Proceedings of the International Congress of Mathematicians
August 16–24, 1983, Warszawa

R. TYRRELL ROCKAFELLAR*

Differentiability Properties of the Minimum Value in an Optimization Problem Depending on Parameters

1. A central topic in optimization theory is the study of the optimal value and optimal solution set

$$p(v) = \inf_{x \in A(v)} f(v, x), \quad X(v) = \operatorname*{argmin}_{x \in A(v)} f(v, x), \tag{1}$$

in an optimization problem over $x \in R^n$ which depends on a parameter vector $v \in R^d$. Often this is a prelude to minimizing or maximizing $p(v)$ subject to further constraints on v, as is the case for instance in decomposition schemes in mathematical programming and various problems of approximation or engineering design. The question of the possible continuity and differentiability properties of the function p is then very important. Such properties also turn out to be critical in the derivation of optimality conditions which characterize the points $x \in X(v)$.

Let us normalize by focusing on behavior around $v = 0$. Assume that $A(0) \neq \emptyset$, the function f: $R \times R^n$ is locally Lipschitz continuous, the set $\operatorname{gph} A = \{(v, x) \mid x \in A(v)\} \subset R^d \times R^n$ is closed, and that for some $\varepsilon > 0$ the set

$$\{(v, x) \mid |v| \leqslant \varepsilon, x \in A(v), f(v, x) \leqslant a\}$$

is bounded for every $a \in R$. Then p is lower semicontinuous on a neighborhood of $v = 0$ with $p(0)$ finite and $X(0)$ nonempty and compact. Our aim is to clarify the circumstances under which p is actually Lipschitz continuous in a neighborhood of $v = 0$ and has directional derivatives of various sorts.

* Research supported by the Air Force Office of Scientific Research, U.S.A.F., under grant No. F49620-82-K-0012.

2. Ordinary one-sided directional derivatives

$$p'(v; h) = \lim_{t \to 0_+} [p(v+th) - p(v)]/t \tag{2}$$

exist only in rather special cases. One such case, among the first to be identified, is that in which $f \in \mathscr{C}^1$ and $A(v)$ is a fixed set B for all v. Then

$$p'(0; h) = \min_{x \in X(0)} \nabla_v f(v, x) \cdot h.$$

The result can be attributed to Danskin [5], although the form in which we have stated it is somewhat different. If B is convex, the condition $x \in X(0)$ is equivalent, of course, to $-\nabla_x f(v, x) \in N_B(x)$ where $N_B(x)$ is the normal cone to B in the sense of convex analysis ([16]).

An example where this applies is

$$f(v, x) = \sum_{j=1}^{n} x_j g_j(v), \quad A(v) \equiv B = \Big\{x = (x_1, \ldots, x_n) \mid x_j \geqslant 0, \sum_{j=1}^{n} x_j = 1\Big\},$$

with $g_j \in \mathscr{C}^1$. Then $p(v) = \min\{g_1(v), \ldots, g_n(v)\}$.

The case where f is a convex function and $\operatorname{gph} A$ is a convex set has also received attention. Then p is a convex function, so the derivatives $p'(0; h)$ do exist. It has been shown by Golshtein [10] (see also Hogan [13]) that for any choice of $x \in X(0)$ one has

$$p'(v; h) = \inf_{k \in A'(v, x; h)} f'(v, x; h, k),$$

where $\operatorname{gph} A'$ is the tangent cone to $\operatorname{gph} A$ at (v, x). In terms of the subgradients of convex analysis ([16]), the equivalent formula is

$$\partial p(v) = \{z \in R^d \mid (z, 0) \in \partial f(0, x) + N_{\operatorname{gph} A}(0, x)\}. \tag{4}$$

Generalizations of (3) to nonconvex cases have been given by Dem'janov *et al.* [6], [7], under rather stringent assumptions. Other results along these lines are those of Hiriart-Urruty [11], the marginal value theorem of Golshtein [10] for nonconvexly parameterized convex programming, and certain extensions of the latter by Rockafellar [17], [21, Theorem 4].

