of directional derivatives,* Quaderni della Sezione di Matematica Applicata, Gruppo di Ottimizzazione e Ricerca Operativa, Università di Pisa, N. 155.

G. M. EWING (1977), *Sufficient conditions for global minima of suitable convex functionals from variational and control theory,* SIAM Rev., **19,** 202-220.

K. FAN, I. GLICKSBERG and A. J. HOFFMAN (1957), *Systems of inequalities involving convex functions,* American Mathematical Society Proceedings, **8,** 617-622.

W. FENCHEL (1953), Convex Cones, Sets and Functions, Lecture Notes, Princeton University.

L. A. FERNANDEZ (1997), *On the limits of the Lagrange multiplier rule,* SIAM Review, 39, 292-297.

A. V. FIACCO (1983a), Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, Academic Press, New York.

A. V. FIACCO (1983b), *Optimal value continuity and differential stability bounds under the Mangasarian-Fromovitz constraint qualification;* in A. V. Fiacco (Ed.), Mathematical Programming with Data Perturbations II, Marcel Dekker Inc., New York, 65-90.

A. V. FIACCO and W. P. HUTZLER (1982), *Optimal value differential stability results for general inequality constrained differentiable mathematical programs;* in A. V. Fiacco (Ed.), Mathematical Programming with Data Perturbations I, Marcel Dekker Inc., New York, 29-43.

J. GAUVIN (1980), *Shadow prices in nonconvex mathematical programming,* Mathematical Programming, **19,** 300-312.

J. GAUVIN and F. DUBEAU (1982), *Differential properties of the marginal function in mathematical programming,* Mathematical Programming Study, **19,** 101-119.

J. GAUVIN and F. DUBEAU (1984), *Some examples and counterexamples for the stability analysis of nonlinear programming problems,* Mathematical Programming Study, **21,** 69-78.

J. GAUVIN and R. JANIN (1990), *Directional derivatives of the value function in parametric optimization,* Annals of Operations Research, **27,** 237-252.

J. GAUVIN and J. W. TOLLE (1977), *Differential stability in nonlinear programming,* SIAM J. on Control and Optimization, **15,** 294-311.

F. GIANNESSI (1989), *Semidifferentiable functions and necessary optimality conditions,* J. Optim. Theory Appl., **60,** 191-241.

F. GIANNESSI (1995), *A common understanding or a common misunderstanding?,* Numer. Funct. Anal. Optim., **16,** 1359-1363.

F. GIANNESSI (2005), Constrained Optimization and Image Space Analysis. Volume 1: Separation of Sets and Optimality Conditions, Springer, New York.

I. GINCHEV, A. GUERRAGGIO and M. ROCCA (2006), *From scalar to vector optimization,* Appl. Math., **51,** 5-36.

G. GIORGI and A. GUERRAGGIO (1992a), *On a characterization of Clarke's tangent cone,* J. Optim. Theory Appl., **74,** 369-372.

G. GIORGI and A. GUERRAGGIO (1992b), *On the notion of tangent cone in mathematical programming,* Optimization, **25,** 11-23.

G. GIORGI and A. GUERRAGGIO (2002), *Characterizations, computations, algebraic and topological properties of tangent cones,* Journal of Statistics & Management Systems, **5,** 275-294.

G. GIORGI, A. GUERRAGGIO and J. THIERFELDER (2004), Mathematics of Optimization: Smooth and Nonsmooth Case, Elsevier, Amsterdam.

G. GIORGI, B. JIMENEZ and V. NOVO (2004), *On constraint qualifications in directionally differentiable multiobjective optimization,* RAIRO Oper. Res., **38,** 255-274.

G. GIORGI and S. KOMLOSI (1993a), *Dini derivatives in optimization. Part I,* Rivista di Matematica per le Scienze Economiche e Sociali, **15,** n. 1, 3-30.

G. GIORGI and S. KOMLOSI (1993b), *Dini derivatives in optimization. Part II,* Rivista di Matematica per le Scienze Economiche e Sociali, **15,** n. 2, 3-24.

G. GIORGI and S. KOMLOSI (1995), *Dini derivatives in optimization. Part III,* Rivista di Matematica per le Scienze Economiche e Sociali, **18,** n. 1, 47-63.

G. GIORGI and C. ZUCCOTTI (2016), *Multiplier rules in optimization under "minimal" differentiability assumptions,* Annals of the University of Bucharest (mathematical series), **LXV,** 1-16.

B. M. GLOVER (1984), *Generalized convexity in nondifferentiable programming,* Bull. Austral. Math. Soc., **30,** 193-218.

