DIRECTIONAL DIFFERENTIABILITY OF A CONTINUAL MAXIMUM FUNCTION OF QUASIDIFFERENTIABLE FUNCTIONS

V.F. Demyanov
I.S. Zabrodin

June 1983
WP-83-58

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

PREFACE

Much recent work in optimization theory has been concerned with the problems caused by nondifferentiability. Some of these problems have now been at least partially overcome by the definition of a new class of nondifferentiable functions called quasidifferentiable functions, and the extension of classical differential calculus to deal with this class of functions. This has led to increased theoretical research in the properties of quasidifferentiable functions and their behavior under different conditions.

In this paper, the problem of the directional differentiability of a maximum function over a continual set of quasidifferentiable functions is discussed. It is shown that, in general, the operation of taking the "continual" maximum (or minimum) leads to a function which is itself not necessarily quasidifferentiable.

Andrzej Wierzbicki
Chairman
System & Decision Sciences

DIRECTIONAL DIFFERENTIABILITY OF A CONTINUAL MAXIMUM FUNCTION OF QUASIDIFFERENTIABLE FUNCTIONS

V.F. Demyanov
I.S. Zabrodin

1. INTRODUCTION

Optimization problems involving nondifferentiable functions are recognized to be of great theoretical and practical significance. There are many ways of approaching the problems caused by nondifferentiability, some of which are now quite well developed while others still require much further work. A comprehensive bibliography of publications concerned with nondifferentiable optimization has recently been compiled [1]--major contributors in this field include J.P. Aubin, F.H. Clarke, Yu.M. Ermoliev, J.B. Hiriart-Urruty, A.Ya. Kruger, S.S. Kutateladze, C. Lemarèchal, B.S. Morduchovich, E.A. Nurminski, B.N. Pshenichniy, R.T. Rockafellar, and J. Warga.

The notion of subgradient has been generalized to nonconvex functions in a number of different ways. One of these involves the definition of a new class of nondifferentiable functions (quasidifferentiable functions) which has been shown to represent a linear space closed with respect to all algebraic operations as well as to the taking of pointwise maximum and minimum [2,3]. This has led to the development of quasidifferential calculus--a generalization of classical differential calculus--which may be used to solve many new optimization problems involving nondifferentiability [4].

This paper deals with the problem of the directional differentiability of a maximum function over a continual set of quasidifferentiable functions. It will be shown that in general the operation of taking the "continual" maximum (minimum) leads to a function which is itself not necessarily quasidifferentiable.

2. AUXILIARY RESULTS

Let us consider a mapping $G : E_n \to 2^{E_m}$, where 2^{E_m} denotes the set of all subsets of E_m . Fix $x_o \in E_n$ and $g \in E_n$, $\|g\|=1$. Choose $y \in G(x_o)$ and introduce the set

$$\gamma(y) \equiv \gamma(x_o,g,y) = \{v \in E_m \mid \exists\, \alpha_o > 0 : y + \alpha v \in G(x_o+\alpha g) \quad \forall \alpha \in [0,\alpha_o]\}.$$

We shall denote the closure of $\gamma(y)$ by $\Gamma(y)$, i.e.,

$$\Gamma(y) \equiv \Gamma(x_o,g,y) = \text{cl}\ \gamma(y) .$$

The set $\Gamma(y)$ is called the set of first-order feasible directions at the point $y \in G(x_o)$ in the direction g .

Remark 1. In the case where G does not depend on x , the set $\Gamma(y) = \Gamma(x_o,g,y)$ does not depend on x_o and g , and is a cone called the cone of feasible directions at y .

A mapping G is said to allow first-order approximation at a point x_o in the direction $g \in E_n$, $\|g\|=1$, if, for an arbitrary convergent sequence $\{y_k\}$ such that

$$y_k \to y \ , \ y_k \in G(x_o + \alpha_k g) \ , \ \alpha_k \to +0 \quad \text{as} \quad k \to \infty ,$$

the following representation holds:

$$y_k = y + \alpha_k v_k + o(\alpha_k)$$

where $v_k \in \Gamma(x_o,g,y)$, $\alpha_k v_k \to 0$, $y \in G(x_o)$.

