J. Austral. Math. Soc. (Series A) 41 (1986), 64–78

ON QUASIDIFFERENTIABLE OPTIMIZATION

B. D. CRAVEN

(Received 25 September 1984; revised 5 June 1985)

Communicated by B. Mond

Abstract

Lagrangian necessary conditions for optimality, of both Fritz John and Kuhn Tucker types, are obtained for a constrained minimization problem, where the functions are locally Lipschitz and have directional derivatives, but need not have linear Gâteaux derivatives; the variable may be constrained to lie in a nonconvex set. The directional derivatives are assumed to have some convexity properties as functions of direction; this generalizes the concept of quasidifferentiable function. The convexity is not required when directional derivatives are replaced by Clarke generalized derivatives. Sufficient Kuhn Tucker conditions, and a criterion for the locally solvable constraint qualification, are obtained for directionally differentiable functions.

1980 *Mathematics subject classification (Amer. Math. Soc.)*: 90 C 30.

1. Introduction and preliminaries

Consider the constrained minimization problem

$$(1)\qquad \operatorname*{Minimize}_{x\in X} f(x) \text{ subject to } -g(x)\in S,\ h(x)\in T,$$

in which X, Y, Z are normed vector spaces, $S \subset Y$ is a closed convex cone with interior, $T \subset Z$ is a closed convex cone, and $f\colon X \to \mathbf{R}$, $g\colon X \to Y$, $h\colon X \to Z$ are locally Lipschitz functions having (one-sided) directional derivatives. It is assumed that the functions are *quasidifferentiable* (in a more general sense than that of Pshenichnyl [16], applicable to vector-valued functions), then necessary conditions for a minimum are obtained in terms of Lagrange multipliers and directional derivatives. These conditions are of Fritz John type; conditions of Kuhn Tucker type are also obtained, using a constraint qualification generalizing that of Robinson [17] to functions possessing only directional derivatives. A further extension holds with Clarke generalized derivatives [2] when the functions are no longer assumed to have directional derivatives.

Denote by X', Y' and Z' the (topological) dual spaces of X, Y, and Z, respectively, *each equipped with its weak * topology*. Denote by $S^* = \{y' \in Y'\colon (\forall s \in S)\langle y', s\rangle \geqslant 0\}$ the dual cone of S. The *directional derivative* of g at x in direction v is

$$(2) \qquad g'(x; v) = \lim_{\alpha \downarrow 0} \alpha^{-1}[g(x + \alpha v) - g(x)],$$

where the limit is taken in the norm topology of Y. If $V \subset X$, let $g'(x; V) = \{g'(x; v)\colon\ v \in V\}$. The function g is called *strongly arcwise differentiable* at x(*s.a.d.* at x) if

$$(3) \qquad g'(x; v) = \lim_{\alpha \downarrow 0} \alpha^{-1}[g(x + \omega(\alpha)) - g(x)]$$

holds for each continuous arc $\alpha \mapsto \omega(\alpha)$ in X such that $\omega(0) = 0$ and $\omega'(0+) = v$. (In [7], *arcwise directionally differentiable* was similarly defined, but in the weak topology of Y.) The function g is *locally Lipschitz* at x if there is some ball N_x with centre x and some finite constant c such that $\|g(y) - g(z)\| \leqslant c\|y - z\|$ whenever $y, z \in N_x$ [9]. The function g will now be called *S-quasidifferentiable* at x if g has a directional derivative at x in each direction v such that $g'(x; \cdot)$ is S-convex. The latter means that, whenever $0 < \beta < 1$ and $u, v \in X$, then

$$(4) \qquad \beta g'(x; u) + (1 - \beta)g'(x; v) - g'(x; \beta u - (1 - \beta)v) \in S.$$

(When $Y = \mathbf{R}$ and $S = \mathbf{R}_+ \equiv [0, \infty)$, this definition reduces to that of Pshenichnyi [16], for then $g'(x; v) = \sup_{w \in K}\langle w, v\rangle$ for a compact set K; for general Y and S it generalizes that of S^*-nearly convex in Craven and Mond [5]; consequently, the sum of an S-convex function and a linearly Gâteaux differentiable function is S-quasidifferentiable [7]. The *tangent cone* of a set $\Delta \subset X$ at $a \in \Delta$ is the set

$$\{d \in X\colon \{d_n\} \to d, \{\alpha_n\} \downarrow 0, a + \alpha_n d_n \in \Delta\}.$$

LEMMA 1. *Let $g\colon X \to Y$ be locally Lipschitz and possess a directional derivative in each direction at x. Then g is s.a.d. at x.*

PROOF. Consider the arc $\alpha \mapsto \omega(\alpha) = \alpha v + \eta(\alpha)$, with $\eta(\alpha) = o(\alpha)$ as $\alpha \downarrow 0$. Then

$$(5) \quad \alpha^{-1}[g(x + \alpha v + \eta(\alpha)) - g(x)] = \alpha^{-1}[g(x + \alpha v + \eta(\alpha)) - g(x + \alpha v)] + \alpha^{-1}[g(x + \alpha v) - g(x)].$$

The second term on the right of (5) tends to $g'(x; v)$ as $\alpha \downarrow 0$, and the first term on the right is $O(\eta(\alpha))/\alpha = o(\alpha)/\alpha$ by the local Lipschitz hypotheses, and so $\to 0$.

https://doi.org/10.1017/S1446788700028081 Published online by Cambridge University Press

This result is a one-sided analog to the result in [14], [4], that a locally Lipschitz function which is linearly Gâteaux differentiable is then also Hadamard differentiable.

