BULL. AUSTRAL. MATH. SOC. 90C48, 49D37
VOL. 30 (1984), 193-218.

GENERALIZED CONVEXITY IN NONDIFFERENTIABLE PROGRAMMING

BEVIL M. GLOVER

For an abstract mathematical programming problem involving quasidifferentiable cone-constraints we obtain necessary (and sufficient) optimality conditions of the Kuhn-Tucker type without recourse to a constraint qualification. This extends the known results to the non-differentiable setting. To obtain these results we derive several simple conditions connecting various concepts in generalized convexity not requiring differentiability of the functions involved.

1. Introduction

A great deal of recent research in mathematical programming has been devoted to the problem of establishing optimality conditions of the Kuhn-Tucker type for abstract programming problems without requiring a constraint qualification to be explicitly satisfied (see, for example, [13], [5], [10]). Craven and Zlobec [13] and Borwein and Wolkowicz [5] considered this problem for abstract differentiable convex programs and obtained necessary Lagrangian conditions for a minimum. Zlobec and Jacobson [36] obtained related results for problems involving a finite number of convex constraints and an arbitrary differentiable objective function. These results were further extended to abstract programming problems with arbitrary differentiable objective function and pseudoconvex constraints in Craven, Glover and Zlobec [10]. The basic tool which has

Received 16 March 1984.

enabled these results to be obtained is a 'stability' condition for the system of constraints, similar to the Slater constraint qualification, which is inherent in the programming problem by virtue of the mild convexity assumptions on the system. This approach was first suggested for convex programs in [5] and extended in [10].

The object of this paper is to further extend the results outlined above to a large class of nondifferentiable programming problems. We derive both necessary and sufficient optimality conditions in a Lagrangian form without requiring a constraint qualification. We discuss these conditions for both finitely constrained and abstract programming problems. In order to obtain these results in the nondifferentiable setting it was necessary to derive several results connecting various concepts of generalized convexity; these appear in section 3 and extend and simplify several results from Diewert [18]. In the final section of the paper we consider an application of these results to fractional programming problems which includes a duality result not requiring a constraint qualification.

2. Notation

Throughout this paper X and Y shall denote normed vector spaces with $X_o \subset X$ an open convex set; X' (respectively Y') denotes the continuous dual space of X (Y) equipped with the weak* topology. For a set $V \subset X$ let $cl\ V$, coV, $\overline{co}V$, and $int\ V$ denote the closure, convex hull, closed convex hull, and interior of V, respectively; the set cone $V \equiv \{\lambda x : \lambda \in R_+,\ x \in V\}$ is the cone generated by V and we shall say V is a convex cone if $V + V \subset V$ and $(\forall \lambda \in R_+)\ \lambda V \subset V$. The dual cone of V is the (weak*) closed convex cone $V^* \equiv \{v \in X' : (\forall x \in V)\ v(x) \geq 0\} \subset X'$. Since X' is endowed with the weak* topology we have $(V^*)^* = V^{**} \subset X$. The tangent cone to V at a point $a \in V$ is the set

$$T(V,a) \equiv \left\{x \in X : \exists (x_n) \subset V,\ (\lambda_n) \subset R_+\backslash\{0\} \text{ with } x_n \to a, \text{ and } \lambda_n(x_n - a) \to x\right\}$$

(for related definitions see Guignard [22] and Dempster and Wets [17]). For convenience we shall let $N(V,a) \equiv (T(V,a))^*$.

A function $f : X_o \to (-\infty,+\infty]$ is convex on X_o if

(1) $(\forall x,\ a \in X_0)(\forall \lambda \in (0,1))\ f(a+\lambda(x-a)) \le f(a) + \lambda(f(x)-f(a))$

The essential domain of f is the set $\operatorname{dom} f \equiv \{x \in X_0 : f(x) \text{ is finite}\}$.

The subdifferential of f at a point $a \in \operatorname{dom} f$ is denoted by $\partial f(a)$ where

$$\partial f(a) \equiv \left\{v \in X' : (\forall x \in X_0)\ f(x) \ge f(a) + v(x-a)\right\}.$$

Clearly $\partial f(a)$ is a weak* closed convex set; if f is continuous at a then $\partial f(a)$ is non-empty and weak* compact (see [32]).

A function is said to be sublinear if it is convex and positively homogeneous (of degree one). Let $W \subset X'$ be a weak* closed convex set, then we define the support function of W, denoted by $s(.,W) : X \to (-\infty,+\infty]$, as

$$s(x,W) \equiv \sup\left\{v(x) : v \in W\right\}, \text{ for } x \in X.$$

Clearly $s(\cdot,W)$ is a sublinear function

We can extend the concept of convexity to vector-valued functions as follows; let $S \subset Y$ be a closed convex cone then $f : X_0 \to Y$ is S-convex on X_0 if

$$(\forall x,\ a \in X_0)(\forall \lambda \in (0,1)\ \ f(a + \lambda(x-a)) - f(a) - \lambda(f(x) - f(a)) \in -S.$$

It is easily shown that f is S-convex if and only if $v \circ f$ is convex for each $v \in S^*$. For convenience we shall denote composition of mappings by juxtaposition, i.e. $v \circ f$ as vf.

A function $f : X_0 \to Y$ is locally Lipschitz on X_0 if for each point in X_0 there exists a neighbourhood U of this point and a positive real number k such that $\|f(x) - f(y)\| \le k\|x - y\|$ for all $x,\ y \in U$.

We now consider several concepts of the derivative for real-valued functions. Let $f : X_0 \to R$, then the (upper right) Dini derivative of f at $a \in X_0$ in the direction $x \in X$ is given by

$$D^+f(a,x) \equiv \limsup_{\lambda \downarrow 0} \lambda^{-1}[f(a + \lambda x) - f(a)]$$

The Clarke derivative (or generalized directional derivative) of f at a in the direction x (see Clarke [6]) is given by

$$f^o(a,x) = \limsup_{\substack{\lambda \downarrow 0 \\ y \to a}} \lambda^{-1}[f(y + \lambda x) - f(y)]$$

Clearly $D^+f(a,x) \le f^o(a,x)$, for each $a \in X_o$ and any $x \in X$. Clarke has shown that, for each $a \in X_o$, $f^o(a,\cdot)$ is a sublinear function and finite and continuous when f is locally Lipschitz. Thus if f is locally Lipschitz the subdifferential of $f^o(a,\cdot)$ is a non-empty weak* compact convex set, this set, denoted by $\partial_C f(a)$, is the Clarke subgradient of f at a.

The function $f : X_o \to R$ is said to be in class C on X_o if, for each $a \in X_o$, $f^o(a,\cdot)$ is lower semi-continuous (l.s.c) at 0 and $f^o(a,x) > -\infty$, for every $x \in X$. Thus if f is in class C it follows that (using Zalinescu [34, Prop. 1] and Rockafellar [32]) $\partial_C f(a)$ is a non-empty weak* closed convex set with $f^o(a,x) = s(x, \partial_C f(a))$. This class C is more general than the class of locally Lipschitz functions defined on X_o.