B. GOLLAN (1984), *On the marginal function in nonlinear programming,* Math. Oper. Res., **9,** 208-221.

J. GWINNER (1981), *Bibliography on non-differentiable optimization and nonsmooth analysis,* Journal of Computational and Applied Mathematics, **7,** 277-285.

H. HALKIN (1974), *Implicit functions and optimization without continuous differentiability of the data,* SIAM J. Control Optim., **12,** 229-236.

J.-B. HIRIART-URRUTY (1979), *New concepts in nondifferentiable programming,* Mémoires de la S. M. F., **60,** 57-85.

J.-B. HIRIART-URRUTY (1982), *Limiting behavior of the approximate first-order and second-order directional derivatives for a convex function,* Nonlinear Analysis: Theory, Methods and Applications, **6,** 1309-1326.

J.-B. HIRIART-URRUTY and C. LEMARECHAL (1993), Convex Analysis and Minimization Algorithms, Vol. I and Vol. II, Springer-Verlag, Berlin and New York.

W. HOGAN (1973), *Directional derivatives for extremal-value functions with applications to the completely convex case,* Op. Res., **21,** 188-209.

R. B. HOLMES (1972), A Course on Optimization and Best Approximation, Springer-Verlag, Berlin.

R. B. HOLMES (1975), Geometrical Functional Analysis and Its Applications, Springer-Verlag, Berlin.

R. HORST (1984a), *Shadow prices and equilibrium prices in linear and nonlinear decision models,* Methods of Operations Research, **48,** 525-543.

R. HORST (1984b), *On the interpretation of optimal dual solutions in convex programming,* J. Operational Res., Soc., **35,** 327-335.

L. R. HUANG (2005), *Separate necessary and sufficient conditions for the local minimum of a function,* J. Optim. Theory Appl., **125,** 241-246.

L. R. HUANG and K. F. NG (1994), *Second-order necessary and sufficient conditions in nonsmooth optimization,* Math. Programming, **66,** 379-402.

A. D. IOFFE (1986) *On the theory of subdifferential;* in J.-B. Hiriart-Urruty (Ed.), Fermat-days 85: Mathematics for Optimization, North Holland, Amsterdam, 183-200.

Y. ISHIZUKA (1992), *Optimality conditions for directionally differentiable multiobjective programming,* J. Optim. Theory Appl., **72,** 91-111.

J. JAHN (2005), Vector Optimization, Springer, Berlin.

R. JANIN (1984), *Directional derivative of the marginal function in nonlinear programming,* Math. Programming Study, **21,** 110-126.

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

V. JEYAKUMAR (1991), *Composite nonsmooth programming with Gâteaux differentiability,* SIAM J. Optim., **1,** 30-41.

B. JIMENEZ and V. NOVO (2002a), *Alternative theorems and necessary optimality conditions for directionally differentiable multiobjective programs,* J. Convex Analysis, **9,** 97-116.

B. JIMENEZ and V. NOVO (2002b), *A finite dimensional extension of Lyusternik theorem with applications to multiobjective optimization,* J. Math. Anal. Appl., **270,** 340-356.

B. JIMENEZ and V. NOVO (2003), *Optimality conditions in directionally differentiable Pareto problems with a set constraint via tangent cones,* Numer. Funct. Anal. Optim., **24,** 557-574.

B. JIMENEZ and V. NOVO (2008), *First order optimality conditions in vector optimization involving stable functions,* Optimization, **57,** 449-471.

B. JIMENEZ, V. NOVO and M. SAMA (2009), *Scalarization and optimality conditions for strict minimizers in multiobjective optimization via contingent derivatives,* J. Math. Anal. Appl., **352,** 788-798.

L. V. KANTOROVIČ and G. P. AKILOV (1980), Analisi Funzionale, Editori Riuniti, Roma.

R. N. KAUL (1985), *Directional derivative of marginal function in mathematical programming,* Indian J. Pure Appl. Math., **16,** 1221-1226.

R. N. KAUL and S. KAUR (1982a), *Generalizations of convex and related functions,* European Journal of Operational Research, **9,** 369-377.

R. N. KAUL and S. KAUR (1982b), *Sufficient optimality conditions using generalized convex functions,* Opsearch (India), **19,** 212-224.

R. N. KAUL and S. KAUR (1984), *Generalized convex functions - properties, optimality and duality,* Technical Report SOL 84-4, Stanford University, Stanford, Ca.

R. N. KAUL and V. LYALL (1988), *Non-convex and semi-differentiable functions,* Indian J. Pure Appl. Math., **19,** 1167-1179.

R. N. KAUL and V. LYALL (1990), *Erratum: "Non-convex and semi-differentiable functions",* Indian J. Pure Appl. Math.,**21,** 1137.

