Mathematical Programming Study 21 (1984) 213–226
North-Holland

DIRECTIONAL DIFFERENTIABILITY OF THE OPTIMAL VALUE FUNCTION IN A NONLINEAR PROGRAMMING PROBLEM

R.T. ROCKAFELLAR

Department of Mathematics, University of Washington, Seattle, WA 98195, USA

Received 14 December 1982
Revised manuscript received 22 July 1983

A parameterized nonlinear programming problem is considered in which the objective and constraint functions are twice continuously differentiable. Under the assumption that certain multiplier vectors appearing in generalized second-order necessary conditions for local optimality actually satisfy the weak sufficient condition for local optimality based on the augmented Lagrangian, it is shown that the optimal value in the problem, as a function of the parameters, is directionally differentiable. The directional derivatives are expressed by a minimax formula which generalizes the one of Gol'shtein in convex programming.

Key words: Marginal Values, Parametric Optimization, Second-Order Optimality Conditions.

1. Introduction

For parameter vectors $v=(v_1,\ldots,v_d)$ ranging over an open set $V\subset\mathbb{R}^d$, we consider the problem

$$\text{minimize } f_0(v,x) \quad \text{over all } x\in\mathbb{R}^N \text{ such that}$$

$$f_i(v,x)\begin{cases}\leqslant 0 & \text{for } i=1,\ldots,s,\\ =0 & \text{for } i=s+1,\ldots,m,\end{cases} \tag{P_v}$$

where f_i is for $i=0,1,\ldots,m$ a function of class $\mathscr{C}^2$ on $V\times\mathbb{R}^n$. The infimum in this problem is denoted by $p(v)$ and the set of optimal solutions by $X(v)$:

$$p(v)=\inf(P_v),\qquad X(v)=\operatorname{argmin}(P_v).$$

Our aim is to establish a formula for the directional derivatives of p at v in terms of Lagrange multiplier vectors associated with the elements x of $X(v)$.

To simplify matters and concentrate on the main issues, we make the following *inf-boundedness assumption*: for each $\bar{v}\in V$ and $\alpha\in\mathbb{R}$, there is an $\varepsilon>0$ such that the set of $(v,x)\in V\times\mathbb{R}^n$ satisfying

$$\begin{aligned}&|v-\bar{v}|\leqslant\varepsilon,\qquad f_0(v,x)\leqslant\alpha,\\ &f_i(v,x)\leqslant\varepsilon \text{ for } \quad i=1,\ldots,s,\qquad |f_i(v,x)|\leqslant\varepsilon \quad\text{for } i=s+1,\ldots,m,\end{aligned} \tag{1.1}$$

Research sponsored in part by the Air Force Office of Scientific Research, AFSC, United States Air Force, under grant no. F49620-82-K-0012.

is bounded. This assumption will be useful in several ways, but it ensures in particular that p is a lower semicontinuous function on V which nowhere takes on $-\infty$ (although it may take on $+\infty$, namely for each v such that (P_v) has no feasible solutions), and that $X(v)$ is a nonempty compact set for each v such that $p(v) < \infty$.

The function p is said to have one-sided directional derivatives in the *ordinary sense* if the limits

$$p'(v; k) = \lim_{t \downarrow 0} \frac{p(v+tk) - p(v)}{t} \tag{1.2}$$

exist, and in the *Hadamard sense* if these limits can in fact be taken as

$$\lim_{\substack{k' \to k \\ t \downarrow 0}} \frac{p(v+tk') - p(v)}{t}. \tag{1.3}$$

In the case where (P_v) is a convex programming problem and the Mangasarian–Fromovitz constraint qualification is satisfied, a minimax formula of Gol'shtein [5] gives derivatives $p'(v; k)$ in the ordinary sense. Gauvin and Dubeau [3] have complemented this in the nonconvex case by providing a formula for $p'(v; k)$ when there is a unique first-order multiplier vector associated with each $x \in X(v)$, and by upper and lower bounds for $p'(v; k)$ more generally. Rockafellar [16] has shown that these results remain true under weaker constraint qualifications, and that they actually yield derivatives in the Hadamard sense.

So far, no one has demonstrated the existence of $p'(v; k)$ in the nonconvex case where Lagrange multiplier vectors are not necessarily unique, nor have second-order conditions been used in connection with such vectors in order to strengthen the known formulas. Both tasks will be undertaken here.

To get around nonconvexity, we shall take advantage of the fact that a saddle point expression of optimality is possible without convexity assumptions, if the ordinary Lagrangian for (P_v) is replaced by the (quadratic-type) augmented Lagrangian [14]. For the second-order refinements, we shall make use of new multiplier sets which we have introduced in [17]. These differ from previously defined second-order multiplier sets in being upper semicontinuous in their dependence on x and v, as well as having other valuable properties. For simplicity in this introduction, however, we postpone discussion of such sets until Section 2 and focus on the version of our main result which can be stated in terms of more familiar multiplier conditions.

The *ordinary* Lagrangian for problem (P_v) is

$$l(v, x, y) = f_0(v, x) + \sum_{i=1}^{m} y_i f_i(v, x) \quad \text{for } y = (y_1, \ldots, y_m). \tag{1.4}$$

The first-order multiplier set $Y^1(v, x)$ associated with any $v \in V$ and feasible solution x to (P_v) consists of all the vectors $y \in \mathbb{R}^m$ (if any) such that

$$\nabla_x l(v, x, y) = 0, \tag{1.5}$$

$$y_i \geq 0 \quad \text{for } i = 1, \ldots, s, \quad \text{with } y_i = 0 \text{ if } i \notin I(v, x), \tag{1.6}$$

where

$$I(v, x) = \{i \in [1, m] \mid f_i(v, x) = 0\} \quad \text{(active set of indices).} \tag{1.7}$$

The *singular* first-order multiplier set $Y_0^1(v, x)$ is defined in the same way, except that l is replaced by the *singular* Lagrangian

$$l_0(v, x, y) = \sum_{i=1}^{m} y_i f_i(v, x). \tag{1.8}$$

Both $Y^1(v, x)$ and $Y_0^1(v, x)$ are polyhedral convex sets, of course, and $Y_0^1(v, x)$ is a cone containing 0. (For technical purposes, the two sets are defined to be empty if x is not feasible.)