A function $h : X_o \to Y$ is directionally differentiable at $a \in X_o$ if the limit

$$h'(a,x) = \lim_{\lambda \downarrow 0} \lambda^{-1}[h(a + \lambda x) - h(a)] \tag{2}$$

exists for each $x \in X$, in the strong topology of Y. If the limit (2) exists in the weak topology of Y then h is said to be weakly directionally differentiable at a, and we denote the limit by $h'_w(a,x)$ for each $x \in X$. If h is directionally differentiable at $a \in X_o$ with $h'(a,\cdot)$ a continuous linear function then f is linearly Gâteaux differentiable at a. If $Y = R$ and h is convex then $h'(a,\cdot)$ exists and is sublinear for each $a \in X_o$ ([32]).

In section 4 we consider a class of nondifferentiable programming problems, the objective and constraint functions involved will satisfy a quasidifferentiability condition. Namely, let $S \subset Y$ be a closed convex cone then $h : X_o \to Y$ is S*-quasidifferentiable at $a \in X_o$ if h is weakly directionally differentiable at a, and for each $v \in S^*$ there is a non-empty weak* compact convex set $\tilde{\partial}(vh)(a)$ such that $(vh)'(a,x) = s(x, \tilde{\partial}(vh)(a))$ for each $x \in X$. If h is S-convex then $\tilde{\partial}(vh)(a) = \partial(vh)(a)$. The set $\tilde{\partial}(vh)(a)$ is known as the quasidifferential of vh at a. Clearly every continuous S-convex function and every linearly Gâteaux differentiable function is S*-quasidifferentiable, for further examples of classes of nondifferentiable functions satisfying this condition see Pshenichnyi [31], Borwein [4], Craven and Mond [11], and Glover [20]. If $Y = R$ and $S = R_+$ then quasidifferentiable refers to R_+-quasidifferentiable.

It should be noted that if $h : X_o \to Y$ is weakly directionally differentiable at $a \in X_o$ and $v \in X'$ then $v(h_w'(a,x)) = (vh)'(a,x)$, for each $x \in X$.

We conclude this section by giving a slightly generalized version of a result due to Clarke [7] for locally Lipschitz functions, we extend the result to functions in class C.

PROPOSITION 2.1. *Let $f : X_o \to R$ and $g : X_o \to R$ be functions, with $co\{f,g\} \subset C$. Let $a \in X_o$, then*

(3) $$\partial_C(f+g)(a) \subset cl\left[\partial_C f(a) + \partial_C g(a)\right].$$

Proof. Let $h = f + g$, then $h \in C$, and, clearly, for each $x \in X$,

(4) $$h^o(a,x) \le f^o(a,x) + g^o(a,x).$$

For convenience let $F(x) = f^o(a,x)$ and $G(x) = g^o(a,x)$ for $x \in X$. Now let $v \in \partial_C h(a)$ and $x \in X$, then using (4) above we have $v(x) \le F(x) + G(x)$; thus $v \in \partial(F+G)(0)$ (since F and G are l.s.c sublinear functions).

Now,
$$\begin{aligned}(F+G)(x) &= \sup\left\{v(x) : v \in \partial(F+G)(0)\right\} \\ &= \sup\left\{v(x) : v \in \partial F(0)\right\} + \sup\left\{v(x) : v \in \partial G(0)\right\} \\ &= \sup\left\{v(x) : v \in \partial F(0) + \partial G(0)\right\} \\ &= \sup\left\{v(x) : v \in cl\left[\partial F(0) + \partial G(0)\right]\right\}\end{aligned}$$

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

$$\begin{aligned} \text{Hence}\quad \partial_C(f+g)(a) &= \partial_C h(a) \\ &\subset \partial(F+G)(0) \\ &= cl\left[\partial F(0) + \partial G(0)\right] \\ &= cl\left[\partial_C f(a) + \partial_C g(a)\right]. \end{aligned}$$ □

It is worth noting that every linearly Gâteaux differentiable function is in class C, and, even more generally, every quasi-differentiable function is in this class. These results follow since $f'(a,.) \le f^O(a,.)$ for a directionally differentiable function $f : X_O \to R$ at any $a \in X_O$, thus if f is quasidifferentiable then $\emptyset \ne \tilde{\partial} f(a) \subset \partial_C f(a)$. Hence the Clarke subgradient, which corresponds to the subdifferential of $f^O(a,\cdot)$, is non-empty which, by [34], implies $f^O(a,\cdot)$ is l.s.c at 0.

A function $f : X_O \to Y$ is said to be radially continuous (resp. l.s.c) if, for each x, $a \in X_O$, the function $\rho(\lambda) = f(a + \lambda x)$ is continuous (l.s.c) as a function $\rho : R \to Y$.

For convenience, for a function $\theta : R \to R$ we shall denote the Dini derivative at α by $D^+\theta(\alpha)$ which is equivalent to $D^+\theta(\alpha, 1)$.

3. Generalized convexity

We begin this section by establishing a mean-value theorem for real-valued functions without requiring the existence of (one-sided) directional derivatives.

THEOREM 3.1. *Let* $\tau : [a,b] \to R$ *be a continuous function on the compact interval* $[a,b]$, $a \ne b$. *Then there is a* $\beta \in (a,b)$ *such that*

(5) $$D^+\tau(\beta)(b - a) \ge \tau(b) - \tau(a)$$

Proof. Since τ is continuous it attains its maximum and minimum at some points c and d (resp.) in $[a,b]$. We assume τ is not constant on $[a,b]$, or we find $D^+\tau(\alpha) = 0$ for all $\alpha \in (a,b)$ and (5) follows immediately.

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

We begin by assuming $\tau(a) = \tau(b)$; it then follows that for some $\beta \in (a,b)$, $D^{+}\tau(\beta) \geq 0$. Consider the following two cases:

(i) $d \neq a$. Thus $d \in (a,b)$ and $\tau(d) < \tau(a) = \tau(b)$. Hence, $\tau(d+\lambda) \geq \tau(d)$, for all $\lambda \in [0,b-d)$; consequently $D^{+}\tau(d) \geq 0$.

(ii) $d = a$. Thus $\tau(x) \geq \tau(a) = \tau(b)$, for all $x \in (a,b)$ and $\tau(c) > \tau(a)$, since τ is not constant on $[a,b]$. Now define the function

$$\mu(\gamma) = \min\left\{\tau(\alpha) : \alpha \in [\gamma,c]\right\}, \quad \text{for } \gamma \in [a,c].$$

Now consider the following cases:

(a) $\mu(\gamma) = \tau(\gamma)$, for all $\gamma \in (a,c)$. Thus, if $c \geq x \geq y \geq a$ then $x \in [y,c]$, hence $\tau(x) \geq \tau(y)$. Thus τ is non-decreasing on $[a,c]$. Hence if $\alpha \in (a,c)$ then $\tau(\alpha+\lambda) \geq \tau(\alpha)$, for all $\lambda \in [0,c-\alpha)$ and consequently $D^{+}\tau(\alpha) \geq 0$.

(b) $\mu(\gamma) \neq \tau(\gamma)$, for some $\gamma \in (a,c)$. Since τ is continuous $\mu(\gamma) = \tau(\delta)$ for some $\delta \in (\gamma,c)$. Thus $\tau(\delta+\lambda) \geq \tau(\delta)$, for all $\lambda \in (0,c-\delta)$ and so $D^{+}\tau(\delta) \geq 0$.

