International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

Quasidifferentiable Functions: Necessary Conditions and Descent Directions

Demyanov, V.F.

IIASA Working Paper

WP-83-064

June 1983

QUASIDIFFERENTIABLE FUNCTIONS: NECESSARY CONDITIONS AND DESCENT DIRECTIONS

V.F. Demyanov

June 1983
WP-83-64

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

The System and Decision Sciences group at IIASA has a long tradition of research in the theory and practice of mathematical optimization. Necessary conditions play a very important role in optimization theory: they provide a means of checking the optimality of a given point and in many cases enable a direction of descent to be found.

In this paper the author studies the necessary conditions for an extremum when either the function to be optimized or the function describing the set on which optimization must be carried out is nondifferentiable. The author's main concern is with quasidifferentiable functions but smooth and convex cases are also discussed.

Andrzej Wierzbicki
Chairman
System and Decision Sciences

QUASIDIFFERENTIABLE FUNCTIONS: NECESSARY CONDITIONS AND DESCENT DIRECTIONS

V.F. Demyanov

1. INTRODUCTION

To solve optimization problems in practice it is necessary to be able to check whether a given point is an extreme point or not, and if it is not, to find a point which is in some sense "better". This is generally achieved through the specification of conditions necessary for optimality. This paper is concerned with extremal problems involving a new class of nondifferentiable functions - the so-called quasidifferentiable functions. Only minimization problems are discussed, without loss of generality.

Different forms of necessary conditions yield different descent directions which can be used to develop a variety of numerical algorithms. Subsections 1.1 and 1.2 provide a brief summary of related problems in mathematical programming and convex analysis.

1.1 *Mathematical programming problems*

Let $\Omega \subset E_n$, $x \in \mathrm{cl}\Omega$ where $\mathrm{cl}\Omega$ denotes the closure of Ω. Set

$$\Gamma(x) = \left\{ v \in E_n \mid \exists \lambda > 0, \{x_i\} : x_i \to x, x_i \neq x, x_i \in \Omega, \frac{x_i - x}{\|x_i - x\|} \to g, v = \lambda g \right\} . \tag{1.1}$$

It is clear that $\Gamma(x)$ is a closed cone. $\Gamma(x)$ is called the set of feasible (in a broad sense) directions of the set Ω at the point x.

Now consider the problem of minimizing a continuously differentiable function f on the set Ω. Let $f^* = \inf_{x\in\Omega} f(x)$.

<u>*Theorem 1*</u>. *For a point* $x^* \in \mathrm{cl}\Omega$ *to be an infimum of* f *on* Ω *it is necessary that*

$$(f'(x^*), v) \geq 0 \qquad \forall v \in \Gamma(x^*) \tag{1.2}$$

where (a,b) *denotes the scalar product of* a *and* b, *and* f'(x) *represents the gradient of* f *at* x.

Unfortunately it is difficult to use this trivial condition in practice.

Let $A \subset \Gamma(x)$ be a convex cone and let $A(x)$ be a family of convex cones such that

$$A \subset \Gamma(x) \qquad \forall A \in A(x) \quad , \qquad \bigcup_{A\in A(x)} A = \Gamma(x) \quad . \tag{1.3}$$

In [1] cones of this type are called "tents". It is always possible to find a family $A(x)$ defined as above (take, for example, $A(x) = \{\ell \mid \ell = \{v = \lambda v_0 \mid \lambda > 0\},\ v_0 \in \Gamma(x)\}$). We denote by A^+ the cone conjugate to A: $A^+ = \{w \in E_n \mid (v,w) \geq 0\ \forall v \in A\}$.

<u>*Theorem 2*</u>. *Condition (1.2) is equivalent to*

$$f'(x^*) \in A^+ \qquad \forall A \in A(x^*) \quad . \tag{1.4}$$

A point $x^* \in \mathrm{cl}\Omega$ *which satisfies (1.4) (or, equivalently, (1.2)) is called a stationary point of* f *on* Ω.

In what follows we shall suppose that Ω is a closed set.

Assume that $x \in \Omega$ is not a stationary point of f on Ω. Then there exists $A \in A(x)$ such that

$$f'(x) \notin A^+ \quad .$$

Let us find

$$\min_{w\in A^+} \|f'(x) - w\| = \|f'(x) - w(A)\| \quad . \tag{1.5}$$

It is not difficult to see that

$$v(A) = w(A) - f'(x) \in A$$

and that v(A) is a descent direction of f on Ω at x, i.e.,

$$(f'(x), v(A)) < 0 \quad .$$

It is also clear that the direction $g_0 = \frac{v_0}{\|v_0\|}$, where $\|v_0\| = \max_{A\in A(x)} \|v(A)\|$, is a direction of steepest descent of the function f on the set Ω at x, i.e.,

$$\frac{\partial f(x)}{\partial g_0} = \inf_{g\in S_1\cap\Gamma(x)} \frac{\partial f(x)}{\partial g} \quad .$$

Here

$$S_1 = \{g \in E_n \mid \|g\| = 1\}, \quad \frac{\partial f(x)}{\partial g} = \lim_{\alpha\to+0} \frac{f(x+\alpha g) - f(x)}{\alpha} \quad .$$

A steepest descent direction may not be unique. Note that

$$\frac{\partial f(x)}{\partial g(A)} = \min_{g\in S_1\cap A} \frac{\partial f(x)}{\partial g} \tag{1.6}$$

where

$$g(A) = \frac{v(A)}{\|v(A)\|} \quad .$$

<u>Remark 1</u>. Condition (1.4) is equivalent to

$$f'(x^*) \in L(x^*) \tag{1.7}$$

where

$$L(x) = \bigcap_{A \in A(x)} A^+ \quad .$$

If $L(x^*) = \{0\}$ then we obtain the well-known condition

$$f'(x^*) = 0 \quad .$$

Example 1. Let

$$x = \left(x^{(1)}, x^{(2)}\right) \in E_2 \quad , \qquad x_0 = (0,0) \quad ,$$

$$\Omega = \ell_1 \cup \ell_2 \cup \ell_3 \quad ,$$

where

$$\ell_1 = \{x = (\alpha, 0) \mid \alpha \geq 0\} \quad ,$$

$$\ell_2 = \{x = (0, \alpha) \mid \alpha \geq 0\} \quad ,$$

$$\ell_3 = \{x = (-\alpha, -\alpha) \mid \alpha \geq 0\} \quad .$$

It is clear that $\Gamma(x_0) = \Omega$ and $A(x_0) = \{\ell_1, \ell_2, \ell_3\}$, i.e., $A(x_0) = \{A_1, A_2, A_3\}$, where $A_1 = \ell_1$, $A_2 = \ell_2$, $A_3 = \ell_3$. Now we have

$$A_1^+ = \{x \in E_2 \mid (x, \overline{\ell}_1) \geq 0\} \quad , \qquad \overline{\ell}_1 = (1,0) \quad ,$$

$$A_2^+ = \{x \in E_2 \mid (x, \overline{\ell}_2) \geq 0\} \quad , \qquad \overline{\ell}_2 = (0,1) \quad ,$$

$$A_3^+ = \{x \in E_2 \mid (x, \overline{\ell}_3) \geq 0\} \quad , \qquad \overline{\ell}_3 = (-1,-1) \quad .$$

It can be seen from Figure 1 that $L(x_0) = \bigcap_{i \in 1:3} A_i^+ = \{0\}$ and therefore $f'(x_0) = 0$ is a necessary condition at x_0.

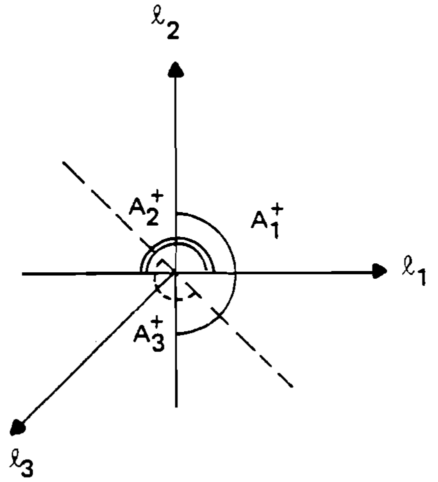

Figure 1

Remark 2. If $x \in \Omega$ is not a stationary point then

$$\min_{v \in L(x_0)} \|v - f'(x)\| = \|v(x) - f'(x)\| > 0 \quad .$$

However, note that the direction

$$g = \frac{v(x) - f'(x)}{\|v(x) - f'(x)\|}$$

has nothing to do with descent directions (it may not even be feasible). Thus, the necessary condition (1.7) provides no information about descent directions if x_0 is not a stationary point. In contrast, condition (1.4) is more workable because it allows us to construct descent and even steepest descent directions.

