BULL. AUSTRAL. MATH. SOC. 90C30
VOL. 16 (1977), 325-339.

Lagrangean conditions and quasiduality

B.D. Craven

For a constrained minimization problem with cone constraints, lagrangean necessary conditions for a minimum are well known, but are subject to certain hypotheses concerning cones. These hypotheses are now substantially weakened, but a counter example shows that they cannot be omitted altogether. The theorem extends to minimization in a partially ordered vector space, and to a weaker kind of critical point (a quasimin) than a local minimum. Such critical points are related to Kuhn-Tucker conditions, assuming a constraint qualification; in certain circumstances, relevant to optimal control, such a critical point must be a minimum. Using these generalized critical points, a theorem analogous to duality is proved, but neither assuming convexity, nor implying weak duality.

1. Introduction

A local minimum of a constrained differentiable minimization problem may be described by lagrangean necessary conditions [10], which extend to objective functions taking values in a partially ordered space. The necessary conditions still hold for a critical point, called a *quasimin* in [6], weaker than a local minimum; and they are also sufficient [3], [11] under additional convexity hypotheses. However, [10] and [6] assume that a cone S, in a constraint $-g(x) \in S$, has an interior; this excludes the cone L_+^p of non-negative functions in an L^p-space, important for optimal control.

Received 6 December 1976.

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

A Fritz John necessary condition for a quasimin is now proved (Theorem 1), with a weakened hypothesis on S ; but a counter example shows that some restriction is necessary (and $S = L_+^p$ is still excluded, unless g is restricted). A *quasimin* was defined in [6] using differentiable arcs, which limits its applicability to optimal control problems; it is now reformulated more generally. A quasimin is necessary for the Kuhn-Tucker conditions to hold (generalized to an objective function taking values in a partially ordered space), and is also sufficient if an extended Kuhn-Tucker constraint qualification is assumed (Theorem 2). While a quasimin does not generally imply a local minimum, it does for a substantial class of problems occurring in optimal control (Theorem 5); optimal control applications will be discussed elsewhere. For real objective functions, a kind of duality relation exists, called *quasiduality* (Theorem 3), between a quasimin of a minimization problem (which need not be convex) and a quasimax of a related maximization problem; to each quasimin of the given problem, there corresponds a quasimax of the quasidual, with the same objective value. No convexity assumptions are made, but there is no global weak duality property.

The following simple example, with $x, u, \lambda, \mu \in \mathbb{R}$, illustrates the phenomena. Applying to the nonconvex problem

(a) Minimize $x - x^2$ subject to $x \geq 0$,

the construction which yields the dual for a convex problem generates here a "dual" with objective function $u - u^2 - \lambda u$ and constraints $\lambda \geq 0$ and $1 - 2u - \lambda = 0$; so the "dual" is equivalent to the problem:

(b) Maximize u^2 subject to $u \leq \frac{1}{2}$,

after substituting for λ . Now (a) has a minimum of 0 at $x = 0$; correspondingly, at $u = 0$, (b) has a quasimax described by $u^2 - 0^2 \leq o(|u-0|)$. (A maximum would require $u^2 - 0^2 \leq 0$. This instance of a quasimax happens also to be a local minimum.) Also (a) has a quasimin of $\frac{1}{4}$ at $x = \frac{1}{2}$, described by $(x-x^2) - (\frac{1}{2}-(\frac{1}{2})^2) \geq o(|x-\frac{1}{2}|)$; correspondingly, at $u = \frac{1}{2}$, (b) has a quasimax (in fact a maximum) of $\frac{1}{4}$. Thus the critical points of (a) and (b) correspond in pairs, with zero "duality gaps"; this is the typical situation, for nonconvex problems.

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

But there is no weak duality: $x \geq 0$ and $u \leq \frac{1}{2}$ do not imply that $(x-x^2) \geq u^2$.

2. Preliminary results

Let X, Y, Z, W be real normed spaces, and X_0 an open subset of X ; X' denotes the dual space of X , and $L(X, Y)$ denotes the space of continuous linear maps from X into Y ; $R_+ = [0, \infty)$. For a function $\omega : X_1 \to Y$, where $0 \in X_1 \subset X$, $\omega(\xi) = o(\|\xi\|)$ means that $\|\omega(\xi)\|/\|\xi\| \to 0$ as $\|\xi\| \to 0$, $\xi \in X_1$; if instead $X_1 = R_+$, $\omega(\alpha) = o(\alpha)$ means $\|\omega(\alpha)\|/\alpha \to 0$ as $\alpha \downarrow 0$. The function $g : X_0 \to Y$ is *Fréchet differentiable* at $a \in X_0$ if there is $g'(a) \in L(X, Y)$ for which

