Control

and Cybernetics

VOL. 18 (1989) No. 1

Sufficient optimality conditions in terms of the usual gradients for nondifferentiable programming problems

by

MARCIN STUDNIARSKI

Institute of Mathematics
University of Łódź
ul. S. Banacha 22
90-238 Łódź, POLAND

A sufficient optimality condition for a general nondifferentiable programming problem involving a locally Lipschitzian function is presented. It is similar to the second-order sufficiency theorems of R.W. Chaney, but also includes conditions of order higher than two. Contrary to Chaney's results, our theorem is stated in terms of usual (not generalized) gradients of the given function at these points at which they exist. A comparison with the classical higher-order sufficient conditions is also given.

AMS 1980 subject classification. Primary: 90C 30. Secondary: 49B34.

KEY WORDS: higher-order sufficient conditions, nonsmooth optimization, Lipschitzian functions.

1. Introduction

During the last few years, a considerable attention has been paid to higher-order optimality conditions in nonsmooth optimization. In particular, Chaney [2], [3] proved severial variants of sufficient optimality conditions for nondifferentiable programming problems in the n-dimensional space R^n. All those conditions are formulated in terms of certain limits involving generalized gradients a locally Lipschitzian function connected with the given problem. Thus, to apply any of Chaney's results, one needs to know the generalized gradients of the relevant function at all points of same neighbourhood of the point being examined (or, at all points of the intersection of this neighbourhood and a certain cone). In turn, the generalized gradient (in the sense of Clarke) can be defined by means of limits of usual gradients. More precisely, given a locally Lipschitzian function $f: R^n \to R$, we have, by [4, Theorem 2.5.1],

$$\partial f(x) = \mathrm{co}\{\lim \nabla f(x_k) \mid x_k \to x \text{ and } x_k \notin \Omega_f \text{ for all } k\}, \qquad (1.1)$$

where "co" denotes the convex hull and Ω_f is the set of points at which f fails to be differentiable (by Rademacher's theorem, Ω_f is of Lebesgue measure zero).

It is natural to ask whether Chaney's results can be reformulated so as to replace the generalized gradients occurring in the sufficient conditions by usual gradients calculated only at those points at which the function considered is differentiable. The aim of the present paper is to give, at least partially, a positive answer to this question.

Our main result is presented in § 2. Although in the proof we apply clarke's theory of generalized gradients the theorem itself does not contain any notion pertaining to that theory. In § 3 we derive another sufficiency theorem which is formulated in terms of some notions used by Chaney [3] and provides sufficient conditions for a point to be a stable local minimum. We also show that this result implies one particular case of Chaney's general sufficiency theorem [3, Theorem 2.5]. Finally, in § 4 we show that some classical higher-order sufficiency theorems for differentiable programming problems follow from our results.

Let us now set some notation. Given x and y in R^n and $\varepsilon > 0$, we denote by $|x|$ the Euclidean norm of x, by $x \cdot y$ the usual inner product of x and y, and by $B(x, \varepsilon)$ the set $\{z \in R^n |\, |z-x| \leqslant \epsilon\}$. If $x \neq y$, we denote by $]x, y[$ the open line segment joining x and y (the same notation is used for open intervals in $\bar{R} = R \cup \{\pm\infty\}$).

In the paper we make use of some notions and theorems of nonsmooth analysis which can be found in [4, Chapter 2]. In particular, if f is a locally Lipschitzian function, we define

$$f^\circ(x\,;\,d) := \limsup_{y \to x\,;\,\lambda \downarrow 0} \lambda^{-1} (f(y + \lambda d) - f(y)).$$

By [4, Proposition 2.1.2], we have

$$f^\circ(x\,;\,d) = \max\{w \cdot d \mid w \in \partial f(x)\} \text{ for all } d \in R^n\,. \tag{1.2}$$

We also use the notation

$$f'(x\,;\,d) := \lim_{\lambda \downarrow 0} \lambda^{-1} (f(x + \lambda d) - f(x))$$

provided this limit exists. We recall that f is said to be subdifferentiably regular at x if $f'(x\,;\,d)$ exist for all $d \in R^n$, and $f'(x\,;\,\cdot) = f^\circ(x;\,\cdot)$.

If S is a subset of R^n, we denote by $K(S, x)$ the contingent cone to S at x. It may be defined in some equivalent manners. We shall use the following definition: $d \in K(S, x)$ if and only if there exist sequences $\{x_k\}$ in S and $\{\lambda_k\}$ of positive numbers, such that $\{x_k\}$ converges to x and $\{\lambda_k(x_k - x)\}$ converges to d. (Let us note that $K(S, x)$ is empty whenever x lies outside the closure of S.) For other characterizations of the contingent cone and some of its properties, see [10, pp. 13-16].