For a continuously differentiable function f

$$\frac{\partial f(x)}{\partial g} = (f'(x), g) \quad .$$

Thus the problem of finding steepest descent directions of f on Ω at x is reduced to that of solving (1.6) (a quadratic programming problem which however becomes linear if the m-norm is used instead of the Euclidean norm) for all $A \in A(x)$. For this reason we are interested in constructing a family $A(x)$ containing as few cones as possible. If Ω is a convex set the cone $\Gamma(x)$ is convex and therefore $A(x)$ consists of only one set.

Let Ω be described by inequalities

$$\Omega = \{x \in E_n \mid h_i(x) \leq 0 \ \forall i \in I\} \tag{1.8}$$

where the h_i's are from C_1, $I = 1:N$.

If $x \in \Omega$ and

$$0 \notin \text{co}\{h_i'(x) \mid i \in Q(x)\} \quad , \tag{1.9}$$

where

$$Q(x) = \{i \in I \mid h_i(x) = 0\} \quad ,$$

then (see, e.g., [2])

$$\Gamma^{+}(x) = \text{cone } \{-h_i^{'}(x) \mid i \in Q(x)\} \quad .$$

Here cone B is the conic hull of B.

It is an easy exercise to show that if a convex cone A contains an interior point then the condition (see (1.4))

$$f^{'}(x^{*}) \in A^{+}$$

is equivalent to the condition

$$0 \in \text{co } \{f^{'}(x^{*}) \cup T_{\eta}(A)\} \quad \forall \eta > 0$$

where

$$T_{\eta}(A) = \{v \in E_n \mid v \in [-A^{+}], \|v\| = \eta\} \ .$$

Assume that $x \in \Omega$ is not a stationary point of f on Ω and suppose that int $A \neq \emptyset$. Then there exists $A \in \mathcal{A}(x)$ such that

$$f^{'}(x) \notin A^{+} \quad .$$

Then, from the above condition,

$$0 \notin \text{co } \{f^{'}(x) \cup T_{\eta}(A)\} \equiv L_{\eta}(A) \quad . \tag{1.10}$$

Let us find

$$\min_{v \in L_{\eta}(A)} \|v\| = \|v_{\eta}(A)\| \quad .$$

From (1.10) we deduce that

$$\|v_{\eta}(A)\| > 0 \quad .$$

It is easy to see that the direction

$$g_{\eta}(A) = - \frac{v_{\eta}(A)}{\|v_{\eta}(A)\|} \tag{1.11}$$

is such that

$$(f'(x), g_\eta(A)) < 0 \quad , \quad g_\eta(A) \in \text{int } A \quad .$$

Hence, $g_\eta(A)$ is a descent direction leading strictly inside the cone A. The fact that $g_\eta(A)$ is an interior direction is important -- the direction g(A) (see (1.6)) may be tangential even though it is the steepest descent direction of f on A (see (1.6)). This feature may be crucial if Ω is described by (1.8) and condition (1.9) holds, since in this case $\Gamma(x)$ is a convex cone and therefore $A(x)$ consists of only one set (namely $\Gamma(x)$). Thus, on the one hand it is possible to find the steepest descent direction g(A) (see (1.7)) but this direction may not be feasible if the h_i's are not linear; on the other hand the descent direction $g_\eta(\Gamma(x))$ is feasible for any $\eta > 0$, where

$$g_\eta(\Gamma(x)) = -\frac{v_\eta}{\|v_\eta\|}$$

and

$$\|v_\eta\| = \min_{v \in L_\eta} \|v\| , \qquad L_\eta = \text{co } \{f'(x); \eta h_i'(x) \mid i \in Q(x)\} \quad .$$

The foregoing analysis reveals the importance of having several (possibly equivalent) necessary conditions, in that this enables us to develop different numerical methods.

Remark 3. It is not difficult to show that, in (1.11), $g_\eta(A) \underset{\eta \to +\infty}{\to} g(A)$, where g(A) is the steepest descent direction of f on A at x.

1.2 *Convex programming problems*

Similar considerations can be applied to constrained non-differentiable convex programming problems of the form

$$\min\{f(x) \mid x \in \Omega\}$$

where

$$\Omega = \{x \in E_n | h(x) \leq 0\}$$

and functions f and h are finite and convex (but not necessarily differentiable) on E_n.

Suppose that there exists a point $\bar{x}$ such that

$$h(\bar{x}) < 0 \quad . \tag{1.12}$$

(This is called the Slater condition.) It follows from convex analysis (see [3]) that

$$\Gamma^+(x) = \begin{cases} \{0\}, \text{ if } h(x) < 0 \ , \\ \text{cone } \{\partial h(x)\}, \text{ if } h(x) = 0 \end{cases}$$

where $\partial h(x)$ is the subdifferential of h at x, i.e.,

$$\partial h(x) = \{v \in E_n | f(z) - f(x) \geq (v, z-x) \ \forall z \in E_n\} \quad . \tag{1.13}$$

Theorem 3 *(see [4]). For* $x^* \in \Omega$ *to be a minimum point of* f *on* Ω *it is necessary and sufficient that*

$$\partial f(x^*) \cap \Gamma^+(x^*) \neq \emptyset \quad . \tag{1.14}$$

Theorem 4 *(see [5]). Let* $h(x^*) = 0$. *Condition (1.14) is equivalent to the condition*

$$0 \in \text{co } \{\partial f(x^*) \cup T_\eta(x^*)\} \equiv L_\eta(x^*) \qquad \forall \eta > 0 \tag{1.15}$$

where

$$T_\eta(x) = \{v \in [-\Gamma^+(x)] \,|\, \|v\| = \eta\} \quad .$$

If $x \in \Omega$ is not a minimum point of f on Ω then the direction

$$g(x) = - \left(\frac{v(x) - w(x)}{\|v(x) - w(x)\|} \right) ,$$

where

$$\|v(x) - w(x)\| = \min_{\substack{v\in\partial f(x)\\ w\in\Gamma^+(x)}} \|v-w\| ,$$

is the steepest descent direction of f on Ω at x.

Let us find

$$g_\eta(x) = - \frac{v_\eta(x)}{\|v_\eta(x)\|} \tag{1.16}$$

where

$$\|v_\eta(x)\| = \min_{v\in L_\eta(x)} \|v\| .$$

The direction $g_\eta(x)$ given by (1.16) is a descent direction and it can be shown that

$$g_\eta(x) \in \text{int } \Gamma(x) .$$

Thus condition (1.15) enables us to find a "feasible" direction (i.e., a direction leading strictly inside Ω), and this can be useful in constructing numerical methods. Some of the methods based on (1.15) are described in Chapter IV of [5].

Note that if x is not a stationary point then

$$g_\eta(x) \xrightarrow[\eta\to+\infty]{} g(x)$$

where g(x) is the steepest descent direction of f on Ω at x.

Theorem 4' (see [5]). Let $h(x^*) = 0$. *Condition (1.14) is equivalent to the condition*

$$0 \in \text{co } \{\partial f(x^*) \cup [\eta\partial h(x^*)]\} \equiv L_{1\eta}(x^*) \quad \forall\eta > 0 . \tag{1.15'}$$

Proof. Consider a function

$$\phi_\eta(x) = \max \{f(x) - f^* , \eta h(x)\}$$

where

$$f^* = \min_{x \in \Omega} f(x) \quad .$$

Since $\phi_\eta(x) \geq 0 \quad \forall x \in E_n$, and $\phi_\eta(x^*) = 0$, x^* is a minimum point of ϕ_η on E_n. However, ϕ_η is a convex function and so

$$\partial\phi_\eta(x^*) \quad = \quad \text{co}\,\{\partial f(x^*) \cup [\eta h(x^*)]\} \quad .$$

Applying a necessary and sufficient condition for an unconstrained minimum of a convex function, we immediately obtain (1.15').

Assume that $x \in \Omega$ is not a minimum point of f on Ω, and find the direction

$$g_{1\eta}(x) \quad = \quad - \frac{v_{1\eta}(x)}{\|v_{1\eta}(x)\|} \tag{1.16'}$$

where

$$\|v_{1\eta}(x)\| \quad = \quad \min_{v \in L_{1\eta}(x)} \|v\| \quad .$$

It can be shown that the direction $g_{1\eta}(x)$ defined by (1.16') is a descent direction and

$$g_{1\eta}(x) \in \text{int}\ \Gamma(x) \quad .$$

Note also that $g_{1\eta}(x) \xrightarrow[\eta \to +\infty]{} g(x)$, where g(x) is the steepest descent direction of f on Ω.

<u>Remark 4</u>. Condition (1.15') is applicable even if Ω is an arbitrary convex compact set (not necessarily described explicitly by a convex function).