$$(*) \qquad g(a+\xi) - g(a) = g'(a)\xi + \omega(\xi) \quad \text{where} \quad \omega(\xi) = o(\|\xi\|) \ ;$$

continuously Fréchet differentiable if also $g'(.)$ is continuous on X_0 ; *Hadamard differentiable* at $a \in X_0$ if (*) is replaced by

$$\|g \circ \zeta(a)-g(a)-g'(a) \circ \zeta'(0)\alpha\|/\alpha \to 0 \quad \text{as} \quad \alpha \downarrow 0 \ ,$$

for each continuous arc $\alpha \mapsto \zeta(\alpha)$ $(\alpha \in R_+)$ such that $\zeta(0) = a$ and the Fréchet derivative $\zeta'(0)$ exists. Clearly Fréchet implies Hadamard.

Let $S \subset Y$, $T \subset Z$, and $P \subset W$ be convex cones. The *dual cone* of S is the convex cone $S^* = \{y' \in Y' : y'(S) \subset R_+\}$; int S denotes the interior (perhaps empty) of S . A set $B \subset S^*$ is a *compact base* for S^* if B is weak * compact in Y' , $0 \notin B$, and $S^* = \{\alpha b : \alpha \in R_+, b \in B\}$. The cone S^* will be called *representable* if S^* possesses a convex weak * compact base. This is so, in particular, if int S is nonempty (see Lemma 3 below). More generally, S^* is representable, by [13, Theorem 3], if S^* is locally compact in the relative weak * topology of Y' .

Assume that int $P \neq \emptyset$; let $f : X_0 \to W$ be continuous; let $Q \subset X_0$. Following the definition in [2], $f(x)$ has a (local) *minimum* at $x = a \in Q$, subject to the constraint $x \in Q$, if $f(x) - f(a) \notin -\text{int } P$ whenever $x \in Q$ and $\|x-a\|$ is sufficiently small. (If $W = R$ and $P = R_+$, this reduces to $f(x) - f(a) \geq 0$.) The point $a \in Q$ will be called a *quasimin* of $f(x)$, subject to $x \in Q$, if for some $\theta(x) = o(\|x-a\|)$ (as $x \to a, x \in Q$),

$$f(x) - f(a) - \theta(x) \notin -\text{int } P .$$

If $P = \mathbb{R}_+$, an equivalent requirement is that

$$\lim_{x \to a, x \in Q} \inf \; [f(x)-f(a)]/\|x-a\| \geq 0 .$$

The present definition supersedes a more complicated, and restricted, definition, given in [6] in terms of arcs. A *quasimax* of $f(x)$ occurs if and only if $-f(x)$ has a quasimin, subject to the same constraint.

Let $h : X_0 \to Z$ be Hadamard differentiable. The system $-h(x) \in T$ is *locally solvable* at the point a (see [6]) if $-h(a) \in T$ and, for some $\delta > 0$, whenever the direction d satisfies

$$\|d\| < \delta \text{ and } h(a) + h'(a)d \in -T ,$$

there exists a solution $x = a + \alpha d + o(\alpha)$ to $-h(x) \in T$, valid for all sufficiently small $\alpha > 0$. If $-h(x) \in T$ consists of finitely many scalar equations and inequalities, then local solvability of $-h(x) \in T$ is readily shown to be equivalent to the Kuhn-Tucker constraint qualification. Thus local solvability generalizes the Kuhn-Tucker constraint qualification to more general (cone and infinite-dimensional) constraints. Suppose that $h(a)\beta + h'(a)d \in -T$ for some $\beta \in \mathbb{R}$, and that $-h(x) \in T$ is locally solvable. For sufficiently large $\gamma > 0$, $\beta + \gamma > 0$ and $\|d'\| < \delta$, where $d' = (\beta+\gamma)^{-1}d$; also $(\beta+\gamma)h(a) + h'(a)d \in -T$, so $h(a) + h'(a)d' \in -T$. Hence $-h(x) \in T$ has a solution $x = a + \alpha d' + o(\alpha)$. Hence $x = a + \alpha d + o(\alpha)$ is a solution.

Let B be a (weak $*$) compact subset of Y'. Denote by $C(B)$ the space of continuous (from the weak $*$ topology of B) real functions on B, with the supremum norm. It is readily shown that the cone of non-negative functions in $C(B)$ has nonempty interior.

Let $E \subset X$ be convex, and let $S \subset Y$ be a convex cone; then the function $f : E \to Y$ is *S-convex* if, whenever $u, v \in E$ and $0 < \lambda < 1$,

$$\lambda f(u) + (1-\lambda)f(v) - f(\lambda u+(1-\lambda)v) \in S .$$

In particular, a linear function is S-convex.