In what follows it is assumed that the mapping G is continuous (in the Hausdorff metric) at a point x_o and allows first-order approximation at x_o in any direction $g \in E_n$, $\|g\| = 1$.

It is also assumed that for every

$$x \in S_\delta(x_o) = \{x \in E_n \mid \|x - x_o\| \leqslant \delta\}, \quad \delta > 0,$$

the set $G(x)$ is closed and sets $G(x)$ are jointly bounded on $S_\delta(x_o)$, i.e., there exists an open bounded set $B \subset E_m$ such that

$$G(x) \subset B \quad \forall x \in S_\delta(x_o).$$

Let us consider the function

$$f(x) = \max_{y \in G(x)} \phi(x,y)$$

where function $\phi(z) = \phi(x,y)$ is continuous in $z = [x,y]$ on $S_\delta(x_o) \times B$ and differentiable on Z_o in any direction $\eta = [g,q] \in E_{n+m}$, i.e., there exists a finite limit

$$\frac{\partial \phi(x_o,y)}{\partial \eta} \equiv \frac{\partial \phi(x_o,y_o)}{\partial [g,q]} = \lim_{\alpha \to +0} \frac{1}{\alpha}[\phi(x_o+\alpha g, y_o+\alpha q) - \phi(x_o,y_o)]$$

$$\forall z_o = [x_o,y_o] \in Z_o.$$

Here

$$Z_o = \{x_o\} \times \{R(x_o)\}$$

$$R(x) = \{y \in G(x) \mid \phi(x,y) = f(x)\}.$$

Suppose that the following conditions hold:

Condition 1. If $q_k \to q$ then

$$\frac{\partial \phi(x_o,y)}{\partial [g,q]} \leqslant \overline{\lim_{k \to +\infty}} \frac{\partial \varphi(x_o,y)}{\partial [g,q_k]} \qquad \forall y \in R(x_o).$$

<u>Condition 2</u>. Let $y_k \in R(x_o + \alpha_k g)$, $y_k \to \bar{y}$. Since G allows first-order approximation at x_o , then

$$y_k = \bar{y} + \alpha_k q_k + o(\alpha_k) \ , \ \bar{y} \in R(x_o) \ .$$

It is assumed that the q_k's are bounded.

<u>Condition 3</u>. Function ϕ is Lipschitzian in some neighborhood of the set Z_o .

Then the following result holds.

Theorem 1. The function f *is differentiable at the point* x_o *in the direction* g *and*

$$\frac{\partial f(x_o)}{\partial g} = \sup_{y \in R(x_o)} \ \sup_{q \in \Gamma(y)} \frac{\partial \phi(x_o, y)}{\partial [g, q]} \ . \qquad (1)$$

Proof. Let us denote by A the right-hand side of (1). Fix $y \in R(x_o)$ and $q \in \gamma(y)$. Then $y + \alpha q \in G(x_o + \alpha g)$ for sufficiently small $\alpha > 0$ and

$$f(x_o + \alpha g) \geqslant \phi(x_o + \alpha g, y + \alpha q) = f(x_o) + \alpha \left(\frac{\partial \phi(x_o, y)}{\partial [g, q]} \right) + o(\alpha) \ .$$

Hence

$$\underline{\lim}_{\alpha \to +0} h(\alpha) \equiv \underline{\lim}_{\alpha \to +0} \frac{1}{\alpha} [f(x_o + \alpha g) - f(x_o)] \geqslant \frac{\partial \phi(x_o, y)}{\partial [g, q]} \ .$$

Since $y \in R(x_o)$ and $q \in \gamma(y)$ are arbitrary then

$$\underline{\lim}_{\alpha \to +0} h(\alpha) \geqslant \sup_{y \in R(x_o)} \ \sup_{q \in \gamma(y)} \frac{\partial \phi(x_o, y)}{\partial [q, g]} \qquad (2)$$

Let $q_k \to q$, $q_k \in \gamma(y)$. Then $q \in \Gamma(y)$. It follows from Condition 1 that

$$\frac{\partial\phi(x_o,y)}{\partial[g,q]} \leq \overline{\lim_{k\to+\infty}} \frac{\partial\phi(x_o,y)}{\partial[g,q_k]} . \quad (3)$$