3. More general results of the kind just mentioned involve additional structure for the constraint set $A(v)$. In escaping from assumptions of either classical differentiability or convexity, such results also rely on new developments in subgradient analysis.

Suppose henceforth that

$$A(v) = \{x \in R^n \mid F(v, x) \in C,\ (v, x) \in D\},$$

where $C \subset R^m$ and $D \subset R^d \times R^n$ are closed sets and $F: R^d \times R^n \to R^m$ is locally Lipschitz continuous. A typical case in mathematical programming is

$$C = \{(u_1, \dots, u_m) \mid u_i \leqslant 0 \text{ for } i = 1, \dots, s,$$
$$u_i = 0 \text{ for } i = s+1, \dots, m\}. \qquad (6)$$

For a locally Lipschitz continuous function $g: R^n \to R$, Clarke ([2]) introduced the directional derivatives

$$g^0(x; k) = \limsup_{\substack{x' \to x \\ t \downarrow 0}} [g(x' + tk) - g(x')]/t$$

and showed there was a unique, nonempty, compact convex set $\partial g(x)$ (whose elements may be called "subgradients") such that

$$g^0(x; k) = \max_{w \in \partial g(x)} k \cdot w.$$

A detailed calculus has grown out of this concept; see Clarke [2], [3], [4], Hiriart-Urruty [11], and Rockafellar [15], [18], [19], [20], [21] in particular. It is known that $g^0(x; k) = g'(x; k)$ when $g \in \mathscr{C}^1$ or g is convex; in the first case $\partial g(x)$ reduces to the gradient $\nabla g(x)$, while in the second case it is the usual subgradient set of convex analysis.

Corresponding geometrically to Clarke's notion of "subgradient" is his definition of the normal cone $N_B(x)$ to an arbitrary closed set $B \subset R^n$ at any point $x \in B$; see [2], [15].

These concepts have been used by Clarke [1] to derive optimality conditions for mathematical programming problems with objective and constraint functions that are locally Lipschitz continuous, and Clarke's result has been sharpened by Hiriart-Urruty [12] and Rockafellar [21]. As background for the marginal value theorem that will be stated below, we first formulate a version of this result for the more general constraint structure in (5). Let

$$K(x) = \{(y, z) \in R^m \times R^d \mid y \in N_C(F(0, x)),$$
$$(z, 0) \in \partial(f + y \cdot F)(x) + N_D(0, x)\},$$

$$K_0(x) = \{(y, z) \in R^m \times R^d \mid y \in N_C(F(0, x)), (z, 0) \in \partial(y \cdot F)(x) + N_D(0, x)\}.$$

THEOREM 1 (*Multiplier Rule*). *Suppose* $x \in X(0)$ *is such that* $K_0(x)$ *contains just* $(0, 0)$. *Then there is a pair* $(y, z) \in K(x)$, *in fact* $K(x)$ *is a nonempty compact set.*

The constraint qualification $K_0(x) = \{(0, 0)\}$ reduces in the case of

$$f \in \mathscr{C}^1, \quad F \in \mathscr{C}^1, \quad D = R^d \times R^n, \ C \text{ as in } (6),$$

to the well-known one of Mangasarian and Fromovitz [14].

Theorem 1 may be derived from Theorem 1 of Rockafellar [21] by applying the latter to the constraints

$$0 = G(v, x, w) = F(v, x) - w, \quad (v, x, w) \in D \times C.$$

By the same route one obtains the following as a special case of Theorem 2 of Rockafellar [21].