LEMMA 2 (The *basic alternative theorem* of [4], page 31). *Let X and Y be normed (or locally convex) spaces; let $S \subset Y$ be a convex cone with nonempty interior; let $\Gamma \subset X$ be convex; let $g\colon \Gamma \to Y$ be S-convex. Then exactly one of the two following systems has a solution*:

$$(\exists x \in X) - g(x) \in \operatorname{int} S; \tag{6}$$

$$(\exists p \in Y')(p \circ g)(\Gamma) \subset \mathbf{R}_+, \qquad 0 \neq p \in S^*. \tag{7}$$

LEMMA 3. *Let X and Y be normed spaces, $S \subset Y$ a closed convex cone with interior, $\Gamma \subset X$, and $g\colon \Gamma \to Y$ a function. Let $F(x, b) = \langle b, g(x) \rangle$, for $x \in \Gamma$ and $b \in Y'$. Then S^* has a convex weak $*$ compact base B (thus $S^* = \{\beta b\colon \beta \in \mathbf{R}_+,\ b \in B\}$ with $0 \notin B$), and*

$$(6) \Leftrightarrow (\exists x \in \Gamma)(\forall b \in B) F(x, b) < 0 \Leftrightarrow \inf_{x \in \Gamma} \sup_{b \in B} F(x, b) < 0; \tag{8}$$

$$(7) \Leftrightarrow (\exists b \in B)(\forall x \in \Gamma) F(x, b) \geqslant 0 \Leftrightarrow \text{NOT}\left\{ \sup_{b \in B} \inf_{x \in \Gamma} F(x, b) < 0 \right\}. \tag{9}$$

PROOF. Since $\operatorname{int} S \neq \varnothing$, S^* has a convex weak $*$ compact base [12], [3]. Since nonzero p exists in S^* if and only if $p = \beta b$ for some $b \in B$ and some $\beta > 0$, and $-g(x) \in \operatorname{int} S$ if and only if $F(x, b) = \langle b, g(x) \rangle < 0$ for each $b \in B$, the left equivalences in (8) and (9) are immediate. Since B is compact and $F(x, \cdot)$ continuous, we have

$$\begin{aligned}(\exists x \in \Gamma)(\forall b \in V) F(x, b) < 0 &\Leftrightarrow (\exists x \in \Gamma) \sup_{b \in B} F(x, b) < 0 \\ &\Leftrightarrow \inf_{x \in \Gamma} \sup_{b \in B} F(x, b) < 0,\end{aligned}$$

and so (8) holds. To prove (9), let $\phi(b) = \inf_{x \in \Gamma} F(x, b)$. Then ϕ is upper semicontinuous (usc) on B since each $F(x, \cdot)$ is continuous and hence usc; moreover, $\sup_{b \in B} \phi(b)$ is attained since ϕ is usc and B compact. (See [13], [1].) Then

$$\begin{aligned}\text{NOT}(7) &\Rightarrow (\forall b \in B)(\exists x \in \Gamma) F(x, b) < 0 \Leftrightarrow (\forall b \in B) \phi(b) < 0 \\ &\Leftrightarrow \sup_{b \in B} \phi(b) < 0,\end{aligned} \tag{10}$$

which proves (8).

The proofs given in [13] are finite-dimensional, but, given compactness, they also hold for infinite dimensions. The continuity of $F(x, \cdot)$ for each x is not fully used, only the upper semicontinuity. Convexity of Γ is not required.

If, in particular, $F(x,b) = \langle b, x\rangle$, if B is compact, and if Γ is convex, then Lemmas 2 and 3 show that exactly one of the two following systems hold:

$$(11) \quad (\exists x \in \Gamma)(\forall b \in B)\langle b, x\rangle < 0; \qquad (\exists b \in B)(\forall x \in \Gamma)\langle b, x\rangle \geqslant 0.$$

This was proved otherwise in [7], Lemma 2, with the additional assumption that Γ is a convex cone.

Under the hypotheses of Lemma 3, it follows from Lemmas 2 and 3 that exactly one of (6) and (7) holds, under any further hypothesis which makes

$$(12) \qquad \inf_{x\in\Gamma}\sup_{b\in B} F(x,b) < 0 \Leftrightarrow \sup_{b\in B}\inf_{x\in\Gamma} F(x,b) < 0.$$

In particular, various hypotheses are known which ensure that

$$(13) \qquad \inf_{x\in\Gamma}\sup_{b\in B} F(x,b) = \sup_{b\in B}\inf_{x\in\Gamma} F(x,b)$$

when B is compact. In particular, a theorem of Ky Fan [8] (see Pomérol [15]) shows that (13) holds when B is compact, $F(x,b) = \langle b, g(x)\rangle$, and g has the property

(14)

$$(\forall \rho \in (0,1))(\forall x, x' \in \Gamma)(\exists x'' \in \Gamma) - g(x'') + \rho g(x) + (1-\rho)g(x') \in S.$$

(It follows that $F(x,b)$ satisfies the "convex-like" property, required for Ky Fan's theorem, namely property (14) with $F(\cdot, b)$ replacing g and with $\mathbf{R}_+$ replacing S. A related property for $F(x,\cdot)$ holds automatically, since here $F(x,\cdot)$ is linear. Note that convexity of Γ is not assumed here.)

In Lemma 2, g is assumed to be S-convex. An equivalent hypothesis is that $F(x,b) = \langle b, g(x)\rangle$ is convex in x for each $b \in B$. The convexity requirement can now be weakened as follows. The function g: $\Gamma \to Y$ will be called *S-metaconvex* if, for some positive weighting function θ: $\Gamma \to \mathbf{R}_+$ which is bounded away from zero and ∞, property (14) holds for the function $\theta(\cdot)g(\cdot)$. Assuming the hypotheses of Lemma 3, (13) then holds with $F(x,b)$ replaced by $\theta(x)\langle b, g(x)\rangle$, and this implies (12), since θ is bounded away from 0 and ∞. The following result has then been proved.