Let d be a unit vector in R^n. We shall say that a sequence $\{x_k\}$ converges to x in the direction d if $\{x_k\}$ converges to x and $\{(x_k - x) / | x_k - x |\}$ converges to d. (Thus, a unit vector d belongs to $K(S, x)$ if and only if there exists a sequence $\{x_k\}$ in S converging to x in the direction d.)

A locally Lipschitzian function f is said to be semismooth at x if the sequence $\{v_k \cdot d\}$ is always convergent whenever $\{x_k\}$ and $\{v_k\}$ are sequences such that $\{x_k\}$ converges to x in the direction d, and $v_k \in \partial f(x_k)$ for all k. (This definition is taken from [3] and is easily seen to be equivalent to the original definition of Mifflin [9].)

2. The main result

In this section we assume that S and W are two subsets of R^n, the set W is open and $f : W \to \overline{R}$ is an extended real-valued function. Next, we assume that $\bar{x} \in S \cap W$, and $f(\bar{x})$ is finite. Similarly as in [3], we consider the following optimization problem:

$$\text{minimize } f(x) \text{ subject to } x \in S \cap W. \tag{P}$$

We shall say that $\bar{x}$ is an isolated local solution of problem (P) if there exists a neighbourhood V of $\bar{x}$ such that $f(x) > f(\bar{x})$ for all $x \in S \cap V$, $x \neq \bar{x}$.

We can now formulate our main result. Let us define

$$E_f := \{x \in W \mid x \neq \bar{x} \text{ and } f(x) \leqslant f(\bar{x})\}. \tag{2.1}$$

THEOREM 2.1. *Suppose that one of the two conditions holds:*

(i) *The cone $C := K(S, \bar{x}) \cap K(E_f, \bar{x})$ contains no nonzero element.*

(ii) *C contains nonzero elements and there exists a neighbourhood U of $\bar{x}$ $(U \subset W)$*

and two functions: $g : U \to R$ and $\psi : U \setminus \{\bar{x}\} \to]0, +\infty[$, such that

(a) *g is Lipschitzian, $g(\bar{x}) = f(\bar{x})$ and $g(x) \leqslant f(x)$ for all $x \in S \cap U$;*

(b) *ψ is continuous and we have*

$$\limsup \nabla g(x_k) \cdot \psi(x_k - \bar{x}) > 0$$

whenever $\{x_k\}$ is a sequence in U such that g is differentiable at x_k for all k and $\{x_k\}$ converges to $\bar{x}$ in some direction $d \in C$.

Then $\bar{x}$ is an isolated local solution of problem (P).

Before proceeding to the proof of the theorem, let us give some comments.

REMARKS 2.2. (a) Under some regularity assuptions on f, the cone $K(E_f, \bar{x})$ can be characterized in terms of the generalized directional derivative $f^\circ(\bar{x}\,;\cdot)$ (see [4, Theorems 2.4.7 and 2.9.10]). However, in the general case, this cone may be difficult to determine. Therefore, it may prove convenient to find a larger cone for which the theorem still holds, and which can be calculated more easily. An example of such a cone is the cone $L(f, \bar{x})$ defined in § 3.

(b) The role of the function g in Theorem 2.1 is the same as in [3, Theorem 2.5]. As Chaney noticed, g may be chosen either to be f itself (if f is locally Lipschitzian), or a certain Lagrangian associated with (P) (see § 4) or a function whose generalized gradient is simpler than that of f.

(c) The function ψ may be chosen arbitrarily, but the most natural way is to choose $\psi(x) = |x-\bar{x}|^{-m}$ where m is any positive integer. We shall show in § 4 that this choice of ψ leads to a generalization of the classical m-th order sufficient conditions. For the case when S is convex, a sufficient condition involving a general function ψ of this type was proved in [12, § 5].

Proof of theorem 2.1 Suppose that the desired conclusion is false. Then there exist a sequence $\{z_k\}$ in $S \cap W$ converging to x, such that $z_k \neq \bar{x}$ and $f(z_k) \leqslant f(\bar{x})$ for all k. By passing to a subsequence, we may assume that $\{z_k\}$ converges to $\bar{x}$ in some direction d. Since $z_k \in S \cap E_f$ for all k, we infer that $d \in C$. If (i) holds, then we reach a contradiction, Hence suppose that (ii) holds. By assumption (a), we have $g(z_k) \leqslant f(z_k) \leqslant f(\bar{x}) = g(\bar{x})$ for all k. According to the mean value therem of Lebourg [4, Theorem 2.3.7.], for every k, there exist $u_k \in]\, \bar{x}, z_k[$ and $v_k \in \partial g(u_k)$ such that $g(\bar{x}) - g(z_k) = v_k \cdot (\bar{x} - z_k)$. Hence, by (1.2) , we obtain $0 \leqslant g(\bar{x}) - g(z_k) \leqslant g^\circ(u_k; \bar{x} - z_k)$. Since $g^\circ(u_k;)$ is positively homogeneous and