The condition $Y_0^1(v, x) = \{0\}$ is a constraint qualification equivalent by duality to the one of Mangasarian and Fromovitz [9]. For a locally optimal solution x to (P_v), it holds if and only if $Y^1(v, x)$ is nonempty and compact, as is well known [2]. Thus in particular, a *necessary condition* for the local optimality of x in (P_v), if the constraint qualification in question is satisfied at x, is the existence of a vector $y \in Y^1(v, x)$. Second-order multiplier sets with parallel properties, except for convexity, will be described in Section 2.

The *augmented* Lagrangian for (P_v) involves a penalty parameter $r > 0$ and is expressed by

$$L(v, x, y, r) = l(v, x, y) + \frac{r}{2} \sum_{i=s+1}^{m} [f_i(v, x)]^2 + \frac{1}{2r} \sum_{i=1}^{s} ([y_i + rf_i(v, x)]_+^2 - [y_i + rf_i(v, x)]^2), \tag{1.9}$$

where

$$[a]_+ = \max\{a, 0\} \quad \text{for } a \in \mathbb{R}. \tag{1.10}$$

(This formula is slightly different in appearance from the one introduced in [13, 14], but amounts to the same thing.) A *locally augmentable* multiplier vector for (P_v) at x is a vector $y \in \mathbb{R}^m$ such that, for r sufficiently large, (x, y) is a local saddle point of $L(v, \cdot, \cdot, r)$ at (x, y). The set of all these will be denoted by $Y^a(v, x)$; one has $y \in Y^a(v, x)$ if and only if $y \in Y^1(v, x)$ and, for r sufficiently large, the function

$$l(v, \cdot, y) + \frac{r}{2}\left(\sum_{i \in I_0(v,x,y)} [f_i(v, \cdot)]_+^2 + \sum_{i \in I_1(v,x,y)} [f_i(v, \cdot)]^2 \right) \tag{1.11}$$

has a local minimum at x, where

$$\begin{aligned} I_0(v, x, y) &= \{i \in I(v, x) \mid 1 \le i \le s, f_i(v, x) = 0, y_i = 0\}, \\ I_1(v, x, y) &= I(v, x) \setminus I_0(v, x, y). \end{aligned} \tag{1.12}$$

Thus $Y^a(v, x)$ is a convex subset of $Y^1(v, x)$, but $Y^a(v, x)$ is not necessarily closed. A *sufficient condition* for the local optimality of x in (P_v) is the existence of a vector $y \in Y^a(v, x)$ [14].

When (P_v) is a convex programming problem, i.e. the function $f_i(v, \cdot)$ is convex on $\mathbb{R}^n$ for $i = 0, 1, \ldots, s$ and affine (linear + constant) for $i = s+1, \ldots, m$, one simply has $Y^a(v, x) = Y^1(v, x)$, and this set is the same for every $x \in X(v)$, coinciding namely with the set $Y(v)$ optimal solutions y to the ordinary *dual* of (P_v) [13]. For the nonconvex case, it will help in clarifying the nature of the results below if we recall briefly the relationship between $Y^a(v, x)$ and the most familiar kind of second-order optimality conditions. Let us associate with any feasible solution x to (P_v) the polyhedral convex cone

$$W(v, x) = \{w \in \mathbb{R}^n \mid \nabla_x f_i(v, x) \cdot w \leq 0 \text{ for } i = 0 \text{ and } i \in I(v, x) \cap [1, s],$$
$$\nabla_x f_i(v, x) \cdot w = 0 \text{ for } i = s+1, \ldots, m\}, \tag{1.13}$$

which for any $y \in Y^1(v, x)$ can also be expressed equivalently as

$$W(v, x) = \{w \in \mathbb{R}^n \mid \nabla_x f_i(v, x) \cdot w \leq 0 \text{ for } i = 0 \text{ and } i \in I_0(v, x, y),$$
$$\nabla_x f_i(v, x) \cdot w = 0 \text{ for } i \in I_1(v, x, y)\}. \tag{1.14}$$

Let

$$Y^2(v, x) = \{y \in Y^1(v, x) \mid w \cdot \nabla^2_{xx} l(v, x, y) w \geq 0 \text{ for all } w \in W(v, x)\}, \tag{1.15}$$

$$Y^2_0(v, x) = \{y \in Y^1_0(v, x) \mid w \cdot \nabla^2_{xx} l_0(v, x, y) w \geq 0 \text{ for all } w \in W(v, x)\}, \tag{1.16}$$

$$Y^2_+(v, x) = \{y \in Y^1(v, x) \mid w \cdot \nabla^2_{xx} l(v, x, y) w > 0 \text{ for all nonzero } w \in W(v, x)\}. \tag{1.17}$$