2. QUASIDIFFERENTIABLE FUNCTIONS

2.1 *Definitions and some properties*

A function f is called quasidifferentiable (q.d.) at a point $x \in E_n$ if it is directionally differentiable at x and if there exist convex compact sets $\underline{\partial} f(x) \subset E_n$ and $\overline{\partial} f(x) \subset E_n$ such that

$$\frac{\partial f(x)}{\partial g} \equiv \lim_{\alpha \to +0} \frac{f(x+\alpha g) - f(x)}{\alpha} = \max_{v \in \underline{\partial} f(x)} (v,g) + \min_{w \in \overline{\partial} f(x)} (w,g) \quad .$$

The pair of sets $Df(x) = [\underline{\partial} f(x), \overline{\partial} f(x)]$ is called the quasi-differential of f at x.

Quasidifferentiable functions were introduced in [6] and have been studied in more detail in [7,8]. A survey of results concerning this class of functions is presented in [9]. It turns out that q.d. functions form a linear space closed with respect to all algebraic operations and, more importantly, to the operations of taking pointwise maximum and minimum. A new form of calculus (quasidifferential calculus) has been developed to handle these functions, and both a chain rule for composite functions and an inverse function theorem have been established [5,9]. In what follows we shall use only two results from quasidifferential calculus (see below).

If $D_1 = [A_1, B_1]$, $D_2 = [A_2, B_2]$ are pairs of convex sets (i.e., $A_i \subset E_n$, $B_i \subset E_n$ are convex sets) we put

$$D_1 + D_2 = [A_1 + A_2 \, , B_1 + B_2]$$

and if $D = [A,B]$ then

$$\lambda D = \begin{cases} [\lambda A \, , \lambda B], \text{ if } \lambda \geq 0 \quad , \\ [\lambda B \, , \lambda A], \text{ if } \lambda < 0 \quad . \end{cases}$$

The following is then true:

1. If functions f_i $(i \in I \equiv 1:N)$ are q.d. at x and $Df_i(x) = [\underline{\partial} f_i(x), \overline{\partial} f_i(x)]$ is a quasidifferential of f_i at x then a function $f = \sum_{i \in I} \lambda_i f_i$ (where the λ_i's are real numbers) is q.d. at x and

$$Df(x) = \sum_{i \in I} \lambda_i Df_i(x) \quad .$$

2. If functions f_i $(i \in I \equiv 1:N)$ are q.d. at x then

$$f = \max_{i \in I} f_i$$

is a q.d. function and

$$Df(x) = [\underline{\partial} f(x), \overline{\partial} f(x)] \tag{2.1}$$

where

$$\underline{\partial} f(x) = \text{co} \left\{ \underline{\partial} f_k(x) - \sum_{\substack{i \in R(x) \\ i \neq k}} \overline{\partial} f_i(x) \mid k \in R(x) \right\} \quad ,$$

$$\overline{\partial} f(x) = \sum_{k \in R(x)} \overline{\partial} f_k(x) \quad , \quad R(x) = \{ i \in I \mid f_i(x) = f(x) \} \quad .$$

L.N. Polyakova [7] has discovered necessary conditions for an unconstrained optimum of f on E_n:

Theorem 5. For $x^* \in E_n$ to be a minimum point of a q.d. function f on E_n it is necessary that

$$- \overline{\partial} f(x^*) \subset \underline{\partial} f(x^*) \quad . \tag{2.2}$$

For $x^{**} \subset E_n$ to be a maximum point of a q.d. function on E_n it is necessary that

$$- \underline{\partial} f(x^{**}) \subset \overline{\partial} f(x^{**}) \quad . \tag{2.3}$$

Conditions (2.2) and (2.3) represent generalizations of the classical necessary conditions for an extreme point of a smooth function f on E_n (in this case $\overline{\partial} f(x) = \{0\}$, $\underline{\partial} f(x) = \{f'(x)\}$ and from (2.2) it follows that $f'(x^*) = 0$. From (2.3) it also follows that $f'(x^{**}) = 0$, i.e., the necessary conditions for a maximum and for a minimum coincide.)

If f is convex on E_n then $\overline{\partial}f(x) = \{0\}$, $\underline{\partial}f(x) = \partial f(x)$, where $\partial f(x)$ is the subdifferential of f at x (see (1.13)), and (2.2) becomes the well-known condition [3,4]

$$0 \in \partial f(x^*) \quad .$$

2.2 *Quasidifferentiable sets. Necessary conditions for constrained optimality*

A set Ω is called quasidifferentiable if it can be represented in the form

$$\Omega = \{x \in E_n \mid h(x) \le 0\}$$

where h is quasidifferentiable on E_n.

The properties of q.d. sets and the necessary conditions for optimality of a q.d. function on a q.d. set are discussed in [8] (see also [5, Chap. II]).

Take $x \in \Omega$ and introduce cones

$$\gamma(x) = \left\{ g \in E_n \;\middle|\; \frac{\partial h(x)}{\partial g} < 0 \right\} \quad ,$$

$$\gamma_1(x) = \left\{ g \in E_n \;\middle|\; \frac{\partial h(x)}{\partial g} \le 0 \right\} \quad .$$

Let $h(x) = 0$. We say that the nondegeneracy condition is satisfied at x if

$$\text{cl}\,[\gamma(x)] = \gamma_1(x) \qquad (2.4)$$

where cl A denotes the closure of A.

Lemma 1 (see [5,8]). If $h(x) < 0$ *then* $\Gamma(x) = E_n$. *If* $h(x) = 0$ *and the nondegeneracy condition (2.4) is satisfied at* x *and* $h(x)$ *is Lipschitzian in some neighborhood of* x *then*

$$\Gamma(x) = \gamma_1(x) \tag{2.5}$$

where $\Gamma(x)$ *is the set of feasible (in a broad sense) directions of* Ω *at* x *(see (1.1)).*

The following two theorems and lemma are proved in [8].

Theorem 6. Let a function f *be Lipschitzian and quasidifferentiable in some neighborhood of a point* $x^* \in \Omega$. *If* $h(x^*) = 0$ *then let* h *be Lipschitzian and q.d. in some neighborhood of* x^* *and the nondegeneracy condition (2.4) be satisfied at* x^*. *For the function* f *to attain its smallest value on* Ω *at* x^* *it is necessary that*

$$-\overline{\partial} f(x^*) \subset \underline{\partial} f(x^*) \qquad \text{if} \quad h(x^*) < 0 \tag{2.6}$$

and

$$(\underline{\partial} f(x^*) + w) \cap [-\text{cl}(\text{cone}(\underline{\partial} h(x^*) + w'))] \neq \emptyset \quad \text{if} \quad h(x^*) = 0 \tag{2.7}$$

for every $w \in \overline{\partial} f(x^*)$, $w' \in \overline{\partial} h(x^*)$.

Theorem 7. Condition (2.7) is equivalent to the condition

$$-\overline{\partial} f(x^*) \subset L(x^*) \tag{2.8}$$

where

$$L(x) = \bigcap_{w \in \overline{\partial} h(x)} [\underline{\partial} f(x) + \text{cl}(\text{cone}\ (\underline{\partial} h(x) + w))] \ . \tag{2.9}$$

A point $x^* \in \Omega$ which satisfies (2.7) when $h(x^*) = 0$ and (2.6) when $h(x^*) < 0$ is called a stationary point of f on Ω.

Note that $L(x)$ is a convex set (and nonempty, since $\underline{\partial} f(x) \subset L(x)$).

<u>*Corollary.*</u> *If* f *and* h *are convex functions it follows from (2.8) that*

$$0 \in \partial f(x^*) - \Gamma^+(x^*) \tag{2.10}$$

where $\partial f(x)$ *is the subdifferential of* f *at* x *(see (1.13)) and* $\Gamma(x)$ *is the cone of feasible directions of* Ω *at* x.

This condition is both necessary and sufficient for $x^* \in \Omega$ to be a minimum point of f on Ω (in the case where $h(x^*) = 0$ it is also assumed that the Slater condition (1.12) holds).

Necessary conditions for a maximum of a q.d. function on a q.d. set can be derived in an analogous fashion [8,5].

2.3 *Descent and steepest descent directions*

Take $x \in \Omega$ and suppose that x is not a stationary point of f on Ω. We shall now consider in more detail the case where $h(x) = 0$ and condition (2.7) is not satisfied. For every $w \in \overline{\partial} f(x)$ and $w' \in \overline{\partial} h(x)$ we calculate

$$\min_{\substack{z \in \underline{\partial} f(x) + w \\ z' \in \mathrm{cl}(\mathrm{cone}(\underline{\partial} h(x) + w'))}} \| z + z' \| = \| z(w,w') + z'(w,w') \| = d(w,w') \ . \tag{2.11}$$

Then we find

$$\rho(x) = \max_{\substack{w \in \overline{\partial} f(x) \\ w' \in \overline{\partial} h(x)}} d(w,w') = d(w_0,w_0') \ . \tag{2.12}$$

Since (2.7) does not hold, $\rho(x) > 0$.