$$\bar{x} - u_k = \lambda_k(\bar{x} - z_k) \text{ for some } \lambda_k > 0, \tag{2.2}$$

we have $\psi(u_k)g^\circ(u_k;\bar{x} - u_k) = \lambda_k \psi(u_k)g^\circ(u_k;\bar{x} - z_k) \geqslant 0$. By passing again to subsequences, we may assume that the limit

$$\lim \psi(u_k)\, g^\circ(u_k; \bar{x} - u_k) \geqslant 0 \tag{2.3}$$

exists (it may either be finite or equal to $+\infty$). By (2.2), we have $(u_k - \bar{x})/|u_k - \bar{x})| = (z_k - \bar{x})/|z_k - \bar{x}|$ for all k, and so, $\{u_k\}$ converges to $\bar{x}$ in the direction d.

For every k, let us denote by $Z(u_k)$ the set of all limits of the form $\lim_{r \to \infty} \nabla g(x_{k,r})$ where $\{x_{k,r}\}_{r=1}^{\infty}$ is a sequence converging to u_k, such that g is differentiable at $x_{k,r}$ for all r. In view of (1.1), we have $\partial g(u_k) = \operatorname{co} Z(u_k)$. Then, it is easy to verify that $\max\,\{w \cdot y | w \in \partial g(u_k)\} = \max\,\{w \cdot y | w \in Z(u_k)\}$ for each $y \varepsilon R^n$(the latter maximum exists since $Z(u_k)$ is nonempty and compact). It follows from this equality and from (1.2) that, for every k, there exists $w_k \in Z(u_k)$ such that

$$g^{\circ}(u_k; \bar{x} - u_k) = w_k \cdot (\bar{x} - u_k). \tag{2.4}$$

Let $K>0$ be the Lipschitz constant for g on U; then $|\nabla g(x)| \leqslant K$ whenever $x \in U$ and g is differentiable at x. Next, let us note that, for every k, the function

$$\varphi_k(x) := \left| \frac{x - \bar{x}}{|x - \bar{x}|} - \frac{u_k - \bar{x}}{|u_k - \bar{x}|} \right| \tag{2.5}$$

is continuous on some neighbourhood of u_k, and $\varphi_k(u_k) = 0$. Hence, for every k, we can choose a positive number δ_k such that

$$\varphi_k(x) \leqslant 1/k \text{ for all } x \in B(u_k, \delta_k) \subset U \setminus \{\bar{x}\}. \tag{2.6}$$

Since ψ is also continuous, the number $M_k := \sup\{\psi(x) \mid x \in B(u_k, \delta_k)\}$ is finite. Now, using the fact that $w_k \in Z(u_k)$ and again the continuity of ψ, we find that, for every k, there exists a point x_k satisfying the conditions:

$$|u_k - x_k| \leqslant \min(\delta_k, |\bar{x} - u_k| / KM_k, 1/k); \tag{2.7}$$

$$g \text{ is differentiable at } x_k \text{ and}$$

$$|\nabla g(x_k) - w_k| \leqslant 1/M_k; \tag{2.8}$$

$$|\psi(u_k) - \psi(x_k)| \leqslant 1/w_k| \text{ provided } w_k \neq 0. \tag{2.9}$$

It follows from (2.5) — (2.7) that the sequences $\{(x_k - \bar{x})/|x_k - \bar{x}|\}$ and $\{(u_k - \bar{x})/|u_k - \bar{x}|\}$ have the same limit d. Using the third estimate in (2.7), we see that $\{x_k\}$ converges to $\bar{x}$. We have thus verified that $\{x_k\}$ converges to $\bar{x}$ in the direction d.

Applying successively (2.4) and (2.7) — (2.9), we obtain

$$\begin{aligned}
&|g(x_k) \cdot \psi(x_k)(\bar{x} - x_k) - \psi(u_k)\, g^{\circ}(u_k; \bar{x} - u_k)| = |\psi(x_k)\nabla g(x_k) \cdot \bar{x} - x_k) - \\
&- \psi(u_k) w_k \cdot (\bar{x} - u_k)| = |\psi(x_k)[\nabla g(x_k) \cdot (u_k - x_k) + (\nabla g(x_k) - w_k) \cdot \bar{x} - u_k)] - \\
&- (\psi(u_k) - \psi(x_k))\, w_k \cdot (\bar{x} - u_k)| \\
&\leqslant M_k[K\,|u_k - x_k| + |\nabla g(x_k) - w_k||\bar{x} - u_k|] + |\psi(u_k) - \psi(x_k|\,|\bar{x} - u_k| \\
&\leqslant 3\,|\bar{x} - u_k| \xrightarrow[k \to \infty]{} 0.
\end{aligned}$$

This and (2.3) allow us to conclude that

$$\lim \nabla g(x_k) \cdot \psi(x_k)(x_k - \bar{x}) = -\lim \nabla g(x_k) \cdot \psi(x_k)(\bar{x} - x_k) \leqslant 0,$$

which contradicts assumption (b).