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

LEMMA 1. *Let* X *and* Y *be normed spaces,* $S \subset Y$ *a convex cone with* $\text{int}\, S \neq \emptyset$ *,* $E \subset X$ *convex, and let* $f : E \to Y$ *be* S*-convex. Then either* $-f(x) \in \text{int}\, S$ *for some* $x \in E$ *, or* $(p \circ f)(E) \subset \mathrm{R}_+$ *for some nonzero* $p \in S^*$ *, but not both.*

Proof. If both systems have solutions, x respectively p , then both $(p \circ f)(x) < 0$ and $(p \circ f)(x) \geq 0$, a contradiction. Assume that there is no $x \in E$ with $-f(x) \in \text{int}\, S$. Then $H = f(E) + \text{int}\, S$ is an open convex set with $0 \notin H$, so by the separation theorem for convex sets ([17], page 64), there is a nonzero $p \in Y'$ with $p(H) \subset \mathrm{R}_+$. If $s \in \text{int}\, S$ and $x \in E$, then $s - \lambda^{-1}f(x) \in \text{int}\, S$ for λ large enough, so $\lambda s \in H$, so $p(s) \geq 0$. Since p is continuous, $p(S) \subset \mathrm{R}_+$. Also, for each $\varepsilon > 0$, $f(x) + \varepsilon s \in H$, so $(p \circ f)(x) \geq -p(\varepsilon s) \to 0$ as $\varepsilon \downarrow 0$.

LEMMA 2 (Generalized Motzkin alternative theorem [5]). *Let* X, Y, Z *be normed spaces,* $A \in \mathrm{L}(X, Z)$ *and* $B \in \mathrm{L}(X, Y)$ *,* $S \subset Y$ *and* $T \subset Z$ *convex cones, with* $\text{int}\, S \neq \emptyset$ *,* T *closed, and* $A^T(T^*)$ *weak* $*$ *closed. Then either*

(i) $-Ax \in T$ *,* $-Bx \in \text{int}\, S$ *, for some* $x \in X$ *, or*

(ii) $p \circ B + q \circ A = 0$ *for some* $q \in T^*$ *and some nonzero* $p \in S^*$ *, but not both.*

Proof. Set $f = B$ and $E = -A^{-1}(T)$. By Lemma 1, *(i)* does not hold if and only if $(\exists 0 \neq p \in S^*)$ $(p \circ B)(E) \subset \mathrm{R}_+$, thus if and only if $-Ax \in T \Rightarrow (p \circ B)(x) \in \mathrm{R}_+$. But this is equivalent, by the generalized Farkas Theorem (see [14], and [8], Theorem 6) since T and $A^T(T^*)$ are closed, to $p \circ B = q \circ (-A)$ for some $q \in T^*$, which is *(ii)*.

LEMMA 3. *Let* S *be a closed convex cone in the normed space* Y *; let* $\text{int}\, S \neq \emptyset$ *. Then the dual cone* S^* *has a convex (weak* $*$ *) compact base.*

Proof. Let $h \in \text{int}\, S$; then $h + N \subset S$ for some neighbourhood N of zero in Y . Let $0 \neq v \in S^*$; then $vh \geq 0$ and, if $vh = 0$, then $v(N) = v(h+N) \subset \mathrm{R}_+$; but, given $v \neq 0$, $vn < 0$ for some $n \in N$. The contradiction shows that $vh > 0$ for each nonzero $v \in S^*$. Setting $B = \{v \in S^* : vh = 1\}$, it follows that $S^* = \{\alpha b : \alpha \in \mathrm{R}_+, b \in B\}$; also $0 \notin B$, and B is convex and weak $*$ closed. If B is also bounded in

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

norm, then B is weak $*$ compact, from the Banach-Steinhaus Theorem. If $b \in B$, then $bh = 1$ and $b(h+N) \subset [0, \infty)$; hence $b(N) \subset [-1, \infty)$. So, for each $n \in N$, $bn \geq -1$ and $b(-n) \geq -1$; hence $\|b\| \leq \beta$ where β depends only on N .