LEMMA 4. *Let X and Y be normed spaces: let $S \subset Y$ be a closed convex cone with nonempty interior; let $\Gamma \subset X$; let g: $\Gamma \to Y$ be S-metaconvex. Then exactly one of the systems* (6) *and* (7) *has a solution.*

2. Lagrangian necessary conditions

Consider now the constrained minimization problem

$$(15) \qquad \operatorname*{Minimize}_{x\in X} f(x) \text{ subject to } -g(x) \in S,\ x \in \Delta,$$

with f, g, and S as for problem (1), and with $\Delta \subset X$.

https://doi.org/10.1017/S1446788700028081 Published online by Cambridge University Press

THEOREM 1. *Let x and Y be normed vector spaces, $S \subset Y$ a closed convex cone with interior, $\Delta \subset X$, and f: $X \to \mathbf{R}$ and g: $X \to Y$ locally Lipschitz functions possessing (one-sided) directional derivatives at each point in each direction. If we assume that Δ has a convex tangent cone V at a and that $f'(a;\cdot)$ is convex and $g'(a;\cdot)$ is S-convex, then necessary conditions for* (15) *to attain a local minimum at $x = a \in \Delta$ are that Lagrange multipliers $\tau \in \mathbf{R}_+$ and $\lambda \in S^*$ exist, not both zero, satisfying*

$$(16) \qquad (\tau f + \lambda g)'(a; V) \subset \mathbf{R}_+, \qquad \langle \lambda, g(a) \rangle = 0.$$

There also exists $\sigma \in V^*$ such that

$$(17) \qquad (\tau f + \lambda g)'(a;\cdot) \geqslant \langle \sigma, \cdot \rangle, \qquad \langle \lambda, g(a) \rangle = 0.$$

The result still holds if the convexity hypotheses on the directional derivatives are weakened to $(\mathbf{R}_+ \times S)$-metaconvexity of the function $\psi(\cdot) = (f'(a;\cdot), g(a) + g'(a;\cdot))$ on V, where V need not be convex.

PROOF. Suppose that the system

$$(18) \qquad f'(a;d) < 0, \qquad -g(a) - g'(a;d) \in \operatorname{int} S$$

has a solution $d \in V$. Since V is the tangent cone to Δ at a, the constraint $x \in \Delta$ is satisfied by $x = a + \alpha d + \eta(\alpha)$ for some sequence of values of $\alpha \downarrow 0$, where $\eta(\alpha) = o(\alpha)$. Since G is locally Lipschitz and directionally differentiable, we have

$$(19) \quad \begin{aligned} g(a + \alpha d + \eta(\alpha)) - g(a) & \\ &= g(a + \alpha d) - g(a) + g(a + \alpha d + \eta(\alpha)) - g(a + \alpha d) \\ &= \alpha g'(a;d) + o(\alpha) + o(\eta(\alpha)) \\ &= \alpha g'(a;d) + o(\alpha) + o(\alpha). \end{aligned}$$

So, for sufficiently small positive α, it follows that

$$(20) \quad \begin{aligned} -g(a + \alpha d + \eta(\alpha)) &= -g(a) - \alpha g'(a;d) + o(\alpha) \\ &= (1-\alpha)(-g(a)) + \alpha(-g'(a;d) - g(a)) + o(\alpha) \in S \end{aligned}$$

since $-g(a) \in S$ and $-g'(a) - g(a;d) \in \operatorname{int} S$. Similarly, since f is locally Lipschitz and directionally differentiable, we have

$$(21) \qquad \alpha^{-1}[f(a + \alpha d + \eta(\alpha)) - f(a)] = f'(a;d) + o(\alpha)/\alpha < 0$$

for sufficiently small positive α. Thus (20) and (21) contradicts the local minimum of (15) at a. It follows that the system (18) has no solution $d \in V$. Since, by hypothesis, ψ is convex and V is convex, Lemma 2 shows that there exists nonzero $p = (\tau, \lambda) \in \mathbf{R}_+ \times S^*$ such that $(p \circ \psi)(V) \subset \mathbf{R}_+$. Thus $\tau \in \mathbf{R}_+$ and $\lambda \in S^*$ are not both zero, and (16) is satisfied.

Since $\Phi(\cdot) \equiv (\tau f + \lambda g)'(a; \cdot)$ is positively homogeneous, its directional derivative $\Phi'(0; v) = \Phi(v)$ for each v. The local Lipschitz hypothesis on f and g ensures that $\Phi(\cdot)$ is bounded. Hence, since Φ is convex, we have

$$\Phi'(0; v) = \sup_{w \in C} \langle w, v \rangle, \tag{22}$$

where the subdifferential $C = \partial\Phi(0)$ is convex and weak $*$ compact. From (16), $\Phi(V) \subset \mathbf{R}_+$. Hence

$$(\forall v \in V)(\exists w \in C)\langle w, v \rangle \geqslant 0. \tag{23}$$

Hence there is no solution to the system

$$(\exists v \in V)(\forall w \in C)\langle w, v \rangle < 0. \tag{24}$$

Since C is compact and V convex, the special case (11) of Lemma 2 and 3 shows that the system

$$(\exists w \in C)(\forall v \in V)\langle w, v \rangle \geqslant 0 \tag{25}$$

has a solution, say $w = \sigma$. Then $\sigma \in V^* \cap C$, so that $\Phi(v) = \Phi(v) - \Phi(0) \geqslant \langle \sigma, v \rangle$ for each $v \in X$. Therefore $\Phi(\cdot) \geqslant \langle \sigma, \cdot \rangle$, which proves (17).

In case the convexity of directional derivatives is weakened to metaconvexity, Lemma 4 instead of Lemma 2 may be applied to the system $-\psi(q) \in \operatorname{int}(\mathbf{R}_+ \times S)$, resulting in (17). Lemma 4 does not require convexity of V.