Obviously the sets $Y^2(v, x)$ and $Y^2_0(v, x)$ are closed, and $Y^2_0(v, x)$ is a cone containing 0. The condition $Y^2(v, x) \neq \emptyset$ is *necessary* for the local optimality of x in (P_v) under certain constraint qualifications (cf. [6, 7, 10]), but the assumption that $Y^2_0(v, x) = \{0\}$ is not adequate for this conclusion, in contrast to the situation for $Y^1(v, x)$ and $Y^1_0(v, x)$. (More will be said on this matter in Section 2.) The condition on $Y^2_+(v, x) \neq \emptyset$ is always *sufficient* for the local optimality of x in (P_v), but it goes beyond the condition $Y^a(v, x) \neq \emptyset$ by implying also that x is *strict*, in the sense that there is a neighborhood of x containing no other feasible solution x' to (P_v) with $f_0(v, x') \leq f_0(v, x)$ (cf. [6, 7, 10]). One has

$$Y^2_+(v, x) \subset Y^a(v, x) \subset Y^2(v, x), \quad \text{with cl } Y^2_+(v, x) = \text{cl } Y^a(v, x)$$
$$= Y^2(v, x) \text{ if } Y^2_+(v, x) \neq \emptyset \tag{1.18}$$

(cf. [14, Theorem 6] and [17, Remark following Proposition 5]).

In the result we now state, we denote by ri C the *relative interior* of a convex set C [12, Section 6].

Theorem 1. *Suppose v is a vector such that (P_v) has feasible solutions, and every optimal solution $x \in X(v)$ satisfies the constraint qualification $Y^1_0(v, x) = \{0\}$ and has*

$$\text{ri } Y^1(v, x) \subset Y^a(v, x). \tag{1.19}$$

Then p possesses finite one-sided directional derivatives at v in the Hadamard sense, and in fact for every $k \in \mathbb{R}^d$

$$p'(v; k) = \min_{x \in X(v)} \max_{y \in Y^1(v,x)} \nabla_v l(v, x, y) \cdot k. \tag{1.20}$$

The set ri $Y^1(v, x)$ in assumption (1.10) can be given a more direct description. Let

$$I^*(v, x) = \{i \in I(v, x) \mid \nabla_x f_i(v, x) \cdot w = 0 \text{ for all } w \in W(v, x)\}. \tag{1.21}$$

Then

$$I^*(v, x) = \{i \in [1, s] \mid \exists y \in Y^1(v, x) \text{ with } y_i > 0\} \cup \{s+1, \ldots, m\}, \tag{1.22}$$

and

$$\text{ri } Y^1(v, x) = \{y \in Y^1(v, x) \mid y_i > 0 \text{ for all } i \in I^*(v, x) \cap [1, s]\}. \tag{1.23}$$

Note that the index sets (1.12) in the definition of $Y^a(v, x)$ reduce to

$$I_1(v, x, y) = I^*(v, x) \text{ and } I_0(v, x, y) = I(v, x) \backslash I^*(v, x) \text{ for all } y \in \text{ri } Y^1(v, x). \tag{1.24}$$

Assumption (1.19), together with the constraint qualification, implies of course by (1.18) that $Y^2(v, x) = Y^1(v, x)$.

Corollary 1. *Let v be a vector such that* (P_v) *has feasible solutions, and every optimal solution* $x \in X(v)$ *satisfies the constraint qualification* $Y_0^1(v, x) = \{0\}$ *and has*

$$\text{ri } Y^1(v, x) \subset Y_+^2(v, x). \tag{1.25}$$

Then the conclusions of Theorem 1 *are valid, moreover with* $X(v)$ *consisting of only finitely many points.*

Corollary 1 is obtained by invoking (1.18) in condition (1.19). The reason $X(v)$ must be finite in this case is that any $x \in X(v)$ having $Y_+^2(v, x) \neq \emptyset$ must be an isolated point of $X(v)$, in accordance with the remarks above. Since $X(v)$ is compact under our inf-boundedness assumption, there can be only finitely many such points.

Corollary 2. *Let v be a vector such that* (P_v) *has feasible solutions and is a convex programming problem. Suppose the constraint qualification* $Y_0^1(v, x) = \{0\}$ *is satisfied for some* $x \in X(v)$. *Then the conclusions of Theorem* 1 *hold, moreover with* $X(v)$ *a convex set* (*hence consisting of infinitely many points, unless it is a singleton*), *and* $Y^1(v, x)$ *the same for every* $x \in X(v)$.

This follows from the fact, noted earlier, that in the convex case, $Y^1(v, x)$ and $Y^a(v, x)$ coincide for each $x \in X(v)$ with the set $Y(v)$ of optimal solutions to the ordinary dual to (P_v). Note, incidentally, that the directional derivative formula (1.19) then reduces to a minimax in (x, y) relative to a product of compact convex sets $X(v)$ and $Y(v)$. The expression $\nabla_v l(v, x, y) \cdot k$ turns out to be affine as a function of $x \in X(v)$ for each $y \in Y(v)$ as well as affine as a function of $y \in Y(v)$ for each $x \in X(v)$, so that a saddle point even exists.

Corollary 2 is closely related to the marginal value theorem of Gol'shtein [5, Section 7] (see also Hogan [8]) for convex programming problems. Gol'shtein's result is in some respects more general: the functions f_i do not have to be differentiable in x, and an abstract constraint $x \in C$ can be present. On the other hand, Corollary 2 brings the conclusion that the derivatives exist not just in the ordinary sense but the Hadamard sense. Furthermore, it imposes convexity only on (P_v) for the v in question, not necessarily for neighboring parameter values. For an extension of Gol'shtein's theorem to nonsmooth convex (or linear) programming problems with primal and dual optimal solution sets not necessarily bounded, see [17, Theorem 4].

