Control and Cybernetics

vol. **32** (2003) No. 2

Kuhn-Tucker type optimality conditions for some class of nonsmooth programming problems

by

Marek J. Śmietański

Faculty of Mathematics of Łódź University
Banacha 22, 90-238 Łódź, Poland

Abstract: We consider two nonlinear programming problems with nonsmooth functions. The necessary and sufficient first order optimality conditions use the Dini and Clarke derivatives. However, the obtained Kuhn-Tucker conditions have a rather classical form. The sufficient conditions alone are obtained thanks to some properties of generalized convexity and generalized linearity of functions. The necessary and sufficient optimality conditions are given in the Lagrange form.

Keywords: nonsmooth programming, Dini derivative, Clarke derivative, generalized convexity, generalized linearity, necessary and sufficient optimality conditions.

1. Introduction

In the present paper we consider nonlinear programming problems, where the nonsmooth objective function and nonsmooth constraints are involved. The necessary and sufficient first order optimality conditions use the Dini and Clarke derivatives. However, the obtained Kuhn-Tucker conditions are in the classical form, because neither subdifferentials nor cones appears here, similarly as in Zangwill (1969). The sufficient conditions alone are obtained thanks to some properties of generalized convexity and generalized linearity of functions. The role of these properties seems to be increasingly promising and appreciated, because now and again they are used in the applications of optimization theory and nonsmooth analysis — see e.g. Giorgi and Komlósi (1995) (and preceding parts). The necessary and sufficient optimality conditions are given in the Lagrange form. This work is some extension of the results given by Glover (1984) to problems with mixed constraints. The results of present paper are not surprising, but the purpose of the present paper was, in particular, exactly like this: to formulate the optimality conditions without cones and subdifferentials.

Mititelu (1987) as the Karush-Kuhn-Tucker conditions, but there, only differentiable programming problems are considered. If the generalized derivatives are sublinear, then the Lagrange multiplier can be 'hidden' behind a subdifferential condition as it is in Gianessi (1989), Komlósi (1993) and Mititelu (1994).

2. Definitions and remarks

Let X be a normed vector space, $X_0 \subset X$ an open convex set and $f : X_0 \to R$.

DEFINITION 1 *A function f is said to be locally Lipschitz on X_0 if, for each point in X_0, there exist a neighborhood U of this point and a positive real number L such that*

$$\|f(x) - f(y)\| \leq L\|x - y\|, \text{ for all } x, y \in U.$$

DEFINITION 2 *A function f is said to be convex on X_0 if, for all $x, y \in X_0$ and $t \in [0, 1]$,*

$$f(y + t(x - y)) \leq f(y) + t(f(x) - f(y)).$$

DEFINITION 3 *The upper Dini derivative of a function f at a point $x \in X_0$ in the direction $d \in X$ is given by*

$$D^+ f(x, d) := \limsup_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}.$$

The lower Dini derivative of a function f at a point $x \in X_0$ in the direction $d \in X$ is given by

$$D_+ f(x, d) := \liminf_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}.$$

REMARK. *Naturally, if, for any point $x \in X_0$ and for any direction $d \in X$ the equality $D^+ f(x, d) = D_+ f(x, d)$ holds, then there exists the directional derivative of f at x, and*

$$f'(x, d) = D^+ f(x, d) = D_+ f(x, d),$$

see Komlósi (1994).

DEFINITION 4 *The Clarke derivative (generalized directional derivative) of a function f at a point $x \in X_0$ in the direction $d \in X$ is given by*

$$f^\circ(x, d) := \limsup \frac{f(z + td) - f(z)}{}.$$

REMARK. *(i) Clearly, $D_+f(x,d) \le D^+f(x,d) \le f^\circ(x,d)$ for each $x \in X_0$ and any $d \in X$.*
(ii) Moreover, for each $x \in X_0$, $f^\circ(x,v)$ is a finite, positively homogeneous, subadditive and convex function of v and an upper semicontinuous function of (x,v) when f is locally Lipschitz (see Clarke 1975, 1976).

DEFINITION 5 *A function f is said to be radially continuous if, for all $x, h \in X_0$, the function $\rho(t) = f(x+th)$ is continuous as a function $\rho : R \to R$.*

3. Generalized convexity and generalized linearity

First, we establish several properties of generalized convexity of functions. The concept of the pseudoconvexity for nondifferentiable functions has been introduced in Diewert (1981), who extended Mangasarian's definition from Mangasarian (1969). The below-mentioned concepts can be found also in Glover (1984) and in Ellaia, Hassouni (1991):

DEFINITION 6 *For a locally Lipschitz function $f : X_0 \to R$, we shall say:*

(i) f is Clarke-pseudoconvex (PCX-Clarke) if, for all $x, y \in X_0$

$$f^\circ(y, x-y) \ge 0 \ \Rightarrow \ f(x) \ge f(y),$$

(ii) f is Dini-pseudoconvex (PCX-Dini) if, for all $x, y \in X_0$

$$D^+f(y, x-y) \ge 0 \ \Rightarrow \ f(x) \ge f(y),$$

(iii) f is Clarke-quasiconvex (QCX-Clarke) if, for all $x, y \in X_0$

$$f^\circ(y, x-y) > 0 \ \Rightarrow \ f(x) > f(y),$$

(iv) f is Dini-quasiconvex (QCX-Dini) if, for all $x, y \in X_0$

$$D^+f(y, x-y) > 0 \ \Rightarrow \ f(x) > f(y),$$

(v) f is strict quasiconvex (SQCX) if, for all $x, y \in X_0$

$$f(x) < f(y) \ \Rightarrow \ \forall_{t \in (0,1)} \ f(y + t(x-y)) < f(y),$$

(vi) f is quasiconvex (QCX) if, for all $x, y \in X_0$

$$f(x) \le f(y) \ \Rightarrow \ \forall_{t \in (0,1)} \ f(y + t(x-y)) \le f(y).$$

We can consider all the above properties at a given point x by fixing x in the above definitions or locally by fixing x and considering a neighborhood G of x instead of the whole set X_0 (see Komlósi, 1993).