But

$$\frac{\partial\phi(x_o,y)}{\partial[g,q_k]} \leq \sup_{q\in\gamma(y)} \frac{\partial\phi(x_o,y)}{\partial[g,q]} .$$

Since $\Gamma(y) = \text{cl}\ \gamma(y)$, then from (3)

$$\sup_{q\in\gamma(y)} \frac{\partial\phi(x_o,y)}{\partial[g,q]} \leq \sup_{q\in\Gamma(y)} \frac{\partial\phi(x_o,y)}{\partial[g,q]} \leq \sup_{q\in\gamma(y)} \frac{\partial\phi(x_o,y)}{\partial[g,q]} .$$

Hence

$$\sup_{q\in\Gamma(y)} \frac{\partial\phi(x_o,y)}{\partial[g,q]} = \sup_{q\in\gamma(y)} \frac{\partial\phi(x_o,y)}{\partial[g,q]} . \quad (4)$$

From (2) and (4) it follows that

$$\underline{\lim}_{\alpha\to+0}\ h(\alpha) \geq A . \quad (5)$$

Now let us choose sequences $\{y_k\}$ and $\{\alpha_k\}$ such that

$$\frac{1}{\alpha_k} [f(x_o + \alpha_k g) - f(x_o)] \longrightarrow \overline{\lim_{\alpha\to+0}}\ h(\alpha) , \quad (6)$$

$$y_k \in R(x_o + \alpha_k g) , \quad y_k \longrightarrow \bar{y} , \quad \alpha_k \longrightarrow +0 .$$

The conditions imposed on the mapping G and the continuity of the function ϕ ensure that the function f is continuous at x_o. Hence, from the equality $f(x_o+\alpha_k g) = \phi(x_o+\alpha_k g, y_k)$, one can conclude that $f(x_o) = \phi(x_o,\bar{y})$, i.e., $\bar{y} \in R(x_o)$.

Since the mapping G allows first-order approximation at x_o, then $y_k = \bar{y} + \alpha_k q_k + o(\alpha_k)$, where $q_k \in \Gamma(\bar{y})$, $\alpha_k q_k \longrightarrow +0$. From Conditions 2 and 3 the q_k's are bounded and the function ϕ is Lipschitzian around Z_o. Without loss of generality one can assume that $q_k \longrightarrow q$. It is clear that $q \in \Gamma(\bar{y})$. Hence

$$f(x_o + \alpha_k g) - f(x_o) = \phi(x_o+\alpha_k g, y_k) - \phi(x_o,\bar{y})$$

$$= \phi(x_o+\alpha_k g, \bar{y}+\alpha_k q_k+o(\alpha_k)) - \phi(x_o,\bar{y}) = Q_1 + Q_2 \qquad (7)$$

where

$$Q_1 = \phi(x_o+\alpha_k g, \bar{y}+\alpha_k q) - \phi(x_o,\bar{y}) = \alpha_k \left(\frac{\partial\phi(x_o,y)}{\partial[g,q]}\right) + o(\alpha_k)$$

$$Q_2 = \phi(x_o+\alpha_k g, \bar{y}+\alpha_k q_k+o(\alpha_k)) - \phi(x_o+\alpha_k g, \bar{y}+\alpha_k q) .$$

Since ϕ is a Lipschitzian function, then

$$|Q_2| \leqslant L\ \alpha_k \|q_k - q + o(\alpha_k)\| . \qquad (8)$$

It follows from (6)-(8) that

$$\overline{\lim_{\alpha\to+0}}\ h(\alpha) = \lim_{k\to\infty} \frac{1}{\alpha_k}[f(x_o + \alpha_k g) - f(x_o)] = \frac{\partial\phi(x_o,\bar{y})}{\partial[g,q]}$$

from which it is clear that

$$\overline{\lim_{\alpha\to+0}}\ h(\alpha) \leqslant \sup_{y\in R(x_o)}\ \sup_{q\in\Gamma(y)} \frac{\partial\phi(x_o,y)}{\partial[g,q]} = A . \qquad (9)$$

Comparison of (5) and (9) now shows that $\lim_{\alpha\to+0} h(\alpha)$ exists and is equal to A, thus completing the proof.