2. Second-order results

The first-order multiplier multifunction $Y^1 : (v, x) \to Y^1(v, x)$ has *closed graph*:

$$\text{if } y^j \in Y^1(v^j, x^j) \text{ and } (v^j, x^j, y^j) \to (v, x, y),\ v \in V, \quad \text{then } y \in Y^1(v, x).$$

The same holds for Y_0^1. Furthermore, at any (v, x) where the constraint qualification $Y_0^1(v, x) = \{0\}$ is satisfied, Y^1 is not only nonempty-compact-valued, as already noted, but also *locally bounded* (cf. [3]). Local boundedness of Y^1 at (v, x) means that there is a bounded subset of $\mathbb{R}^m$ which, for every (v', x') in some neighborhood of (v, x), includes $Y^1(v', x')$. No such properties hold, however, for the second-order multifunctions Y^2 and Y_0^2, or for Y^a.

We proceed now to describe second-order multiplier multifunctions which do have such properties. In terms of these we will be able to derive new upper bounds for Hadamard directional derivatives and eventually prove a tighter result than the one stated as Theorem 1. The modified second-order conditions are based on developments in [17].

We shall make use of the concept of a sequence of subspaces M^j of $\mathbb{R}^n$ *converging* to a subspace M, in the sense that

$$\lim_{j\to\infty} \operatorname{dist}(M^j, z) = \operatorname{dist}(M, z) \quad \text{for all } z \in \mathbb{R}^n,$$

where 'dist' denotes Euclidean distance. (See [18, 19] for results on this kind of convergence.) For each feasible solution x to (P_v), let $\mathcal{M}(v, x)$ denote the collection of all subspaces M expressible in this way as limits of subspaces M^j of the following form:

$$M^j = \{w \in \mathbb{R}^n \mid \nabla_x f_i(v^j, x^j) \cdot w = 0 \text{ for all } i \in I(v, x)\}, \quad \text{where } (v^j, x^j) \to (v, x). \tag{2.1}$$

Every $M \in \mathcal{M}(v, x)$ has dimension at least $n - |I(v, x)|$, where $|I(v, x)|$ is the number of indices in $I(v, x)$. Furthermore, every $M \in \mathcal{M}(v, x)$ is included in the subspace

$$Z(v, x) = \{w \in \mathbb{R}^n \mid \nabla_x f_i(v, x) \cdot w = 0 \text{ for all } i \in I(v, x)\}, \tag{2.2}$$

and

$$Z(v, x) \subset W(v, x) \quad \text{if } Y^1(v, x) \neq \emptyset. \tag{2.3}$$

The multiplier sets we shall be concerned with are

$$\tilde{Y}^2(v, x) = \{y \in Y^1(v, x) \mid \exists M \in \mathcal{M}(v, x) \text{ with } w \cdot \nabla^2_{xx} l(v, x, y) w \geq 0 \text{ for all } w \in M\}, \tag{2.4}$$

$$\tilde{Y}^2_0(v, x) = \{y \in Y^1_0(v, x) \mid \exists M \in \mathcal{M}(v, x) \text{ with } w \cdot \nabla^2_{xx} l_0(v, x, y) w \geq 0 \text{ for all } w \in M\}. \tag{2.5}$$

It is evident from (2.3) that

$$Y^2(v, x) \subset \tilde{Y}^2(v, x) \text{ and } Y^2_0(v, x) \subset \tilde{Y}^2_0(v, x) \quad \text{when } Y^2(v, x) \neq \emptyset. \tag{2.6}$$

The set $\tilde{Y}^2_0(v, x)$ contains 0 (when x is feasible for (P_v)), but neither $\tilde{Y}^2(v, x)$ nor $\tilde{Y}^2_0(v, x)$ need be convex.

Theorem 2. *The second-order multiplier sets $\tilde{Y}^2(v, x)$ and $\tilde{Y}^2_0(v, x)$ are closed, and in fact the multifunctions $\tilde{Y}^2$ and $\tilde{Y}^2_0$ are of closed graph. For a locally optimal solution x to (P_v) one has $\tilde{Y}^2(v, x)$ nonempty and compact if $\tilde{Y}^2_0(v, x) = \{0\}$.*

In particular, a necessary condition for the local optimality of x in (P_v), if the constraint qualification $\tilde{Y}^2_0(v, x) = \{0\}$ is satisfied, is the existence of a vector $y \in \tilde{Y}^2(v, x)$.

Proof. We derive this from results in [17] by way of a simple reformulation. Consider the problem

$$\text{minimize } f_0(x', x) \text{ over all } (x', x) \in V \times \mathbb{R}^n \text{ satisfying}$$

$$f_i(x', x) + u_i \begin{cases} \leq 0 & \text{for } i = 1, \ldots, s, \\ = 0 & \text{for } i = s+1, \ldots, m, \end{cases} \tag{$\bar{P}_{u,v}$}$$

$$g_j(x', x) + v_j = 0 \quad \text{for } j = 1, \ldots, d,$$

where $(u_1, \ldots, u_m) = u \in \mathbb{R}^n$, $(v_1, \ldots, v_d) = v \in \mathbb{R}^d$, and

$$g_j(x', x) = -x'_j \quad \text{for } x' = (x'_1, \ldots, x'_d) \in \mathbb{R}^d. \tag{2.7}$$

