INVEXITY IN NONLINEAR PROGRAMS AND CONTROL PROBLEMS

I. Smart

ABSTRACT

Subject to certain regularity hypotheses, the Kuhn-Tucker conditions are necessary for optimality in nonlinear programs. These conditions become sufficient under assumptions of invexity. This paper presents some known results and new observations on invexity, with extension to optimal control problems.

1980 Amer.Math.Soc. Classification (1985 revision)
90C30 (Secondary 90C48).

1. INTRODUCTION

Consider the general nonlinear program:

(P) Minimize $f(x)$

subject to $g(x) \in S$

where $f\colon C \to \mathbb{R}$ and $g\colon C \to Y$, with C an open subset of the normed space X, and S a polyhedral cone in the finite dimensional normed space Y.

For the sake of simplicity, it is assumed that $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $S = \mathbb{R}^m_+$, and f and g are Fréchet differentiable. Nevertheless, the ensuing results may all be considered in a more abstract context (see, for example, Craven and Glover [4]).

Under certain regularity conditions, or constraint qualifications, such as the Kuhn-Tucker, Arrow-Hurwicz-Uzawa, or reverse convex constraint qualification (Mangasarian [8]), the Kuhn-Tucker conditions are necessary for optimality of (P). That is, if x^* is optimal for (P), then there exists $\lambda^* \in \mathbb{R}^m$ such that:

$$\nabla f(x^*) - \nabla(\lambda^{*T} g(x^*)) = 0$$

$$\lambda^{*T} g(x^*) = 0$$

$$\lambda^* \geq 0.$$

In general, these conditions are not sufficient for optimality. We will present a recently discovered class of functions, called invex functions, which yield sufficiency, and lead to the extension of duality results in a variety of nonlinear programs.

2. INVEXITY

Let $h\colon C \to \mathbb{R}$ be differentiable. Then h is:

a) convex if $h(x) - h(u) \geq (x-u)^T \nabla h(u)$ for all x, $u \in C$;

b) quasi-convex if $h(x) \leq h(u) \Rightarrow (x-u)^T \nabla h(u) \leq 0$;

c) pseudo-convex if $(x-u)^T \nabla h(u) \geq 0 \Rightarrow h(x) \geq h(u)$.

Note that convexity ⇒ pseudo-convexity ⇒ quasi-convexity. (Mangasarian [8]).

It is well established that the Kuhn-Tucker conditions at (x^*,λ^*) are sufficient for optimality at x^* whenever x^* is feasible for (P), and either:

1) (Kuhn-Tucker [7]) f convex, g_i concave, $i = 1,\dots,m$; or

2) (Mangasarian [8]) f pseudo-convex, g_i quasi-concave for all $i \in I$, where $I = \{i \,|\, g_i(x^*) = 0\}$; or

3) (Mond [10]) f pseudo-convex, $\lambda^{*T}g$ quasi-concave. (2) ⇒ 3) as quasi-concavity is not additive); or

4) (Mond [10]) $f - \lambda^{*T}g$ pseudo-convex.

Hanson [6] noted that the convexity requirements could be weakened as there was no explicit dependence on the linear term $(x-u)$ in proving sufficiency; the linear bounds imposed by the notions of convexity could be replaced by arbitrary non-linear bounds.

DEFINITION 2.1

The differentiable function $h\colon C \to \mathbb{R}$ is invex if there exists a vector function $\eta\colon C \times C \to \mathbb{R}^n$ such that $h(x) - h(u) \geq \eta(x,u)^T \nabla h(u)$ for all x, $u \in C$.

The term invex stems from invariant convex (Craven [3]) -: if $q: \mathbb{R}^n \to \mathbb{R}$ is differentiable and convex, and $\phi: \mathbb{R}^r \to \mathbb{R}^n$ $(r \geq n)$ is differentiable with $\nabla\phi$ of rank n, then $f = q \circ \phi$ is invex. Indeed, f will be invex when q is invex, but with respect to a different η.

Further generalization is possible: h is quasi-invex if there exists $\eta: C \times C \to \mathbb{R}^n$ such that $h(x) \leq h(u) \Rightarrow \eta(x,u)^T \nabla h(u) \leq 0$; and pseudo-invex if there exists $\eta: C \times C \to \mathbb{R}^n$ such that

$$\eta(x,u)^T \nabla h(u) \geq 0 \quad \Rightarrow \quad h(x) \geq h(u).$$

This allows a restatement of the conditions sufficient for optimality at a feasible x^*, with (x^*,λ^*) satisfying the Kuhn-Tucker conditions.