Let

$$g_0 = - \frac{v_0 + w(v_0)}{\| v_0 + w(v_0) \|} \ . \tag{2.13}$$

Lemma 2. If $h(x) = 0$ *and the nondegeneracy condition (2.4) is satisfied then the direction* g_0 *(see (2.13)) is a steepest descent direction of* f *on* Ω *at* x *and* $d(x) = \|v_0 + w(v_0)\|$ *is the rate of steepest descent, i.e.,*

$$\frac{\partial f(x)}{\partial g_0} = \min_{g \in \Gamma(x) \cap S_1} \frac{\partial f(x)}{\partial g} = - d(x) \quad . \qquad (2.14)$$

Remark 5. Since there may exist several w_0, w_0' satisfying (2.12), there may exist several (or infinitely many) directions of steepest descent. (This is impossible for convex sets and convex or continuously differentiable functions.)

Remark 6. Let $K(w') = \mathrm{cl}\,(\mathrm{cone}(\underline{\partial} h(x^*) + w'))$.

If $\mathrm{int}\, K^+(w') \neq \emptyset$, then condition (2.7) is equivalent to

$$0 \in \mathrm{co}\, \{[\partial f(x^*) + w] \cup T_\eta(w')\} \equiv L_\eta(w,w')$$

where

$$T_\eta(w) = \{v \in K(w') \mid \|v\| = \eta\} \quad , \quad \eta > 0 \quad .$$

If for some $x \in \Omega$ and $w \in \overline{\partial} f(x)$, $w' \in \overline{\partial} h(x)$ we have $h(x) = 0$ and $0 \notin L_\eta(w,w')$, then

$$g_\eta(w,w') = - \frac{z_\eta(w,w')}{\|z_\eta(w,w')\|}$$

where

$$\|z_\eta(w,w')\| = \min_{z \in L_\eta(w,w')} \|z\|$$

is a descent direction of f on Ω at x and, above all, is feasible, i.e.,

$$\frac{\partial f(x)}{\partial g} < 0 \quad \text{and} \quad \frac{\partial h(x)}{\partial g} < 0 \quad .$$

Remark 7. If $x \in \Omega$ is not a stationary point of f on Ω conditions (2.6) and (2.7) allow us to find steepest descent directions (see Lemma 2), but in the case where $h(x) = 0$ the directions thus obtained may not necessarily be feasible.

Condition (2.8) is similar to (2.2) and if x is not a stationary point we have

$$- \bar{\partial} f(x) \not\subset L(x) \quad . \tag{2.15}$$

Let us find

$$\max_{v \in \bar{\partial} f(x)} \rho(v) = \rho(v(x))$$

where

$$\rho(x) = \min_{w \in L(x)} \|v+w\| = \|v+w(v)\| \quad .$$

It follows from (2.15) that $\rho(v(x)) > 0$ but it is not clear whether

$$g_0 = - \frac{v(x) + w(v(x))}{\|v(x) + w(v(x))\|}$$

is a descent direction.

Let $h(x) = 0$. The problem of finding a steepest descent direction is equivalent to the following problem:

$$\min \theta \tag{2.16}$$

subject to

$$\frac{\partial f(x)}{\partial g} \leq \theta \quad , \tag{2.17}$$

$$\frac{\partial h(x)}{\partial g} \leq 0 \quad , \tag{2.18}$$

$$\|g\| \leq 1 \quad . \tag{2.19}$$

Since f and h are quasidifferentiable functions, problem (2.16)-(2.19) can be rewritten as

$$\min \{\theta \mid \theta \in E_1, g \in E_n, [\theta, g] \in \Omega_1\} \tag{2.16'}$$

where $\Omega_1 \subset E_{n+1}$ is described by inequalities

$$\max_{v \in \underline{\partial} f(x)} (v,g) + \min_{w \in \overline{\partial} f(x)} (w,g) \leq \theta \quad , \tag{2.17'}$$

$$\max_{v' \in \underline{\partial} h(x)} (v',g) + \min_{w' \in \overline{\partial} h(x)} (w',g) \leq 0 \quad , \tag{2.18'}$$

$$\|g\| \leq 1 \quad . \tag{2.19'}$$

Let $\theta(w,w') \equiv \theta(w,w',x)$, $g(w,w') \equiv g(w,w',x)$ be a solution to the problem

$$\min \{\theta \mid \theta \in E_1, g \in E_n, [\theta, g] \in \Omega_1(w,w')\} \tag{2.20}$$

where $w \in \overline{\partial} f(x)$, $w' \in \overline{\partial} h(x)$, and $\Omega_1(w,w') \subset E_{n+1}$ is described by inequalities

$$\max_{v \in \underline{\partial} f(x)} (v,g) + (w,g) \leq \theta \quad , \tag{2.21}$$

$$\max_{v' \in \underline{\partial} h(x)} (v',g) + (w',g) \leq 0 \quad , \tag{2.22}$$

$$\|g\| \leq 1 \quad . \tag{2.23}$$

Let $[\theta^*(x), g^*(x)]$ denote a solution to problem (2.16')-(2.19'). It is clear that

$$\theta^*(x) = \theta(w^*, w'^*) \quad , \quad g^*(x) = g(w^*, w'^*)$$

where

$$[w^*, w'^*] = \arg\min \{\theta(w,w') \mid w \in \overline{\partial} f(x), w' \in \overline{\partial} h(x)\} \quad . \tag{2.24}$$

It is also clear that $\theta^* = -\rho(x)$ (see (2.12)). Here $g^*(x)$ is a steepest descent direction, and $\theta^*(x)$ is the rate of steepest descent; it has already been pointed out that direction $g^*(x)$ (as well as directions $g(w,w'))$ may not be feasible.

We shall therefore consider the following problem:

$$\min \{\theta \mid \theta \in E_1, g \in E_n, [\theta, g] \in \Omega_{1\eta}\} \tag{2.25}$$

where $\eta > 0$, and $\Omega_{1\eta} \subset E_{n+1}$ is described by

$$\max_{v \in \underline{\partial} f(x)} (v,g) + \min_{w \in \overline{\partial} f(x)} (w,g) \leq \theta \quad , \tag{2.26}$$

$$\max_{v' \in \underline{\partial} h(x)} (v',g) + \min_{w' \in \overline{\partial} h(x)} (w',g) \leq \eta\theta \quad , \tag{2.27}$$

$$\|g\| \leq 1 \quad . \tag{2.28}$$

Let $(\theta_\eta(x), g_\eta(x))$ be a solution to problem (2.25)-(2.28).

Now let us also consider the following problem:

$$\min \{\theta \mid \theta \in E_1, g \in E_n, [\theta, g] \in \Omega_{1\eta}(w,w')\} \tag{2.29}$$

where $w \in \overline{\partial} f(x)$, $w' \in \overline{\partial} h(x)$, and $\Omega_{1\eta}(w,w') \subset E_{n+1}$ is described by inequalities

$$\max_{v \in \underline{\partial} f(x)} (v,g) + (w,g) \leq \theta \quad , \tag{2.30}$$

$$\max_{v' \in \underline{\partial} h(x)} (v',g) + (w',g) \leq \eta\theta \quad , \tag{2.31}$$

$$\|g\| \leq 1 \quad . \tag{2.32}$$

If $\theta_\eta(w,w') \equiv \theta_\eta(w,w',x)$, $g_\eta(w,w') \equiv g_\eta(w,w',x)$ is a solution to problem (2.29)-(2.32) then

$$\theta_\eta(x) = \theta_\eta(w_\eta, w'_\eta) \quad ; \qquad g_\eta(x) = g_\eta(w_\eta, w'_\eta)$$

where

$$[w_\eta, w'_\eta] = \arg\min \{\theta_\eta(w,w') \mid w \in \bar{\partial} f(x), w' \in \bar{\partial} h(x)\} . \quad (2.33)$$

Direction $g_\eta(x)$ is feasible for any $\eta > 0$.

<u>Remark 8</u>. When solving problem (2.24) (as well as (2.33)) it is sufficient to consider only boundary points of the sets $\bar{\partial} f(x)$ and $\bar{\partial} h(x)$. Furthermore, if each of these sets is a convex hull of a finite number of points, it is sufficient to solve only a finite number of problems of the form (2.20)-(2.23) (or, for problem (2.33), of the form (2.29)-(2.32)). These become linear programming problems if the Euclidean norm in (2.23) (or (2.32)) is replaced by the m-norm:

$$\|g\|_m = \max \{|g_i| \mid i \in 1:n\}$$

where

$$g = (g_1, \ldots, g_n) \quad .$$

<u>Remark 9</u>. Let $\eta_k \xrightarrow[k\to\infty]{} \infty$. Without loss of generality we can assume that $g_{\eta_k}(x) \longrightarrow g^*$. It is possible to show that g^* is a steepest descent direction of f on Ω at x and that $\theta_{\eta_k}(x) \to \theta^*(x)$, where $\theta^*(x)$ is the rate of steepest descent.