REMARK. *By Definition 6 it easily follows that every PCX-Clarke function is PCX-Dini. Similarly, every QCX-Clarke function is QCX-Dini.*

In the following lemma we give some relationships between the above intro-

LEMMA 1 *Let $f : X_0 \to R$ be radially continuous. Then:*
(i) f is QCX-Dini if and only if f is QCX;
(ii) If f is PCX-Dini, it is also QCX-Dini, SQCX and QCX.

Proof. See Diewert (1981). ∎

REMARK. *(i) By Lemma 1 and remark after Definition 6, it immediately follows that every radially continuous function, which is PCX-Clarke, is also QCX and SQCX.*

(ii) Other relationships between the generalized convexity properties can be proved similarly to these, which are presented by Komlósi in (1993).

We now quote another characterization of quasiconvexity of a function (by means of convexity of the lower level sets):

LEMMA 2 *A function $f : X_0 \to R$ is QCX if and only if the set*

$$H^\gamma := \{x \in X_0 : \ f(x) \le \gamma\}$$

is convex for any real number $\gamma \in R$.

Proof. The proof is an immediate consequence of (vi) in Definition 6. We leave it to the reader. ∎

Now, we establish several properties of generalized linearity of functions (defined e.g. by Komlósi in 1993):

DEFINITION 7 *For a locally Lipschitz function $f : X_0 \to R$, we shall say:*
(i) f is Clarke-quasilinear (QL-Clarke) *if f is simultaneously Clarke quasiconvex and Clarke-quasiconcave (i.e. f and $-f$ are Clarke-quasiconvex),*
(ii) f is Dini-quasilinear (QL-Dini) *if f is simultaneously Dini-quasiconvex and Dini-quasiconcave (i.e. f and $-f$ are Dini-quasiconvex),*
(iii) f is Clarke-pseudolinear (PL-Clarke) *if f is simultaneously Clarke-pseudoconvex and Clarke-pseudoconcave (i.e. f and $-f$ are Clarke-pseudoconvex),*
(iv) f is Dini-pseudolinear (PL-Dini) *if f is simultaneously Dini-pseudoconvex and Dini-pseudoconcave (i.e. f and $-f$ are Dini-pseudoconvex).*

REMARK. *By remark after Definition 6 and part (ii) of Lemma 1 it easily follows that every PL-Clarke function is PL-Dini, every QL-Clarke function is QL-Dini and every PL-Dini function is QL-Dini.*

The following lemma is natural extension of the fact, which was proved by Kortanek and Evans (1967) and Chew and Choo (1984) for differentiable functions:

LEMMA 3 *If a directionally differentiable function $f : X_0 \to R$ is Dini-quasilinear at a point $y \in X_0$, then*

Proof. Assume that f is QL-Dini at the point $y \in X_0$. Then f and $-f$ are QCX-Dini functions at this point. This means that the following conditions are satisfied (by Definition 6):

$$D^+ f(y, x-y) > 0 \Rightarrow f(x) > f(y), \text{ for all } x \in X_0,$$
$$D^+(-f)(y, x-y) > 0 \Rightarrow -f(x) > -f(y), \text{ for all } x \in X_0,$$

or, equivalently:

$$f(x) \le f(y) \Rightarrow D^+ f(y, x-y) \le 0, \text{ for all } x \in X_0,$$
$$-f(x) \le -f(y) \Rightarrow -D_+ f(y, x-y) \le 0, \text{ for all } x \in X_0.$$

Now, from the above implications and remarks after Definitions 3 and 4 we obtain the necessity. The sufficiency is obvious. ■

4. Problems with inequality constraints

Now, we consider applications of the above concepts in nonsmooth programming. We establish first order optimality conditions for some constrained problems of nonlinear programming with nondifferentiable non-convex objective function and constraints.

Consider the following problem:

$$\begin{cases} f(x) \to \min\limits_{x \in X_0}, \\ g_i(x) \le 0, \ i \in I = \{1, \ldots, n\}, \end{cases} \tag{ICP}$$

where $R^n \supset X_0$ is a nonempty open convex set, $f : X_0 \to R$, $g_i : X_0 \to R$, for each $i \in I$, are locally Lipschitz.

Now, we define some sets which will be needed in further considerations:

DEFINITION 8 *The set*

$$F := \{x \in X_0 : g_i(x) \le 0, \ i \in I\}$$

is said to be the all feasible points set (feasible set). The set

$$K_x := \{d \in R^n : \underset{\varepsilon > 0}{\exists} \ \underset{0 \le t \le \varepsilon}{\forall} \ x + td \in F\}$$

is said to be the all feasible directions set for a point x.

Geometrically: a direction d is feasible for some point x if sufficiently small moves from the point x along the direction d do not lead outside the feasible set.

The following lemma describes the behavior of the objective function's values along feasible directions from the set $K_{\tilde{x}}$, where $\tilde{x}$ is an optimal point of

LEMMA 4 *Let $\tilde{x}$ be a local minimum point of problem (ICP). Then:*

(i) $D_+f(\tilde{x},d) \geq 0$ for all $d \in K_{\tilde{x}}$;

(ii) $f^0(\tilde{x},d) \geq 0$ for all $d \in \bar{K}_{\tilde{x}}$ (where $\bar{K}_{\tilde{x}}$ denotes the closure of $K_{\tilde{x}}$).