Clearly (P_v) can be identified with $(\bar{P}_{0,v})$. The inf-boundedness condition (1.1) corresponds to a similar blanket condition in [17], specialized to $(\bar{P}_{0,v})$. The theory in [17] associates with each feasible solution (x', x) to $(\bar{P}_{u,v})$ (with $x' = v$) a pair of multiplier sets $K^2(u, v, x', x)$ and $K^2_0(u, v, x', x)$ which for $u = 0$ work out to

$$\begin{aligned} K^2(0, v, x', x) &= \{(y, \nabla_v l(v, x, y)) \mid y \in \tilde{Y}^2(v, x)\}, \\ K^2_0(0, v, x', x) &= \{(y, \nabla_v l_0(v, x, y)) \mid y \in \tilde{Y}^2_0(v, x)\}. \end{aligned} \tag{2.8}$$

We need only apply [17, Proposition 6 and Theorem 7] to draw the desired conclusions. □

Of course the necessary condition in Theorem 2 has a form tailored to the parameterization in terms of v, which is the subject of attention here. This parameterization could be suppressed, and one would then have a tighter result on local optimality for a fixed problem. See [17, Section 5], where a much weaker constraint qualification is developed for that purpose and connections with second-order conditions involving the set $Y^2(v, x)$ in Section 1 are explored. Of course $\tilde{Y}^2(v, x)$ coincides with $Y^2(v, x)$ and $Y^1(v, x)$ when (P_v) is a convex programming problem, and this is also true (by [17, Theorem 9] as applied to the reformulated problem $(\bar{P}_{0,v})$ above) when the following condition is satisfied: x is a locally optimal solution to (P_v) such that for each vector w satisfying $|w| = 1$ and (for the index set $I^*(v, x)$ in (1.21)–(1.24))

$$\nabla_x f_i(v, x) \cdot w \begin{cases} <0 & \text{for } i \in I(v, x) \backslash I^*(v, x), \\ =0 & \text{for } i \in I^*(v, x), \end{cases}$$

there is a sequence $x^j \to x$ with $f_i(v, x^j) = 0$ for all $i \in I^*(v, x)$ and $(x^j - x)/|x^j - x| \to w$. A simple example in [17, Section 5] shows, on the other hand, that $\tilde{Y}^2(v, x)$ can be much smaller than $Y^1(v, x)$ and provide much sharper information about properties of the optimal value $p(v)$.

The following theorem, which will be proved in Section 3, can be seen as a sharpened form of Theorem 1.

Theorem 3. *Suppose v is a vector such that (P_v) has feasible solutions, and every optimal solution $x \in X(v)$ satisfies the constraint qualifications $\tilde{Y}^2_0(v, x) = \{0\}$ and has*

$$\tilde{Y}^2(v, x) \subset \text{cl } Y^a(v, x). \tag{2.9}$$

Then p possesses finite one-sided directional derivatives at v in the Hadamard sense, and in fact for every $k \in \mathbb{R}^d$

$$p'(v; x) = \min_{x \in X(v)} \max_{y \in \tilde{Y}^2(v,x)} \nabla_v l(v, x, y) \cdot k. \tag{2.10}$$

This implies Theorem 1, because $\tilde{Y}^2(v, x) \subset Y^1(v, x)$ and $\tilde{Y}^2_0(v, x) \subset Y^1_0(v, x)$. Assumption (1.19) is more restrictive than (2.9), because it implies

$$\text{cl } Y^a(v, x) \supset \text{cl}[\text{ri } Y^1(v, x)] = Y^1(v, x).$$

(For a closed convex set C, one always has cl[ri C] = C, cf. [12, Section 6].)

Corresponding to Corollary 1 of Theorem 1, we have the following.

Corollary. *Let v be a vector such that (P_v) has feasible solutions, and every optimal solution $x \in X(v)$ satisfies the constraint qualification $\tilde{Y}^2_0(v, x) = \{0\}$ and has*

$$\tilde{Y}^2(v, x) \subset \text{cl } Y^2_+(v, x). \tag{2.11}$$

Then the conclusions of Theorem 3 are valid, moreover with $X(v)$ consisting of only finitely many points.

We shall also prove in Section 3 a complementary result.

Theorem 4. *Let v be a vector such that (P_v) has feasible solutions. Suppose that for each $x \in X(v)$ the constraint qualification $\tilde{Y}_0^2(v, x) = \{0\}$ is satisfied, and the set $\tilde{Y}^2(v, x)$ actually consists of a single vector $y(x)$. Then p has finite one-sided derivatives at v in the Hadamard sense, and in fact for every $k \in \mathbb{R}^d$*

$$p'(v; k) = \min_{x \in X(v)} \nabla_v l(v, x, y(x)) \cdot k. \tag{2.12}$$

Theorem 4 tightens a series of results of Gauvin and Tolle [4], Gauvin [2], and Gauvin and Dubeau [3]. These deal with the case where $y(x)$ is not just the unique element of $Y^2(v, x)$, but of $Y^1(v, x)$.

Corollary. *Let v be a vector such that (P_v) has a unique optimal solution x and a unique multiplier vector $y \in \tilde{Y}^2(v, x)$. If the constraint qualification $\tilde{Y}_0^2(v, x) = \{0\}$ is satisfied, then p is differentiable at v with gradient*

$$\nabla p(v) = \nabla_v l(v, x, y). \tag{2.13}$$

The corollary is the case of Theorem 2 where $X(v)$ is a singleton. It may be compared with the classical result that if the 'strong' second-order optimality conditions for (P_v) are satisfied at x, then (2.13) holds, provided the optimal value function is redefined in terms of local optima relative to a certain neighborhood of x (cf. Robinson [11]). The 'strong' optimality conditions require the existence of some $y \in Y_+^2(v, x)$ which has $y_i > 0$ for every active inequality constraint; in addition the vectors $\nabla_x f_i(v, x)$ for $i \in I(v, x)$ must be linearly independent. Under these circumstances x is an isolated point of $X(v)$, and y is the unique element not only of $\tilde{Y}^2(v, x)$, but of $Y^1(v, x)$. A stronger conclusion can then be drawn: the (redefined) function p is of class C^2 in a neighborhood of v.