Thus the theorem is established for the special case $\tau(a) = \tau(b)$. For the general case define the function

$$\rho(x) = \tau(x) - \tau(b) - \left[\frac{\tau(b) - \tau(a)}{b-a}\right](x-a)$$

Clearly ρ is continuous and $\rho(a) = \rho(b) = 0$ also, for each $\alpha \in (a,b)$,

$$D^{+}\rho(\alpha) = D^{+}\tau(\alpha) - \left[\frac{\tau(b) - \tau(a)}{b-a}\right].$$

Thus, by the special case above, there is a $\beta \in (a,b)$ with $D^{+}\rho(\beta) \geq 0$, which is equivalent to (5). □

REMARK 3.2. Theorem 3.1 remains valid provided τ is l.s.c and satisfies the following condition:

(6) $$(\exists c \in [a,b])\ \tau(c) = \sup\{\tau(\alpha) : \alpha \in [a,b]\}$$

It is easily shown (see [18]) that (6) is satisfied when τ is quasi-convex; Thus (6) is true when $L_{\tau}(\alpha) \equiv \{x \in R : \tau(x) \leq \alpha\}$ is closed and

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

convex for each $\alpha \in R$. Theorem 3.1 improves upon the mean-value theorem of Diewert [18] in which (5) was shown to hold for some $\beta \in [a,b)$ (under hypotheses closely related to condition (6) above). If τ possesses a right derivative at each point $\alpha \in [a,b)$ (possibly infinite) then $D^{+}\tau(\alpha) = \tau'_{+}(\alpha)$ and the result of Theorem 3.1 can be found in Flett [19].

We now establish the Generalized Mean-Value Theorem for functions defined on abstract spaces.

THEOREM 3.3 (Generalized Mean-Value Theorem)

Let $f : X_o \to R$ *be radially continuous. Let* $x, a \in X_o$. *Then there exists a* $\tilde{\lambda} \in (0,1)$ *such that*

$$f(x) - f(a) \le D^{+}f(a + \tilde{\lambda}(x - a), x - a). \tag{7}$$

Proof. Define the function $F : [0,1] \to R$ by $F(\lambda) = f(a + \lambda(x - a))$. Since f is radially continuous F is a continuous function. Thus, by Theorem 3.1, there is a $\tilde{\lambda} \in (0,1)$ such that

$$F(1) - F(0) \le D^{+}F(\tilde{\lambda}) \tag{8}$$

Now, $D^{+}F(\tilde{\lambda}) = \limsup_{\lambda \downarrow 0} \lambda^{-1}(f(a + \tilde{\lambda}(x - a) + \lambda(x - a)) - f(a + \tilde{\lambda}(x - a)))$

$$= D^{+}f(a + \tilde{\lambda}(x - a), x - a).$$

Thus (7) and (8) are equivalent and the result follows. □

REMARK 3.4. By Remark 3.2, (7) remains valid provided f is radially l.s.c and satisfies the following line segment maximum property (see [18])

$$(\forall x, a \in X_o)(\exists \tilde{\lambda} \in [0,1]) \qquad f(a + \tilde{\lambda}(x - a)) = \sup\left\{ f(a + \lambda(x - a)) : \lambda \in [0,1] \right\} \tag{9}$$

If f possesses a two-sided directional derivative on X_o then Theorem 3.3 provides the generalized mean-value theorems of Nashed [30] and Yamamuro [33]; in this case $F(\cdot)$ is differentiable and (consequently) continuous on (0,1). The radial continuity assumption on f is automatically satisfied if f is finitely directionally differentiable (see [30]).

For the special case in which f is locally Lipschitz on X_o, Theorem 3.3 yields the following:

$$(10) \quad x,a \in X_o \Rightarrow (\exists\tilde{\lambda} \in (0,1))\ f(x) - f(a) \le f^o(a + \tilde{\lambda}(x - a), x - a)$$

This follows easily from (7) since it is always true that $D^+f \le f^o$. Lebourg [26] established the mean-value theorem represented by (10) in a more general form as follows:

$$(11) \quad \begin{cases} x,a \in X_o \Rightarrow (\exists\tilde{\lambda} \in (0,1),\ v \in \partial f_C(a + \tilde{\lambda}(x - a))) \\ \qquad\qquad\qquad\qquad\qquad\qquad f(x) - f(a) = v(x - a) \end{cases}$$

It is clear by the definition of $\partial f_C(x)$ that (11) implies (10).

We now define several concepts of generalized convex functions.

DEFINITION 3.5. For a function $f: X_o \to R$ we define the following properties:

(i) (Hiriart-Urruty [24] <u>pseudoconvex (Clarke)</u> (PCX(Clarke))

$$(\forall x,a \in X_o)\ f^o(a,x - a) \ge 0 \Rightarrow f(x) \ge f(a)$$

(ii) (Diewert [18]) <u>pseudoconvex (Dini)</u> (PCX(Dini))

$$(\forall x,a \in X_o)\ D^+f(a,x - a) \ge 0 \Rightarrow f(x) \ge f(a)$$

(iii) (Diewert [18]) <u>quasiconvex (Dini)</u> (QCX(Dini))

$$(\forall x,a \in X_o)\ D^+f(a,x - a) > 0 \Rightarrow f(x) > f(a)$$

(iv) (Mangasarian [27]) <u>strictly quasiconvex</u> (SQCX)

$$(\forall x,a \in X_o)\ f(x) < f(a) \Rightarrow (\forall\lambda \in (0,1))\ f(a + \lambda(x - a)) < f(a).$$

(v) (Mangasarian [27]) <u>quasiconvex</u> (QCX)

$$(\forall x,a \in X_o)\ f(x) \le f(a) \Rightarrow (\forall\lambda \in (0,1))\ f(a + \lambda(x - a)) \le f(a).$$

We can consider these concepts locally (about $a \in X_o$) by replacing X_o by $X_o \cap N$ for a suitable open neighbourhood N. Similarly the properties can be considered at a point, this follows by allowing x to vary with a fixed. It follows easily that every PCX(Clarke) function is

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

PCX(Dini) (since $D^+f \le f$). Similarly every QCX(Clarke) function (where this concept is defined in a manner analogous to (iii)) is QCX(Dini). It has been shown by Karamardian [25] that every radially l.s.c SQCX function is QCX (and, consequently, l.s.c, by Crouzeix [16, p.114]). In the following theorem we establish relationships between and characterizations for the concepts described above analogous to those established by Mangasarian [27] for differentiable functions.

THEOREM 3.6. *Let* $f : X_o \to R$ *be a radially continuous function. Then,*

(i) f *is QCX(Dini) if and only if* f *is QCX;*

(ii) if f *is PCX(Dini) it is also QCX, QCX(Dini) and SQCX.*

(iii) f *is PCX(Dini) if and only if*

(12) $$[(\forall x,a \in X_o)\ f(x) < f(a) \Rightarrow (\exists b(x,a) > 0)(\forall\lambda \in (0,1))\ f(a + \lambda(x - a)) - f(a) + \lambda b(x,a) \le 0]$$

Proof. (i) That QCX implies QCX(Dini) is straightforward. Conversely, suppose f is QCX(Dini), $f(x) \le f(a)$ and that for some $\bar\lambda \in (0,1)$ $f(a + \bar\lambda(x - a)) > f(a) \ge f(x)$; hence f is not QCX. Define the set $U = \{\lambda : f(a + \lambda(x - a)) \le f(a)$ and $\lambda \in [0,\bar\lambda]\}$. Clearly $U \ne \emptyset$ since $0 \in U$. Now, by the radial continuity of f, there is a $\beta \in [0,\bar\lambda)$ such that $\beta \in U$ and $\lambda \notin U$ for all $\lambda \in (\beta,\bar\lambda]$.