1) f and $-g_i$ invex with respect to the same η, $i = 1,\dots,m$;

2) f pseudo-invex, $-g_i$ quasi-invex for $i \in I$, with respect to the same η;

3) f pseudo-invex, $-\lambda^{*T}g$ quasi-invex with respect to the same η;

4) $f - \lambda^{*T}g$ pseudo-invex.

Invexity also allows the weakening of necessary conditions. The aforementioned constraint qualifications require differentiability but not convexity, while there are two other constraint qualifications which do not require differentiability. These are -

Slater's: there exists $x^\circ \in C$ such that $g_i(x^\circ) > 0$, $i = 1,\dots,m$;

and Karlin's: there exists no $p \in \mathbb{R}^m$, $p \geq 0$, $p \neq 0$ such that $p^T g(x) \leq 0$ for all $x \in C$.

By the generalized Gordan theorem, these conditions are equivalent when the g_i are concave, but an analogous theorem of the alternative is not available for invex functions. However, both yield the Kuhn-Tucker conditions if the $-g_i$ are invex. Ben-Israel and Mond [1] deal with the Slater condition, and we prove here the corresponding result for Karlin's condition.

Theorem 2.2. Let x^* be optimal for (P), $-g_i$ invex with respect to the same η for $i = 1,\ldots,m$, and assume Karlin's constraint qualification is satisfied. Then there exists $\lambda^* \in \mathbb{R}^m$ such that (x^*,λ^*) satisfies the Kuhn-Tucker conditions.

Proof. The Fritz-John conditions are satisfied at x^*; that is, there exist $r_0^* \in \mathbb{R}$ and $r^* \in \mathbb{R}^m$ such that

$$r_0^* \nabla f(x^*) - \nabla(r^{*T} g(x^*)) = 0$$

$$r^{*T} g(x^*) = 0$$

$$(r_0^*, r^*) \geq 0, \quad (r_0^*, r^*) \neq 0.$$

We need to show that $r_0^* > 0$, so taking $\lambda^* = r^*/r_0^*$ gives the result. Assume that $r_0^* = 0$.

Then $\nabla(r^{*T} g(x^*)) = 0$, $r^{*T} g(x^*) = 0$, $r^* \geq 0$, $r^* \neq 0$.

By invexity of the $-g_i$ and non-negativity of the r_i^*, $i = 1,\ldots,m$, $r^{*T} g(x) - r^{*T} g(x^*) \leq \eta(x,x^*)^T \nabla(r^{*T} g(x^*))$ for all $x \in C$. i.e. $r^{*T} g(x) \leq 0$ for all $x \in C$. This contradicts Karlin's condition, and hence $r_0^* > 0$. ∎

Martin [9] further relaxed invexity requirements through complementary slackness and feasibility, and defined Kuhn-Tucker invexity of the program (P): there exists a function $\eta : C \times C \to \mathbb{R}^n$ such that if $x, u \in C$, $g(x) \geq 0$, $g(u) \geq 0$, then $f(x) - f(u) \geq \eta(x,u)^T \nabla f(u)$ and $g_i(u) = 0$ implies $\eta(x,u)^T \nabla g_i(u) \geq 0$.

This leads to the next result (Martin [9])

Theorem 2.3. Every Kuhn-Tucker point of (P) is a global minimizer if and only if (P) is Kuhn-Tucker invex.

In the unconstrained problem, this theorem leads to a corollary which gives a characterization of invex functions which, at present, is the best way of identifying such functions.

Corollary 2.4 (Craven and Glover [4]). Let $h: C \to \mathbb{R}$ be differentiable. Then h is invex if and only if every stationary point is a global minimizer.

It is adequate for the purposes of sufficiency, and later, duality, to know that functions are invex without identifying an appropriate η. However, this corollary allows us to find an η when h is known to be invex; viz.

$$\eta(x,u) = \begin{cases} \dfrac{(h(x)-h(u))\nabla h(u)}{\nabla h(u)^T \nabla h(u)} & \text{if } \nabla h(u) \neq 0 \\ 0 & \text{if } \nabla h(u) = 0 \end{cases}$$

A pertinent question is how to distinguish invexity from the previous generalizations to quasi- and pseudo-convexity. A partial answer is available.