Convex hypotheses are required, however, for result (17), since (22) uses convexity. The deduction of (17) from (16) is not trivial, because $(\tau f + \lambda g)'(a; \cdot)$ is not in general linear.

The constraint $x \in \Delta$ may take the form $-h(x) \in T$, where T is a closed convex cone, and where h is T-quasidifferentiable. If this constraint is also locally solvable [4] at a, then its tangent cone $V = \{v \in X\colon -h(a) - h'(a; v) \in T\}$ is convex, as required by Theorem 1. If h is linearly Gâteaux differentiable at a, then Farkas's theorem shows that V^* is the weak $*$ closure of the convex cone $[h(a), h'(a)]^T(T^*)$, or equivalently, of $h'(a)^T(T_{-h(a)})^*$, where $h'(a)$ is the linear derivative, and where $T_c = \{\alpha(t - c)\colon \alpha \in \mathbf{R}_+, t \in T\}$, with here $c = -h(a)$. If H is τ-quasidifferentiable, but not necessarily differentiable, then results of Zalinescu [19] and Glover [10] show that V^* is the weak $*$ closure of the cone $U = -\bigcup_{t \in T^*} \partial(t\zeta)(0)$, where $\zeta(\alpha, x) = \alpha h(a) + h'(a; x)$. Assume a Robinson constraint qualification, namely that $h(a) + h'(a; X) + T$ contains a neighbourhood of 0 in Y (see [17], and [4], page 150), and that X is complete. Then ζ is $(\mathbf{R}_+ \times T)$-convex and positively homogeneous, and then the constraint qualification implies that $\zeta(\mathbf{R}, X) + T = V$. From [10], Lemma 3, the cone U is already weak $*$ closed. Assuming this, the following "directional derivative" version of the Fritz John necessary conditions follows from (17), with some multiplier $\mu \in T^*$:

$$(\tau f + \lambda g + \mu h)'(a; X) \subset \mathbf{R}_+, \quad \langle \lambda, g(a) \rangle = 0, \quad \langle \mu, h(a) \rangle = 0, \tag{26}$$

where $\tau \in \mathbf{R}_+$ and $\lambda \in S^*$ are not both zero, and where $\mu \in T^*$.

These Fritz John conditions, (17) or (26), will lead to corresponding Kuhn Tucker conditions under any constraint qualification which gives $\tau \neq 0$. Slater's constraint qualification is not applicable since g (and h) may lack convex, or pseudoconvex, properties. Consider the problem

$$\text{(27)} \qquad \text{Minimize } f(x) \text{ subject to } -g(x) \in S,\ x \in \Delta,$$

with f and g having directional derivatives, but with the convex cone S not necessarily having interior. If f and g were continuously Fréchet differentiable, the Robinson constraint qualification would require that $g(a) + g'(a)(\Delta - a) + S$ contains a ball with centre 0 in Y, which implies a similar inclusion with $\Delta - a$ replaced by the cone it generates. This requirement may be generalized to directionally differentiable functions, with the tangent (contingent) cone V (not necessarily convex) to Δ at a, so as to assume that

$$\text{(28)} \qquad g(a) + g'(a;V) + S \supseteq N,$$

where N is the ball in Y with centre 0.

THEOREM 2. *For problem* (27), *with functions f and g having directional derivatives, assume that Fritz John conditions*

$$\text{(29)} \qquad \begin{gathered} (\tau f + \lambda g)'(a;\cdot) \geqslant \langle \sigma, \cdot \rangle, \qquad \tau \in \mathbf{R}_+,\ \lambda \in S^*, \\ (\tau, \lambda) \neq (0,0), \quad \sigma \in V^*, \quad \langle \lambda, g(a) \rangle = 0 \end{gathered}$$

hold at the feasible point a, and that the constraint qualification (28) *holds there. Then* $\tau \neq 0$ *so that* (29) *holds with* $\tau = 1$.

PROOF. Suppose, if possible that $\tau = 0$. Then $\lambda \neq 0$, and (29) gives $\langle \lambda, g(a) \rangle = 0$ and $(\lambda g)'(a; v) \geqslant \langle \sigma, v \rangle \geqslant 0$, for all $v \in V$, since $\sigma \in V^*$. From (28), we obtain

$$0 + (\lambda g)'(a;V) + \langle \lambda, S \rangle \supseteq \langle \lambda, N \rangle \supseteq (-\delta, \delta)$$

for some $\delta > 0$. Since $0 \neq \lambda \in S^*$, we have $\langle \lambda, S \rangle = \mathbf{R}_+$. Hence $\mathbf{R}_+ + \mathbf{R}_+ \supseteq (-\delta, \delta)$, which is a contradiction. So $\tau \neq 0$.

Theorem 1 and Theorem 2 extend to the weak minimization of a vector-valued function $f\colon X \to W$ with respect to a convex cone P in the space W; thus $f(x) - f(a) \notin -\operatorname{int} P$ for all feasible x in a neighbourhood of a. Then $\tau \in \mathbf{R}_+$ is replaced by $\tau \in P^*$, and (in the proof of Theorem 1) $f'(a; d) < 0$ by $-f'(a; d) \in \operatorname{int} P$. In Theorem 2, the conclusion is that $0 \neq \tau \in P^*$.