THEOREM 2. *Suppose that for every* $x \in X(0)$, *the set* $K_0(x)$ *contains just* $(0, 0)$. *Then* p *is Lipschitz continuous in a neighborhood of* 0 *and*

$$\partial p(0) \subset \operatorname{co} \bigcup_{x \in X(0)} \{z \mid \exists y, (y, z) \in K(x)\},$$

$$p^0(0; h) \leqslant \max_{\substack{x \in X(0) \\ (y,z) \in K(x)}} z \cdot h.$$

In the case of assumption (7), this result was proved by Gauvin ([8], [9]).

References

[1] Clarke F. H. A New Approach to Lagrange Multipliers, *Math. Oper.Res.* **1** (1976), pp. 165–174.

[2] Clarke F. H., Generalized Gradients and Applications, *Trans. Amer. Math. Soc.* **205** (1975), pp. 247–262.

[3] Clarke F. H., Generalized Gradients of Lipschitz Functionals, *Adv. in Math.* **40** (1981), pp. 52–67.

[4] Clarke F. H., *Optimization and Nonsmooth Analysis*, Wiley, New York 1983.

[5] Danskin J. M., *The Theory Max-Min and Its Applications to Weapons Allocations Problems*, Springer-Verlag, New York 1967.

[6] Dem'janov V. F. and Malozemov V. N., The Theory of Nonlinear Minimal Problems, *Russian Math. Surveys* **26** (1971), pp. 55–115.

[7] Dem'janov V. F. and Pevnyi A. B., First and Second Marginal Values of Mathematical Programming Problems, *Soviet Math. Dokl.* **13** (1972), pp. 1502–1506.

[8] Gauvin J., The Generalized Gradients of a Marginal Function in Mathematical Programming, *Math. Oper. Res.* **4** (1979), pp. 458–463.

[9] Gauvin J. and Dubeau F., Differential Properties of the Marginal Value Function in Mathematical Programming, *Math. Programming Stud.*, to appear.

[10] Golshtein E. G., *Theory of Convex Programming*, Trans. Math. Monographs 26, Amer. Math. Soc., Providence, RI, 1972.

[11] Hiriart-Urruty J.-B., Gradients-généralises de fonctions marginales, *SIAM J. Control and Optim.* **16** (1978), pp. 301–316.

[12] Hiriart-Urruty J.-B., Refinements of Necessary Optimality Conditions in Nondifferentiable Programming, I, *Appl. Math. Optim.* **5** (1979), pp. 63–82.

[13] Hogan W. W., Directional Derivatives for Extremal-Value Functions with Applications to the Completely Convex Case, *Oper. Res.* **21** (1973), pp. 188–209.

[14] Mangasarian O. L. and Fromovitz S., The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints, *J. Math. Anal. Appl.* **17** (1967), pp. 37–47.

[15] Rockafellar R. T., Clarke's Tangent Cones and the Boundaries of Closed Sets in R^n, *Nonlinear Anal.* **3** (1979), pp. 145–154.

[16] Rockafellar R. T., *Convex Analysis*, Princeton University Press, Princeton, N. J., 1970.

[17] Rockafellar R. T., Directional Differentiability of the Optimal Value Function in a Nonlinear Programming Problem, *Math. Programming Stud.*, **21** (1984).

[18] Rockafellar R. T., Directionally Lipschitzian Functions and Subdifferential Calculus, *Proc. London Math. Soc.* (3) **39** (1979), pp. 331–355.

[19] Rockafellar R. T., Generalized Directional Derivatives and Gradients of Nonconvex Functions, *Canad. J. Math.* **32** (1980), pp. 257–280.

[20] Rockafellar R. T., *La theorie des sous-gradients et ses applications: Fonctions convexes et non convexes*, Collection Chaire Aisenstadt Presses de l'Université de Montréal, Montréal, Quebec, 1979. English version: *The Theory of Subgradients and Its Applications: Convex and Nonconvex Functions*, Helderman Verlag, W. Berlin, 1981.

[21] Rockafellar R. T., Lagrange Multipliers and Subderivatives of Optimal Value Functions in Nonlinear Programming, *Math. Programming Stud.* **17** (1982), pp. 28–66.