Lemma 2.5 (Crouzeix and Ferland [5]). Let h be a differentiable quasi-convex function on an open convex set $C \subset \mathbb{R}^n$. Then h is pseudo-convex if and only if h has a minimum at $x \in C$ whenever $\nabla h(x) = 0$.

By the corollary, the last condition in this lemma is equivalent to invexity. Thus, under the assumption of quasi-convexity, invexity and pseudo-convexity coincide; so for an invex function not to be pseudo-convex, it must also not be quasi-convex. Such functions do exist: see the examples at the end of this section.

Another approach to characterization is through associated sets. The epigraph of h is $E_h = \{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R} \mid x \in C,\ h(x) \leq \alpha\}$, and the lower level sets of h are given by $L_h(\alpha) = \{x \in \mathbb{R}^n \mid x \in C,\ h(x) \leq \alpha\}$ for $\alpha \in \mathbb{R}$. Mangasarian [8] showed that h is convex if and only if E_h is convex in $\mathbb{R}^{n+1}$, and h is quasi-convex iff $L_h(\alpha)$ is convex for each $\alpha \in \mathbb{R}$. Before invex functions came into general use, Zang, Choo and Avriel [13] had characterized them in terms of the lower level sets.

If $L_h(\alpha)$ is non-empty, then it is strictly lower semi-continuous if for all $x \in L_h(\alpha)$ and sequences $\{\alpha_i\}$ with $\alpha_i \to \alpha$, $L_h(\alpha_i)$ non-empty, there exist $K \in \mathbb{N}$, a sequence $\{x_i\}$, $x_i \to x$, and $\beta(x) \in \mathbb{R}$, $\beta(x) > 0$, such that $x_i \in L_h[\alpha_i - \beta(x)\|x_i - x\|]$, $i = K, K + 1, \ldots$.

It may then be shown (Zang, Choo and Avriel [13]) that h is invex if and only if $L_h(\alpha)$ is strictly lower semi-continuous for all α satisfying $L_h(\alpha)$ non-empty.

EXAMPLES

1) (Ben-Israel and Mond [1]) $h:\mathbb{R} \to \mathbb{R}$ defined by $h(x) = x^3$ is quasi-convex since $L_h(\alpha) = (-\infty, \sqrt[3]{\alpha}]$ for all $\alpha \in \mathbb{R}$, but is not invex since $x = 0$ is a stationary point but not a global minimum.

2) $h:\mathbb{R}^2 \to \mathbb{R}$ defined by $h(x) = 1 + x_1^2 - e^{-x_2^2}$ is invex as the only stationary point (0,0) is a global minimum. It is not quasi-convex since for $x = (1.12, 2.329)$, $u = (1.31, 1.697)$ one has $h(x) \leq h(u)$ but $(x-u)^T \nabla h(u) > 0$.

3) (Ben-Israel and Mond [1]). Any function with no stationary points is invex, such as $h:\mathbb{R}^2 \to \mathbb{R}$ defined by $h(x) = x_1^3 + x_1 - 10x_2^3 - x_2$, which is also not quasi-convex.

The relationships between these notions of generalized convexity of differentiable functions defined on an open convex set may be represented pictorially.

(Ben-Israel and Mond [1] have observed that as every stationary point of a pseudo-invex function is a global minimum, such a function is also invex. However, this may not be with respect to the same η, although invex functions are certainly pseudo-invex with respect to the same η.)