<u>Remark 10</u>. Let $x \in \Omega$ and $h(x)$ not necessarily equal zero. Consider the problem

$$\min \{\theta \mid \theta \in E_1, g \in E_n, [\theta, g] \in \Omega_{2\eta}\} \quad (2.34)$$

where $\eta > 0$, and $\Omega_{2\eta} \subset E_{n+1}$ is described by

$$\max_{v \in \underline{\partial} f(x)} (v,g) + \min_{w \in \bar{\partial} f(x)} (w,g) \le \theta \ ,$$

$$h(x) + \max_{v' \in \underline{\partial} h(x)} (v',g) + \min_{w' \in \bar{\partial} h(x)} (w',g) \le \eta\theta \ , \quad (2.35)$$

$$\|g\| \le 1 \quad .$$

The replacement of (2.31) by (2.35) enables us to deal with points in Ω close to the boundary. It is hoped that, as in mathematical programming (see, e.g., [10]), it will eventually be possible to develop superlinearly (or even quadratically) convergent algorithms.

A geometric interpretation of problem (2.16)-(2.19) is given by (2.12). For a similar interpretation of problem (2.29)-(2.32) we use the following result (obtained by A. Shapiro [11]):

Theorem 8. Let $x^* \in \Omega$ *and* $h(x^*) = 0$. *Functions* f *and* h *are assumed to be quasidifferentiable on* E_n. *For* x^* *to be a minimum point of* f *on* Ω *it is necessary that*

$$L_1(x^*) \subset L_2(x^*) \tag{2.36}$$

where

$$L_1(x) = -[\overline{\partial} f(x) + \overline{\partial} h(x)] \quad , \tag{2.37}$$

$$L_2(x) = \text{co}\ \{\underline{\partial} f(x) - \overline{\partial} h(x)\ ,\ \underline{\partial} h(x) - \overline{\partial} f(x)\} \quad . \tag{2.38}$$

Proof. Let x^* be a minimum point of f on Ω and let $h(x^*) = 0$. Consider a function

$$F(x) = \max\ \{f(x)-f^*, h(x)\}$$

where

$$f^* = f(x^*) = \min_{x \in \Omega} f(x) \quad .$$

It is clear that $F(x) \geq 0 \quad \forall x \in E_n$. Since $F(x^*) = 0$ it can be concluded that x^* is a minimum point of F on E_n. But F is a q.d. function (because it is the pointwise maximum of q.d. functions $f(x) - f^*$ and $h(x)$).

Applying (2.1) we have

$$DF(x^*) = [\underline{\partial} F(x^*), \overline{\partial} F(x^*)]$$

where

$$\underline{\partial}F(x^*) \quad = \quad co\ \{\underline{\partial}f(x^*) - \overline{\partial}h(x^*)\ ,\ \underline{\partial}h(x^*) - \overline{\partial}f(x^*)\}\quad ,$$

$$\overline{\partial}F(x^*) \quad = \quad \overline{\partial}f(x^*) + \overline{\partial}h(x^*)\quad .$$

Since x^* is a minimum point of F on E_n, (2.2) leads immediately to (2.36). Q.E.D.

Remark 11. Condition (2.36) is equivalent to (2.7) and is applicable even in the case where the nondegeneracy condition (2.4) does not hold. However, it seems that condition (2.6) is always satisfied at a degenerate point.

Now let us consider the case where $x \in \Omega$, $h(x) = 0$ and condition (2.36) does not hold. We first find

$$d(x) \quad = \quad \max_{v \in L_1(x)} \rho(v) \quad = \quad \rho(v(x)) \tag{2.39}$$

where

$$\rho(v) \quad = \quad \min_{w \in L_2(x)} \|v-w\| \quad = \quad \|v-w(v)\| \quad . \tag{2.40}$$

It is clear that $\rho(v(x)) > 0$.

Since sets $L_1(x)$ and $L_2(x)$ are convex there exists for every $v \in L_1(x)$ a unique $w(v)$ which satisfies (2.40), but there is not necessarily a unique $v(w)$ which satisfies (2.39).

Consider a direction

$$g_0 \quad = \quad \frac{v(x) - w(v(x))}{\|v(x) - w(v(x))\|} \quad . \tag{2.41}$$

Lemma 3. The direction g_0 *defined by (2.41) is a descent direction of* f *on* Ω *at* x.

Proof. By definition (see (2.39)-(2.41))

$$\max_{v \in L_1(x)} (v, g_0) \quad > \quad \max_{v \in L_2(x)} (w, g_0) \quad . \tag{2.42}$$

In particular, it follows from (2.42) that

$$\max_{v \in L_1(x)} (v,g_0) > \max_{w \in \underline{\partial} f(x) - \overline{\partial} h(x)} (w,g_0) \quad , \qquad (2.43)$$

$$\max_{v \in L_1(x)} (v,g_0) > \max_{w \in \underline{\partial} h(x) - \overline{\partial} f(x)} (w,g_0) \quad . \qquad (2.44)$$

From (2.43)

$$\max_{v \in [-\overline{\partial} f(x)]} (v,g_0) + \max_{v \in \overline{\partial} h(x)} (v,g_0) > \max_{w \in \underline{\partial} f(x)} (w,g_0) + \max_{w \in [-\overline{\partial} h(x)]} (w,g_0) \quad ,$$

i.e.,

$$- \min_{v \in \overline{\partial} f(x)} (v,g_0) > \max_{w \in \underline{\partial} f(x)} (w,g_0) \quad . \qquad (2.45)$$

But (2.45) implies that

$$\frac{\partial f(x)}{\partial g_0} = \max_{v \in \underline{\partial} f(x)} (v,g_0) + \min_{w \in \overline{\partial} f(x)} (w,g_0) < 0 \ . \qquad (2.46)$$

Analogously, it follows from (2.44) that

$$\frac{\partial h(x)}{\partial g_0} = \max_{v \in \underline{\partial} h(x)} (v,g_0) + \min_{w \in \overline{\partial} h(x)} (w,g_0) < 0 \ . \qquad (2.47)$$

Inequality (2.47) implies that g_0 is feasible; inequality (2.46) shows that it is a descent direction. Q.E.D.

Remark 12. The direction g_0 defined by (2.39)-(2.41) may not be unique.

Observe that since Ω can be described by

$$\Omega = \{x \in E_n | h_\eta(x) \leq 0\} \quad ,$$

where $h_\eta(x) = \eta h(x)$, $\eta > 0$, we can obtain the following necessary condition

$$L_{1\eta}(x^*) \subset L_{2\eta}(x^*) \tag{2.36'}$$

where

$$L_{1\eta}(x) = -[\overline{\partial}f(x) + \eta\overline{\partial}h(x)] \quad , \tag{2.37'}$$

$$L_{2\eta}(x) = \text{co}\ \{\underline{\partial}f(x) - \eta\overline{\partial}h(x); \eta\underline{\partial}h(x) - \overline{\partial}f(x)\} \ . \tag{2.38'}$$

For a nonstationary point x (when $h(x) = 0$) it is possible to obtain a descent direction $g_{0\eta}$ different from g_0.

It is also useful to note that if λ is a quasidifferentiable function strictly positive on Ω then Ω can be given in the form

$$\Omega = \{x \mid \lambda(x)h(x) \leq 0\} \quad .$$

This representation provides a variety of necessary conditions and, consequently, a variety of descent directions at a nonstationary point.

2.4 *Sufficient conditions for a local minimum*

Necessary conditions (2.7), (2.8), (2.36) can be modified in such a way that they become sufficient conditions for a local minimum of f on Ω.

Recall that

$$f(x_0 + \alpha g) = f(x_0) + \alpha\frac{\partial f(x_0)}{\partial g} + o(\alpha,g) \quad , \tag{2.48}$$

$$h(x_0 + \alpha g) = h(x_0) + \alpha\frac{\partial h(x_0)}{\partial g} + o_1(\alpha,g) \quad . \tag{2.49}$$

Functions f and h are assumed to be continuous and quasidifferentiable at $x_0 \in \Omega$; it is also assumed that

$$\frac{o(\alpha,g)}{\alpha} \longrightarrow 0$$

uniformly with respect to $g \in S_1$ in (2.48) and that if $h(x_0) = 0$ then

$$\frac{o_1(\alpha,g)}{\alpha} \longrightarrow 0$$

uniformly with respect to $g \in S_1$ in (2.49). Recall also that

$$S_1 = \{g \in E_n \mid \|g\| = 1\} \quad .$$

<u>Theorem 9</u> *(see [5,8]). If* $h(x_0) < 0$ *and*

$$- \bar{\partial} f(x_0) \subset \text{int } \underline{\partial} f(x_0) \tag{2.50}$$

then x_0 *is a local minimum point of* f *on* Ω.