3. Necessary conditions for a quasimin

THEOREM 1. *Let* X, Y, Z, W *be real Banach spaces,* X_0 *an open subset of* X ; *let* $P \subset W$, $S \subset Y$, $T \subset Z$ *be convex cones, with* $\text{int}\, P \neq \emptyset$, S *closed,* S^* *representable; let the functions* $f : X_0 \to W$, $g : X_0 \to Y$, *and* $h : X_0 \to Z$ *be Hadamard differentiable; let* $-h(x) \in T$ *be locally solvable at* $a \in X_0$, *and let the convex cone* $N = [h'(a) \quad h(a)]^T(T^*)$ *be weak* $*$ *closed in* $X' \times \mathbb{R}$. *Then a necessary condition for* $f(x)$ *to have a quasimin at* $x = a$, *subject to the constraints* $-g(x) \in S$ *and* $-h(x) \in T$, *is that, for some* $u \in P^*$, $v \in S^*$, $w \in T^*$, *with* u *and* v *not both zero,*

(FJ) $$uf'(a) + vg'(a) + wh'(a) = 0 \; ; \quad vg(a) = 0 \; ; \quad wh(a) = 0 \; .$$

Proof. By hypothesis, S^* has a (weak $*$) compact convex base B . From the separation theorem for convex sets,

$$-g(x) \in S \Leftrightarrow (\forall v \in S^*) \; -vg(x) \geq 0 \Leftrightarrow (\forall b \in B) \; -bg(x) \geq 0 \Leftrightarrow -G(x) \in K \; ,$$

where $G : X_0 \to C(B)$ is defined by $(\forall x \in X_0, \forall b \in B) \; G(x)(b) = bg(x)$, and $K = \{\psi \in C(B) : \psi(B) \subset \mathbb{R}_+\}$. Then $\text{int}\, K \neq \emptyset$; and G is Hadamard differentiable.

Suppose that the linear system $-Aq \in T$, $-Bq \in \text{int}\, V$, where

$$A = [h'(a) \quad h(a)] \; , \quad B = \begin{bmatrix} f'(a) & 0 \\ G'(a) & G(a) \end{bmatrix} \; , \quad V = \begin{bmatrix} P \\ K \end{bmatrix} \; ,$$

has a solution $q = (d, \beta) \in X \times \mathbb{R}$. Then $-f'(a)d \in \text{int}\, P$, $-g'(a)d - g(a)\beta \in \text{int}\, S$, $-h'(a)d - h(a)\beta \in T$. From the last, local solvability gives a solution $x = x(\alpha) \equiv a + \alpha d + o(\alpha) \quad (\alpha \downarrow 0)$ to $-h(x) \in T$. Then, for sufficiently small $\alpha > 0$, $-h\big(x(\alpha)\big) \in T$ and

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

$$\begin{aligned}-G\big(x(\alpha)\big) &= -G(a) - \alpha G'(a)d + o(\alpha) \\ &= (1-\alpha\beta)[-G(a)] + \alpha[-G'(a)d-G(a)\beta] + o(\alpha) \\ &\in K + \text{int } K + o(\alpha) \subset K .\end{aligned}$$

The quasimin therefore requires that $f\big(x(\alpha)\big) - f(a) - \sigma(\alpha) \notin -\text{int } P$ for some $\sigma(\alpha) = o(\alpha)$; hence $f'(a)d \notin -\text{int } P$, contradicting $f'(a)d \in -\text{int } P$ obtained above.

Hence the linear system has no solution q . Since also the cone N is closed, Lemma 2 shows that, for some nonzero $y = (u, \lambda) \in V^*$ (thus $u \in P^*$ and $\lambda \in K^*$) and some $w \in T^*$, $wA + yB = 0$. Hence

$$uf'(a) + \lambda G'(a) + wh'(a) = 0 \ ; \quad g(a) = 0 \ ; \quad wh(a) = 0 .$$

If $\lambda = 0$, then $u \neq 0$, and so (FJ) holds with $v = 0$. Suppose that $\lambda \neq 0$. Since $\lambda \in \big(C(B)\big)'$, the Riesz representation theorem represents λ by a signed measure μ , such that $\lambda\psi = \int_B \mu(db)\psi(b)$ for each $\psi \in C(B)$. Then $\lambda \in K^*$ requires that $\mu(E) \geq 0$ for each Borel subset $E \subset B$. Since $\lambda \neq 0$, $\mu(B) > 0$. For each $x \in X$, $G'(a)x$ maps $b \in B$ to $bg'(a)x$. Hence $\lambda G'(a)x = \int_B \mu(db)bg'(a)x = \mu(B)b^*g'(a)x$ where $b^* = \int_B [\mu(B)]^{-1}\mu(db)b$. So b^* is the weak $*$ limit of a net of approximative sums of the form $\sum_i \gamma_i b_i$ where each $b_i \in B$, $\gamma_i > 0$, $\sum_i \gamma_i = 1$. Hence b^* is in the closed convex hull of B , and hence in B , since B is convex compact. Hence $\lambda G'(a) = vg'(a)$ where $v = \mu(B)b^* \in S^*$; $v \neq 0$ since $0 \notin B$ and $\mu(B) > 0$. Similarly $\lambda G(a) = 0$ implies $\int_B \mu(db)bg(a) = 0$, which implies $vg(a) = 0$. Thus (FJ) is proved.