R. N. KAUL, V. LYALL and S. KAUR (1988), *Semilocal pseudolinearity and efficiency,* European Journal of Operational Research, **36,** 402-409.

Q. Q. KHAN and N. D. TUAN (2007), *Optimality conditions for nonsmooth multiobjective optimization using Hadamard directional derivatives,* J. Optim. Theory Appl., **133,** 341-357.

A. N. KOLMOGOROV and S. V. FOMIN (1980), Elementi di Teoria della Funzioni e di Analisi Funzionale, Edizioni MIR, Mosca.

S. KOMLOSI (1983), *Some properties of nondifferentiable pseudoconvex functions,* Mathematical Programming, **26,** 232-237.

S. KOMLOSI (1995), *Generalized monotonicity and generalized convexity,* J. Optim. Theory Appl., **84,** 361-376.

S. KOMLOSI (2005), *Generalized convexity and generalized derivatives;* in N. Hadjisavvas, S. Komlosi and S. Schaible (Eds.), Handbook of Generalized Convexity and Generalized Monotonicity, Springer, New York, 421-463.

S. KOMLOSI and M. PAPPALARDO (1994), *A general scheme for first order approximations in optimization,* Optimization Methods and Software, **3,** 143-152.

D. T. LUC (1989), Theory of Vector Optimization, Lecture Notes in Economics and Math. Systems, 319, Springer, New York.

D. V. LUU and M.-H. NGUYEN (2009), *On alternative theorems and necessary conditions for efficiency,* Optimization, **58,** 49-62.

O. L. MANGASARIAN (1965), *Pseudo-convex functions,* SIAM Journal on Control, **3,** 281-290.

O. L. MANGASARIAN (1969), Nonlinear Programming, McGraw-Hill Book Company, New York.

H. MASSAM and S. ZLOBEC (1974), *Various definitions of the derivative in mathematical programming,* Math. Programming, **7,** 144-161.

H. MASSAM and S. ZLOBEC (1978), *Various definitions of the derivative in mathematical programming - An addendum,* Math. Programming, **14,** 108-111.

K. M. MIETTINEN (1999), Nonlinear Multiobjective Optimization, Kluwer Acad. Pub., Boston.

B. S. MORDUKHOVICH (2006), Variational Analysis and Generalized Differentiation, Vol. 1 and Vol. 2, Springer, Berlin.

R. N. MUKHERJEE and S. K. MISHRA (1996), *Multiobjective programming with semilocally convex functions,* J. Math. Anal. Appl., **199,** 409-424.

R. N. MUKHERJEE and A. K. SINGH (1990), *Multi-objective optimization involving nonconvex semi-differentiable functions,* Indian J. Pure Appl. Math., **21,** 326-329.

M. Z. NASHED (1971), *Differentiability and related properties of nonlinear operators: some basic aspects of the role of differentials in nonlinear functional analysis;* in L. B. Rall (Ed.), Nonlinear Functional Analysis and Applications, Academic Press, New York, 103-309.

M. H. NGUYEN and D. V. LUU (2007), *On necessary conditions for efficiency in directionally differentiable optimization problems,* Nonlinear Funct. Anal. and Applications, **12,** 439-453.

A. NIJENHUIS (1974), *Strong derivatives and inverse mappings,* The Amer. Math. Monthly, **81,** 969-980.

V. NOVO and B. JIMENEZ (2004), *Lagrange multipliers in multiobjective optimization under mixed assumptions of Fréchet and directional differentiability,* Revista Investigacion Operacional, **25,** 34-47.

J. M. ORTEGA and W. C. RHEINBOLDT (1970), Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York.

J. PALATA (1989), *A survey of conical approximations used in optimization,* Optimization, **20,** 147-161.

J.-P. PENOT (1978), *Calcul sous-différentiel et optimization,* J. Funct. Anal., **27,** 248-276.

J.-P. PENOT (1984), *Differentiability of relations and differential stability of perturbed optimization problems,* SIAM J. Control and Optimization, **22,** 529-551.

B. H. POURCIAU (1980), *Modern multiplier rules,* The Amer. Math. Monthly, **87,** 433-452.

V. PREDA (1996), *Optimality conditions and duality in multiple objective programming involving semilocally convex and related functions,* Optimization, **36,** 219-230.

V. PREDA and I. CHITESCU (1999), *On constraint qualifications in multiobjective optimization problems: semidifferentiable case,* J. Optim. Theory Appl., **100,** 417-433.

V. PREDA and I. M. STANCU-MINASIAN (1997), *On optimality and duality in multiobjective nonsmooth programming;* in R. Caballero, F. Ruiz and R. E. Steuer (Eds.), Advances in Multiple Objective and Goal Programming, Springer-Verlag, Berlin, 178-187.