3. Upper and lower estimates

The formulas in Theorems 3 and 4 will be derived from more general estimates for the *upper* and *lower Hadamard derivative*.

$$p^+(v; k) = \limsup_{\substack{k' \to k \\ t \downarrow 0}} \frac{p(v + tk') - p(v)}{t}, \tag{3.1}$$

$$p_+(v; k) = \liminf_{\substack{k' \to k \\ t \downarrow 0}} \frac{p(v + tk') - p(v)}{t}, \tag{3.2}$$

as well as, in cases where p is Lipschitz continuous, the *Clarke derivatives* ([1, 15])

$$p^\circ(v;k)=\limsup_{\substack{v'\to v\\ k'\to k\\ t\downarrow 0}}\frac{p(v'+tk')-p(v')}{t}. \tag{3.3}$$

Lipschitz continuity of p in a neighborhood of v means that for some constant $\alpha>0$ one has

$$|p(v'')-p(v')|\le\alpha|v''-v'|$$

for all v' and v'' in the neighborhood in question. Then the limit (3.3) is bounded above by $\alpha|k|$, and the limit is unaffected if the condition $k'\to k$ is dropped and one simply takes $k'=k$ in the difference quotient.

Theorem 5. *Let v be a vector such that (P_v) has feasible solutions, and every optimal solution $x\in X(v)$ satisfies the constraint qualification $\tilde Y_0^2(v,x)=\{0\}$. Then p is finite and Lipschitz continuous on a neighborhood of v, and for every $k\in\mathbb{R}^d$*

$$p^\circ(v;k)\le\max_{x\in X(v)}\max_{y\in\tilde Y^2(v,x)}\nabla_v l(v,x,y)\cdot k. \tag{3.4}$$

Proof. This is just a matter of applying a corresponding result [17, Corollary to Theorem 6] to the reformulated problem $(\bar P_{u,v})$ introduced in the proof of Theorem 1. Let the optimal value in $(\bar P_{u,v})$ be $\bar p(u,v)$, and the optimal solution set be $\bar X(u,v)$. The cited result says that if every optimal solution (x',x) to $(\bar P_{0,v})$ satisfies $K_0^2(0,v,x',x)=\{(0,0)\}$, then $\bar p$ is Lipschitz continuous on a neighborhood of $(0,v)$ and

$$\bar p^\circ(u,v;h,k)\le\max_{(x',x)\in\bar X(0,v)}\max_{(y,y')\in K^2(0,v,x,x')}[y\cdot h+y'\cdot k]. \tag{3.5}$$

Here $K^2(0,v,x',x)$ and $K_0^2(0,v,x',x)$ are given by (2.8), and

$$(x',x)\in\bar X(0,v)\quad\Leftrightarrow\quad x'=x\quad\text{and}\quad x\in X(v). \tag{3.6}$$

The condition $K_0^2(0,v,x',x)=\{(0,0)\}$ is equivalent, for $(x',x)\in\bar X(0,v)$, to our constraint qualification $\tilde Y_0^2(v,x)=\{0\}$, so we see that $\bar p$ is indeed Lipschitzian around $(0,v)$, (3.5) does hold, and

$$\bar p^\circ(0,v;h,k)\le\max_{v\in X(v)}\max_{y\in\tilde Y^2(v,x)}[y\cdot h+\nabla_v l(v,x,y)\cdot k]. \tag{3.7}$$

Since $\bar p(0,\cdot)=p$, we have p Lipschitz continuous around v and

$$\bar p^\circ(0,v;0,k)=\limsup_{\substack{(u',v')\to(0,v)\\ (h',k')\to(0,k)\\ t\downarrow 0}}\frac{\bar p(u'+th',v'+tk')-p(u',v')}{t}$$

$$\ge\limsup_{\substack{(0,v')\to(0,v)\\ (0,k')\to(0,k)\\ t\downarrow 0}}\frac{\bar p(0,v'+tk')-\bar p(0,v')}{t}=p^\circ(v;k).$$

This, combined with (3.7), yields (3.4). □

The preceding result generalizes one of Gauvin and Dubeau [3] for first-order multiplier conditions. An extension to nonsmooth programming problems may be found in [16].

Theorem 6. *Let v be a vector such that (P_v) has feasible solutions, and every optimal solution $x \in X(v)$ satisfies the constraint qualification $\tilde{Y}_0^2(v, x) = \{0\}$. Then for every $k \in \mathbb{R}^d$ the Hadamard semiderivatives $p^+(v; k)$ and $p_+(v; k)$ are finite, and*

$$p^+(v; k) \leqslant \inf_{x \in X(v)} \max_{y \in \tilde{Y}^2(v,x)} \nabla_v l(v, x, y) \cdot k, \tag{3.8}$$

$$p_+(v; k) \geqslant \min_{x \in X(v)} \min_{y \in \tilde{Y}^2(v,x)} \nabla_v l(v, x, y) \cdot k. \tag{3.9}$$

Proof. The finiteness of $p^+(v; k)$ and $p_+(v; k)$ stems from the Lipschitz continuity of p around v as ensured by Theorem 5. Tackling the proof of (3.9) first, we observe that a change of notation from k to $-k$ turns the task into one of verifying whether

$$-p_+(v; -k) \leqslant \max_{x \in X(v)} \max_{y \in \tilde{Y}^2(v,x)} \nabla_v l(v, x, y) \cdot k$$

holds for all $k \in \mathbb{R}^d$. But this inequality is a consequence of the one in Theorem 5, because

$$-p_+(v; -k) = \limsup_{\substack{k' \to k \\ t \downarrow 0}} \frac{p(v) - p(v - tk)}{t} \leqslant \limsup_{\substack{v' \to v \\ k' \to k \\ t \downarrow 0}} \frac{p(v' + tk') - p(v')}{t} = p^\circ(v; k).$$