By Theorem 3.3 there is a $\gamma \in (\beta,\bar\lambda)$ such that

(13) $$f(a + \bar\lambda(x - a)) - f(a + \beta(x - a)) \le D^+f(a + \gamma(x - a),(\bar\lambda - \beta)(x - a)).$$

By the choice of β, $\gamma \notin U$; thus $f(a + \gamma(x - a)) > f(a) \ge f(x)$. Now, as f is QCX(Dini), it follows that

(14) $$D^+f(a + \gamma(x - a),(1 - \gamma)(x - a)) \le 0$$

(15) $$\Rightarrow D^+f(a + \gamma(x - a),\ x - a) \le 0.$$

Hence combining (13) and (15) yields the following

$$f(a + \bar\lambda(x - a)) \le f(a + \beta(x - a)).$$

However $\beta \in U$, thus $f(a) < f(a + \bar\lambda(x - a)) \le f(a + \beta(x - a)) \le f(a)$;

this gives the required contradiction. Thus QCX(Dini) implies QCX.

(ii) The proof that PCX(Dini) implies QCX can be fashioned on (i) above. If we suppose that f is PCX(Dini) and not QCX (13) will follow as in the proof of (i) with $f(a + \gamma(x - a)) > f(a) \geq f(x)$. Now, since f is PCX(Dini), (14) will follow with a strict inequality; thus the contradiction still remains, and the result easily follows.

Now suppose f is not SQCX, thus there are $x,a \in X_0$ with $f(x) < f(a)$ and, for some $\bar{\lambda} \in (0,1)$, $f(a + \bar{\lambda}(x - a)) \geq f(a)$. Since f is PCX(Dini) it follows that $D^+f(a,x- a) < 0$. Hence there is a $\gamma > 0$ such that $f(a + \lambda(x - a)) < f(a)$, for all $\lambda \in (0,\gamma)$. Since f is radially continuous there is a $\beta \in (0,\gamma)$ with

$f(x) < f(a + \beta(x - a)) < f(a) \leq f(a + \bar{\lambda}(x - a))$. By the above, f is QCX at $z = a + \beta(x - a)$; hence

(16) $$f(z + \alpha(1 - \beta)(x - a)) \leq f(z), \text{ for all } \alpha \in (0,1).$$

Let $\bar{\alpha} = (\bar{\lambda} - \beta)/(1 - \beta)$, then $\bar{\alpha} \in (0,1)$ since $0 < \beta < \bar{\lambda} < 1$. Thus, by putting $\alpha = \bar{\alpha}$ into (16), we find $f(a + \bar{\lambda}(x - a)) \leq f(a + \beta(x - a))$. However this is a contradiction to our choice of β. Hence f is SQCX.

(iii) It is easily shown that (12) implies PCX(Dini).

To prove the converse suppose f is PCX(Dini) and (12) is not satisfied. Hence there are $x,a \in X_0$ with $f(x) < f(a)$ and such that

(17) $$(\forall b > 0)(\exists\lambda(b) \in (0,1))\ f(a + \lambda(b)(x - a)) - f(a) + \lambda(b)b > 0.$$

Thus, for each $i \in N$, there is a $\lambda_i \in (0,1)$ such that

(18) $$(f(a + \lambda_i(x - a)) - f(a))/\lambda_i > -1/i.$$

By taking a subsequence if necessary we can assume $\lambda_i \to \bar{\lambda} \in [0,1]$. Now consider the following two cases:

(a) $\bar{\lambda} = 0$. From (18) it follows that

$$D^+f(a,x - a) \geq \limsup_{i\to+\infty} (f(a + \lambda_i(x - a)) - f(a))/\lambda_i \geq 0.$$

However this contradicts the PCX(Dini) assumption since $f(x) < f(a)$.

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

(b) $\overline{\lambda} > 0$. Now, by assumption, $f(a + \lambda(x - a))$ is continuous in λ. Hence, by (18), $f(a + \overline{\lambda}(x - a)) \geq f(a)$. However, by part (ii) above, f is SQCX, hence $f(a + \overline{\lambda}(x - a)) < f(a)$; thus we have a contradiction.

Hence (12) must be satisfied as required. □

REMARK 3.7. Parts (i) and (ii) of Theorem 3.6 are valid under the more general hypothesis that f is radially l.s.c and satisfies the line segment maximum property (9). It has been shown by Diewert [18] that it is not possible, in general, to weaken the continuity assumption in part (iii) (a counterexample is given [18, p.70]). Parts (i) and (ii) can be found in [18] where they are established using different methods to those adopted here. The proofs given here are simpler than those in [18] without greatly strengthening the continuity assumptions involved. The method of proof used in part (iii) has been adapted from the proof, for differentiable functions over R^n given by Avriel et al [1]. Mifflin [28] has shown that Theorem 3.6(ii) is valid for semiconvex functions, (a function $f : X_o \to R$ is semiconvex if it is locally Lipschitz, directionally differentiable, PCX(Clarke) and $(\forall a \in X_o)\ f'(a,\cdot) = f^o(a,\cdot))$; clearly this is a special case of part (ii).

For locally Lipschitz functions it follows immediately from Theorem 3.6 and Definition 3.5 that every PCX(Clarke) function is QCX, SQCX, and satisfies condition (12). That PCX(Clarke) implies SQCX was stated, without proof, in Hiriart-Urruty [24]. It should be noted that condition (12) would not, in general, be sufficient for a function to be PCX(Clarke); this follows since the latter depends on the nature of the function on a whole neighbourhood whereas the former is a 'line segment' condition.

4. Applications in mathematical programming

In this section we consider applications of the results in the preceding section to mathematical programming problems. In particular we establish Kuhn-Tucker type optimality conditions both for finitely and infinitely constrained problems without assuming a constraint

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

qualification. This extends the work in Craven, Glover and Zlobec [10] to allow non-convex, nondifferentiable constraint and objective functions.

The following lemma forms the basis of the results to follow.

LEMMA 4.1. (Craven, Glover and Zlobec [10])

Let $g : X_o \to Y$ *with* $S \subset Y$ *a closed convex cone. Let* H *be a generating set for* S^* *(hence* $S^* = \text{cone}(\overline{co}\, H)$*); define the set* $H^< = \{v \in H : vg(x) < 0 \text{ for some } x \in g^{-1}(-S)\}$*. Let* Q *be a weak* compact subset of* $H^<$*; also assume that* vg *is l.s.c and SQCX for each* $v \in H$*. Then* $0 \notin \overline{co}\, Q$*,* $\overline{co}\, Q$ *is weak* compact,* $\text{int}\, Q^*$ *is non-empty and there exists an* $\bar{x} \in g^{-1}(-S)$ *with* $g(\bar{x}) \in -\text{int}\, Q^*$*.*

REMARK 4.2. Lemma 4.1 was originally established, under the assumption that g is S-convex, by Borwein and Wolkowitz [5]. Craven, Glover and Zlobec [10] extended this result to differentiable pseudoconvex functions (i.e. vg is PCX(Dini) and differentiable for each $v \in H$); however a close examination of the proof shows that SQCX is sufficient, the lower semi-continuity assumption is necessary to ensure vg is also QCX, for each $v \in H$. It is not difficult to show that Lemma 4.1 remains valid if vg is QCX and SQCX for each $v \in \hat{Q}$ and $g^{-1}(-S)$ is convex.