3. Sufficient conditions for the stability of minimum points

Throughout this section, S, W and $\bar{x}$ will be as in §2, but $f: W \to R$ will be locally Lipschitzian. We shall now prove another sufficiency theorem in which both the assumptions and the conclusion are somewhat stronger than in Theorem 2.1. This result can be derived from Theorem 2.1 if we take $\psi(x) = |x - \bar{x}|^{-m}$ where $m \geqslant 2$ is a positive integer. In this case, we can prove not only sufficient conditions for $\bar{x}$ to be an isolated local solution of problem (P), but also to be an isolated local minimum with order m of (P) (using the terminology of [1], [13]). This, in turn, allows us to infer that the minimum point is stable in the following sense: all functions (of a suitable class) which are sufficiently close to f have local minimum points relative to S within a prescribed distance from $\bar{x}$. A sufficient condition for stability in this sense has been established by Hyers in [7], [8] and is formulated below in the finite-dimensional setting.

THEOREM 3.1 (*Hyers*). *Let ρ: $[0, +\infty[\to R$ be a strictly increasing function with $\rho(0) = 0$. Let E be a closed subset of R^n, let $\bar{x} \in E$, and let $h: E \to R$ be a function such that*

$$h(x) - h(\bar{x}) \geqslant \rho(|x - \bar{x}|) \text{ for all } x \in E. \tag{3.1}$$

For a given $\varepsilon > 0$, let $\tilde{h}: E \to R$ be any lower semicontinuous function satisfying the inequality $|\tilde{h}(x) - h(x)| < \rho(\varepsilon)/2$ for all $x \in E \cap B(\bar{x}, \varepsilon)$. Then $\tilde{h}$ has a minimum value on $E \cap B(\bar{x}, \varepsilon)$ which is taken on at an interior point of $B(\bar{x}, \varepsilon)$.

The strengthening of assumptions in our next sufficiency theorem will consist in replacing the cone $K(E_f, \bar{x})$ by a larger cone $L(f, \bar{x})$ introduced by Chaney in [3]. Let us recall how it is defined: First, for a given unit vector d in R^n, we define $\partial_d f(\bar{x})$ to be the set of all v in R^n for each of which there exist sequences $\{x_k\}$ in W and $\{v_k\}$ in R^n, such that $\{x_k\}$ converges to $\bar{x}$ in the direction d, $\{v_k\}$ converges to v, and $v_k \in \partial f(x_k)$ for all k. Next, we define $L(f, \bar{x})$ to be the set of all points d where $d \in R^n$, $|d| = 1$, $t \geqslant 0$, and $v_0 \cdot d \leqslant 0$ for some $v_0 \in \partial_d f(\bar{x})$.

LEMMA 3.2. *If f is locally Lipschitzian and E_f is defined by formula (2.1), then $K(E_f, \bar{x}) \subset L(f, \bar{x})$.*

Proof. Suppose that $d \in K(E_f, \bar{x})$. Then there exist sequences $\{y_k\}$ in E_f and $\{\lambda_k\}$ of positive numbers, such that $\bar{x} = \lim y_k$ and $d = \lim \lambda_k (y_k - \bar{x})$. By the definition of E_f, we have $y_k \neq \bar{x}$ and $f(y_k) \leqslant f(\bar{x})$ for all k. Let us now consider two cases: (i) $|d| = 1$ and (ii) $d = 0$. We shall show that, in both cases, $d \in L(f, \bar{x})$.

Case (i). We have $\lim |\lambda_k (y_k - \bar{x})| = |d| = 1$, and so,

$$\lim \frac{y_k - \bar{x}}{|y_k - \bar{x}|} = \lim \frac{\lambda_k (y_k - \bar{x})}{|\lambda_k (y_k - \bar{x})|} = d. \tag{3.2}$$

For every k, by Lebourg's mean value theorem [4, Theorem 2.3.7], there exist

$u_k \in]\bar{x}, y_k[$ and $v_k \in \partial f(u_k)$ such that $0 \geqslant f(y_k) - f(\bar{x}) = v_k \cdot (y_k - \bar{x})$. Inasmuch as $u_k - \bar{x}$ is a positive multiple of $y_k - \bar{x}$, we obtain from (3.2) that $\lim (u_k - \bar{x}) / |u_k - \bar{x}| = d$.