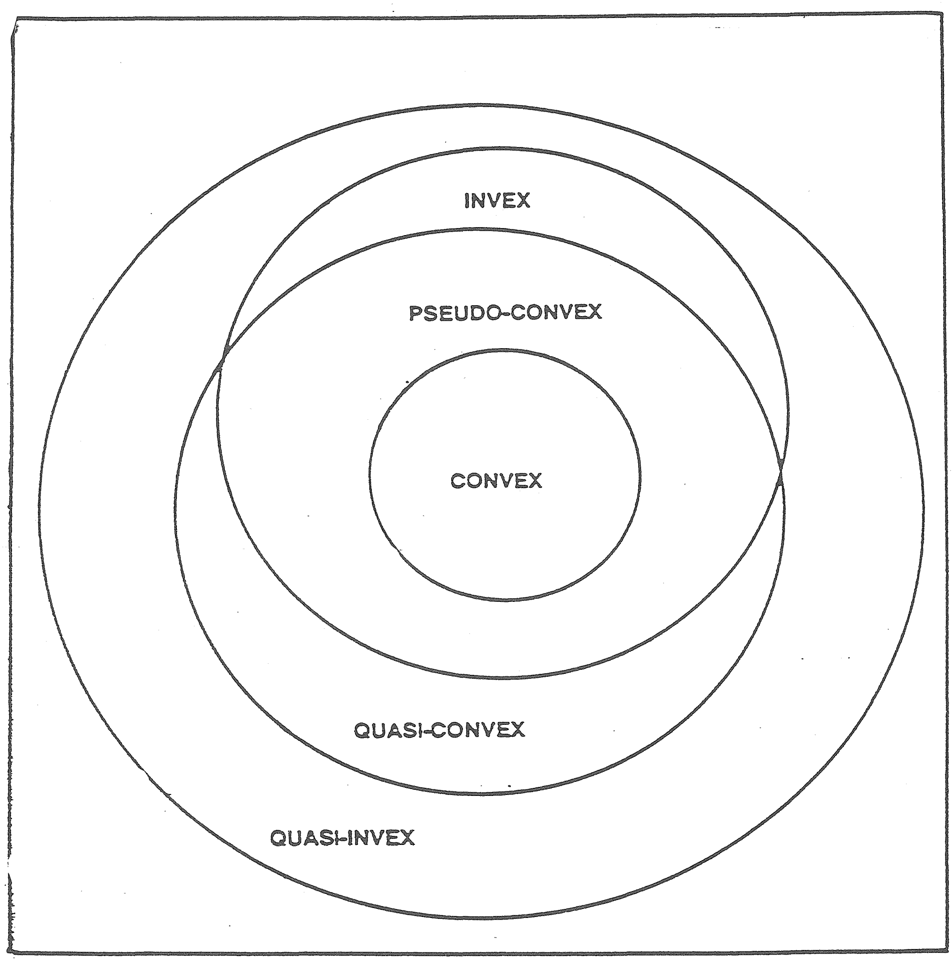

Figure 1

3. DUALITY

The Wolfe dual for (P) when $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $S = \mathbb{R}^m_+$, and f and g are differentiable is:

$$\text{(D)} \qquad \text{Maximize} \quad f(u) - \lambda^T g(u)$$
$$\text{subject to:} \quad \nabla f(u) - \nabla(\lambda^T g(u)) = 0$$
$$\lambda \geq 0.$$

Weak duality (inf (P) ≥ sup (D)) holds with either of the assumptions:

1) (Hanson [6]) f invex, $-g_i$ invex with respect to same η, $i = 1,\ldots,m$; or

2) $f - \lambda^T g$ invex for all $\lambda \in \mathbb{R}^m_+$.

Condition 2) is a weaker requirement, so we prove the result for this case. Let x be feasible for (P), and (u,λ) feasible for (D). Then, by invexity of $f - \lambda^T g$,

$$f(x) - \lambda^T g(x) - (f(u)-\lambda^T g(u)) \geq \eta(x,u)^T[\nabla f(u) - \nabla(\lambda^T g(u))] = 0.$$

But $\lambda \geq 0$ and $g(x) \geq 0$ give $f(x) - \lambda^T g(x) \leq f(x)$, so that $f(x) \geq f(u) - \lambda^T g(u)$.

Martin [9] gave an invexity condition on (P) which is necessary and sufficient for weak duality. The problem (P) is weak duality invex if there exists η: $C \times C \to \mathbb{R}^n$ such that for x, $u \in C$, $g(x) \geq 0$, one has:

if there exists $\lambda \in \mathbb{R}^m_+$ such that $\nabla f(u) - \nabla(\lambda^T g(u)) = 0$

then $f(x) - f(u) \geq \eta(x,u)^T \nabla f(u)$

and $g_i(u) + \eta(x,u)^T \nabla g_i(u) \geq 0$, $i = 1,\ldots,m$;

otherwise $\eta(x,u)^T \nabla f(u) < 0$

and $\eta(x,u)^T \nabla g_i(u) \geq 0$, $i = 1,\ldots,m$.

Lemma 3.1. (Martin [9]). Weak duality holds for problems (P) and (D) if and only if (P) is weak duality invex.