As for the estimate (3.8), we shall also derive it from Theorem 5, but by a localization argument. Fix any $\hat{x} \in X(v)$ and replace f_0 in (P_v) by

$$\hat{f}_0(v, x) = f_0(v, x) + |x - \hat{x}|^3.$$

This modification does not change the multiplier sets $\tilde{Y}^2(v, \hat{x})$, $\tilde{Y}_0^2(v, \hat{x})$, or otherwise upset any of our assumptions, but it makes $\hat{x}$ into the *only* optimal solution and replaces p by another lower semicontinuous function $\hat{p} \geqslant p$ having $\hat{p}(v) = p(v)$. Applying Theorem 5 to the modified problem, we obtain for all $k \in \mathbb{R}^d$

$$\hat{p}^\circ(v; k) \leqslant \max_{y \in \tilde{Y}^2(v,\hat{x})} \nabla_v l(v, \hat{x}, y) \cdot k.$$

Since $\hat{p} \geqslant p$ but $\hat{p}(v) = p(v)$, we also have

$$\hat{p}^\circ(v; k) = \limsup_{\substack{v' \to v \\ k' \to k \\ t \downarrow 0}} \frac{\hat{p}(v' + tk') - \hat{p}(v')}{t} \geqslant \limsup_{\substack{k' \to k \\ t \downarrow 0}} \frac{\hat{p}(v + tk') - \hat{p}(v)}{t}$$

$$\geqslant \limsup_{\substack{k' \to k \\ t \downarrow 0}} \frac{p(v + tk') - p(v)}{t} = p^+(v; k),$$

and it follows that

$$p^+(v; k) \leqslant \max_{y \in \tilde{Y}^2(v,\hat{x})} \nabla_v l(v, x, y) \cdot k.$$

This being true for arbitrary $\hat{x} \in X(v)$, we obtain (3.8). □

Proof of Theorem 4. The hypothesis of Theorem 6 is satisfied with $\tilde{Y}^2(v, x) = \{y(x)\}$ (singleton), so that

$$\max_{y \in \tilde{Y}^2(v,x)} \nabla_v l(v, x, y) \cdot k = \min_{y \in \tilde{Y}^2(u,x)} \nabla_v l(v, x, y) \cdot k = \nabla_v l(v, x, y(x)) \cdot k$$

for all $k \in \mathbb{R}^d$. Therefore

$$p^+(v; k) \leq \inf_{x \in X(v)} \nabla_v l(v, x, y(x)) \cdot k \leq p_+(v; k).$$

This gives the existence of Hadamard derivatives satisfying (2.12). The infimum over $X(v)$ is actually attained, because $X(v)$ is compact and $\nabla_v l(v, x, y(x))$ is continuous relative to $x \in X(v)$. Indeed, $y(x)$ is continuous relative to $x \in X(v)$ by the closedness and local boundedness of $\tilde{Y}^2$ in Theorem 2. □

Theorem 7. *Let v be a vector such that (P_v) has feasible solutions and every $x \in X(v)$ has $Y^a(v, x) \neq \emptyset$. Then for every $k \in \mathbb{R}^d$ there is an $\bar{x} \in X(v)$ such that*

$$p_+(v; k) \geq \nabla_v l(v, \bar{x}, y) \cdot k \quad \textit{for all } y \in Y^a(v, \bar{x}). \tag{3.10}$$

Consequently

$$p_+(v; k) \geq \inf_{x \in X(v)} \sup_{y \in Y^a(v,x)} \nabla_v l(v, x, y) \cdot k. \tag{3.11}$$

Proof. Fix any $k \in \mathbb{R}^d$. We can suppose that $p_+(v; k) < \infty$, for otherwise (3.10) and (3.11) hold trivially. By definition of $p_+(v, k)$ there exist sequences $k^j \to k$ and $t_j \downarrow 0$ such that

$$\infty > [p(v + t_j k^j) - p(v)]/t_j \to p_+(v; k). \tag{3.12}$$

Then in particular

$$v + t_j k^j \to v \quad \text{and} \quad \infty > p(v + t_j k^j) \to p(v), \tag{3.13}$$

so by our inf-boundedness assumption (1.1) the set $X(v + t_j k^j)$ is nonempty for all j. For arbitrarily chosen $x^j \in X(v + t_j k^j)$ we have

$$f_0(v + t_j k^j, x^j) = p(v + t_j k^j) \to p(v),$$
$$f_i(v + t_j k^j, x^j) \begin{cases} \leq 0 & \text{for } i = 1, \ldots, s, \\ = 0 & \text{for } i = s+1, \ldots, m, \end{cases} \tag{3.14}$$

from which it follows (again by our inf-boundedness assumption) that the sequence $\{x^j\}$ is bounded. Passing to subsequences if necessary, we can reduce everything to the case where x^j converges to some $\bar{x}$. Then $\bar{x} \in X(v)$ by virtue of (3.14) and the continuity of every f_i.