Thus, $\{u_k\}$ converges to $\bar{x}$ in the direction d. Since the sequence $\{v_k\}$ is bounded in norm (cf. [4, Proposition 2.1.2 (a)], we may assume that it converges to a point $v_o \in \partial_d f(\bar{x})$. Furthermore, $v_o \cdot d = \lim\ v_k \cdot (y_k - \bar{x}) / |y_k - \bar{x}| = \lim (f(y_k) - f(\bar{x})) / |y_k - \bar{x}| \leqslant 0$, and so, $d \in L(f, \bar{x})$.

Case (ii). By taking a subsequence, we may assume that $\{(y_k - \bar{x}) / |y_k - \bar{x}|\}$ converges to a unit vector d_1. Similarly as in case (i), we can prove that $d_1 \in L(f, \bar{x})$. Hence $d = 0 \cdot d_1 \in L(\mathrm{f}, \bar{x})$.

REMARK 3.3. If f is semismooth at $\bar{x}$, then $f'(\bar{x}, d) = v_o \cdot d$ whenever $v_o \in \partial_d f(\bar{x})$ (see the proof of [2, Theorem 2.16]). Hence $\{d \in R^n | f'(\bar{x}; d) \leqslant 0\} = L(f, \bar{x}) \cup \{0\}$ (note that the set on the left always contains 0, whereas $L(f, \bar{x})$ may be empty). If, moreover, f is subdifferentiably regular at $\bar{x}$, and $0 \notin \partial f(\bar{x})$, then $K(E_f, \bar{x})$ is nonempty and $K(E_f, \bar{x}) = \{d \in R^n | f^{\circ}(\bar{x}; d) \leqslant 0\} = L(f, \bar{x})$ by [4, Theorem 2.4.7]. However, if $0 \in \partial f(\bar{x})$, the inclusion in Lemma 3.2 may be strict even for a smooth f. For instance, let $f: R \to R$ be given by $f(x) = x^2$, and let $\bar{x} = 0$. Then $K(E_f, \bar{x}) = \emptyset$, while $L(f, \bar{x}) = R$.

THEOREM 3.4. *Let f be locally Lipschitzian on W, and let $\bar{x} \in S \cap W$. Take any $\mu \geqslant 0$. Suppose that one of the two conditions holds:*

(i) *The cone $D := K(S, \bar{x}) \cap L(f, \bar{x})$ contains no nonzero element.*

(ii) *D contains nonzero elements and there exist a positive integer $m \geqslant 2$ and a function $g: U \to R$ (where U is open and $\bar{x} \in U \subset W$)* such that

(a) *g is Lipschitzian, $g(\bar{x}) = f(\bar{x})$ and $g(x) \leqslant f(x)$ for all $x \in S \cap U$;*

(b) *we have*

$$\lim\sup \nabla g(x_k) \cdot (x_k - \bar{x}) / |x_k - \bar{x}|^{\,m} > \mu \tag{3.3}$$

whenever $\{x_k\}$ is a sequence in U such that g is differentiable at *x_k for all k and $\{x_k\}$ converges to $\bar{x}$* in some direction $d \in D$.

Then there exists $\varepsilon > 0$ such that

$$f(x) > f(\bar{x}) + (\mu / m)\, |x - \bar{x}|^{\,m} \text{ for all } x \in S \cap B(\bar{x}, \varepsilon),\ x \neq \bar{x}. \tag{3.4}$$

If, moreover, S is closed and $\mu > 0$, then any real valued lower semicontinuous function $\tilde{f}$ defined on $S \cap B(\bar{x}, \varepsilon)$ and satisfying the inequality

$$|\tilde{f}(x) - f(x)| < \mu\varepsilon^m / 2m \text{ for all } x \in S \cap B(\bar{x}, \varepsilon)$$

has a minimum value on $S \cap B(\bar{x}, \varepsilon)$ which is taken on at an interior point of $B(\bar{x}, \varepsilon)$.

P r o o f. Define a function φ on W by $\varphi(x) := f(x) - \eta(x)$ where $\eta(x) := (\mu/m) |x - \bar{x}|^{\,m}$. We shall show that the assuumptions of Theorem 2.1 are satisfied with

f replaced by φ. Since η is continuously differentiable, we have $\partial\eta(x) = \{\nabla\eta(x)\}$ for all x (cf. [4, Proposition 2.2.4]). Moreover, $\nabla\ \eta(x) = \mu|x-\bar{x}|^{m-2}\,(x-\bar{x})$ Hence, it is easy to verify (by using the well-known calculus rules for generalized gradients; cf. [4, p. 38]) that $\partial_d\varphi(\bar{x}) = \partial_d f(\bar{x})$ for all unit vectors d. Consequently, we have

$$L\ (\varphi, \bar{x}) = L\ (f,\bar{x}). \tag{3.5}$$

Suppose that (i) holds. Then, by (3.5), the cone $K(S,\bar{x})\cap L(\varphi,\bar{x})$ contains no nonzero element. Applying Lemma 3.2 to the function φ, we infer that $K(S,\bar{x})\cap K(E_\varphi,\bar{x})$ also contains no nonzero element. Thus, condition (i) of Theorem 2.1 holds with f replaced by φ.