As a consequence of Theorem 3.6 and Remark 3.7 the above result holds under either of the following conditions:

(a) vg is radially continuous and PCX(Dini) for each $v \in H$,

(b) vg is locally Lipschitz and PCX(Clarke) for each $v \in H$.

In [10] two further concepts of generalized convex functions were defined under which an appropriate version of Lemma 4.1 can be stated; however it is not difficult to show that both these concepts imply SQCX and QCX. It was also shown in [10] that the conclusion of Lemma 4.1 does not hold if the condition on the function vg is weakened to quasiconvexity.

(A) Finitely constrained programs

Consider the following program:

(P) $\underset{x \in X_o}{\text{Minimize}}$ $f(x)$ subject to $x \in C$, $g_i(x) \le 0$, $i \in B = \{1,\dots,n\}$;

where $f : X_o \to R$ and (for each $i \in B$) $g_i : X_o \to R$ are locally Lipschitz

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

functions, and C is a non-empty closed subset of X_o.

Define the following sets:

$$F = \{x \in C : (\forall i \in B)\ g_i(x) \le 0\},\ \text{the } \underline{\text{feasible set}};$$

$$I = \{i \in B : (\forall x \in F)\ g_i(x) = 0\}\quad I^< = B\backslash I.$$

Let $a \in F$ be the putative minimum of (P) and define $J = \{i \in B : g_i(a) = 0\}$.

THEOREM 4.3. *Consider program* (P) *with* $a \in F$. *Then a necessary condition for* (P) *to attain a local minimum at* a *is that there are* $\lambda_i \ge 0$, $i \in \{0,1,\ldots,n\}$, *not all zero, such that*

(19) $$0 \in \lambda_o \partial_C f(a) + \sum_{i=1}^{n} \lambda_i \partial_C g_i(a) - N(C,a),\quad \sum_{i=1}^{n} \lambda_i g_i(a) = 0$$

Proof. Let $a \in F$ be a local minimum of (P) . Now suppose, if possible, that the following system has a solution $d \in X$,

(20) $$f^o(a,d) < 0,\ (\forall i \in B)\ g_i(a) + g_i^o(a,d) < 0,\ d \in T(C,a).$$

Since $D^+f(a,d) \le f^o(a,d) < 0$, it easily follows that for some $\gamma > 0$, sufficiently small, $f(a + \lambda d) < f(a)$, for all $\lambda \in (0,\gamma)$. Similarly, since $D^+g_i(a,d) < -g_i(a)$ for each $i \in B$, there are $\gamma_i > 0$ such that $g_i(a + \lambda d) < 0$ for all $\lambda \in (0,\gamma_i)$ and each $i \in B$. Let $\delta = \min\{\gamma,\gamma_1,\ldots,\gamma_n\}$. Then $a + \lambda d \in F$ *and* $f(a + \lambda d) < f(a)$, for all $\lambda \in (0,\delta)$; this is a contradiction to $a \in F$ being a local minimum. Hence system (20) has no solution $d \in X$.

Define the function $F : X \to R^{n+1}$ where

$$F(d) = (f^o(a,d), g_1^o(a,d), \ldots, g_n^o(a,d)).$$

Also let $s = (0, g_1(a), \ldots, g_n(a)) \in R^{n+1}$. Since f and (for each $i \in B$) g_i are locally Lipschitz, F is continuous and R_+^{n+1}-sublinear. Now no solution to the system (20) is equivalent to no solution existing to

(21) $$F(d) + s \in -int\, R_+^{n+1},\ d \in T(C,a).$$

Since $T(C,a)$ is convex this is equivalent, by the Basic Alternative

Theorem ([8, p.31]), to the following

(22) $$(\exists 0 \neq p \in R_+^{n+1})(\forall d \in T(C,a))\ pF(d) \geq -p(s)$$

Hence, by [20, Corollary 1], (22) is equivalent to

(23) $$0 \in \partial(pF)(0) \times \{p(s)\} - (N(C,a) \times R_+)$$

Let $p = (\lambda_0, \lambda_1, \ldots, \lambda_n)$, then it follows that (using [7,Theorem 1])

$$0 \in \lambda_0 \partial_C f(a) + \sum_{i=1}^{n} \lambda_i \partial_C g_i(a) - N(C,a), \quad \sum_{i=1}^{n} \lambda_i g_i(a) \geq 0$$

which is easily seen to be equivalent to (19) since $a \in F$. □

REMARK 4.4. Theorem 4.3 has been established previously by Clarke [6], and Hiriart-Urruty [23], however the method of proof used here differs considerably from the earlier versions which relied critically on the local Lipschitz behavior of the functions. The above theorem can be extended to include functions in the class C; if we assume $co\{f, g_i\ (i=1,..,n)\}$ is contained in C then (23) can be replaced by the following condition:

(24) $$0 \in cl[\partial(pF)(0) \times \{p(s)\} - (N(C,a) \times R_+)].$$

That (24) is equivalent to (22) follows from the Generalized Non-homogeneous Farkas' Theorem, [20,Theorem 2.2]. As we are assuming f and (for each i) g_i are in C the function F may not always take values in R^{n+1} (since the generalized gradients may take the value $+\infty$), however this problem can be overcome by restricting F to the set

$$A = \text{dom}\, f \cap \left(\bigcap_{i=1}^{n} \text{dom}\, g_i\right).$$

Then no solution $d \in A$ to system (20) is easily seen to imply (22) and the result then follows as above. If we let $p = (\lambda_0, \lambda_1, \ldots, \lambda_n)$ then using an induction argument on (3), (24) can be expressed in the following form

(25) $$0 \in cl\left[cl\ (\lambda_0 \partial_C f(a) + \sum_{i=1}^{n} \lambda_i \partial_C g_i(a)) \times \left\{\sum_{i=1}^{n} \lambda_i g_i(a)\right\} - (N(C,a) \times R_+)\right].$$

Thus we have established the following theorem.

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

THEOREM 4.5. *Consider program (P) with* $a \in F$. *Let* f *and (for each* $i \in B$) g_i *be functions in class* C. *Then a necessary condition for (P) to attain a local minimum at* a *is that there are* $\lambda_i \geq 0$, $i \in \{0,1,\ldots,n\}$, *not all zero, such that* (25) *holds.*

The main result of this section is the following Kuhn-Tucker type theorem for program (P). For convenience let

$D = \{x \in X_o : (\forall i \in I)\ g_i(x) \leq 0\}$.

THEOREM 4.6. *Consider program (P) with* $a \in F$. *Let* f *and (for each* $i \in B$) g_i *be locally Lipschitz with* g_i *PCX(Clarke) for each* $i \in I^<$ *and QCX for each* $i \in I$. *Furthermore assume* C *is convex. Then a necessary condition for (P) to attain a local minimum at* a *is that there are* $\lambda_i \geq 0$, *for* $i \in I^<$, *such that*

$$(26) \qquad 0 \in \partial_C f(a) + \sum_{i\in I^<} \lambda_i \partial_C g_i(a) - N(C \cap D, a), \quad \sum_{i\in I^<} \lambda_i g_i(a) = 0.$$

Proof. Program (P) can be expressed in the following form:

(P') $\underset{x\in X_o}{\text{Minimize}}\ f(x)$ subject to $x \in C \cap D$ $g_i(x) \leq 0$, $i \in I^<$.