Strong duality may be established as for convex programming: we require the extra condition that a constraint qualification which guarantees necessity of the Kuhn-Tucker conditions be satisfied.

Theorem 3.2 (Hanson [6]). Let x^* be optimal for (P), and f and $-g_i$ be invex with respect to the same η, $i = 1,\dots,m$. Assume one of the constraint qualifications (Kuhn-Tucker, Arrow-Hurwicz-Uzawa, reverse convex, Slater, Karlin) is satisfied. Then there exists λ^* such that (x^*,λ^*) is optimal for (D) and the respective objective functions are equal.

Strict converse duality readily follows.

Theorem 3.3 Let x^* be optimal for (P), $(\bar{x},\bar{\lambda})$ optimal for (D), and assume a constraint qualification is satisfied. If the invexity conditions of Theorem 3.2 hold, with f strictly invex at $\bar{x}$, then $x^* = \bar{x}$.

[Here, f strictly invex at $\bar{x}$ means that

$$f(x) - f(\bar{x}) > \eta(x,\bar{x})^T\nabla f(\bar{x}) \quad \text{for all } x \in C,\ x \neq \bar{x}]$$

Proof. Assume that $x^* \neq \bar{x}$.

By Theorem 3.2, there exists λ^* such that (x^*,λ^*) is optimal for (D); thus $f(x^*) = f(x^*) - \lambda^{*T}g(x^*) = f(\bar{x}) - \bar{\lambda}^T g(\bar{x})$ (1)

Now strict invexity of f at $\bar{x}$ gives

$$f(x^*) - f(\bar{x}) > \eta(x^*,\bar{x})^T\nabla f(\bar{x}), \tag{2}$$

and invexity of $-g_i$, $i = 1,\dots,m$, with $\bar{\lambda} \geq 0$, gives

$$-\bar{\lambda}^T g(x^*) + \bar{\lambda}^T g(\bar{x}) \geq -\eta(x^*,\bar{x})^T\nabla(\bar{\lambda}^T g(\bar{x})) \tag{3}$$

Adding (2) and (3) gives

$$f(x^*)-f(\bar{x}) - \bar{\lambda}^T g(x^*) + \bar{\lambda}^T g(\bar{x}) > \eta(x^*,\bar{x})^T(\nabla f(\bar{x}) - \nabla(\bar{\lambda}^T g(\bar{x})))$$

$$= 0 \quad \text{as } (\bar{x},\bar{\lambda}) \text{ feasible for (D).}$$

But, from (1), this implies $f(x^*) - \bar{\lambda}^T g(x^*) > f(x^*) - \lambda^{*T} g(x^*)$, that is, $\bar{\lambda}^T g(x^*) < 0$, which is a contradiction.

Therefore, $x^* = \bar{x}$. ∎

Further relaxation of invexity requirements for duality are achieved by the use of a Mond-Weir dual (Mond and Weir [12]).

(MWD) Maximize $f(u)$

$$\text{subject to: } \nabla f(u) - \nabla(\lambda^T g(u)) = 0$$
$$\lambda^T g(u) \le 0$$
$$\lambda \ge 0.$$

The weak, strong and converse duality results previously stated will apply to (P) and (MWD) under the assumptions that f is pseudo-invex and $-\lambda^T g$ is quasi-invex for all $\lambda \ge 0$ with respect to the same η. Strict invexity is replaced by strict pseudo-invexity for converse duality, with f said to be strictly pseudo-invex at $\bar{x}$ if for all $x \in C$, $x \ne \bar{x}$, we have

$$\eta(x,\bar{x})^t \nabla f(\bar{x}) \ge 0 \Rightarrow f(x) > f(\bar{x}).$$

4. OPTIMAL CONTROL PROBLEMS

We now consider a class of mathematical programs on an infinite dimensional function space.

(CP) Minimize $\displaystyle\int_{t_0}^{t_f} f(t,x(t),u(t))dt$

subject to: $x(t_0) = x_0, x(t_f) = x_f$ (fixed boundary conditions)

$G(t,x(t),u(t)) = x'(t)$ (state equations) (4)

$R(t,x(t),u(t)) \ge 0$

where $f: I \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $G: I \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $R: I \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^r$ are assumed to be continuously differentiable with respect to x and u almost everywhere on

$I = [t_0, t_f]$; and $x'(t)$ denotes derivative with respect to t.

x, the state variable, and u, the control variable, are assumed to be piecewise smooth functions on I.