Remark 2. Equation (1) has been proved under some different assumptions elsewhere [5] (see also [6], § 10). The case where ϕ is differentiable was studied by Hogan [7].

3. QUASIDIFFERENTIABLE CASE

Let us consider once again the function

$$f(x_o) = \max_{y\in G(x)} \phi(x,y) \tag{10}$$

where mapping G satisfies the conditions specified earlier and function $\phi(z) = \phi(x,y)$ is continuous in z on $S_g(x_o) \times B$ and quasidifferentiable on Z_o, i.e., for any point $z_o = [x_o,y_o] \in Z_o$ there exist convex compacts $\underline{\partial}\phi(z_o) \subset E_{n+m}$ and $\bar{\partial}\phi(z_o) \subset E_{n+m}$ such that

$$\frac{\partial\phi(x_o,y_o)}{\partial[g,q]} = \lim_{\alpha\to+0} \frac{1}{\alpha} [\phi(x_o+\alpha g,y_o+\alpha q) - \phi(x_o,y_o)]$$

$$= \max_{[v_1,v_2]\in\underline{\partial}\phi(z_o)} [(v_1,g) + (v_2,q)] \tag{11}$$

$$+ \min_{[w_1,w_2]\in\bar{\partial}\phi(z_o)} [(w_1,g) + (w_2,q)] .$$

It is also assumed that Conditions 2 and 3 are satisfied. (Condition 1 follows immediately from (11).) Thus, all the conditions of Theorem 1 are fulfilled and we arrive at

Theorem 2. The function f *defined by (10) is directionally differentiable and, moreover,*

$$\frac{\partial f(x_o)}{\partial g} = \sup_{y\in R(x_o)} \sup_{q\in\Gamma(y)} \Bigg\{ \max_{[v_1,v_2]\in\underline{\partial}\phi(x_o,y)} [(v_1,g) + (v_2,q)]$$

$$+ \min_{[w_1,w_2]\in\bar{\partial}\phi(x_o,y)} [(w_1,g) + (w_2,q)] \Bigg\} . \tag{12}$$

<u>*Remark 3*</u>. Since $y \in R(x_o,y)$, the following relation holds:

$$\sup_{q\in\Gamma(y)} \Bigg\{ \max_{v_2\in\underline{\partial}\phi_y(x_o,y)} (v_2,q) + \min_{w_2\in\bar{\partial}\phi_y(x_o,y)} (w_2,q) \Bigg\} = 0$$

$$\forall\, y \in R(x_o) .$$

Here $\underline{\partial}\phi_y(x_o,y)$ and $\bar{\partial}\phi_y(x_o,y)$ are the projections of sets $\underline{\partial}\phi(x_o,y)$ and $\bar{\partial}\phi(x_o,y)$, respectively, onto E_m .

Remark 4. Pshenichniy [8] considered the case where $G(x)$ does not depend on x and $F_y(x) = \phi(x,y)$ is a directionally differentiable function for every fixed y , i.e., there exists

$$\frac{\partial\phi(x,y)}{\partial g} = \lim_{\alpha\to+0} \frac{1}{\alpha} [\phi(x+\alpha g,y) - \phi(x,y)] .$$

Then

$$\phi(x+\alpha g,y) = \phi(x,y) + \alpha \frac{\partial\phi(x,y)}{\partial g} + o(\alpha,y) . \tag{13}$$

Under an additional assumption about the behavior of $o(\alpha,y)$ in (13), it has been proved that

$$\frac{\partial f(x)}{\partial g} = \max_{y\in R(x)} \frac{\partial\phi(x,y)}{\partial g} . \tag{14}$$

It is clear that equation (14) differs from equation (12).