DISCUSSION. Theorem 4 of [*10*] is applicable since the equivalent constraint $-G(x) \in K$ has $\text{int } K \neq \emptyset$. However, the following counter example shows that some restriction on S is required. (Hence Theorem 5.11 of Dempster [*12*] requires an additional hypothesis.) A similar example is possible with L^2 replacing l^2 .

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

Let l^2 denote real Hilbert sequence space. Define a continuous linear map $M : l^2 \to l^2$ as the map taking $x = (x_1, x_2, \ldots) \in l^2$ to $Mx = (\alpha_1 x_1, \alpha_2 x_2, \ldots)$ where $\alpha_n = n^{-2}$. Note that M is not an open map, and hence the subspace $M(l^2)$ is not closed in l^2 . Let Q denote the convex cone $Q = \{(x_1, x_2, \ldots) \in l^2 : (\forall n)\ x_n \geq 0\}$. Then $Q^* = Q$, identifying $(l^2)'$ with l^2 . It is readily shown that $\text{int}\ Q = \emptyset$. Let $f = (\alpha_1, \alpha_2, \ldots) \in (l^2)'$. Since $Mx \in Q \Leftrightarrow (\forall n)\ x_n \geq 0$, $f(x)$ is minimized, subject to $x \in l^2$ and $Mx \in Q$, at $x = 0 = (0, 0, \ldots)$. If (FJ) holds at this minimum, then there exist $\tau \geq 0$ and $v = (v_1, v_2, \ldots) \in Q^*$, not both zero, for which $\tau f = vM$. Hence $\tau\alpha_n = v_n\alpha_n$ for each $n = 1, 2, \ldots$. Since $\{v_n\} \to 0$ and $\alpha_n > 0$, $\tau = 0$; hence also $(\forall n)\ v_n = 0$, so $v = 0$ and $\tau = 0$. So (FJ) does *not* hold here.

Consider the minimization problem of Theorem 1 with h and T omitted, and with $g(x) = -Mx$, $S = Q$. The example shows that S cannot then be unrestricted. If, instead, g and S are omitted, and $h(x) = -Mx$, $T = Q$, then the linear constraint $Mx \in Q$ is locally solvable; so the example shows that the hypothesis that N is closed cannot be omitted.

4. Conditions necessary and sufficient for a quasimin

Consider now the constraints $-g(x) \in S$ and $-h(x) \in T$ combined into a single constraint $k(x) \in K$. Assume that K is a closed convex cone in $V = Y \times Z$. The problem of minimizing $f(x)$ subject to $k(x) \in K$ satisfies the *generalized Kuhn-Tucker condition* at the point $a \in X_0$ if $k(a) \in K$, and for some $\lambda \in K^*$ and some nonzero $\tau \in P^*$,

$$\tau f'(a) = \lambda k'(a) \ ; \quad \lambda k(a) = 0 \ .$$

In particular, if $W = \mathrm{R}$ and $P = \mathrm{R}_+$, then $\tau = 1$ can be assumed, and the usual Kuhn-Tucker condition is recovered.

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

THEOREM 2. *Let* X, V, W *be real Banach spaces,* X_0 *an open subset of* X *; let* $P \subset W$ *and* $K \subset V$ *be closed convex cones, with* $\text{int } P \neq \emptyset$ *; let* $f : X_0 \to W$ *be Fréchet differentiable at* $a \in X_0$ *, and let* $k : X_0 \to V$ *be continuously Fréchet differentiable. If the generalized Kuhn-Tucker condition holds at* a *, then* $f(x)$ *has a quasimin at* $x = a$ *, subject to the constraint* $k(x) \in K$ *. The converse holds under the additional hypotheses that the convex cone* $N_0 = [k'(a) \;\; k(a)]^T(K^*)$ *is (weak *) closed in* $X' \times \mathrm{R}$ *and that the set* $U = k(a) + k'(a)(X) - K$ *contains a neighbourhood of zero.*