By assumption $C \cap D$ is a closed convex set, hence $N(C \cap D, a) = (C \cap D - a)^*$. By choosing the standard basis vectors in R^n as a generating set for the cone R^n_+ and noting that g_i is continuous and PCX(Clarke) for each $i \in I^<$, we can apply Lemma 4.1. Thus there is a point $\bar{x} \in F$ with $g_i(\bar{x}) < 0$ for each $i \in I^<$. Let $a \in F$ be a local minimum for (P), thus by Theorem 4.3 there are $\lambda_o, \lambda_i \geq 0$, $i \in I^<$, such that

$$(27) \qquad 0 \in \lambda_o \partial_C f(a) + \sum_{i\in I^<} \lambda_i \partial_C g_i(a) - N(C \cap D, a), \quad \sum_{i\in I^<} \lambda_i g_i(a) = 0$$

Hence it suffices if we can show $\lambda_o > 0$. Suppose, if possible, that $\lambda_o = 0$. Thus, by (27), there are $w_i \in \partial_C g_i(a)$, for $i \in I^< \cap J$, with

$$(28) \qquad v \equiv \sum_{i\in I^< \cap J} \lambda_i w_i \in N(C \cap D, a)$$

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

Now, for each $i \in I^{<} \cap J$, $\lambda_i > 0$ and $g_i(\bar{x}) < g_i(a) = 0$. Thus, since g_i is PCX(Clarke), $g_i^o(a,\bar{x} - a) < 0$, for each $i \in I^{<} \cap J$. Hence $v(\bar{x} - a) < 0$; however this is a contradiction to (28) since $\bar{x} - a \in C \cap D - a$, so that $v(\bar{x} - a) \geq 0$. Hence $\lambda_o > 0$ is required and (26) follows immediately from (27). It can be assumed that $I^{<} \cap J \neq \emptyset$ in (27) since, for $\lambda_o = 0$, there is some $\lambda_i > 0$ with $i \in I^{<}$, it easily follows that $i \in I^{<} \cap J$. □

REMARK 4.7. Theorem 4.6 provides a necessary optimality condition for program (P) which does not require a constraint qualification to be satisfied. By virtue of Lemma 4.1 and the expression of (P) in the form of (P') we are able to use a 'Slater-type' constraint qualification which is implicit in (P'). We are tacitly assuming in Theorem 4.6 that $I^{<}$ is non-empty, however this may not always be the case; if $I^{<}$ is empty then we are dealing with a program involving only equality constraints and (26) yields the well-known necessary optimality condition ([23])

$$\partial_C f(a) \cap (F - a)^* \neq \emptyset$$

where F is a closed convex set.

If f and (for each $i \in B$) g_i are finite convex functions defined on R^n, then the Clarke subgradient coincides with the normal convex subdifferential; in this case a similar result to (26) has been established by Ben-Tal, Ben-Israel and Zlobec [3]. This latter result differs from (26) in that $T(C \cap D,a)$ is replaced by the <u>cone of directions of constancy</u>, denoted by $D_I^{=}(a)$, namely

$$D_I^{=}(a) = \bigcap_{i \in I} \left\{d \in R^n : (\exists \alpha > 0)\ g_i(a + \lambda d) = g_i(a) \quad \text{for all} \quad \lambda \in (0,\alpha]\right\}.$$

It is easily shown that $N(D,a) = (T(D,a))^* = (D - a)^* \subset (D_I^{=}(a))^*$, thus the necessary (and sufficient) condition of [3] is a direct corollary of Theorem 4.6. If f is PCX(Clarke) and (for each $i \in B$) g_i is QCX(Clarke) then it is straightforward to show that (26) is a sufficient optimality condition for (P).

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

From a computational viewpoint it is worth noting that Zlobec and Craven [35] have suggested a finite iterative method for calculating the minimal index set of binding constraints (in our notation the set I) for differentiable convex programs.

It is possible to extend Theorem 4.6 to include the case in which the objective and constraint functions are (more generally) in class C. In this case we replace the necessary condition (26) by the following:

$$0 \in cl\left\{\left[(cl(\partial_C f(a) + \sum_{i\in I^<}\lambda_i \partial_C g_i(a))) \times \left\{\sum_{i\in I^<}\lambda_i g_i(a)\right\})\right] - (N(C \cap D\ a) \times R_+)\right\}$$

The proof follows using a similar (although more complicated) argument to that of Theorem 4.6 and, for the sake of brevity, we omit the details.

(B) Mathematical programming in abstract spaces with cone constraints.

Consider the following program:

(P") $\quad \underset{x\in X_o}{\text{Minimize}}\ f(x)$ subject to $-g(x) \in S$;

where $f : X_o \to R$ and $g : X_o \to Y$ are weakly directionally differentiable functions. Let $F = \{x \in X_o : -g(x) \in S\}$, the feasible region for (P").

THEOREM 4.8. (Glover [21])

Consider program (P") with $a \in F$. *At* a, *let* f *be quasi-differentiable and* g *S^*-quasidifferentiable. Let* H *be a generating set for* S^*, *define* $H^<$ *as in Lemma 4.1 and let* $Q \subset H^<$ *be weak* compact. Furthermore assume that* vg *is PCX(Dini) for each* $v \in H$ *and l.s.c for each* $v \in H\backslash Q$. *Then the implications* (a) $\Rightarrow$ (b) $\Rightarrow$ (c) *are satisfied where*

(a) *(P") attains a local minimum at* a.

(b) *There exists a* $\bar{v} \in Q^{**}$ *such that*

(29) $$0 \in \tilde{\partial} f(a) + \tilde{\partial}(\bar{v}g)(a) - (D(Q) - a)^*,\ \bar{v}g(a) = 0$$

where $D(Q) = \{x \in X_o : (\forall v \in H\backslash Q)\ vg(x) \le 0\}$.

(c) *There is no solution* $d \in X$ *to the following system:*

$$f'(a,d) < 0,\ g_w'(a,d) \in -cl\left[\text{cone } (Q^* + g(a))\right],\ d \in cl\left[\text{cone } (D(Q) - a)\right]$$

REMARK 4.9. The proof of Theorem 4.8 is similar to the proof of Theorem 1 in Craven, Glover and Zlobec [10], it requires the Generalized Fritz John Theorem in Glover [20, 21] and relies, critically, on Lemma 4.1. Since g is weakly directionally differentiable and vg is PCX(Dini) for each $v \in H$, it follows that $(vg)'(a,d) = vg'_w(a,d) = D^+(vg)(a,d)$, also by Theorem 3.6 vg is SQCX and so Lemma 4.1 is applicable. Hence there is an $\bar{x} \in F$ such that $-g(\bar{x}) \in \text{int } Q^*$; this supplies the 'stability' result necessary to ensure we obtain the (modified) Kuhn-Tucker conditions (29). It is not difficult to show that (c) implies (a) provided f is PCX(Dini) and vg is QCX(Dini) for each $v \in H$. Theorem 4.8 is more general than Theorem 1 in [10]; the latter result required that f be linearly Gâteaux differentiable and g be either linearly Gâteaux differentiable or S-convex and continuous, clearly these conditions are special cases of the quasidifferentiability conditions of Theorem 4.8. If both f and g are linearly Gâteaux differentiable then (29) becomes $(\exists \bar{v} \in Q^{**})\ f'(a) + vg'(a) \in (D(Q) - a)^*$.

It is possible to further extend Theorem 4.8 to minimization problems involving vector-valued objective functions, with an appropriate extension of the concept of local minima (see [9]).