Proof. (i) The proof is an immediate consequence of the Definitions 3 and 8.

(ii) Using part (i) of remark after Definition 4 and part (i) of this lemma, we find that $f^\circ(\tilde{x},d) \geq 0$ for all $d \in K_{\tilde{x}}$. An arbitrary direction $d_0 \in \bar{K}_{\tilde{x}}$ can be represented as the limit of a sequence of directions from $K_{\tilde{x}}$, namely $d_0 = \lim_{k\to\infty} d_k$. Obviously $f^\circ(\tilde{x},d_k) \geq 0$ for all k, because $d_k \in K_{\tilde{x}}$. After taking the limit as $k \to \infty$, we obtain that $f^\circ(\tilde{x},d_0) \geq 0$ by the continuity of the Clarke derivative with respect to the direction (see remark (ii) after Definition 4) for an arbitrary direction $d_0 \in \bar{K}_{\tilde{x}}$. ■

REMARK. *By part (i) of Lemma 4 and remark after Definition 4 it immediately follows that $D^+f(\tilde{x},d) \geq 0$ for all $d \in K_{\tilde{x}}$, too.*

To prove the existence of the multipliers in a further theorem, we use Generalized Basic Alternative Theorem (proved by Craven and Wang, preprint):

LEMMA 5 *Let X and Y be normed spaces, $\Gamma \subset X$ - a convex set, $S \subset X$ - a convex cone with nonempty interior, $f : \Gamma \to Y$ - a S-convex mapping (i.e., for all $x, y \in \Gamma$ and $t \in (0,1)$, $f(y + t(x-y)) - f(y) - t(f(x) - f(y)) \in -S$) and $E \subset Y$ such that $E + Int\, S$ is convex. If*

$$(f(\Gamma) + E) \cap (-Int\, S) = \emptyset, \tag{i}$$

then there exists a nonzero $\lambda \in S^* \equiv -S^0$ (where $S^0 = \{z \in Y : \forall_{s\in S}\, zs \leq 0\}$) such that, for all $x \in \Gamma$ and $e \in E$,

$$\lambda(f(x) + e) \geq 0. \tag{ii}$$

If $0 \in E$ and $\omega \in f(\Gamma)$, then $\lambda f(\Gamma) \subset R_+$ and $\lambda E \subset -\lambda\omega + R_+$.

Proof. (after Craven and Wang):
Since f is S-convex, the set $f(\Gamma) + Int\, S$ is open convex (see McCormick, 1967), hence the set $K := f(\Gamma)+E+Int\, S = [f(\Gamma)+Int\, S]+[E+Int\, S]$ is open convex. From (i), $0 \notin K$, hence there exists a nonzero $\lambda \in S^*$ such that (ii) holds. The remaining statements follow by setting $e = 0$, or by setting $f(x) = \omega$. ■

The theorem below gives first order optimality conditions for (ICP). First, we note that the Kuhn-Tucker conditions alone can serve only for the statement of the non-optimality of feasible points, because they are only necessary conditions. They turn out to be sufficient only when we make additional assumptions.

Considering the properties of the objective function, we establish two versions of optimality conditions. For $x \in F$, we denote

$$I_x := \{i \in I : g_i(x) = 0\},$$

THEOREM 1 *Let $\tilde{x}$ be a local minimum point of problem (ICP), where all functions are locally Lipschitz. Assume that the set $K_{\tilde{x}}$ satisfies the following regularity conditions:*

$$\mathcal{K}_{\tilde{x}} := \{d \in R^n : D^+ g_i(\tilde{x}, d) \leq 0, \text{ for all } i \in I_{\tilde{x}}\} \subset K_{\tilde{x}}, \tag{1}$$

and for some $d \in K_{\tilde{x}}$

$$D^+ g_i(\tilde{x}, d) \neq 0, \text{ for all } i \in I_{\tilde{x}}, \tag{2}$$

and suppose that f and g_i, $i \in I_{\tilde{x}}$, have convex upper Dini derivatives at $\tilde{x}$. Then there exist multipliers $\lambda_1, \dots, \lambda_n$, such that:
(a) $\lambda_i \geq 0$ for $i \in I$;
(b) $\lambda_i g_i(\tilde{x}) = 0$ for $i \in I$;
(c) $D^+ f(\tilde{x}, d) + \sum_{i \in I} \lambda_i D^+ g_i(\tilde{x}, d) \geq 0$, for all $d \in K_{\tilde{x}}$.
Conversely, if we assume that the following situation holds:
(d) f is Dini-pseudoconvex at $\tilde{x}$, and g_i, $i \in I$, are Dini-quasiconvex, and conditions (a), (b), (c) are satisfied, then the Kuhn-Tucker conditions (a), (b), (c) are sufficient for global minimality of $\tilde{x}$ in problem (ICP).

Proof. Examine exactly the set K_x for an arbitrary feasible point x. We can divide constraints into active, i.e. $g_i(x) = 0$, and passive, i.e. $g_i(x) < 0$. We will try to express the set K_x by means of constraints. It is easy to show that only active constraints are needed. Indeed, if $g_i(x) < 0$, we can move from the point x along an arbitrary feasible direction without violation of this constraint (since g_i are radially continuous as locally Lipschitz). Hence, there is no effect of passive constraints on the form of the set K_x. We will prove the inclusion $K_x \subset \mathcal{K}_x$, which means that if d is a feasible direction for the point x (i.e. $d \in K_x$) then $D^+ g_i(x, d) \leq 0$ for active constraints, i.e. for $i \in I_x$. Let $d \in K_x$, and take any i such that $g_i(x) = 0$. Hence, it follows that $g_i(x + td) \leq 0$ for sufficiently small $t \geq 0$ (otherwise, $x + td \notin F$). Therefore

$$\frac{g_i(x + td) - g_i(x)}{t} \leq 0 \text{ for sufficiently small } t \geq 0.$$

By taking the upper limit as $t \downarrow 0$ we obtain

$$0 \geq \limsup_{t \downarrow 0} \frac{g_i(x + td) - g_i(x)}{t} = D^+ g_i(x, d), \text{ for all } i \in I_x,$$

so that $d \in \mathcal{K}_x$. Hence, it follows that $K_x \subset \mathcal{K}_x$.