Proof. Let the generalized Kuhn-Tucker condition hold; let $k(x) \in K$; then $\lambda k(x) \geq 0$; setting $z = x - a$,

$$\tau f'(a)z = \lambda k'(a)z = \lambda k(x) - \lambda k(a) + \phi(z) \geq \phi(z)$$

where $\phi(z) = o(\|z\|)$. Since $0 \neq \tau \in W'$, there is $w \in W$ with $\tau w \neq 0$; setting $\psi = -(\tau w)^{-1}\phi w$, $\tau\psi = -\phi$, so that $\tau[f'(a)z+\psi(z)] \geq 0$, where $\psi(z) = o(\|z\|)$. If $x = a$ is not a quasimin, then there is some sequence $\{z_n\} \to 0$ for which $k(a+z_n) \in K$, and whenever $\theta(z) = o(\|z\|)$,

$$f(a+z_n) - f(a) - \theta(z_n) \in -\text{int } P .$$

Now $f(a+z_n) - f(a) = f'(a)z_n + o(\|z_n\|)$; so, choosing θ suitably, $f'(a)z_n + \psi(z_n) \in -\text{int } P$ as $n \to \infty$, hence $\tau|f'(a)z_n+\psi(z_n)| < 0$ as $n \to \infty$; the contradiction shows that $x = a$ is a quasimin.

Conversely, assume a quasimin, let N_0 be closed, and let U contain a neighbourhood. The hypothesis on U , and continuous differentiability of k , imply ([16], Corollary 1, and [6], Theorem 3) that $k(x) \in K$ is locally solvable. Then the generalized Kuhn-Tucker condition follows from Theorem 1, with g and S omitted; since v is absent, $\tau \equiv u \neq 0$. (For this converse, f need only be Hadamard differentiable.)

5. Quasiduality

In this section only, let $W = \mathrm{R}$ and $P = \mathrm{R}_+$. Consider the two problems:

(A) Minimize $F(x)$ subject to $x \in \mathrm{A}$;

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

(B) Maximize $\Phi(y)$ subject to $y \in B$.

Problem (B) will be called a *quasidual* of (A) if the following condition holds:

if (A) has a quasimin at $x = \xi \in A$, then (B) has a quasimax at some $y = \eta \in B$, and $F(\xi) = \Phi(\eta)$.

Under additional hypotheses of convexity (or related properties), which are *not* made here, a quasimin is necessarily a minimum, and a quasimax is a maximum, and quasiduality implies the usual duality.

Consider the following pair of problems:

(QP) $\text{quasimin}_x \; f(x)$ subject to $k(x) \in K$;

(QD) $\text{quasimax}_{u,v} \; f(u) - vk(u)$ subject to $v \in K^*$,

$f'(u) - vk'(u) = 0$.

THEOREM 3. *Let* $f : X_0 \to R$ *be Hadamard differentiable; let* k *be continuously Fréchet differentiable; as in Theorem* 2, *let* N_0 *be closed and let* U *contain a neighbourhood of zero. Let* (QP) *have a quasimin at* $x = a \in X_0$. *Then* (QD) *is a quasidual of* (QP).

Proof. Let (u, v) satisfy the constraints of (QD); let (QP) have a quasimin at $x = a$; from Theorem 2, the Kuhn-Tucker condition holds for (QP) at $x = a$, for some $\lambda \in K^*$. Set $u = a + p$ and $v = \lambda + q$. Then

$$\begin{aligned} f(a) - [f(u)-vk(u)] &= f(a) - f(a+p) - vk(a+p) \\ &= -f'(a)p - o(\|p\|) + vk(a) + (\lambda+q)\big(k'(a)p+o(\|p\|)\big) \\ &= -[f'(a)-\lambda k'(a)]p - o(\|p\|) + vk(a) + o(\|p\|+\|q\|) \\ &\geq o(\|p\|+\|q\|) \; . \end{aligned}$$

Hence (QD) has a quasimax at $(u, v) = (a, \lambda)$. Since also $\lambda k(a) = 0$, by the Kuhn-Tucker condition, $f(a) - \lambda k(a)$, so (QD) is a quasidual of (QP).

There is also a *converse quasiduality* result, analogous to the converse duality results of [9], and [*4*, Theorem 3.1]. These cited results however assume convexity, which is not required here. Note that (A) is a quasidual of (B) if, whenever (B) has a quasimax, (A) has a corresponding quasimin, with equal values of the two objective functions.

THEOREM 4. *Let* f *and* k *be twice continuously Fréchet*

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

differentiable; let (QD) *have a quasimax at* $(u, v) = (a, \lambda)$ *; let the adjoint* M^T *of the linear map* $M = f''(a) - \lambda k''(a)$ *be bijective. Then* (QP) *is a quasidual of* (QD).

Proof. Since f and k are twice continuously differentiable, and M^T is invertible, the constraints of (QD) are locally solvable, and the cone-closure hypothesis of Theorem 1 is fulfilled for (QD). Hence (FJ) holds for (QD) at (a, λ) . The calculation of [4, Lemma 3.1], then applies, given M^T bijective, showing that $k(a) \in K$ and the Kuhn-Tucker condition holds for (QP). From Theorem 2, the Kuhn-Tucker condition implies a quasimin for (QP). Since also $f(a) = f(a) - \lambda k(a)$, (QP) is a quasidual to (QD).