3. Necessary conditions with generalized derivatives

Consider now the problem

$$\text{(30)} \qquad \underset{x \in X}{\text{Minimize}}\ f(x) \text{ subject to } -g(x) \in S,\ x \in \Delta,$$

in which X and Y are normed spaces, Y is a Banach space, $S \subset Y$ is a closed convex cone with interior, and f: $X \to \mathbf{R}$ and g: $X \to Y$ are locally Lipschitz functions, now not necessarily possessing directional derivatives. The set $\Delta \subset X$ will still be assumed to have a convex tangent cone at a local minimum point, a, of (30). Since it has the local Lipschitz property, f possesses a generalized derivative in Clarke's sense (e.g. [2]):

$$\text{(31)} \qquad f^0(a; d) = \limsup_{y \to a,\, \alpha \downarrow 0} \alpha^{-1}[f(y + \alpha d) - f(y)];$$

moreover, $f^0(a; \cdot)$ is convex, positively homogeneous and continuous [15]. Since definition (31) does not apply to the vector-valued function g, the constraint $-g(x) \in S$ is replaced equivalently by $(\forall b \in B) bg(x) \equiv \langle b, g(x) \rangle \leqslant 0$, where B is a convex weak $*$ compact base for S^*. Note, however, that the mapping $b \mapsto (bg)^0(a; d)$ can no longer be assumed linear. Denote by $C(B)$ the space of continuous real functions on B, and by $\mathbf{M}(B)$ its dual space, the space of (nonnegative Radon) measures on B. Since B is weak $*$ compact and Y is complete, we have $k_0 \equiv \sup_{b \in B} \|b\| < \infty$ as a consequence of the uniform boundedness theorem.

LEMMA 5. *If g: $X \to Y$ is locally Lipschitz, then the mapping π, given by $\pi(b) = (bg)^0(a; d)$ for $b \in B$, is continuous, for each fixed a and d.*

PROOF. From (31), applied to the real function bg, we obtain

$$(bg)^0(a; d) = \lim_{\rho \downarrow 0} \sup_{\|y - s\| \leqslant \rho,\, 0 < \alpha < \rho} \alpha^{-1}[bg(y + \alpha d) - bg(y)].$$

Hence, for b, $b' \in B$ we have

$$\begin{aligned} &((b + b')g)^0(a; d) \\ &\quad \leqslant \lim_{\rho \downarrow 0} \left\{ \sup \alpha^{-1}[bg(y + \alpha d) - bg(y)] + \sup \alpha^{-1}[b'g(y + \alpha d) - b'g(y)] \right\} \\ &\quad = (bg)^0(a; d) + (b'g)^0(a; d). \end{aligned}$$

https://doi.org/10.1017/S1446788700028081 Published online by Cambridge University Press

Also, for $\beta > 0$, $(\beta bg)^0(a; d) = \beta(bg)^0(a; d)$. Hence the mapping π is convex. Its values are also bounded, since $|\pi(b)| \leqslant k_0 k_1 \|d\|$, where k_0 is the bound on $\sup_{b \in B}\|b\|$, and where k_1 is a Lipschitz constant for g in a neighbourhood of a. Hence, by [18, Theorem 8], π is continuous.

THEOREM. 3. *Let X be a normed vector space, Y a Banach space, $S \subset Y$ a closed convex cone with interior, $\Delta \subset X$, and f: $X \to \mathbf{R}$ and g: $X \to Y$ locally Lipschitz functions. If we assume that Δ has a convex tangent cone V at a, then necessary conditions for* (30) *to attain a local minimum at $x = a \in \Delta$ are that there exist $\tau \geqslant 0$ and $\rho \in \mathbf{M}(B)$, not both zero, such that*

$$\Phi(v) \equiv \tau f^0(a; v) + \int_B (bg)^0(a; v)\rho(db) \quad \text{and} \quad \lambda = \int_B b\rho(db)$$

satisfy

$$\Phi(V) \subset \mathbf{R}_+, \quad \langle \lambda, g(a) \rangle = 0. \tag{32}$$

There also exists $\sigma \in V^*$ such that

$$\Phi(\cdot) \geqslant \langle \sigma, \cdot \rangle, \quad \langle \lambda, g(a) \rangle = 0. \tag{33}$$

PROOF. Suppose that the system

$$f^0(a; d) < 0, \quad (\forall q \in B)\beta(qg)(a) + (qg)^0(a; d) < 0 \tag{34}$$

has a solution $(d, \beta) \in V \times \mathbf{R}$. Since V is the tangent cone to Δ at a, the constraint $x \in \Delta$ is satisfied by $x = a + \alpha d + \eta(\alpha)$ for some sequence of values of $\alpha \downarrow 0$, where $\eta(\alpha) = o(\alpha)$. Then $f^0(a; d) < -2\gamma$ for some $\gamma > 0$. From the definition [2] of the Clarke generalized directional derivative $f^0(a; d)$, it follows that $f(y + \alpha d) - f(y) \leqslant \alpha f^0(a; d) + \gamma$ whenever α and $\|y - a\|$ are sufficiently small. Hence

$$\begin{aligned} (35)\quad & f(a + \alpha d + \eta(\alpha)) - f(a) \\ &= f(a + \alpha d + \eta(\alpha)) - f(y + \alpha d) + f(y + \alpha d) - f(y) + f(y) - f(a) \\ &\leqslant k\|y - a - \eta(\alpha)\| + \alpha\left[f^0(a; d) + \gamma\right] + k\|y - a\| \end{aligned}$$

where k is the Lipschitz constant for f, when α and $\|y - a\|$ are sufficiently small

$$\begin{aligned} &\leqslant 2k\|y - a\| + o(\alpha) - 2\alpha\gamma + \alpha\gamma \\ &< -\alpha\gamma/3 + o(\alpha) \quad \left(\text{whenever } \|y - a\| < \alpha\gamma/(3k)\right) \\ &< 0 \quad (\text{whenever } \alpha \text{ is sufficiently small positive}). \end{aligned}$$