(Necessity). Let $\tilde{x}$ be a local minimum point of problem (ICP). By condition (1) and the inclusion $K_{\tilde{x}} \subset \mathcal{K}_{\tilde{x}}$, we have that:

$$K_{\tilde{x}} = \{d \in R^n : D^+ g_i(\tilde{x}, d) \leq 0, \text{ for all } i \in I_{\tilde{x}}\}. \tag{3}$$

This equality and part (i) of Lemma 4 imply

Since $g_i(\tilde{x}) = 0$ for all $i \in I_{\tilde{x}}$, it follows that

$$D^+ f(\tilde{x}, d) \geq 0, \text{ for each } d \text{ such that } g_i(\tilde{x}) + D^+ g_i(\tilde{x}, d) \leq 0, \text{ for all } i \in I_{\tilde{x}}. \quad (4)$$

If we denote:

$$G(d) = (D^+ f(\tilde{x}, d), D^+ g_{i_1}(\tilde{x}, d), \ldots, D^+ g_{i_k}(\tilde{x}, d)),$$

where $\{i_1, \ldots, i_k\} = I_{\tilde{x}}$ (so $G : R^n \to R^{k+1}$) and

$$e = (0, g_{i_1}(\tilde{x}), \ldots, g_{i_k}(\tilde{x})) = (0, \ldots, 0) \in R^{k+1},$$

then, by (4), it follows that

$$(G(K_{\tilde{x}}) + \{e\}) \cap (-Int\, R_+^{k+1}) = \emptyset, \text{ for all } d \in K_{\tilde{x}}.$$

Since f and g_i, $i \in I_{\tilde{x}}$, have convex upper Dini derivatives at $\tilde{x}$, then $G(\cdot)$ is a R_+^{k+1}-convex mapping, so, by using Lemma 5, we can state that there exists $\Lambda' \in R_+^{k+1}$ such that, for all $d \in K_{\tilde{x}}$, the inequality

$$\Lambda'(G(d) + e) \geq 0 \quad (5)$$

holds. Let $\Lambda' = (\lambda'_0, \lambda'_{i_1}, \ldots, \lambda'_{i_k})$, then, by (5), it follows that

$$\lambda'_0 D^+ f(\tilde{x}, d) + \sum_{i \in I_{\tilde{x}}} \lambda'_i (g_i(\tilde{x}) + D^+ g_i(\tilde{x}, d)) \geq 0, \text{ for all } d \in K_{\tilde{x}}.$$

In this way we obtain the existence of the multipliers, $\lambda'_i \in R_+$, $i \in I_{\tilde{x}} \cup \{0\}$, not all zero, such that

$$\lambda'_0 D^+ f(\tilde{x}, d) + \sum_{i \in I_{\tilde{x}}} \lambda'_i D^+ g_i(\tilde{x}, d) \geq -\sum_{i \in I_{\tilde{x}}} \lambda'_i g_i(\tilde{x}) = 0, \text{ for all } d \in K_{\tilde{x}}. \quad (6)$$

Now, suppose that $\lambda'_0 = 0$. Then

$$\sum_{i \in I_{\tilde{x}}} \lambda'_i D^+ g_i(\tilde{x}, d) \geq 0, \text{ for all } d \in K_{\tilde{x}}.$$

Then, using (1), since $\lambda'_i \geq 0$, $i \in I_{\tilde{x}}$, we have that

$$\sum_{i \in I_{\tilde{x}}} \lambda'_i D^+ g_i(\tilde{x}, d) = 0, \text{ for all } d \in K_{\tilde{x}},$$

which contradicts assumption (2), hence $\lambda'_0 > 0$. Dividing (6) by λ'_0, we get

$$D^+ f(\tilde{x}, d) + \sum_{i \in I_{\tilde{x}}} \lambda_i D^+ g_i(\tilde{x}, d) \geq 0, \text{ for all } d \in K_{\tilde{x}},$$

Now, if $i \notin I_{\tilde{x}}$, then $g_i(x) < 0$ and thus without loss of generality we can set $\lambda_i = 0$, and, for $i \in I_{\tilde{x}}$, we have that $g_i(x) = 0$ and $\lambda_i \in R_+$, so that the Kuhn-Tucker conditions are satisfied.