In order to prove the theorem, it will suffice to show that (3.10) holds for this $\bar{x}$. In view of (3.12), this amounts to demonstrating for arbitrary $\bar{y} \in Y^a(v, \bar{x})$ that

$$\lim_{j \to \infty} [p(v + t_j k^j) - p(v)]/t_j \geq \nabla_v l(v, \bar{x}, \bar{y}) \cdot k. \tag{3.15}$$

We know from the definition of $Y^a(v, \bar{x})$ in Section 1 that the condition $\bar{y} \in Y^a(v, \bar{x})$ entails

$$\bar{y}_i \geq 0 \quad \text{for } i = 1, \ldots, s \quad \text{and} \quad \bar{y}_i f_i(v, \bar{x}) = 0 \text{ for } i = 1, \ldots, m \tag{3.16}$$

and the existence of some $r > 0$ and neighborhood U of $\bar{x}$ such that

$$L(r, v, x, \bar{y}) \geq L(r, v, \bar{x}, \bar{y}) \quad \text{for all } x \in U. \tag{3.17}$$

Since $\bar{x} \in X(v)$ and $x^j \in X(v + t_j k^j)$, we have

$$L(r, v, x, \bar{y}) = f_0(v, \bar{x}) = p(v),$$

$$L(r, v + t_j k^j, x^j, \bar{y}) \leq f_0(v + t_j k^j, x^j) = p(v + t_j k^j).$$

Therefore

$$\begin{aligned}[p(v + t_j k^j) - p(v)]/t_j &\geq [L(r, v + t_j k^j, x^j, \bar{y}) - L(r, v, x, \bar{y})]/t_j \\ &\geq [L(r, v + t_j k^j, x^j, \bar{y}) - L(r, v, x^j, \bar{y})]/t_j \end{aligned} \tag{3.18}$$

for all j large enough that x^j belongs to the neighborhood U in (3.17). But L is a continuously differentiable function of all its variables, according to its formula (1.9) (because each f_i is continuously differentiable). The mean value theorem can therefore be used to write the last difference quotient in (3.18) as

$$\nabla_v L(r, v + \theta_j k^j, x^j, \bar{y}) \cdot k^j \quad \text{for } \theta_j \in (0, t_j).$$

Therefore

$$\lim_{j \to \infty} [p(v + t_j k^j) - p(v)]/t \geq \nabla_v L(r, v, \bar{x}, \bar{y}) \cdot k, \tag{3.19}$$

where from (1.9) and (3.16) one calculates

$$\nabla_v L(r, v, \bar{x}, \bar{y}) = \nabla_v l(v, \bar{x}, \bar{y}).$$

In view of (3.12), inequality (3.19) is now seen to be the same as the desired inequality (3.11). □

Proof of Theorem 3. Since $\tilde{Y}_0^2(v, x) = \{0\}$, we have $\tilde{Y}^2(v, x) \neq \emptyset$ by Theorem 2. Assumption (2.9) then implies $Y^a(v, x) \neq \emptyset$. Theorems 6 and 7 are both applicable. The desired formula (2.10) is the combination of (3.8) and (3.10). □

References

[1] F.H. Clarke, "Generalized gradients and applications", *Transactions of the American Mathematical Society* 205 (1975) 247–262.

[2] J. Gauvin, "The generalized gradient of a marginal value function in mathematical programming", *Mathematics of Operations Research* 4 (1979) 458–463.

[3] J. Gauvin and F. Dubeau, "Differential properties of the marginal function in mathematical programming", *Mathematical Programming Study*, to appear.

[4] J. Gauvin and J.W. Tolle, "Differential stability in nonlinear programming", *SIAM Journal on Control and Optimization* 15 (1977) 294–311.

[5] E.G. Gol'shtein, *Theory of convex programming*, Translations of Mathematical Monographs 36 (American Mathematical Society, Providence, RI, 1972).

[6] M.R. Hestenes, *Calculus of variations and optimal control theory* (Wiley, New York, 1966).

[7] M.R. Hestenes, *Optimization theory: The finite-dimensional case* (Wiley, New York, 1975).

[8] W.W. Hogan, "Directional derivatives for extremal-value functions with applications to the completely convex case", *Operations Research* 21 (1973) 188–209.

[9] O.L. Mangasarian and S. Fromovitz, "The Fritz John necessary conditions in the presence of equality and inequality constraints", *Journal of Mathematical Analysis and Applications* 17 (1967) 37–47.

[10] G.P. McCormick, "Second-order conditions for constrained minima", *SIAM Journal on Applied Mathematics* 15 (1967) 641–652.

[11] S.M. Robinson, "Stability theory for systems of inequalities, part II: Differentiable nonlinear systems", *SIAM Journal on Numerical Analysis* 13 (1976) 497–513.

[12] R.T. Rockafellar, *Convex analysis* (Princeton University Press, Princeton, NJ, 1970).

[13] R.T. Rockafellar, "A dual approach to solving nonlinear programming problems by unconstrained optimization", *Mathematical Programming* 5 (1973) 345–373.

[14] R.T. Rockafellar, "Augmented Lagrange multiplier functions and duality in nonconvex programming", *SIAM Journal on Control* 12 (1974) 268–285.

[15] R.T. Rockafellar, *The theory of subgradients and its applications: Convex and nonconvex functions* (Haldermann Verlag, West Berlin, 1981).

[16] R.T. Rockafellar, "Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming", *Mathematical Programming Study* 17 (1982) 28–66.

[17] R.T. Rockafellar, "Marginal values and second-order necessary conditions for optimality", *Mathematical Programming* 26 (1983) 245–286.

[18] G. Salinetti and R.J.B. Wets, "On the convergence of convex sets in finite dimensions", *SIAM Review* 21 (1979) 18–33.

[19] R.J.B. Wets, "On a compactness theorem for epiconvergent sequences of functions", forthcoming.