Example 1. Let $x\in E_1$, $y\in E_1$, $G(x) \equiv G = [-2,2]$, $\phi(x,y) = x - 2|y - x|$, and

$$f(x) = \max_{y\in[-2,2]} (x - 2|y - x|) . \tag{15}$$

It is clear that

$$f(x) = x , \ R(x) = \{x\} \quad \forall x \in (-2,2) . \tag{16}$$

Choose $x \in (-2,2)$ and verify equation (14). We shall now compute the right-hand side of (14). Since $\phi(x,y) = x - 2 \max \{y - x , -y + x\}$, then for $y \in R(x) = \{x\}$ it follows [9] that

$$\frac{\partial\phi(x,y)}{\partial g} = g - 2 \max \{-g,g\} .$$

Hence for $g_1 = +1$

$$\max_{y \in R(x)} \frac{\partial \phi(x,y)}{\partial g_1} = 1 - 2 = -1 ,$$

and for $g_2 = -1$

$$\max_{y \in R(x)} \frac{\partial \phi(x,y)}{\partial g_2} = -1 - 2 = -3 .$$

But from (16) it is clear that

$$\frac{\partial f(x)}{\partial g} = g \qquad \forall x \in (-2,2) . \tag{17}$$

Thus equation (14) does not hold for any direction g (in E_1 there are only two directions g such that $\|g\| = 1 : g = +1$ and $g = -1$).

Now let us verify equation (12). Denote by D the right-hand side of (12). The function $\phi(x,y)$ is quasidifferentiable. From quasidifferential calculus [2-4] it follows that if $y = x$ then one can choose $\underline{\partial}\phi(x,y) = \{(1,0)\}$, $\bar{\partial}\phi(x,y) = \text{co}\ \{(-2,2) , (2,-2)\}$.

For the function f described by (15) we have

$$\Gamma(y) \equiv \Gamma(x,y) = E_1 \quad \forall x \in (-2,2) .$$

Computing D :

$$D = \sup_{q \in E_1} \left\{ (1.g) + (0.q) + \min_{[w_1,w_2] \in \text{co}[(-2,2),(2,2)]} [(w_1.g) + (w_2.q)] \right\}$$

$$= g + \sup_{q \in E_1} \min_{[w_1,w_2] \in [(-2,2),(2,-2)]} [(w_1.g) + (w_2 q)] . \tag{18}$$

It is clear from Figure 1 that for any g the second term on the right-hand side of (18) is equal to zero, i.e., $D = g$. (The supremum in (18) is attained at $q = g$.)