(Sufficiency) Note that the all feasible points set can be represented in the form of the following intersection of sets:

$$F = \{x : g_i(x) \leq 0,\ i \in I\} = \bigcap_{i \in I} \{x : g_i(x) \leq 0\} = \bigcap_{i \in I} H_i^0.$$

Since the functions g_i are QCX-Dini, then by Lemma 1(i) they are QCX as well, so that the sets H_i^0 are convex by Lemma 2. Hence, the set F is convex as the intersection of a finite number of convex sets. Let $y \in F$ be an arbitrary feasible point. Since $\tilde{x} \in F$, then by convexity of F, it follows that

$$\tilde{x} + t(y - \tilde{x}) \in F, \text{ for all } t \in [0, 1].$$

This means, by definition of the set $K_{\tilde{x}}$, that $\tilde{d} = y - \tilde{x} \in K_{\tilde{x}}$ (so the direction $\tilde{d}$ is feasible). Hence, using (3), we obtain

$$D^+ g_i(\tilde{x}, \tilde{d}) \leq 0 \text{ for } i \in I_{\tilde{x}}. \tag{7}$$

By condition (b), it follows that $\lambda_i = 0$ for $i \notin I_{\tilde{x}}$. Using this fact together with conditions (a), (c) and (7), we obtain that $D^+ f(\tilde{x}, \tilde{d}) \geq 0$, which means that $D^+ f(\tilde{x}, y - \tilde{x}) \geq 0$. Since the function f is PCX-Dini at the point $\tilde{x}$, we can state that $f(y) \geq f(\tilde{x})$. And since y is an arbitrary feasible point, this means that $\tilde{x}$ is a global minimum point of problem (ICP). ■

It appears that the assumption of convexity of upper Dini derivatives is rather strong. However, there are similar assumptions made openly (see e.g. Komlósi, 1993) or hidden behind the constraint qualification (see e.g. Mititelu, 1994) in the known versions of the Kuhn-Tucker optimality conditions.

REMARK. *(i) It is well known that there exist various types of regularity conditions. Condition (1) with gradients instead of upper Dini derivatives is known as the Zangwill constraint qualification (see Bazaraa, Sherali and Shetty, 1993). Giorgi and Guerraggio (1994) give other constraint qualifications, which are stronger:*

- the Cottle constraint qualification

$$\tilde{x} \in X_0,\ \mathcal{K}_{\hat{x}} \subseteq \bar{K}_1 := \{d \in R^n : D^+ g_i(\tilde{x}, d) < 0,\ \textit{for all } i \in I_{\tilde{x}}\};$$

- the Arrow-Hurwicz-Uzawa I constraint qualification

$$\tilde{x} \in X_0,\ \mathcal{K}_{\hat{x}} \subseteq \bar{K}_p := \{d \in R^n : D^+ g_i(\tilde{x}, d) \leq 0,\ \textit{for all } i \in J;$$
$$D^+ g_i(\tilde{x}, d) < 0,\ \textit{for all } i \in I_{\tilde{x}} \setminus J\},$$

For problem (ICP), Bazaraa and Shetty (1976) present the following implications for the above-mentioned constraint qualifications:

$$\text{Cottle C.Q.} \implies \text{Arrow-Hurwicz-Uzawa I C.Q.} \implies \text{Zangwill C.Q.}$$

(ii) Condition (2) is similar to the first condition of generalized Slater's constraint qualification in Giorgi and Mititelu (1983).

If in Theorem 1 we make changes in some assumptions, then we obtain:

THEOREM 2 *If, in Theorem 1, we assume that the following regularity conditions hold:*

$$\mathcal{K}_{\tilde{x}}^{\circ} := \{d \in R^n : g_i^{\circ}(\tilde{x}, d) \leq 0, \text{ for all } i \in I_{\tilde{x}}\} \subset K_{\tilde{x}}, \tag{8}$$

and for some $d \in K_{\tilde{x}}$

$$g_i^{\circ}(\tilde{x}, d) \neq 0, \text{ for all } i \in I_{\tilde{x}}, \tag{9}$$

and suppose that g_i, $i \in I_{\tilde{x}}$, are Clarke-quasiconvex, then condition (c) has the form:
(c′) $f^{\circ}(\tilde{x}, d) + \sum_{i \in I} \lambda_i g_i^{\circ}(\tilde{x}, d) \geq 0$, for all $d \in \bar{K}_{\tilde{x}}$.
And, conversely, if we assume that the following situation holds:
(d') f is Clarke-pseudoconvex at $\tilde{x}$, and conditions (a), (b), (c') are satisfied, then the Kuhn-Tucker conditions (a), (b), (c') are sufficient for global minimality of $\tilde{x}$ in problem (ICP).

Proof. In the same way as in Theorem 1 we can express the set $K_{\tilde{x}}$ by means of constraints. Let $d \in K_x$, and take any $i \in I_x$. This means that $g_i(x) = 0$. Hence, it follows that $g_i(x + td) \leq 0$ for sufficiently small $t \geq 0$ (otherwise, $x + td \notin F$). Every direction from the set K_x can be expressed as $y - x$ for some $y \in F$. Then $x + t(y - x) \in F$ by convexity of F. Hence, it follows that

$$g_i(x + t(y - x)) \leq g_i(x) \text{ for sufficiently small } t \geq 0.$$

By assumption (d'), g_i are QCX-Clarke, so that

$$g_i^{\circ}(x, x + t(y - x) - x) \leq 0 \text{ for sufficiently small } t \geq 0.$$

Hence, $g_i^{\circ}(x, td) \leq 0$ and since g_i° is a positively homogeneous function (by part (ii) of remark after Definition 4), then $tg_i^{\circ}(x, d) \leq 0$ for sufficiently small $t \geq 0$. This means that $g_i^{\circ}(x, d) \leq 0$ for $d \in K_x$. Hence $K_x = \mathcal{K}_x^{\circ}$, if the regularity condition (8) holds. The rest of the proof is analogous to the proof of Theorem 1. ∎

REMARK. *(i) Instead the regularity condition (8) in the theorem above we can assume the regularity condition (1), Dini (or Clarke)-quasiconvexity of g_i, $i \in I \setminus I_{\tilde{x}}$ and convexity of the upper Dini derivatives of f and g_i, $i \in I_{\tilde{x}}$. Then all conditions follow by Theorem 1, remark after Definition 4 and remark after*

(ii) We can write various types of regularity conditions mentioned in remark after Theorem 1, i.e. Zangwill constraint qualification, Cottle constraint qualification, Arrow-Hurwicz-Uzawa I constraint qualification, with the Clarke derivatives instead of the upper Dini derivatives, like for Theorem 1.
(iii) Let the objective function f have the same properties as constraints, i.e. f is QCX-Dini and we have $D^+f(\tilde{x},d) \neq 0$ for all $d \in K_{\tilde{x}}$ in the case (c) or f is QCX-Clarke and we have $f^\circ(\tilde{x},d) \neq 0$ for all $d \in K_{\tilde{x}}$ in the case (c'). Then $\tilde{x}$ is a strict global minimum point for problem (ICP). This follows by the definitions of the respective type of quasiconvexity (parts (iii) and (iv) of Definition 6).