V. PREDA, I. M. STANCU-MINASIAN and A. BATATORESCU (1996), *Optimality and duality in nonlinear programming involving semilocally preinvex and related functions,* J. Infor. & Optim. Sciences, **17,** 585-596.

B. N. PSHENICHNYI (1971), Necessary Conditions for an Extremum, Marcel Dekker, New York.

H. RADEMACHER (1919), *Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale*, Mathematische Annalen, **79,** 340-359.

D. RALPH and S. DEMPE (1995), *Directional derivatives of the solution of a parametric nonlinear program,* Math. Programming, **70,** 159-172.

A. W. ROBERTS and D. E. VARBERG (1973), Convex Functions, Academic Press, New York.

S. M. ROBINSON (1987), *Local structure of feasible sets in nonlinear programming. Part III: stability and sensitivity*, Mathematical Programming Study, **30,** 45-66.

R. T. ROCKAFELLAR (1970), Convex Analysis, Princeton University Press, Princeton.

R. T. ROCKAFELLAR (1980), *Generalized directional derivatives and subgradients of nonconvex functions,* Canadian Journal of Mathematics, **32,** 257-280.

R. T. ROCKAFELLAR (1981), The Theory of Subgradients and Its Applications to Problems of Optimization. Convex and Nonconvex Functions, Heldermann-Verlag, Berlin.

R. T. ROCKAFELLAR (1984), *Directional differentiability of the optimal value function in a nonlinear programming problem*, Mathematical Programming Study, **21,** 213-226.

R. T. ROCKAFELLAR and R. J.-B. WETS (2009), Variational Analysis, Springer Verlag, Berlin.

Y. SAWARAGI, H. NAKAYAMA and T. TANINO (1985), Theory of Multiobjective Optimization, Academic Press, New York.

A. SHAPIRO (1990), *On concepts of directional differentiability,* J. Optim. Theory Appl., **66,** 477-487.

K. SHIMITZU, Y. ISHIZUKA and J. F. BARD (1997), Nondifferentiable and Two-Level Mathematical Programming, Kluwer Academic Publishers, Boston.

M. STUDNIARSKI (1986), *Necessary and sufficient conditions for isolated local minima of nonsmooth functions,* SIAM J. Control and Optimization, **24,** 1044-1049.

M. STUDNIARSKI (1991), *Second-order necessary conditions for optimality in nonsmooth nonlinear programming,* J. Math. Anal. Appl., **154,** 303-317.

A. TAA (1999), *Optimality conditions for vector mathematical programming via a theorem of the alternative,* J. Math. Anal. Appl., **233,** 233-245.

W. A. THOMPSON and D. W. PARKE (1973), *Some properties of generalized convex functions,* Operations Research, **21,** 305-313.

C. URSESCU (1982), *Tangent set's calculus and necessary conditions for extremality,* SIAM J. Control Optim., **20,** 563-574.

H. UZAWA (1958a), *A theorem on convex polyhedral cones;* in K. J. Arrow, L. Hurwicz and H. Uzawa (Eds.), Studies in Linear and Nonlinear Programming, Stanford Univ. Press, Stanford, 23-31.

H. UZAWA (1958b), *A note on the Menger-Wieser theory of imputation,* Zeitschrift für Nationalökonomie, **18,** 318-334.

M. VAINBERG (1956), Variational Methods for the Study of Nonlinear Operators, Holden-Day, San Francisco, Ca.

S. YAMAMURO (1974), Differential Calculus in Topological Spaces, Springer, Berlin.

X. M. YANG (1994), *Two counterexamples of T. Weir's paper,* J. Math. Anal. Appl., **188,** 717-718.

X. Q. YANG (1996), *On second-order directional derivatives,* Nonlinear Analysis, Theory, Methods and Applications, **26,** 55-66.

X. Q. YANG (1999), *On relations and applications of generalized second-order directional derivatives,* Nonlinear Analysis, **36,** 595-614.

D. E. WARD (1987), *Isotone tangent cones and nonsmooth optimization,* Optimization, **18,** 769-783.

D. E. WARD (1988), *The quantificational tangent cones,* Canad. J. Math., **40,** 666-694.

D. E. WARD (1989), *Directional derivative calculus and optimality conditions in nonsmooth mathematical programming,* Journal of Information & Optimization Sciences, **10,** 81-96.

T. WEIR (1992), *Programming with semilocally convex functions,* J. Math. Anal. Appl., **168,** 1-12.

T. WEIR and B. MOND (1988), *Preinvex functions in multiple-objective optimization,* J. Math. Anal. Appl., **136,** 29-38.