6. When does a quasimin imply a minimum?

Let I be a compact subset of R^p . Let $X = L^1(I, \mathrm{R}^n)$, the space of measurable functions x from I into R^n , having finite $L^1(I)$-norm $\|x\| = \int_I |x(t)|dt$, where $|\cdot|$ denotes euclidean norm in R^n , and dt denote Lebesgue measure on I . Let X_0 be an open subset of the space X . Define $f : X_0 \to \mathrm{R}^r$ by $f(x) = \int_I h(x(t), t)dt$, where the function $h : \mathrm{R}^n \times I \to \mathrm{R}^r$ is continuous. Define minimum and quasimin of $f(x) \in \mathrm{R}^r$ in terms of the cone R^r_+ , the nonnegative orthant in R^r . Let k map X_0 into a space of continuous M-valued functions on I , where M is a normed space; let S be a convex cone in M . Denote by K the convex cone consisting of those continuous functions $\psi : I \to M$ for which $\psi(t) \in S$ for each $t \in I$. Then $k(x) \in K$ iff $(\forall t \in I)\ k(x)(t) \in S$.

Consider the minimization problem:

(P*) $$\underset{x \in X_0}{\text{minimize}}\ f(x) \quad \text{subject to} \quad k(x) \in K ,$$

with f, k, K as specified above. This is an abstract version of an optimal control problem (see, for example, [10]).

The following measure properties will be required (see [15, Section

42]). A point $t_0 \in I$ is a *point of density* of a measurable set $E \subset I$ if $\sup_{J_k \to t_0} \left[\limsup_k m(J_k \cap E)/m(J_k)\right] = 1$, taking limits over sequences $\{J_k\}$ of intervals containing t_0. A function $g : I \to \mathbb{R}$ is *approximately continuous* at t_0 if there is a measurable set $E_0 \subset I$ such that t_0 is a point of density of E_0 (and hence $m(E_0) > 0$), and also

$$\lim_{t \to t_0, t \in E_0} g(t) = g(t_0) .$$

THEOREM 5. *Let* (P^*) *have a quasimin at* $x = \eta \in k^{-1}(K)$; *let* X *have the* $L^1(I)$*-norm; let* h *satisfy a Lipschitz condition*

$$|h(u, t)-h(v, t)| \le c|u-v| \quad (u, v \in \mathbb{R}^n) .$$

Then $h(x(t), t)$ *is minimized, almost everywhere in* I , *with respect to* $x \in k^{-1}(K)$, *at* $x = \eta$. *Consequently* (P^*) *has a minimum at* $x = \eta$.

Proof. If the conclusion does not hold, then for some $x^* \in k^{-1}(K)$ and some $A^\# \subset I$, with measure $m(A^\#) > 0$,

(i) $$(\forall t \in A^\#) \quad h(x^*(t), t) - h(\eta(t), t) \in -\text{int}\, \mathbb{R}^r_+ .$$

The Lipschitz hypothesis shows that, for each component h_i of h , there is a bounded measurable function ϕ_i such that

$$(\forall t \in I) \quad h_i(x^*(t), t) - h(\eta(t), t) = \phi_i(t)|x^*(t)-\eta(t)| .$$

(Where $x^*(t) = \eta(t)$, $\phi_i(t) = 0$.) From [15, Theorem 42.3], ϕ_i is approximately continuous almost everywhere on I , and [15, Theorem 42.2], shows that the points of a measurable set $E \subset I$, with $m(E) > 0$, are almost everywhere points of density of E . Deleting from $A^\#$ the finitely many subsets on which ϕ_i is not approximately continuous $(i = 1, 2, \ldots, r)$, and the set of points which are not points of density of $A^\#$, leaves a set A , where $m(A) = m(A^\#)$. Let $t_0 \in A_0$. Then

https://doi.org/10.1017/S0004972700023431 Published online by Cambridge University Press

(ii) $$\lim_{t \to t_0, t \in A} \phi_i(t) = \phi_i(t_0) < 0 ,$$

by the approximate continuity, and also using (i). Consequently, for some $\delta > 0$ and each $t \in B = \{t \in A : |t-t_0| < \delta\}$, and each i , $-\phi_i(t) \geq \beta \equiv \frac{1}{2} \min_i |-\phi_i(t_0)| > 0$.