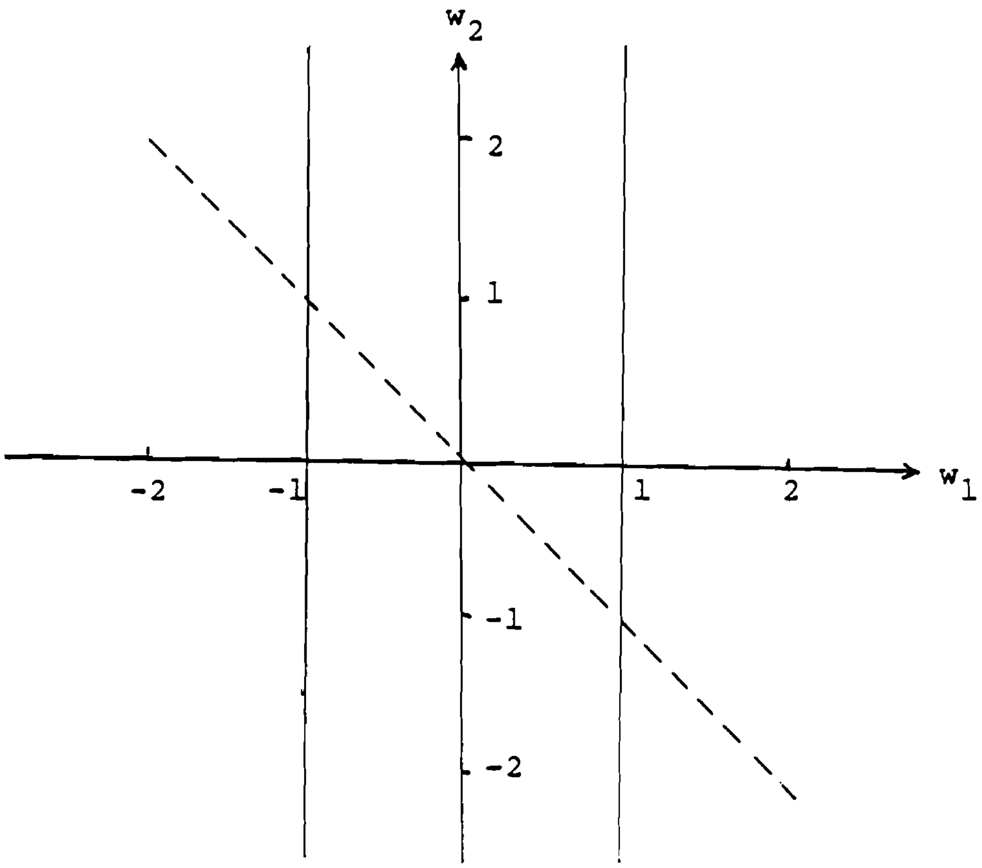

Figure 1.

Thus, from (17), equation (12) is correct in this case.

Remark 5. When solving practical problems in which it is required to minimize a max function over a continual set of points, this maximum function is often discretized (the continual set replaced by a grid of points). In many cases this operation is a legitimate one [10], but we shall show that in the case where ϕ is a quasidifferentiable function this replacement may be dangerous.

Let f again be described by (15). Define f_N as

$$f_N(x) = \max_{y \in \sigma_N} (x - 2|y - x|) ,$$

where $\sigma_N = \{x_1, \ldots, x_N\}$, $x_k \in [-2,2]$.

This function has N local minima (see Figure 2), although the original $f = x$ has no local minimum which is not also global on $[-2,2]$. This demonstrates that the discretization of the max-type function must be carried out very cautiously.