We now consider Theorem 4.8 applied to program (P).

THEOREM 4.10. *Consider program* (P) *with* $a \in F$. *At* a, *let* f *and (for each* $i \in B$) g_i *be quasidifferentiable. Let* g_i *be* PCX(Dini) *for each* $i \in I^<$ *and* QCX *for each* $i \in I$. *Furthermore assume* C *is convex. Then a necessary condition for* (P) *to attain a local minimum at* a *is that there are* $\lambda_i \geq 0$, $i \in I^<$, *with*

$$(30) \qquad 0 \in \tilde{\partial} f(a) + \sum_{i \in I^<} \lambda_i \tilde{\partial} g_i(a) - (C \cap D - a)^*, \quad \sum_{i \in I^<} \lambda_i g_i(a) = 0$$

REMARK 4.11. It is worth noting that Theorems 4.6 and 4.10 present similar but independent results. The directional differentiability assumption in Theorem 4.10 is not required in Theorem 4.6 where we use, instead, the Lipschitz properties of the functions. Clarke [7] has shown that if $f: X_o \to R$ is quasidifferentiable with the multifunction $\partial f(\cdot)$ (strongly) upper semi-continuous then f is locally Lipschitz with $\partial_C f(a) = \tilde{\partial} f(a)$, for each $a \in X_o$. If f and g_i are differentiable in program (P) then (30) becomes

$$(31) \qquad f'(a) + \sum_{i \in I^<} \lambda_i g_i'(a) \in (C \cap D - a)^*, \quad \sum_{i \in I^<} \lambda_i g_i(a) = 0 \qquad ,$$

However (31) does not follow from Theorem 4.6 unless the functions are continuously differentiable (this follows since the Clarke subgradient may contain continuous linear functionals other than the derivative, see [7], [26, Theorem 2.2]).

(C) Application to fractional programming

In this section we consider the following fractional programming problem:

$$\text{(P1)} \quad \underset{x \in X_o}{\text{Minimize}}\ f_1(x)/f_2(x) \quad \text{subject to} \quad x \in C,\ g_i(x) \le 0,\ i \in B;$$

where f_1, $f_2 : X_o \to R$ and (for each $i \in B$) $g_i : X_o \to R$ are convex functions with $f_1(x) \le 0$ and $f_2(x) > 0$ for each $x \in X_o$ feasible for (P1). Further we assume C is closed and convex. We shall also assume all functions involved are continuous.

Borwein [4] (and others) have shown that $f = f_1/f_2$ is quasi-differentiable with $\tilde{\partial} f(a) = (f_2(a)\partial f_1(a) - f_1(a)\partial f_2(a))/(f_2(a))^2$. Similarly, Bector [2] has shown that f is PCX(Dini). Since f_i $(i = 1,2)$ is a continuous convex function it is locally Lipschitz ([14]), and consequently f is also locally Lipschitz ([4]). It should be noted that f is not in general convex or differentiable, hence the results in [10] would not be directly applicable to problem (P1). In this case $\tilde{\partial} f(a)$ coincides with $\partial_C f(a)$ so that Theorem 4.10 and 4.6 yield identical optimality conditions. Thus, using (26), a necessary optimality condition for (P1) is the following with $a \in F$ the putative minimum:

$$(32)\quad \begin{cases} 0 \in f_2(a)\partial f_1(a) - f_1(a)\partial f_2(a) + \sum_{i\in I^<} \lambda_i \partial g_i(a) - N(D,a), \\ \qquad\qquad\qquad\qquad\qquad\qquad \sum_{i\in I^<} \lambda_i g_i(a) = 0 \end{cases}$$

with $\lambda_i \geq 0$, for each $i \in I^<$, and $D = \{x \in X_o : (\forall i \in I)\ g_i(x) \leq 0\}$. It follows immediately that the above condition is also sufficient for optimality since f is PCX(Dini); consequently this condition characterizes optimality for (P1) without requiring a constraint qualification to be satisfied (and also without recourse to equivalent convex programs).

Mond and Zlobec [29] have presented a duality result for a finitely constrained, nondifferentiable, convex program without a constraint qualification being required using the work of [3]. This result can easily be extended to include non-convex programs such as (P1). Thus a dual program to (P1) (related to the Wolfe dual) can be given as

$$\text{(D1)}\qquad \text{Maximize } \left[f_1(u) + \sum_{i\in B} \lambda_i g_i(u)\right] / f_2(u)$$

subject to $(\forall i \in B)\ \lambda_i \geq 0$, and $g_i(u) \leq 0$, for $i \in I$, with

$$0 \in f_2(u)\partial f_1(u) - f_1(u)\partial f_2(u) + \sum_{i\in B} \lambda_i (f_2(u)\partial g_i(u) - g_i(u)\partial f_2(u)) - N(D,u).$$

A related dual, assuming a constraint qualification, was derived by Craven and Mond [12]. The proof of duality follows easily from the following lemma which we include without proof (similar results can be found in [21], [15], and [37]).

LEMMA 4.12. *Let* $h : X_o \to R$ *and* $k : X_o \to R$ *be continuous convex functions on a convex set* $U \subset X_o$, *furthermore assume* $h(x) \leq 0$ *and* $k(x) > 0$ *for each* $x \in U$. *Then, for each* x, $a \in U$, *we have*

$$\beta(x,a)(H(x) - H(a)) \geq H'(a, x - a)$$

where $\beta(x,a) = k(x)/k(a)$ and $H = h/k$.

https://doi.org/10.1017/S0004972700001908 Published online by Cambridge University Press

THEOREM 4.13. *(D1) is a dual program to (P1).*

Proof. (Weak duality) Let $x \in X_o$ be feasible for (P1) and $(u,\lambda_1,\ldots,\lambda_n)$ be feasible for (D1). For convenience let $F = f_1/f_2$ and $G_i = g_i/f_2$ (for $i \in B$). Thus it is easily seen that the main constraint of (D1) is equivalent to

(33) $$0 \in \tilde{\partial}F(u) + \sum_{i\in B}\lambda_i\tilde{\partial}G_i(u) - N(D,u)$$

Hence, by (33), there are $v \in \tilde{\partial}F(u)$, and $w_i \in \tilde{\partial}G_i(u)$ such that

(34) $$v + \sum_{i\in B}\lambda_i w_i \in N(D,u).$$

Now consider the following:

$$F(x) - (F(u) + \sum_{i\in B}\lambda_i G_i(u)) \geq F(x) - F(u) + \sum_{i\in B}\lambda_i(G_i(x) - G_i(u))$$

(since $x \in F$ and $f_2(x) > 0$)

$$\geq \alpha(F'(u,x-u) + \sum_{i\in B}\lambda_i G_i'(u,x-u))$$

(using Lemma 4.12, with $\alpha = f_2(u)/f_2(x)$)

$$\geq \alpha((v + \sum_{i\in B}\lambda_i w_i)(x-u))$$

$$\geq 0 \text{ by (34) since } x - u \in D - u.$$

In this final step we require that $N(D,u) = (D-u)^*$, which follows by the convexity conditions and the fact that $u \in D$ from the constraints of (D1). Hence since $F = f_1/f_2$ is the objective function of (P1), weak duality is established.