Suppose now that (ii) holds, and let $\tilde{g} : U\to R$ be given by $\tilde{g}(x) := g(x) - \eta(x)$. We shall verify that condition (ii) of Theorem 2.1 holds with $\psi(x) = |x-\bar{x}|^{-m}$ and with f and g replaced by φ and $\tilde{g}$, respectively. Obviously, we have $\tilde{g}(\bar{x}) = \varphi(\bar{x})$ and $\tilde{g}(x) \leqslant \varphi(x)$ for all $x\in S\cap U$. Next, let $\{x_k\}$ be any sequence in U such that $\tilde{g}$ is differentiable at x_k for all k and $\{x_k\}$ converges to $\bar{x}$ in some direction $d\in K(S,\bar{x})\cap K(E_\varphi,\bar{x})$. Then g is also differentiable at x_k for all k. Using Lemma 3.2 and equality (3.5), we find that $d\in D$, and so, by our assumption, inequality (3.3) holds. Hence

$$\begin{aligned}&\limsup \nabla\tilde{g}(x_k)\cdot(x_k-\bar{x})/|x_k-\bar{x}|^{m} =\\ &= \limsup\,[\nabla g(x_k) - \mu|x_k-\bar{x}|^{m-2}\,(x_k-\bar{x})]\cdot(x_k-\bar{x})/|x_k-\bar{x}|^{m} =\\ &= \limsup \nabla g(x_k)\cdot(x_k-\bar{x})/|x_k-\bar{x}|^{m} - \mu > 0.\end{aligned}$$

Thus, we can apply Theorem 2.1 to the function φ. We get that there exists $\varepsilon>0$ such that $\varphi(x) > \varphi(\bar{x})$ for all $x\in S\cap B(\bar{x},\varepsilon)$, $x \neq \bar{x}$. Hence (3.4) holds. The final statement of the theorem follows directly from (3.4) and Theorem 3.1 (in which one should take $\rho(t) = \mu t^m/m$, $E = S \cap B(\bar{x},\varepsilon)$ and $h = f\,|S \cap B(\bar{x},\varepsilon)$). ■

REMARK 3.5. For $m = 2$, Theorem 3.4 implies a particular case of a general second-order sufficiency theorem due to Chaney [3, Theorem 2.5], namely, the case where $C(d^*) = \{td^* | \ t \geqslant 0\}$ for all d^*. To see this, let us first observe that, with this choice for $C(d^*)$, condition (b) (iv) of [3, Theorem 2.5] may be omitted. In fact, if $\{x_k\}$, $\{w_k\}$ and $d = d^*$ satisfy conditions (b) (i) — (b) (iii) of Chaney's theorem, then the sequence $\{w_k\}$ is bounded (cf. [4, Proposition 2.1.2 (a)]). Hence, there exists a subsequence of $\{w_k\}$ converging to some $w \in \partial_d\ g(x^*)$. By assumption (a) of Chaney's theorem, we have $w\cdot d \geqslant 0$, and so, (b) (iv) holds for this subsequence. Therefore, if assumption (b) of [3, Theorem 2.5] is valid as stated, it continues to be valid when condition (iv) is deleted. Next, let us note that assumption (b) of Theorem 3.4 is weaker than assumption (b) of [3, Theorem 2.5] with the restrictions (i) — (iii) only. This follows from the fact that, in view of (1.1), we have $\nabla g(x) \in \partial g(x)$ whenever g is differentiable at x.

4. The case of differentiable functions

We shall consider here problem (P) in the case when S is given by

$$S = \{x \in W \mid g_i(x) \leqslant 0,\ i \in I;\ g_j(x) = 0,\ j \in J\} \tag{4.1}$$

where I and J are given finite sets and W is, as before, an open subset of R^n. Let $m \geqslant 2$ be a fixed integer. We shall assume that the functions f and g_i, $i \in I \cup J$, are m times (Fréchet) differentiable on W. The aim of this section is to derive the classical higher-order sufficient optimality conditions from Theorem 3.4.