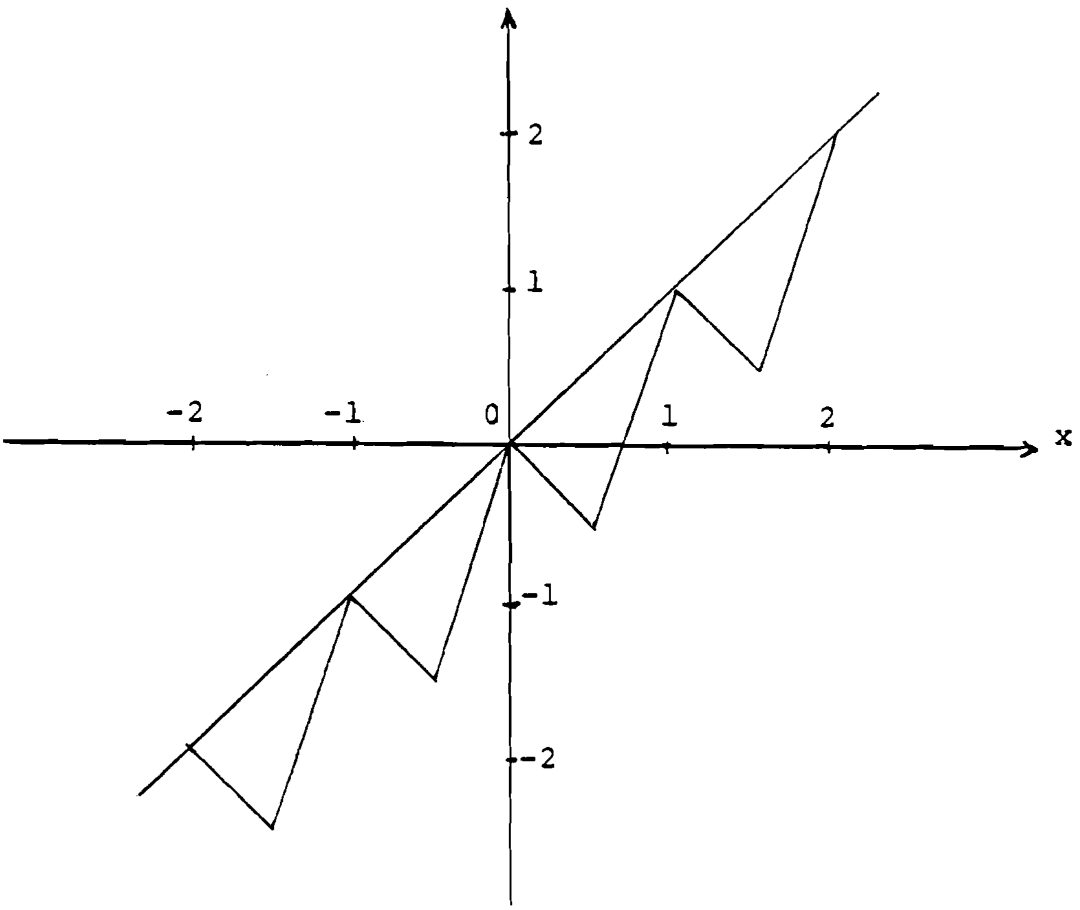

Figure 2

Example 2. This example illustrates that Condition 2 is essential to our argument.

$$\phi(x,y) = x - 2 \min_{t\in[-2,2]} \sqrt{(x - t^3)^2 + (y - t)^2}$$

$$f(x) = \max_{y\in[-2,2]} \phi(x,y) \ . \tag{19}$$

It is clear that for $x \in (-2,2)$,

$$R(x) = \{y \mid y = \sqrt[3]{x}\} , \quad f(x) = x . \tag{20}$$

If $y = \sqrt[3]{x}$, then the minimum in (19) is achieved at $t = x$. Take $x_o = 0$. Then $R(0) = \{0\}$, $\Gamma(y) = E_1$. Construct a quasidifferential of the function ϕ at the point $(0,0)$. By the rules of quasidifferential calculus one can choose

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

Let us denote by D the right-hand side of equation (12) and evaluate it.

$$D \equiv D(g) = \sup_{q \in E_1} \left\{ 1.g + 0 + \min_{[w_1,w_2] \in \text{co}\,[(-2,0),(2,0)]} (w_1 \cdot g + w_2 \cdot q) \right\}$$

$$= g - 2|g| .$$

If $g_1 = 1$ then $D(g_1) = -1$; if $g_2 = -1$ then $D(g_2) = -3$.

But it is clear from (20) that $\partial f(0)/\partial g = g$. Thus equation (12) does not hold, and the reason is that Condition 2 is not satisfied. Indeed, taking an arbitrary sequence $x_k = x_o + \alpha_k g$ where $\alpha_k \to +0$, putting, for example, $g = 1$, $x_o = 0$, we obtain $y_k = \bar{y} + \alpha_k v_k$, $y_k \in R(x_k)$.

For $\bar{y} = 0$ this leads to $R(x_k) = \{\sqrt[3]{\alpha_k}\}$. But

$$v_k = \frac{\sqrt[3]{\alpha_k}}{\alpha_k} \xrightarrow[k \to \infty]{} +\infty .$$

REFERENCES

1. E.A. Nurminski (Ed.). Progress in Nondifferentiable Optimization. CP-82-S8, IIASA, Laxenburg, 1982.

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

3. L.N. Polyakova. Necessary conditions for an extremum of quasidifferentiable functions. Vestnik Leningrad University, No. 13, pp. 57-62, 1980.

4. V.F. Demyanov and L.V. Vasiliev. Nondifferentiable Optimization. Nauka, Moscow, 1981. (In Russian)

5. T.K. Vinogradova, V.F. Demyanov, and A.B. Pevnyi. On the directional differentiability of functions of maximum and minimum. Optimization (Novosibirsk), 10(27), pp. 17-21, 1973. (In Russian)

6. V.F. Demyanov. Minimax: Directional Differentiability. Leningrad University Press, Leningrad, 1974. (In Russian)

7. W.W. Hogan. Directional derivatives for extremal value functions. West. Manage. Sci. Inst., Working Paper No. 177, Los Angeles, 1971.

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

9. V.F. Demyanov and V.N. Malozemov. Introduction to Minimax. Wiley, New York, 1974.

10. V.N. Malozemov. On the convergence of a grid method in the best polynomial approximation problem. Vestnik Leningrad University, No. 19, pp. 138-140, 1970. (In Russian)