Theorems 1 and 2 are not a special cases of some more general results early mentioned. Most of the theorems established optimality conditions for a local solution. Mititelu (1994) presented the Kuhn-Tucker sufficient condition using some additional inequality with the Clarke derivatives of constraint functions. Komlósi (1993) established only the necessary conditions. Moreover, his optimality conditions cannot be applied for general problems using the Clarke derivative and require some regularity and usual quasiconvexity (but not generalized). Next, Gianessi (1989) proved sufficient condition but only for convex functions.

5. Problems with mixed constraints

In this part we extend further the above results. Namely, we establish the optimal conditions for the nonlinear programming problems with equality and inequality constraints.

Consider the following problem:

$$\begin{cases} f(x) \to \min\limits_{x \in X_0}, \\ g_i(x) \leq 0, \ i \in I = \{1, \ldots, m\}, \\ h_j(x) = 0, \ j \in J = \{1, \ldots, k\}, \end{cases} \qquad (MCP)$$

where $R^n \supset X_0$ is a nonempty open convex set, $f : X_0 \to R$, $g_i : X_0 \to R$, for each $i \in I$, $h_j : X_0 \to R$, for each $j \in J$, are locally Lipschitz.

For problem (MCP), we define the feasible points set as follows:

$$F := \{x \in X_0 : g_i(x) \leq 0, \ i \in I; \ h_j(x) = 0, j \in J\}.$$

Assume that the functions h_j, $j \in J$ are directionally differentiable. For $x \in F$, we denote

$$I^*(d) := \{i \in I_x : D^+g_i(x,d) \neq 0\} \text{ for } d \in K_x$$

and

The theorem below gives the optimality conditions for (MCP). Note that, just as ever, the Kuhn-Tucker conditions alone may serve only to state nonoptimality of feasible points, because they are only necessary conditions. Clearly, they turn out to be sufficient, too, when we make additional assumptions with respect to the objective function and all constraints.

THEOREM 3 *Let $\tilde{x}$ be a local minimum point for problem (MCP). Assume that the set $K_{\tilde{x}}$ satisfies the following regularity condition:*

$$\mathcal{K}_{\tilde{x}} := \{d \in R^n : D^+ g_i(\tilde{x}, d) \leq 0,\ i \in I_{\tilde{x}}; h'_j(\tilde{x}, d) = 0,\ j \in J\} \subset K_{\tilde{x}}, \quad (10)$$

and for some $d \in K_{\tilde{x}}$

$$I^*(d) = I_{\tilde{x}}, \quad (11)$$

and suppose that the functions $f, g_i,\ i \in I_{\tilde{x}}$, have convex upper Dini derivatives at $\tilde{x}$, and the functions $h_j,\ j \in J$ have directional derivatives at $\tilde{x}$.

Then there exist multipliers $\lambda_1, \ldots, \lambda_m$ and $\mu_1, \ldots, \mu_k$, such that:
(a) $\lambda_i \geq 0$ for $i \in I$;
(b) $\lambda_i g_i(\tilde{x}) = 0$ for $i \in I$;
(c) $D^+ f(\tilde{x}, d) + \sum_{i \in I} \lambda_i D^+ g_i(\tilde{x}, d) + \sum_{j \in J} \mu_j h'_j(\tilde{x}, d) \geq 0$ for all $d \in K_{\tilde{x}}$.
Conversely, if problem (MCP) satisfies the following conditions:
(d) $I^*(d) \cup J^*(d) \neq \emptyset$;
(e) $g_i,\ i \in I^*(d)$, are Dini-quasiconvex at $\tilde{x}$ and $h_j,\ j \in J^*(d)$ are Dini-quasilinear at $\tilde{x}$;
and we assume that one of the following situations holds:
(f_1) f is Dini-pseudoconvex at $\tilde{x}$;
(f_2) f is Dini-quasiconvex at $\tilde{x}$ and $D^+ f(\tilde{x}, d) \neq 0$ for all $d \in K_{\tilde{x}}$,
then the Kuhn-Tucker conditions are also sufficient for $\tilde{x}$ to be an optimal solution of problem (MCP): global in the case (f_1) or strictly global in the case (f_1).