We denote the r-th differential of f at $x \in W$ with the incerement d by $f^{(r)}(x)\,d^r$ (where $r \leqslant m$ and $d^r = (d,...,d) \in (R^n)^r$). For a given $x \in S \cap W$, we denote by $I(x)$ the set $\{i \in I \mid g_i(x) = 0\}$.

THEOREM 4.1. *Let $\bar{x} \in S \cap W$. Suppose that one of the two conditions holds:*

(i) *There is no solution d to the system*

$$\begin{aligned} &\nabla f(\bar{x}) \cdot d \leqslant 0,\ \nabla g_i(\bar{x}) \cdot d \leqslant 0 \text{ for } i \in I(\bar{x}), \\ &\nabla g_j(\bar{x}) \cdot d = 0 \text{ for } j \in J,\ d \neq 0. \end{aligned} \tag{4.2}$$

(ii) *The set of solutions to* (4.2) *is nonempty and there exist multipliers λ_i, $i \in I \cup J$, such that*

(a) $\lambda_i \geqslant 0$ *and* $\lambda_i\, g_i(\bar{x}) = 0$ *for* $i \in I$;

(b) *the Lagrangian* $L := f + \sum_{i \in I \cup J} \lambda_i g_i$ *satisfies the following conditions:*

$$L^{(r)}(\bar{x}) = 0 \text{ for } r = 1,...,m-1, \tag{4.3}$$

$$L^{(m)}(\bar{x})\,d^m > 0 \text{ for all } d \text{ satisfying (4.2).} \tag{4.4}$$

Then there exist $\beta > 0$ and a neighbourhood V of $\bar{x}$, such that $f(x) > f(\bar{x}) + \beta|x - \bar{x}|^m$ for all $x \in S \cap V$, $x \neq \bar{x}$.

Proof. Since f is continuously differentiable, we have $\partial f(x) = \{\nabla f(x)\}$ for all $x \in W$, and so, $L(f,\bar{x}) \subset \{d \in R^n \mid \nabla f(\bar{x}) \cdot d \leqslant 0\}$. Next, it can easily be shown (cf. [6, pp. 221-222]) that $K(S,\bar{x})$ is contained in the set of directions d satisfying the conditions $\nabla g_i(\bar{x}) \cdot d \leqslant 0$ for $i \in I(\bar{x})$ and $\nabla g_j(\bar{x}) = 0$ for $j \in J$. Suppose that (i) holds. Then it follows from the inclusions just stated that condition (i) of Theorem 3.4 is also fulfilled. In this case, we can always choose $\mu > 0$ in Theorem 3.4, and so, the desired conclusion holds.

Suppose now that (ii) is true. In order to verify condition (ii) of Theorem 3.4, we put $U = W$ and $g = L$. Since the set of all unit vectors d satisfying (4.2) is nonempty and compact, the function $d \to L^{(m)}(\bar{x})\, d^m/(m-1)!$ attains its minimal value γ on this set. In view of (4.4), we have $\gamma > 0$. Let us choose μ in Theorem 3.4 to be an arbitrary number in $[0, \gamma]$. It follows from (4.1) and assumption (a) of the present theorem that $L(\bar{x}) = f(\bar{x})$ and $L(x) \leqslant f(x)$ for all $x \in S \cap W$. Suppose further that $\{x_k\}$ is a sequence in W converging to $\bar{x}$ in some direction $d \in K(S, \bar{x}) \cap L(f, \bar{x})$. Then d is a unit vector satisfying (4.2), and so,

$$L^{(m)}(\bar{x})\, d^m / (m-1)\,! > \mu. \tag{4.5}$$

Applying Taylor's formula to the mapping $\nabla L : W \to R^n$ and using (4.3), we obtain that, for each $x \in W$,

$$\nabla L(x) = (\nabla L)^{(m-1)}(\bar{x})\,(x - \bar{x})^{m-1} / (m-1)\,! + |x - \bar{x}|^{m-1} R(x)$$

where $\lim_{x \to \bar{x}} R(x) = 0$. Hence

$$\nabla L(x) \cdot (x - \bar{x}) / |x - \bar{x}|^m$$

$$= \frac{1}{(m-1)\,!} L^{(m)}(\bar{x}) \left(\frac{x - \bar{x}}{|x - \bar{x}|} \right)^m + R(x) \cdot \left(\frac{x - \bar{x}}{|x - \bar{x}|} \right).$$

Substituting x_k for x and using the fact that $\{(x_k - \bar{x}) / |x_k - \bar{x}|\}$ converges to d, we get

$$\lim \nabla L(x_k) \cdot (x_k - \bar{x}) / |x_k - \bar{x}|^m = L^{(m)}(\bar{x})\, d^m / (m-1)\,!$$

This and (4.5) give us inequality (3.3) with $g = L$. Applying Theorem 3.4, we get the desired conclusion.