If $h(x_0) = 0$ *and*

$$r = \min_{\substack{w \in \bar{\partial} f(x_0) \\ w' \in \bar{\partial} h(x_0)}} r(w,w') > 0 \quad , \tag{2.51}$$

where $r(w,w')$ *is the radius of the maximal sphere centered at the origin that can be inscribed in the set*

$$L(w,w') = \underline{\partial} f(x_0) + w + \text{cl } (\text{cone}(\underline{\partial} h(x_0) + w')) \quad ,$$

then x_0 *is a strict local minimum point of* f *on* Ω *and*

$$r = \min_{g \in \Gamma(x_0) \cap S_1} \frac{\partial f(x_0)}{\partial g} \quad .$$

<u>Theorem 10</u>. *If* $h(x_0) = 0$ *and*

$$- \bar{\partial} f(x_0) \subset \text{int } L(x_0) \quad , \tag{2.52}$$

where $L(x)$ *is defined by (2.9), then* x_0 *is a strict local minimum point of* f *on* Ω.

The proof of this theorem is analogous to that of Theorem 9 (see, e.g., [5,§7, Chap. II]).

Theorem 11. *If* $h(x_0) = 0$ *and*

$$L_1(x_0) \subset \text{int } L_2(x_0) \quad , \tag{2.53}$$

where $L_1(x_0)$ *and* $L_2(x_0)$ *are defined by (2.37) and (2.38), then* x_0 *is a strict local minimum point of* f *on* Ω.

Proof. From (2.53) it follows that there exists an $r > 0$ such that

$$\max_{v \in L_1(x_0)} (v,g) \le \max_{w \in L_2(x_0)} (w,g) - r \quad \forall g \in S_1 \quad ,$$

i.e.,

$$\max_{v \in [-\overline{\partial} f(x_0) - \overline{\partial} h(x_0)]} (v,g) \le M - r \quad \forall g \in S_1 \tag{2.54}$$

where

$$M = \max_{w \in co\{\underline{\partial} f(x_0) - \overline{\partial} h(x_0), \underline{\partial} h(x_0) - \overline{\partial} f(x_0)\}} (w,g) \quad .$$

Since

$$\max_{v \in co\{A \cup B\}} (v,g) = \max \left\{ \max_{v \in A} (v,g) \ , \ \max_{v \in B} (v,g) \right\}$$

then from (2.43)

$$- \min_{w \in \overline{\partial} f(x_0)} (w,g) - \min_{w \in \overline{\partial} h(x_0)} (w,g) \le \max \left\{ \max_{v \in \underline{\partial} f(x_0)} (v,g) - \min_{w \in \overline{\partial} h(x_0)} (w,g) ; \right.$$

$$\left. \max_{v \in \underline{\partial} h(x_0)} (v,g) - \min_{w \in \overline{\partial} f(x_0)} (w,g) \right\} - r \quad \forall g \in S_1 \quad . \tag{2.55}$$

Two cases are possible:

1. $$M = \max_{w \in \underline{\partial} f(x_0) - \overline{\partial} h(x_0)} (w,g) = \max_{v \in \underline{\partial} f(x_0)} (v,g) - \min_{w \in \overline{\partial} h(x_0)} (w,g)$$

2. $$M = \max_{w \in \underline{\partial} h(x_0) - \overline{\partial} f(x_0)} (w,g) = \max_{v \in \underline{\partial} h(x_0)} (v,g) - \min_{w \in \overline{\partial} f(x_0)} (w,g) \quad .$$

In case 1 it follows from (2.55) that

$$\max_{v\in\underline{\partial}f(x_0)} (v,g) + \min_{w\in\overline{\partial}f(x_0)} (w,g) = \frac{\partial f(x_0)}{\partial g} \geq r \quad . \qquad (2.56)$$

In case 2 it follows from (2.55) that

$$\max_{v\in\underline{\partial}h(x_0)} (v,g) + \min_{w\in\overline{\partial}h(x_0)} (w,g) \geq r \quad . \qquad (2.57)$$

Since $\frac{o(\alpha,g)}{\alpha} \longrightarrow 0$ uniformly with respect to $g \in S_1$ in (2.48) and $\frac{o_1(\alpha,g)}{\alpha} \longrightarrow 0$ uniformly with respect to $g \in S_1$ in (2.49), then (2.56) and (2.57) suggest that there exists an $\alpha > 0$ such that for any $x \in S_\alpha(x_0) = \{x \in E_n \mid \|x-x_0\| \leq \alpha\}$ and $x \neq x_0$ either

$$f(x) > f(x_0) \qquad (2.58)$$

(in case 1), or

$$h(x) > h(x_0) = 0 \qquad (2.59)$$

(in case 2).

If (2.59) holds, then $x \notin \Omega$.

Thus, it follows from (2.58) and (2.59) that

$$f(x) > f(x_0) \qquad \forall x \in \Omega \cap S_\alpha(x_0) \quad , \quad x \neq x_0 \quad ,$$

i.e., x_0 is a strict local minimum point of f on Ω. Q.E.D.

Remark 13. Theorem 11 is stated by A. Shapiro in [11].

Example 2. Let $x = \left(x^{(1)}, x^{(2)}\right) \in E_2$, $x_0 = (0,0)$; $f(x) = |x^{(1)}| - |x^{(2)}| + x^{(2)}$; $h(x) = -\frac{1}{2}|x^{(1)}| - x^{(2)}$; $\Omega = \{x \in E_2 \mid h(x) \leq 0\}$.

From quasidifferential calculus we have

$$\underline{\partial} f(x_0) = \text{co } \{(1,1),(-1,1)\} , \quad \overline{\partial} f(x_0) = \text{co } \{(0,1),(0,-1)\} ,$$

$$\underline{\partial} h(x_0) = \{(0,-1)\} , \quad \overline{\partial} h(x_0) = \text{co } \{(\tfrac{1}{2},0),(-\tfrac{1}{2},0)\} .$$

It is shown in [5, §5, Chap. II] that the nondegeneracy condition (2.4) is satisfied at x_0 (see Fig. 2).

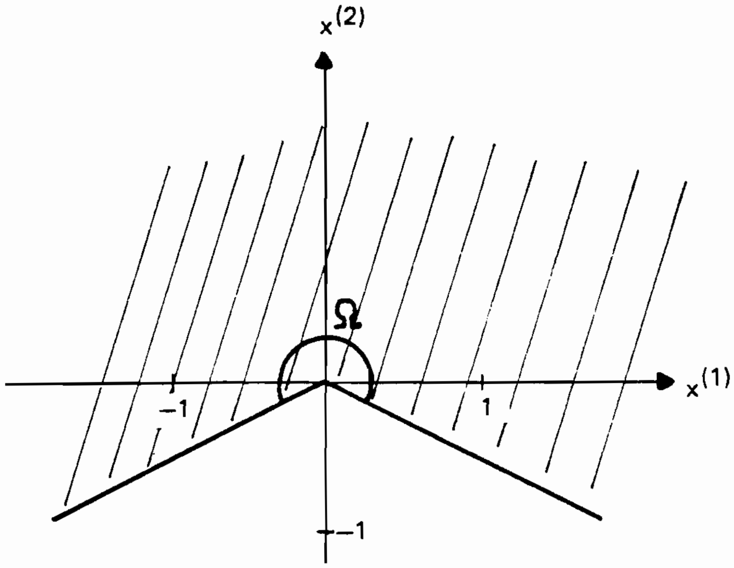

Figure 2

We shall now verify the necessary conditions for a minimum. Construct sets $L(x_0)$, $L_1(x_0)$, $L_2(x_0)$ (see (2.9), (2.37), (2.38)):

$$L(x_0) = \bigcap_{w \in \overline{\partial} h(x_0)} [\underline{\partial} f(x_0) + \text{cl}(\text{cone}(\underline{\partial} h(x_0) + w))] =$$

$$= \text{co } \{(-1,1),(1,1),(0,-1)\} ,$$

$$L_1(x_0) = -[\overline{\partial} f(x_0) + \overline{\partial} h(x_0)] = \text{co}\{(\tfrac{1}{2},1),(\tfrac{1}{2},-1),(-\tfrac{1}{2},-1),(-\tfrac{1}{2},1)\} ,$$

$$\underline{\partial} f(x_0) - \overline{\partial} h(x_0) = \text{co } \{(\tfrac{3}{2},1),(-\tfrac{3}{2},1)\} ,$$

$$\underline{\partial} h(x_0) - \overline{\partial} f(x_0) = \text{co } \{(0,0),(0,-2)\} ,$$

$$L_2(x_0) = \text{co}\{\underline{\partial} f(x_0) - \overline{\partial} h(x_0);\ \ \underline{\partial} h(x_0) - \overline{\partial} f(x_0)\} =$$