Since B is compact and, by Lemma 4, since $b \mapsto (bg)^0(a; \cdot)$ is usc, there exists $\gamma' > 0$ for which $(\forall b \in B)(bg)(a) + (bg)^0(a; d) \leqslant -2\gamma' < 0$. An argument similar to that of (35), applied to bg, then shows that

$$(bg)(a + \alpha d + \eta(\alpha)) \leqslant (bg)(a) + \alpha(bg)^0(a; d) + 5\alpha\gamma/3 + o(\alpha) \tag{36}$$

https://doi.org/10.1017/S1446788700028081 Published online by Cambridge University Press

for all sufficiently small positive α. (Here K is replaced by $k' \sup_{b \in B} \|b\|$, where k' is the Lipschitz constant for g, and γ is replaced by γ'.) Consequently, for each $b \in B$, we have

$$(37)\quad (bg)(a + \alpha d + \eta(\alpha)) \leqslant (1 - \alpha\beta)(bg)(a) + \alpha\left[(bg)^0(a; d) + \beta(bg)(a)\right] + 5a\gamma'/3 + o(\alpha) \leqslant 0 + (-2\alpha\gamma') + 5\alpha\gamma'/3 + o(\alpha) < 0$$

for all sufficiently small positive α. Hence $-g(a + \alpha d + \eta(\alpha)) \in S$ for such α. Thus the local minimum of (30) is contradicted. Hence (34) has no solution $(d, \beta) \in V \times \mathbf{R}$.

Now define a function $\psi = (\psi_1, \psi_2)$: $V \to \mathbf{R} \times C(B)$ by $(\forall d \in V)\psi_1(d) = f^0(a; d)$ and $(\forall b \in V)\psi_2(d)(b) = (bg)(a) + (bg)^0(a; d)$. By Lemma 5, $b \mapsto (bg)^0(a; d)$ is continuous, so $\psi_2(d) \in C(B)$. Since $f^0(a; \cdot)$ and each $(bg)^0(a; \cdot)$ is a convex function, ψ is K-convex, where K denotes the convex cone of nonnegative functions on $\{0\} \times B$. Then the system $-\psi(q) \in \operatorname{int} K$ has no solution $d \in V$. Lemma 2, or 4, then shows that there exists $\tau \geqslant 0$ and $\rho \in \mathbf{M}(B)$, not both zero, such that (32) holds.

The proof of (33) from (32) is the same as the proof of (18) from (16) in Theorem 1.

In Theorems 1, 3, and 4, $\langle \lambda, g(a) \rangle = 0$, $\lambda \in S^*$, and $-g(a) \in S$, so that $\lambda \in (S_{-g(a)})^*$, where S_c denotes the convex cone $\{\alpha(s - c)\colon s \in S, \alpha \in \mathbf{R}_+\}$ here $c = -g(a)$). However, since $(bg)^0(s; v)$ is not generally a linear function of b, it does not follow from (32) that $\tau f^0(a; v) + (\lambda g)^0(a, v) \in \mathbf{R}_+$ for each $v \in V$. (This holds if, in particular, $S = \mathbf{R}^m_+$.) Note that λ is represented (perhaps not uniquely) by the measure ρ, and ρ is supported by the closure in the weak * topology of the set $\{b \in B\colon bg(a) \neq 0\}$.

4. Conditions for local solvability

A Lagrangian necessary condition (26) for a minimum was obtained under the assumption that a constraint $-h(x) \in T$ was locally solvable at the point a. The known (Robinson) sufficient condition for local solvability [7], [4, page 150] assumes that h is continuously Fréchet differentiable. This suggests the conjecture that a generalized Robinson condition such as (28) may suffice for local solvability when the function is only directionally differentiable, but when some bound exists on the remainder terms arising in the definition of directional derivative.

https://doi.org/10.1017/S1446788700028081 Published online by Cambridge University Press

Consider then a function h: $X \to Z$ of the form

$$h(x) = h(a) + \phi(x - a) + \psi(x - a), \tag{38}$$

where ϕ is positively homogeneous, but not necessarily linear (so that $\phi(x - a)$ could be a directional derivative $h'(a; x - a)$), and where

$$\|\psi(x) - \psi(y)\| < \varepsilon\|x - y\| \quad \text{whenever } \|x - a\|, \|y - x\| < \delta(\varepsilon). \tag{39}$$

The construction of the implicit function, which local solvability requires, requires the solvability of an approximation to $-h(x) \in T$. In particular, when $T = \{0\}$, the Newton's method construction of the implicit function used in [4] assumes that $h'(a)$ has a bounded right inverse (which is implied by the Robinson condition when $T = \{0\}$). This assumption may be generalized to require ϕ to have a bounded right inverse.

A partial answer to this question is given by the following theorem. The hypothesis required is that the system

$$p + h'(a; u) + s = y \tag{40}$$

can be solved for $u \in X$ and $s \in T$ satisfying the bounds $\|u\| \leqslant \rho\|y\|$ and $\|s + p\| \leqslant \rho\|y\|$ for some ρ independent of the given p and y, whenever $\|y\|$ and $\|p - b\|$ are sufficiently small, where $b = -h(a)$. The system $-h(x) \in T$ will be called *tractable* at the point a if h is directionally differentiable if h satisfies (38) and (39) with $\phi(x - a) = h'(a; x - a)$, and if h satisfies the stated bounded solvability property. In case h is continuously Fréchet differentiable, this solvability of (40) follows from the Robinson criterion, namely that $b + h'(a)(X) + T$ contains a ball with centre 0. In fact, the less restrictive hypothesis that $b + h'(s; X) + T$ contains a ball with centre 0 is stable to small perturbations of b away from $-h(a)$, and so implies the solvability of (40), though not necessarily the boundedness of solutions u and $s + p$.