Proof. (Necessity) Let $\tilde{x}$ be an optimal point of problem (MCP). In a very simple way we can prove the necessary conditions using Theorem 1. First, we note that problem (MCP) with mixed constraints is a particular case of problem (ICP) with inequality constraints only in the following form:

$$\begin{cases} f(x) \to \min\limits_{x \in X_0}, \\ g_i(x) \leq 0,\ i \in I = \{1, \ldots, m\}, \\ h_j(x) \leq 0,\ j \in J = \{1, \ldots, k\}, \\ -h_j(x) \leq 0,\ j \in J = \{1, \ldots, k\}. \end{cases} \quad (MCP')$$

Then, the regularity conditions (1) and (2) for problem (MCP′) are satisfied. Thus, we obtain the existence of multipliers $\lambda'_i \geq 0,\ i \in I,\ \lambda'_j \geq 0,\ j \in J,\ \lambda''_j \geq 0,\ j \in J$ such that

$$\lambda'_i g_i(\tilde{x}) = 0,\ i \in I,$$
$$\lambda'_j h_j(\tilde{x}) = 0,\ j \in J,$$

and

$$D^+f(\tilde{x},d)+\sum_{i\in I}\lambda_i' D^+g_i(\tilde{x},d)+\sum_{j\in J}\lambda_j' D^+h_j(\tilde{x},d)$$
$$+\sum_{j\in J}\lambda_j'' D^+(-h_j)(\tilde{x},d)\geq 0.$$

By the last condition, it follows that

$$0\leq D^+f(\tilde{x},d)+\sum_{i\in I}\lambda_i' D^+g_i(\tilde{x},d)+\sum_{j\in J}\lambda_j' D^+h_j(\tilde{x},d)-\sum_{j\in J}\lambda_j'' D_+h_j(\tilde{x},d)$$
$$=D^+f(\tilde{x},d)+\sum_{i\in I}\lambda_i D^+g_i(\tilde{x},d)+\sum_{j\in J}\mu_j h_j'(\tilde{x},d),$$

where $\lambda_i=\lambda_i'$, $i\in I$, $\mu_j=\lambda_j'-\lambda_j''\in R$, $j\in J$. So, we obtain conditions (a)–(c).

(Sufficiency) Assume that conditions (a)-(e) hold. Let $\tilde{d}=x-\tilde{x}$. Since $\tilde{d}\in K_{\tilde{x}}$ and the functions g_i, $i\in I^*(\tilde{d})$, are QCX-Dini at $\tilde{x}$ (by condition (e)), then inequality

$$g_i(x)\leq g_i(\tilde{x}),\ \text{for all } x\in F,\ i\in I^*(\tilde{d})$$

implies that

$$D^+g_i(\tilde{x},x-\tilde{x})\leq 0,\ \text{for all } x\in F,\ i\in I^*(\tilde{d}).$$

Next, since $\lambda_i\geq 0$ by condition (a), then

$$\lambda_i D^+g_i(\tilde{x},x-\tilde{x})\leq 0,\ \text{for all } x\in F,\ i\in I^*(\tilde{d}).$$

By the definition of the set $I^*(\tilde{d})$, we have that

$$D^+g_i(\tilde{x},\tilde{d})=0 \text{ for } i\in I_{\tilde{x}}\setminus\bigcup_{\tilde{d}\in K_{\tilde{x}}} I^*(\tilde{d}),$$

so that the above condition holds for all $i\in I_{\tilde{x}}$. Moreover, by the condition (b) it follows that $\lambda_i=0$ for $i\in I\setminus I_{\tilde{x}}$ i.e.,

$$\lambda_i D^+g_i(\tilde{x},x-\tilde{x})\leq 0,\ \text{for all } x\in F,\ i\in I. \tag{12}$$

Since the functions h_j, $j\in J^*(\tilde{d})$, are QL-Dini at $\tilde{x}$ (by condition (e)), then, by Lemma 3, the equality $h_j(x)=h_j(\tilde{x})$ implies that

$$h_j'(\tilde{x},x-\tilde{x})=0,\ \text{for all } x\in F,\ j\in J^*(\tilde{d}).$$

By the definition of the set $J^*(\tilde{d})$, we have that

$$h_j'(\tilde{x},\tilde{d})=0 \text{ for } j\in J\setminus\bigcup J^*(\tilde{d}),$$

so that the above condition holds for all $j \in J$. Hence obviously,

$$\mu_j h_j'(\tilde{x}, x - \tilde{x}) = 0, \text{ for all } x \in F, \ j \in J. \tag{13}$$

Now, including (12) and (13) in condition (c), we obtain that

$$D^+ f(\tilde{x}, x - \tilde{x}) \geq 0, \text{ for all } x \in F.$$

We consider now the case (f_1). Since f is PCX-Dini at $\tilde{x}$, then we obtain the inequality $f(x) \geq f(\tilde{x})$, for all $x \in F$, which means that $\tilde{x}$ is a global minimum point of problem (MCP). In the case (f_2), since $D^+ f(\tilde{x}, d) \neq 0$, for all $d \in K_{\tilde{x}}$, we have $D^+ f(\tilde{x}, x - \tilde{x}) > 0$, for all $x \in F$. Next, since f is QCX-Dini at $\tilde{x}$, the above inequality implies that $f(x) > f(\tilde{x})$, for all $x \in F$, which means that this time $\tilde{x}$ is a strict global minimum point of problem (MCP). ∎

REMARK. *If condition (10) is written in the form:*

$$\mathcal{K}_{\tilde{x}} := \{d \in R^n : g_i^\circ(\tilde{x}, d) \leq 0, \, i \in I_{\tilde{x}}; \ h_j'(\tilde{x}, d) = 0, \, j \in J\} \subset K_{\tilde{x}},$$

the assumption of convexity of the upper Dini derivatives is omitted, generalized properties in the conditions (e), (f_1) and (f_2) are in the Clarke sense, and condition (c) has the form:

$$f^\circ(\tilde{x}, d) + \sum_{i \in I} \lambda_i g_i^\circ(\tilde{x}, d) + \sum_{j \in J} \mu_j h_j'(\tilde{x}, d) \geq 0, \text{ for all } d \in K_{\tilde{x}},$$

then we obtain the second version of the optimality conditions for problem (MCP'), just as in Theorem 2.