$$= \text{co}\{(\tfrac{3}{2},1),(-\tfrac{3}{2},1),(0,-2)\} \quad .$$

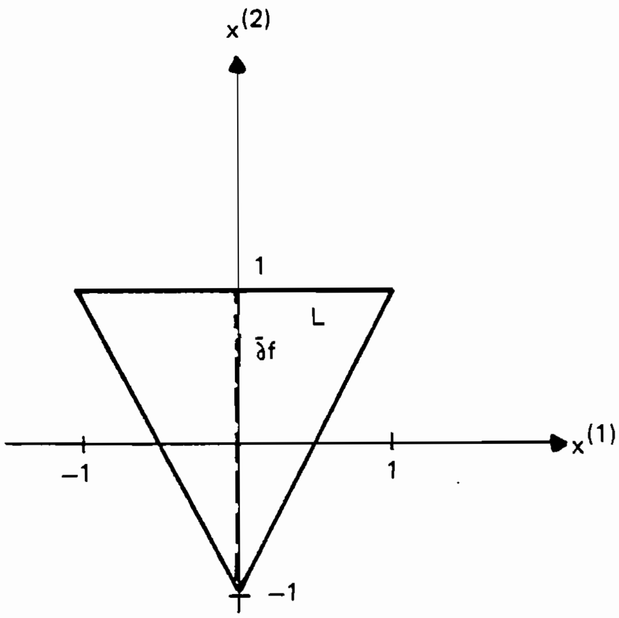

Figure 3

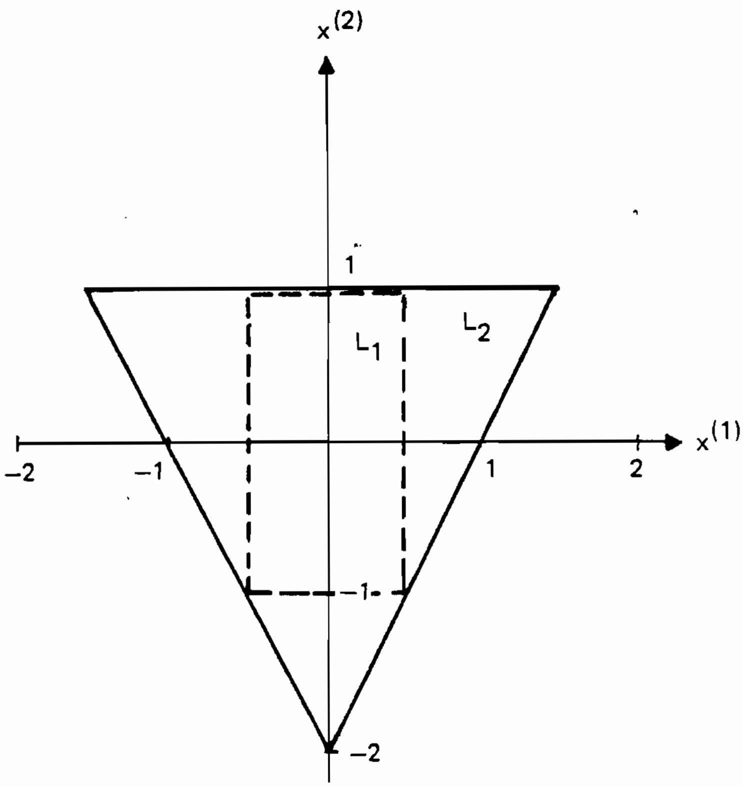

Figure 4

From Figs. 3 and 4 it is clear that the necessary conditions for a minimum (2.8) and (2.36) are satisfied.

Example 3. Let $x \in E_2$, $x_0 = (0,0)$; Ω and h be the same as in Example 2 and

$$f(x) = |x^{(1)}| - \tfrac{1}{2}|x^{(2)}| + x^{(2)} \quad .$$

Now we have

$$\underline{\partial} f(x_0) = \text{co}\ \{(1,1),(-1,1)\}, \quad \overline{\partial} f(x_0) = \text{co}\ \{(0,\tfrac{1}{2}),(0,-\tfrac{1}{2})\} \quad .$$

$L(x_0)$ remains the same:

$$L(x_0) = \text{co}\ \{(-1,1),(1,1),(0,-1)\} \quad .$$

Let us find $L_1(x_0)$ and $L_2(x_0)$:

$$L_1(x_0) = -[\overline{\partial}f(x_0) + \overline{\partial}h(x_0)] = \text{co}\ \{(\tfrac{1}{2},\tfrac{1}{2}),(\tfrac{1}{2},-\tfrac{1}{2}),(-\tfrac{1}{2},-\tfrac{1}{2}),(-\tfrac{1}{2},\tfrac{1}{2})\} \ ,$$

$$\underline{\partial}f(x_0) - \overline{\partial}h(x_0) = \text{co}\ \{(\tfrac{3}{2},1),(-\tfrac{3}{2},1)\} \ ,$$

$$\underline{\partial}h(x_0) - \overline{\partial}f(x_0) = \text{co}\ \{(0,-\tfrac{1}{2}),(0,-\tfrac{3}{2})\} \ ,$$

$$L_2(x_0) = \text{co}\ \{\underline{\partial}f(x_0) - \overline{\partial}h(x_0), \underline{\partial}h(x_0) - \overline{\partial}f(x_0)\}$$

$$= \text{co}\ \{(\tfrac{3}{2},1),(-\tfrac{3}{2},1),(0,-\tfrac{3}{2})\} \quad .$$

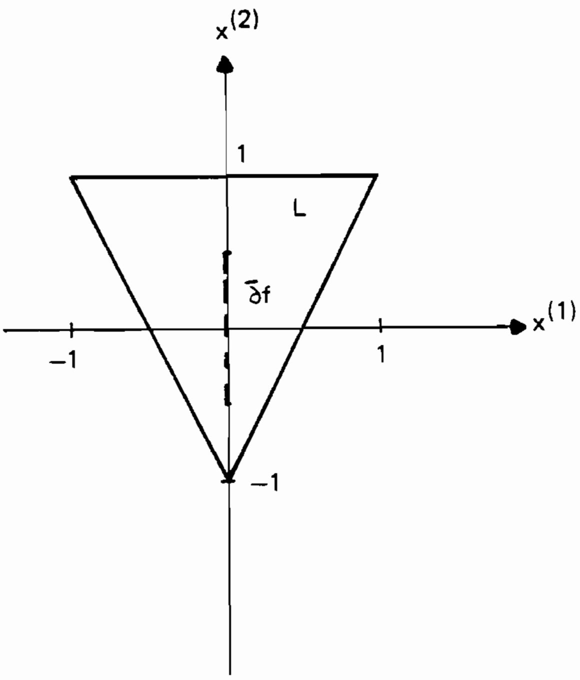

Figure 5

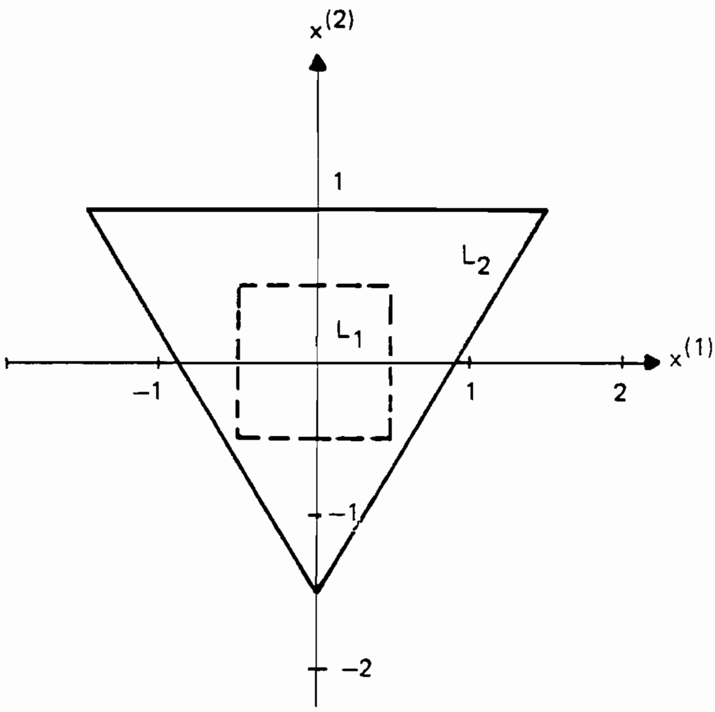

Figure 6

It is clear from Figure 5 that

$$-\ \overline{\partial}f(x_0) \subset \text{int}\ L(x_0)$$

i.e., the sufficient condition for a strict local minimum (2.52) is satisfied. Figure 6 shows that the sufficient condition (2.53) is also satisfied.

Remark 14. In Example 2 x_0 was in fact a minimum point but we cannot deduce this from the necessary conditions alone.

Example 4. Let $x = (x^{(1)}, x^{(2)}) \in E_2$, $x_0 = (0,0)$; $f(x) = |x^{(1)}| - |x^{(2)}| + x^{(2)}$; $h(x) = -|x^{(1)}| - x^{(2)}$; $\Omega = \{x | h(x) \le 0\}$.