THEOREM 4. *Let T be a closed convex cone; let X be complete; let h: $X \to Y$ be directionally differentiable; let the system $-h(x) \in T$ be tractable at the point a, where $b = -h(a) \in T$. Then $-h(x) \in T$ is locally solvable at a; thus, $-h(x) \in T$ has a solution $x = a = \alpha c + o(\alpha)$ as α_0 whenever c satisfies $-h(a) - h'(a; c) \in T$.*

PROOF. A Newton-type iterative sequence $\{x_n\}$ constructed as follows. Assume $\|c\| \leqslant 1$. Let $x_0 = a + \alpha c$, $b_0 = h(a) + \alpha h'(a; c)$ for sufficiently small $\alpha \in (0, 1)$. For $n = 0, 1, 2, \ldots,$ the points $x_{n+1} \in X$ and $s_n \in T$ are required to satisfy

$$b_n + h'(a; x_{n+1} - x_n) + s_{n+1} = y_n \equiv -[h(x_n) + a_n], \tag{41}$$

https://doi.org/10.1017/S1446788700028081 Published online by Cambridge University Press

where $b_n = -s_n$. The *tractable* hypothesis ensures that, for some constant ρ independent of n, we have $\|x_{n+1} - x_n\| \leqslant \rho\|y_n\|$ and $\|s_{n+1} + b_n\| \leqslant \rho\|y_n\|$, provided that $\|y_n\|$ and $\|b_n - b\|$ are sufficiently small. It then follows that

$$\begin{aligned}(42)\qquad h(x_{n+1}) &= h(x_n) + h'(a; x_{n+1} - x_n) + \xi(x_n; x_{n+1} - x_n) \\ &\qquad \left(\text{where } \|\xi(x_n; x_{n+1} - x_n)\| < \varepsilon\|x_{n+1} - x_n\|\right) \\ &= -y_n + b_n + h'(a; x_{n+1} - x_n) + \xi(x_n; x_{n+1} - x_n) \\ &= s_{n+1} + \xi(x_n; x_{n+1} - x_n)\end{aligned}$$

by (41). Here $0 < \varepsilon < 1/(2\rho)$ and $0 < \alpha < \frac{1}{2}\delta(\varepsilon)$ from (39). If $\|y_n\|$ is small, then $\|x_{n+1} - x_n\|$ is small, and so (42) shows that $\|y_{n+1}\| = -h(x_{n+1}) + s_{n+1}\|$ is then small. An inductive proof shows that $\|x_n - x_{n-1}\| \leqslant \alpha(\rho\varepsilon)^n$ for $n = 1, 2, \ldots$; the details are omitted, since they are the same as in [4], pages 148–149, for the Fréchet differentiable case. Note, however, that the cited proof does not use the linearity of the derivative, but only its positive homogeneity in direction. Consequently, $\{x_n\}$ is a Cauchy sequence in the Banach space X, so it converges to a limit $\bar{x}(\alpha) \in X$; then $\{s_n\}$ is a Cauchy sequence in Y, so it converges to some $\bar{s}(\alpha) \in T$, since T is closed. Then $-g(\bar{x}(\alpha)) = \bar{s}(\alpha) \in T$, and $\bar{x}(\alpha) = a + \alpha c + o(\alpha)$ as in [4].

5. Sufficient Kuhn Tucker conditions

Assume that problem (30) satisfies the Kuhn Tucker conditions, in the form given by Theorem 1, at the feasible point a, namely that

$$(43)\qquad \begin{gathered}(\exists \lambda \in S^*, \sigma \in V^*)(f + \lambda g)'(a; \cdot) \geqslant \langle \sigma, \cdot \rangle, \\ \langle \lambda, g(a) \rangle = 0, \qquad a \in \Delta,\ -g(a) \in S.\end{gathered}$$

Under suitable convex, or generalized convex, hypotheses, it follows from (43) that the point a minimizes problem (30) (see [4], [6]). For (Fréchet) differentiable functions, the generalized convex hypotheses have been weakened to *invex* [11], [5]. An analogous result can be given for *directionally differentiable* functions. For problem (30), let

$$(44)\qquad \Phi(s) = \begin{bmatrix} f(x) \\ g(x) \end{bmatrix}, \quad K = \mathbf{R}_+ \times S,\ \Lambda = [1, \lambda].$$

THEOREM 5. *Assume that a function η: $X \times X \to V$ exists, satisfying*

$$(45)\qquad (\forall x \in V)\Phi(x) - \Phi(a) - \Phi'(a; \eta(x, a)) \in K,$$

where a is a feasible point for problem (30) *satisfying the Kuhn Tucker conditions* (43), *where f and g are s.a.d. functions. Then a minimizes* (30).

PROOF. Let x be any feasible point for (30), so that $x \in \Delta$, $-g(x) \in S$. Denote by $\geqslant_K$ the partial ordering induced by the convex cone K, so that $p \geqslant_K q \Leftrightarrow p - q \in K$. Then $p \geqslant_K q$ and $\Lambda \in K^*$ imply that $\langle \Lambda, p \rangle \geqslant \langle \Lambda, q \rangle$. It follows that

$$(46) \qquad \begin{aligned} f(x) - f(a) &\geqslant_K \langle \Lambda, \Phi(x) - \Phi(a) \rangle \\ &\geqslant_K \langle \Lambda, \Phi'(a; \eta(x,a)) \rangle \quad (\text{by } (45) \text{ and } \Lambda \in K^*) \\ &\geqslant \langle \sigma, \eta(x,a) \rangle \quad (\text{since } \sigma \in V^*,\ \eta(x,a) \in V) \\ &\geqslant 0. \end{aligned}$$

So the point a is a minimum for (30).

REMARKS. Property (45) generalizes the cone-index property defined in [5] to functions which are s.a.d. rather than Fréchet differentiable. However, no necessary or sufficient conditions for (45), similar to those in [5], are yet available.

6. Quasimin

The point a will be called an *arcwise quasimin* of problem (30) if, for each nonzero $\xi \in X$, there exist $\theta(d, \xi) \in X$ such that, for each continuous arc $\alpha \mapsto \omega(\alpha) \in \Delta - a$ satisfying $\omega(0) = 0$, $\omega'(0+) = \xi$, and $-g(a + \omega(\alpha)) \in S$, we have

$$(47) \qquad f(a + \omega(\alpha)) - f(a) - \theta(\alpha) \geqslant 0 \quad \text{as } \alpha \downarrow 0.$$

This definition generalizes the definition of *quasimin* [3], [4], which requires that $f(a+z) - f(a) - o(\|z\|) \geqslant 0$ as $z \to a$ through feasible pionts, to continuous arcs considered separately.