Define the continuous (nondifferentiable) arc $\lambda \mapsto \xi_\lambda$ $(\lambda \in \mathbb{R}_+)$ by

$$\xi_\lambda(t) = x^*(t) \quad \text{for } t \in B_\lambda \equiv \{t \in B : |t-t_0| < \psi(\lambda)\} ,$$

$$\xi_\lambda(t) = \eta(t) \quad \text{otherwise,}$$

where $\psi(\lambda)$ is chosen so that, as $\lambda \downarrow 0$, the $L^1(I)$-norm $\|\xi_\lambda - \eta\| = \lambda$. Then $\xi_0 = \eta$, and the form chosen for the constraint $k(x) \in K$ ensures that $\xi_\lambda \in k^{-1}(K)$. Then, for each i ,

(iii) $$-f_i(\xi_\lambda) + f_i(\eta) \geq \int_{B_\lambda} \beta|x^*(t) - \eta(t)|dt \geq \beta\|\xi_\lambda - \eta\| .$$

Hence $f(\xi_\lambda) - f(\eta)$ lies in $-\text{int } \mathbb{R}_+^r$, and is distant at least $\beta\|\xi_\lambda - \eta\|$ from the boundary of $-\text{int } \mathbb{R}_+^r$. This contradicts the quasimin of (P^*) at $x = \eta$. Thus (i) cannot hold.

Integrating h then shows that (P^*) is minimized at $x = \eta$.

REMARKS. This theorem depends on the choice of the $L^1(I)$-norm. It extends a result given by Berkovitz [1, p. 288]. The case $p > 1$ corresponds to an optimal control problem involving a partial differential equation; this will be detailed elsewhere.

The Lipschitz hypothesis need only hold almost everywhere. The theorem also holds, with slight change to the proof, if instead f is Hadamard differentiable.

References

[1] L.D. Berkovitz, *Optimal control theory* (Applied Mathematical Sciences, 12. Springer-Verlag, New York, Heidelberg, Berlin, 1974).

[2] B.D. Craven, "Nonlinear programming in locally convex spaces", *J. Optimization Theory Appl.* 10 (1972), 197-210.

[3] B.D. Craven, "Sufficient Fritz John optimality conditions", *Bull. Austral. Math. Soc.* 13 (1975), 411-419.

[4] B.D. Craven, "Converse duality in Banach space", *J. Optimization Theory Appl.* 17 (1975), 229-238.

[5] B.D. Craven, "A generalized Motzkin alternative theorem" (Research Report, 13. School of Mathematical Sciences, University of Melbourne, Parkville, 1975).

[6] B.D. Craven, "Lagrangean conditions for optimization in Banach spaces" (Research Report, 6. School of Mathematical Sciences, University of Melbourne, Parkville, 1976).

[7] B.D. Craven, "Alternative theorems with convex cones" (Research Report, 15. School of Mathematical Sciences, University of Melbourne, Parkville, 1976).

[8] B.D. Craven and J.J. Koliha, "Generalizations of Farkas's theorem", *SIAM J. Math. Appl.* (to appear).

[9] B.D. Craven and B. Mond, "On converse duality in nonlinear programming", *Operations Res.* 19 (1971), 1075-1078.

[10] B.D. Craven and B. Mond, "Transposition theorems for cone-convex functions", *SIAM J. Appl. Math.* 24 (1973), 603-612.

[11] B.D. Craven and B. Mond, "Sufficient Fritz John optimality conditions for nondifferentiable convex programming", *J. Austral. Math. Soc. Ser. B* (to appear).

[12] M.H.A. Dempster, "Lectures on abstract optimization and its applications" (Department of Mathematics, School of Mathematical Sciences, University of Melbourne, Parkville, 1975).

[13] D.A. Edwards, "On the homeomorphic affine embedding of a locally compact cone into a Banach dual space endowed with the vague topology", *Proc. London Math. Soc.* (3) 14 (1964), 399-414.

[14] Leonid Hurwicz, "Programming in linear spaces", *Studies in linear and non-linear programming*, 38-102 (Stanford Mathematical Studies in Social Sciences, II. Stanford University Press, Stanford, California, 1958).

[15] M.E. Munroe, *Introduction to measure and integration* (Addison-Wesley, Cambridge, Massachusettes, 1953).

[16] Stephen M. Robinson, "Stability theory for systems of inequalities, Part II: Differentiable nonlinear systems" (MRC Technical Summary Report, 1388. University of Wisconsin-Madison, Mathematics Research Center, Madison, Wisconsin, 1974).

[17] Helmut H. Schaefer, *Topological vector spaces* (Graduate Texts in Mathematics, 3. Springer-Verlag, New York, Heidelberg, Berlin, 1971).

Department of Mathematics,
University of Melbourne,
Parkville,
Victoria.