(Strong duality) Let $a \in X_o$ be optimal for (P1). Thus a is optimal for the following equivalent program,

(P1') Minimize $F(x)$ subject to $G_i(x) \leq 0$, for $i \in B$,

(we are assuming $C = X_o$, for the sake of brevity). By the convexity assumptions in (P1) F and G_i (for $i \in B$) are quasidifferentiable, PCX(Dini) functions; thus, by Theorem 4.10, there are $\lambda_i \geq 0$, (for $i \in I^{<}$), such that

$$(35) \qquad 0 \in \tilde{\partial}F(a) + \sum_{i \in I^{<}} \lambda_i \tilde{\partial}G_i(a) - N(D,a), \quad \sum_{i \in I^{<}} \lambda_i G_i(a) = 0.$$

Thus, by letting $\lambda_i = 0$ for $i \in B \backslash I^{<}$, $(a, \lambda_1, \ldots, \lambda_n)$ is feasible for (D1) with $\sum_{i \in B} \lambda_i G_i(a) = 0$. Now strong duality follows, using weak duality, since

$$\text{Min (P1)} = F(a) = F(a) + \sum_{i \in B} \lambda_i G_i(a) = \text{Max (D1)}. \qquad \square$$

REMARK 4.14. A dual program closely resembling that of [29] can be deduced from Theorem 4.13 and (D1) by letting $f_2(x) = 1$ for each $x \in X_0$.

Acknowledgements

The author wishes to thank Dr. B.D. Craven for his helpful advice during the long preparation of this paper, a great amount of the research for which was completed while the author was a research student at the University of Melbourne.

References

[1] M. Avriel, W.E. Diewert, S. Schaible and W.T. Ziemba, "Introduction to concave and generalized concave functions", in *Generalized Concavity in Optimization and Economics*, S. Schaible and W.T. Ziemba (eds.), Academic Press, New York, 1981, 21-50.

[2] C.R. Bector, "Programming problems with convex fractional functions," *Operations Research* 16 (1968), 383-389.

[3] A. Ben Tal, A. Ben-Israel and S. Zlobec, "Characterization of optimality in convex programming without a constraint qualification", *J.O.T.A.* 20 (1976), 417-437.

[4] J.M. Borwein, "Fractional programming without differentiability", *Math. Prog.* 11 (1976), 283-290.

[5] J.M. Borwein and H. Wolkowicz, "Characterization of optimality without constraint qualification for the abstract convex program", *Math. Programming Study* 19 (1982), 77-100.

[6] F.H. Clarke, "A new approach to Lagrange multipliers", *Math. Oper. Res.* 1, (1976), 165-174.

[7] F.H. Clarke, "Generalized gradients of Lipschitz functionals", *Advances in Math.* 40 (1981), 52-67.

[8] B.D. Craven, *Mathematical Programming and Control Theory*, (Chapman and Hall, London, 1978).

[9] 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, 661-680.

[10] B.D. Craven, B.M. Glover and S. Zlobec, "On minimization subject to cone constraints", *Numer. Func. Analysis Optimiz.*, 6(4), (1983), 363-378.

[11] B.D. Craven and B. Mond, "Lagrangean conditions for quasi-differentiable optimization" in *Survey of Mathematical Programming*, A. Prékopa (ed.), Akadémiai Kiadó, Budapest, 1, 1976.

[12] B.D. Craven and B. Mond, "On duality for fractional programming", *Z.A.M.M.* 58 (1959), 278-279.

[13] B.D. Craven and S. Zlobec, "Complete characterization of optimality for convex programming in Banach spaces", *Applicable Analysis* 11 (1980), 61-78.

[14] S. Cobzas, "On the Lipschitz properties of continuous convex functions", *Mathematica (Cluj)* 21(44) (1979), 123-125.

[15] S. Chandra, "Strong pseudo-convex programming", *Indian J. of Pure and Applied Math.* 3 (1972), 178-182.

[16] J.-P. Crouziex, "Continuity and differentiability properties of quasiconvex functions on R^n", in *Generalized Concavity in Optimization and Economics*, S. Schaible and W.T. Ziemba (eds.), Academic Press, New York, 1981, 109-130.

[17] M.A.H. Dempster and R.J.-B. Wets, "On regularity conditions in constrained optimization", Center for Advanced Study in the Behavioral Sciences, Stanford University (1975).

[18] W.E. Diewert, "Alternative characterizations of six kinds of quasi-concavity in the nondifferentiable case with applications to non-smooth programming", in *Generalized Concavity in Optimization and Economics*, S. Schaible and W.T. Ziemba (eds.), Academic Press, New York (1981), 51-94.

[19] T.M. Flett, *Differential Analysis*, (Cambridge University Press Cambridge, 1980).

[20] B.M. Glover, "A generalized Farkas lemma with application to quasi-differentiable programming", *Z. Oper. Res.* 26 (1982), 125-142.

[21] B.M. Glover, *Nondifferentiable Programming* (M.Sc. Thesis, University of Melbourne, 1982).

[22] M. Guignard, "Generalized Kuhn-Tucker conditions for mathematical programming problems in Banach space", *SIAM J. Control* 7 (1969).

[23] J.B. Hiriart-Urruty, "On optimality conditions in nondifferentiable programming", *Math. Prog.* 14 (1978), 73-86.

[24] J.B. Hiriart-Urruty, "New concepts in nondifferentiable programming", *Bull. Soc. Math. France, Mémoire* 60 (1979), 57-85.

[25] S. Karamardian, "Strictly quasiconvex (concave) functions and duality in mathematical programming", *J. Math. Analysis and Applns.* 20 (1967), 344-358.

[26] G. Lebourg, "Generic differentiability of lipschitzian functions", *Trans. Amer. Math. Soc.* 256 (1979), 125-144.

[27] O.L. Mangasarian, *Nonlinear Programming*, (McGraw-Hill, New York, 1969).

[28] R. Mifflin, "Semismooth and semiconvex functions in constrained optimization", *SIAM J. Control and Optimiz.* 15 (1977), 959-972.

[29] B. Mond and S. Zlobec, "Duality for nondifferentiable programming without a constraint qualification", *Utilitas Mathematica* 15 (1979), 291-302.

[30] M.Z. Nashed, "Differentiability and related properties of nonlinear operators", in *Nonlinear Functional Analysis and Applications*, L.B. Hall (ed.), (Academic Press, New York, 1971, 103-309).

[31] B.N. Pshenichnyi, *Necessary Conditions for an Extremum*, (Marcel Dekkar, New York, 1971).

[32] R.T. Rockafellar, *Conjugate Duality and Optimization*, SIAM Regional Conference Series in Applied Mathematics 16 (1974), Philadelphia.

[33] S. Yamamuro, *Differential Calculus in Topological Linear Spaces*, (Lecture Notes in Mathematics 374 Springer-Verlag, 1974).

[34] C. Zalinescu, "On an abstract control problem", *Numer. Func. Analysis Optimiz.* 2 (1980), 531-542.

[35] S. Zlobec and B.D. Craven, "Stabilization and determination of a set of minimal binding constraints in convex programming", *Math. Operation-forschung und Statistik, series Optimization* 12 (1981), 203-220.

[36] S. Zlobec and D.H. Jacobson, "Minimizing an arbitrary function subject to convex constraints", *Utilitas Mathematica* 17 (1980), 239-257.

[37] B.D. Craven and B.M. Glover, "Invex functions and duality", *J. Austral. Math. Soc.*, (to appear).

43 Wirbill Street,
Cobram, Victoria, 3644, Australia.