THEOREM 6. *The necessary conditions of Theorem* 1 *remain valid when the local minimum of* (30) *at a is weakened to arcwise quasimin at a. Conversely, if the feasible point a satisfies the Kuhn Tucker conditions* (17) *with* $\tau = 1$, *then a is an arcwise quasimin of* (30).

PROOF. The necessary conditions are proved similarly to those of Theorem 1, if we note that when $f'(a; d) < 0$ from (18), then an extra $o(\alpha)$ term may be added to (19), so that (19) also contradicts an arcwise quasimin.

For the converse, let a be a feasible point satisfying (17) with $\tau = 1$. Suppose that a is not an arcwise quasimin. Then there exists an arc $\alpha \mapsto \omega(\alpha) \in \Delta - a$ with $\omega(0) = 0$ and $\omega'(0+) = \xi$ such that, whenever $\theta(\alpha) = o(\alpha)$, we have

$$f(a + \omega(\alpha)) - f(a) - \theta(\alpha) < 0. \tag{48}$$

It follows that

$$\begin{aligned} \lambda g'(a;\xi)\alpha &= \lambda[g(a+\omega(\alpha)) - g(a) + \rho(\alpha)], \quad \text{where } \rho(\alpha) = o(\alpha) \\ &\leqslant 0 + \lambda\rho(\alpha), \quad \text{since } \langle \lambda, g(a) \rangle = 0. \end{aligned}$$

Then, from (17), we obtain

$$f(a+\omega(\alpha)) - f(a) = f'(a;\xi)\alpha + \sigma(\alpha), \quad \text{where } \sigma = o(\alpha).$$

Choose $\theta(\alpha) = -\sigma(\alpha) + \lambda\rho(\alpha)$. Then $\theta(\alpha)' = o(\alpha)$, and, from (17) with $\tau = 1$, we obtain

$$f(a+\omega(\alpha)) - f(a) \geqslant -\lambda g'(a;\xi)\alpha + \sigma(\alpha) \geqslant -\theta(\alpha),$$

thereby contradicting (48).

Acknowledgment

The author is indebted to a referee, and also to Mr. V. Jeyakumar, for various amendments and improvements to this paper, including a tidier proof for Theorem 1, and an amendment to the constraint qualification used.

References

[1] C. Berge, *Topological spaces* (Oliver and Boyd, Edingburgh and London, 1963).
[2] F. H. Clarke, 'Generalized gradients of Lipschitz functionals,' *Adv. in Math.* **40** (1981), 52–67.
[3] B. D. Craven, 'Lagrangean conditions and quasiduality,' *Bull. Austral. Math. Soc.* **16** (1977), 325–339.
[4] B. D. Craven, *Mathematical programming and control theory* (Chapman and Hall, London, 1978).
[5] B. D. Craven, 'Invex functions and constrained local minima,' *Bull. Austral. Math. Soc.* **24** (1981), 357–366.
[6] B. D. Craven, 'Vector-valued optimization' (in *Generalized Concavity in Optimization and Economics*, S. Schaible and W. T. Ziemba (eds.), Academic Press, New York, 1981, pp. 661–687).
[7] B. D. Craven and B. Mond, 'Lagrangean conditions for quasidifferentiable optimization' (*Survey of Mathematical Programming*, Vol. 1, A. Prékopa (ed.), Akadémiai Kiadó, Budapest, and North-Holland, Amsterdam, 1979, pp. 177–192).
[8] Ky Fan, 'Minimax theorems,' *Proc. Nat. Acad. Sci. U.S.A.* **39** (1953), 42–47.

[9] T. M. Flett, *Differential analysis* (Cambridge Univ. Press., Cambridge, 1980).
[10] B. M. Glover, 'A generalized Farkas lemma with applications to quasidifferentiable programming', *Zeitschrift für Operations Res.* **26** (1982), 125–141.
[11] Morgan A. Hanson, 'On sufficiency of the Kuhn Tucker conditions,' *J. Math. Anal. Appl.* **80** (1981), 545–550.
[12] E. Jameson, *Ordered linear spaces*, (Lecture Notes in Math. 141, Springer-Verlag, Berlin, 1970).
[13] O. L. Mangasarian, *Nonlinear programming* (McGraw-Hill, New York, 1969).
[14] M. Z. Nashed, 'Differentiability and related properties of nonlinear functional analysis' (*Nonlinear Functional Analysis and Applications*, L. B. Rall (ed.), Academic Press, New York, 1971, pp. 103–310).
[15] J. Ch. Pomérol, 'Inequality system and minimax theorems', *J. Math. Anal. Appl.* **103** (1984), 262–292.
[16] P. N. Pshenichnyi, *Necessary conditions for an extremum* (Marcel Dekker, New York, 1971).
[17] S. M. Robinson, 'Stability theory for systems of inequalities; part II: differentiable nonlinear systems', *SIAM J. Numer. Anal.* **13** (1976), 497–513.
[18] R. T. Rockafellar, *Conjugate duality and optimization* (Soc. Indust. Appl. Math., Philadelphia, 1974).
[19] R. T. Rockafellar, 'Generalized directional derivatives and subgradients of nonconvex functions', *Canad. J. Math.* **32** (1980), 257–280.
[20] C. Zalinescu, 'A generalization of the Farkas lemma and applications to convex programming', *J. Math. Anal. Appl.* **66** (1978), 651–678.

Mathematics Department
University of Melbourne
Parkville, Victoria, 3052
Australia