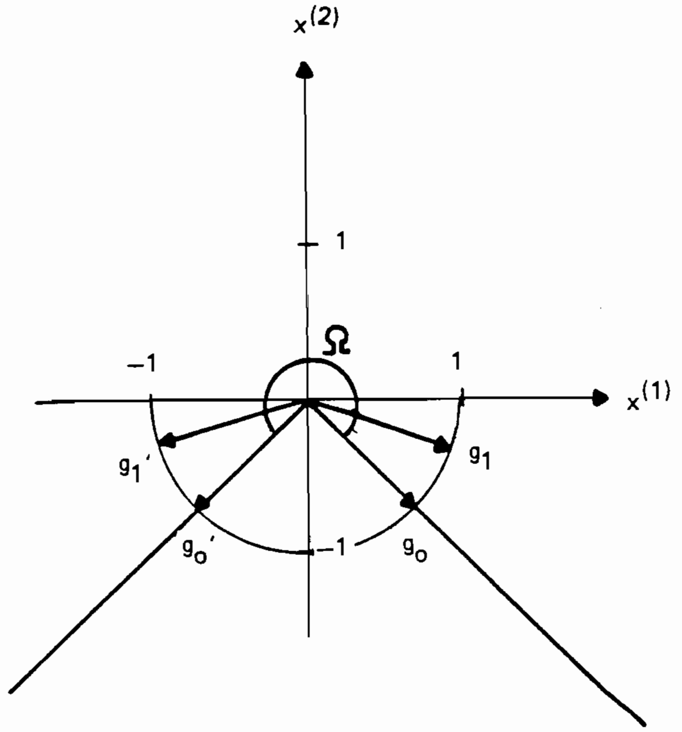

Figure 7

Thus the function f is the same as in Example 2 - only Ω has been changed (see Figure 7). Note that the nondegeneracy condition (2.4) is satisfied and that

$$\underline{\partial} f(x_0) = \text{co}\,\{(1,1),(-1,1)\}\ , \quad \overline{\partial} f(x_0) = \text{co}\,\{(0,1),(0,-1)\}\ ,$$

$$\underline{\partial} h(x_0) = \{(0,-1)\}\ , \quad \overline{\partial} h(x_0) = \text{co}\,\{(1,0),(-1,0)\}\ .$$

Construct sets $L(x_0)$, $L_1(x_0)$, $L_2(x_0)$:

$$L(x_0) = \bigcap_{w \in \bar{\partial} h(x_0)} [\underline{\partial} f(x_0) + \mathrm{cl}(\mathrm{cone}(\underline{\partial} h(x_0) + w))]$$

$$= \mathrm{co}\{(-1,1),(1,1),(0,0)\} \quad ,$$

$$\underline{\partial} f(x_0) - \bar{\partial} h(x_0) = \mathrm{co}\ \{(2,1),(-2,1)\} \quad ,$$

$$\underline{\partial} h(x_0) - \bar{\partial} f(x_0) = \mathrm{co}\ \{(0,0),(0,-2)\} \quad ,$$

$$L_1(x_0) = -[\bar{\partial} f(x_0) + \bar{\partial} h(x_0)] = \mathrm{co}\ \{(1,1),(1,-1),(-1,-1),(-1,1)\},$$

$$L_2(x_0) = \mathrm{co}\ \{-\underline{\partial} f(x_0) - \bar{\partial} h(x_0), \underline{\partial} h(x_0) - \bar{\partial} f(x_0)\}$$

$$= \mathrm{co}\ \{(2,1),(-2,1),(0,-2)\} \quad .$$

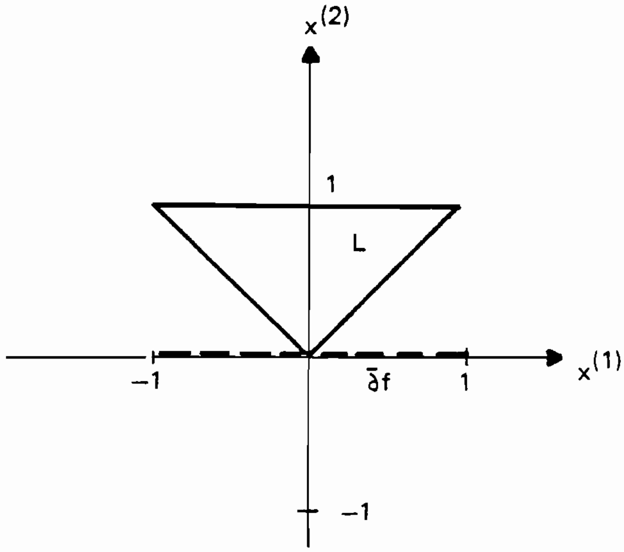

Figure 8

We observe that the necessary condition (2.8) is not satisfied (see Figure 8). We then calculate the steepest descent directions (see [5,§7,Chap.II]), obtaining $g_0 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ and $g_0' = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$.

It can be seen from Figure 8 that the necessary condition (2.36) is also not satisfied (this is hardly surprising since conditions (2.8) and (2.36) are equivalent). We shall now find directions satisfying (2.39)-(2.41). It is clear from Figure 9

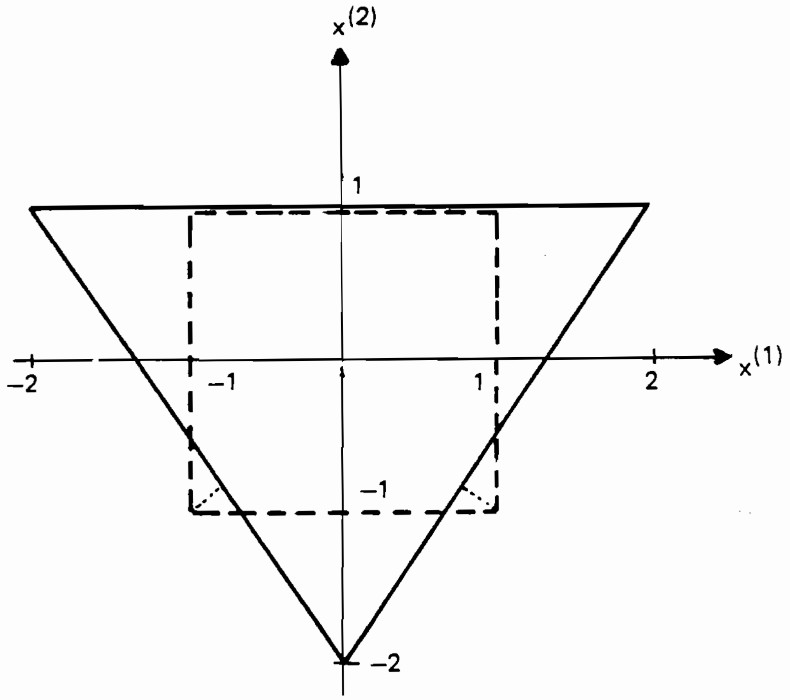

Figure 9

that there exist two directions of this kind:

$$g_1 = \left(\frac{3}{\sqrt{10}}, -\frac{1}{\sqrt{10}}\right) \quad , \quad g_1' = \left(-\frac{3}{\sqrt{10}}, -\frac{1}{\sqrt{10}}\right) .$$

Figure 6 shows that the directions of steepest descent g_0 and g_0' are tangent directions but that the descent directions g_1 and g_1' are interior.

REFERENCES

1. V.G. Boltyanskiy. A method of tents in the theory of extremal problems. *Uspekhi Math. Sci.*, Vol. 30(3), 1975, pp. 3-55.

2. S. Karlin. *Mathematical Methods and Theory in Games, Programming and Economics*, Vols. I, II, Addison-Wesley, Reading, Mass., 1959.

3. R.T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

4. B.N. Pshenichniy. *Necessary Conditions for Extremum Problems*. Marcel Dekker, New York, 1971.

5. V.F. Demyanov and L.V. Vasiliev. *Nondifferentiable Optimization* (in Russian). Nauka, Moscow, 1981.

6. V.F. Demyanov and A.M. Rubinov. On quasidifferentiable functionals. *Soviet Math. Dokl.*, Vol. 21(1), 1980, pp. 14-17.

7. L.N. Polyakova. Necessary conditions for an extremum of quasidifferentiable functions. *Vestnik Leningrad Univ. Math.*, Vol. 13, 1981, pp. 241-247.

8. V.F. Demyanov and L.N. Polyakova. Minimization of a quasidifferentiable function on a quasidifferentiable set. *USSR Comput. Math. Math. Phys.*, Vol. 20(4), 1980, pp. 34-43.

9. V.F. Demyanov and A.M. Rubinov. On some approaches to the nonsmooth optimization problem (in Russian). *Ekonomika i Matimaticheskie Metody*, Vol. 17(6), 1981, pp. 1153-1174.

10. B.N. Pshenichniy and Yu.M. Danilin. *Numerical Methods for Extremum Problems* (in Russian). Nauka, Moscow, 1975.

11. A. Shapiro. On optimality conditions in quasidifferentiable optimization (forthcoming).