REMARK 4.2. For $m = 2$, the assumptions of Theorem 4.1 are the same as in [5, Theorem 3.2], but the conclusion of our theorem is slightly stronger. Suppose now that $\bar{x} \in S \cap W$ is a point at which the Karush-Kuhn-Tucker necessary optimality conditions for problem (P) are satisfied, i.e. there exist multipliers λ_i, $i \in I \cup J$, satisfying condition (ii) (a) of Theorem 4.1, such that $\nabla L(\bar{x}) = 0$. Then, by virtue of [5, Theorem 3.5], conditions (4.2) are equivalent to the following ones:

$$\nabla g_i(\bar{x}) \cdot d \leqslant 0 \text{ for } i \in I(\bar{x}) / I^*,$$

$$\nabla g_i(\bar{x}) \cdot d = 0 \text{ for } i \in I^* \cup J,\ d \neq 0$$

where $I^* := \{i \in I \mid \lambda_i > 0$. This shows that Theorem 4.1 generalizes both [6, Theorem 7.4] and the theorem proved in [11]. Moreover, if the Karush-Kuhn-Tucker optimality conditions hold, then the inequality $\nabla f(\bar{x}) \cdot d$

$\leqslant 0$ in (4.2) can be replaced by the equality $\nabla f(\bar{x}) \cdot d = 0$. This can be deduced from the proof of [5, Theorem 3.5].

Acknowledgment. The author wishes to thank D. Zagrodny (Łódź) for his helpful suggestion concerning the proof of Theorem 3.4.

References

[1] AUSLENDER A. Stability in Mathematical Programming with Nondifferentiable Data. *SIAM J. Control Optim.* **22** (1984), 239-254.

[2] CHANEY R.W. Second-Order Sufficiency Conditions for Nondifferentiable Programming Problems. *SIAM J. Control Optim.* **20** (1982), 20-33.

[3] CHANEY R.W. A General Sufficiency Theorem for Nonsmooth Nonlinear Programming. *Trans. Amer. Math. Soc.* **276** (1983), 235-245.

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

[5] HAN S.P, MANGASARIAN O.L. Exact Penalty Functions in Nonlinear Programming. *Math. Programming* **17** (1979), 251-269.

[6] HESTENES M.R. Optimization Theory. The Finite Dimensional Case. New York. J.Wiley, 1975.

[7] HYERS D. H. On the Stability of Minimum Points. *J.Math. Anal. Appl.* **62** (1978), 530-537.

[8] HYERS D. H. Stability of Minimum Points for Problems with Constraints. Differential Geometry, Calculus of Variations, and their Applications. *Lect. Notes. Pure Appl. Math.* **100** (1985), 283-289.

[9] MIFFLIN R. Semismooth and Semiconvex Functions in Constrained Optimization. *SIAM J. Control Optim.* **15** (1977), 959-972.

[10] ROCKAFELLAR R.T. The Theory of Subgradients and its Applications to Problems of Optimization: Convex and Nonconvex Functions. Berlin, Heldermann, 1981.

[11] ROLEWICZ S. On Sufficient Conditions of Optimality of Second Order. *Ann. Polon. Math.* **42** (1983), 297-300.

[12] STUDNIARSKI M. Mean Value Theorems and Sufficient Optimality Conditions for Nonsmooth Functions. *J. Math. Anal. Appl.* **111** (1985), 313-326.

[13] STUDNIARSKI M. Necessary and Sufficient Conditions for Isolated Local Minima of Nonsmooth Functions. *SIAM J. Control Optim.* **24**(1986), 1044-1049.

Received, July 1988

Warunki dostateczne optymalności w terminach zwykłych gradientów dla niegładkich zadań programowania matematycznego

W pracy przedstawiono warunki dostateczne optymalności dla ogólnego niegładkiego zadania programowania matematycznego, w którym funkcja minimalizowana jest lokalnie lipschitzowska. Twierdzenia te są podobne do znanych wyników R.W. Chaney'a, ale zawierają także warunki dostateczne dowolnego rzędu. W odróżnieniu od wyników Chaney'a, w sformułowaniach przedstawionych tu twierdzeń występują tylko zwykłe (nie uogólnione) gradienty funkcji, obliczone w tych punktach, w których istnieją.

Достаточные условия оптимальности в терминах обыкновенных градиентов для негладких задач математического програмирования

В работе представлены достаточные условия оптимальности для общей негладкой задачи математического програмирования, в которой минимизируемая функция локально липшицева. Эти теоремы похожи на известные результаты Р.В. Ченея, но они включают также достаточные условия произвольного порядка. В отличие от результатов Ченея, в формулировках представленных здесь теорем участвуют только обыкновенные не обобщенные градиенты функций, вычисленные в тех точках, в которых они существуют.