6. Summary

The results obtained in this work are not surprising, but the purpose was fulfilled. The first order optimality conditions were formulated for the Lipschitz programming problems without the cones. The optimality conditions use only the upper Dini derivative or the Clarke derivative for problems with inequality constraints and additionally the usual directional derivative for problems with mixed conditions. In fact, the Dini and Clarke derivatives are not the only useful generalized derivatives in these cases. The sensible optimality conditions can be obtained for many different classes of derivatives. Here, we mention e.g. the upper Dini-Hadamard derivative (or upper Hadamard derivative or upper hypo-derivative), the Rockafellar derivative (or the Clarke-Rockafellar derivative or the circa-derivative), the weak Rockafellar derivative (or incident derivative or inner epiderivative). The properties and relationships between these and other derivatives are given by Komlósi (1995), Penot (1998) and Elster, Thier-

7. Acknowledgments

The author wishes to thank Professor M. Studniarski from the University of Łódź for his advices during the long preparation of this paper, Professor B.D. Craven from the University of Melbourne for his help in the search for the references, and the anonymous referees for the very important and helpful suggestions which have improved the paper.

References

BAZARAA, M.S., SHERALI, H.D. and SHETTY, C.M. (1993) *Nonlinear programming. Theory and Algorithms.* New York, Second Edition, John Wiley & Sons, Inc.

BAZARAA, M.S. and SHETTY, C.M. (1976) *Foundations of Optimization.* Berlin, Springer Verlag.

CHEW, K.L. and CHOO, E.V. (1984) Pseudolinearity and efficiency. *Mathematical Programming*, **28**, 226–239.

CLARKE, F.H. (1975) Generalized gradients and applications. *Transactions of the American Mathematical Society*, **205**, 247–262.

CLARKE, F.H. (1976) A new approach to Lagrange multipliers. *Math. Oper. Res.*, **1**, 165–174.

CRAVEN, B.D. and WANG, S. (preprint) *Second order optimality conditions in multiobjective programming.* Australia, University of Melbourne.

DIEWERT, W.E. (1981) Alternative characterizations of six kinds of quasiconcavity in the nondifferentiable case with applications to non-smooth programming. In: *Generalized Concavity in Optimization and Economics*, S. Schaible and W.T. Ziemba (eds.), New York, Academic Press, 51–94.

ELLAIA, R. and HASSOUNI, A. (1991) Characterizations of nonsmooth functions through their generalized gradients. *Optimization*, **22** (3), 401–416.

ELSTER, K.-H. and THIERFELDER, J. (1989) On cone approximations and generalized directional derivatives. In: *Nonsmooth Optimization and Related Topics*, F.H. Clarke, V.F. Dem'yanov, F. Giannessi (eds.), Plenum Press, New York, 133–154.

GIANESSI, F. (1989) Semidifferentiable functions and necessary optimality conditions. *Journal of Optimization Theory and Applications*, **60**, 2, 191–241.

GIORGI G. and GUERRAGGIO, A. (1994) First order generalized optimality conditions for programming problems with a set constraints. In: *Generalized Convexity Proceedings*, Pécs, Hungary, Lect. Notes in Econ. and Math. Syst., 405, Springer-Verlag, 171–185.

GIORGI, G. and KOMLÓSI, S. (1995) Dini derivatives in optimization - part III. *Rivista di matematica per la scienze economiche e sociali*, 18°, Fasciolo 1°, 47–63.

GIORGI, G. and MITITELU, S. (1983) Extremum conditions in quasi-convex programming. Dipartimento di Richerca Operativa e Scienze Statistiche, GLOVER, B.M. (1984) Generalized convexity in nondifferentiable programming. *Bulletin of The Austral. Math. Soc.*, **30**, 193–218.

KOMLÓSI, S. (1993) Quasiconvex first-order approximations and Kuhn-Tucker type optimality conditions. *European Journal of Operational Research*, **65**, 248–276.

KOMLÓSI, S. (1993) First and second characterizations of pseudolinear functions. *European Journal of Operational Research*, **67**, 278–286.

KOMLÓSI, S. (1994) Generalized monotonicity in non-smooth analysis, in: *Lecture Notes in Econ. and Math. Systems*, Hungary, Springer-Verlag, 263–275.

KOMLÓSI, S. (1995) Monotonicity and quasimonotonicity in nonsmooth analysis. In: D.-Z. Du, L. Qi, R.S. Womersley, eds., *Recent Advances in Nonsmooth Optimization*. World Scientific Publishing Co. Pte Ltd., 193–214.

KORTANEK, K.O. and EVANS, J.P. (1967) Pseudoconcave Programming and Lagrange regularity. *Operations Research*, **15**, 882–892.

MANGASARIAN, O.L. (1969) *Nonlinear programming*. New York, McGraw-Hill Book Co.

MCCORMICK, G.P. (1967) Second-order conditions for constrained minima. *SIAM J. Appl. Math.*, **15**, 3, 641–652.

MITITELU, S. (1987) Optimality, minimax and duality in nonlinear programming. *Optimization* **18**, 501–506.

MITITELU, S. (1994) A survey on optimality and duality in nonsmooth programming. In: *Generalized Convexity*, S. Komlósi, T. Rapcsak and S. Scheible (eds.), Springer, Heidelberg, 211–228.

PENOT, J.P. (1998) Are generalized derivatives useful for generalized convex functions. In: J.-P. Crouzeix, M. Volle, J.-E. Martinez-Legaz, eds., *Generalized Convexity, Generalized Monotonicity*. Kluwer Academic Publishers, 3–39.

ZANGWILL, W.I. (1969) *Nonlinear Programming; a Unified Approach*. New York, Englewood Cliffs, Prentice-Hall.