When a constraint qualification is satisfied, the necessary conditions for (x^*, u^*) to be optimal for (CP) are (Berkovitz [2]):

there exist piecewise smooth multiplier functions $\lambda: I \to \mathbb{R}^n$, $\mu: I \to \mathbb{R}^r$

such that $F \equiv f - \lambda^T[G - x'] - \mu^T R$

satisfies $F_x = d/dt\ F_{x'}$

$$F_u = 0$$

$$\mu_i R_i = 0, \qquad i = 1, \ldots, r$$

almost everywhere on I (except that at t corresponding to discontinuities of $u^*(t)$, $F_x = d/dt\ F_{x'}$ holds for right and left hand limits.)

Here, $F_{x'}$, F_x and F_u denote partial derivatives with respect to x, x' and u respectively.

In order to establish sufficiency and duality, the notion of invexity needs to be extended to a class of functionals.

<u>Definition 4.1</u> (Mond and Smart [11]). For a scalar function $h(t, x(t), x'(t), u(t))$, associate the functional $H(x, x', u) \equiv \int_{t_0}^{t_f} h(t, x(t), x'(t), u(t))dt$. H is said to be invex in x, x' and u on I if there exist functions $\eta(t, x, x^*, x', x^{*\prime}, u, u^*) \in \mathbb{R}^n$ (with $\eta = 0$ at t such that $x(t) = x^*(t)$) and $\zeta(t, x, x^*, x' x^{*\prime}, u, u^*) \in \mathbb{R}^m$ such that

$$H(x, x', u) - H(x^*, x^{*\prime}, u^*)$$
$$\geq \int_{t_0}^{t_f} [\eta^T h_x(t, x^*, x^{*\prime}, u^*) + \frac{d\eta^T}{dt} h_{x'}(t, x^*, x^{*\prime}, u^*)$$
$$+ \zeta^T h_u(t, x^*, x^{*\prime}, u^*)]dt$$

$$(= \int_{t_0}^{t_f} [\eta^T (h_x(t,x^*,x^{*\prime},u^*) - \frac{d}{dt} h_{x'}(t,x^*,x^{*\prime},u^*)) + \zeta^T h_u(t,x^*,x^{*\prime},u^*)]dt)$$

for all piecewise smooth x, x^*, u, u^* defined on I.

Invex functionals have a similar characterization to invex functions. H is invex if and only if every critical point is a global minimizer, where (x^*,u^*) is a critical point if

$$h_x(t,x^*,x^{*\prime},u^*) = \frac{d}{dt} h_{x'}(t,x^*,x^{*\prime},u^*) \quad \text{a.e. on } I$$

and $$h_u(t,x^*,x^{*\prime},u^*) = 0 \quad \text{a.e. on } I.$$

The following theorem is proved in Mond and Smart [11].

Theorem 4.2. If there exists $(x^*,u^*,\lambda^*,\mu^*)$ satisfying the Berkovitz conditions, with (x^*,u^*) feasible for (CP), and $\int_{t_0}^{t_f} f$, $\int_{t_0}^{t_f} -\lambda^{*T}(G-x')$ and $\int_{t_0}^{t_f} -\mu^{*T}R$ are all invex with respect to the same functions η and ζ, then (x^*,u^*) is optimal for (CP).

Note that $\int_{t_0}^{t_f} -\lambda^{*T}(G-x')$ is convex if and only if $\int_{t_0}^{t_f} -\lambda^{*T}G$ is convex, but with invexity it is necessary to include the linear term $\int_{t_0}^{t_f} \lambda^{*T}x'$.

As with static problems, a dual program can be formulated and conditions for weak and strong duality obtained.

Denote by (CD) the Wolfe-type problem:

Maximize $$\int_{t_0}^{t_f} [f(t,x,u) - \lambda(t)^T(G(t,x,u) - x') - \mu(t)^T R(t,x,u)]dt$$

subject to: $x(t_0) = x_0, \quad x(t_f) = x_f$

$$f_x(t,x,u) - G_x(t,x,u)\lambda(t) - R_x(t,x,u)\mu(t) = \lambda'(t) \tag{5}$$

$$f_u(t,x,u) - G_u(t,x,u)\lambda(t) - R_u(t,x,u)\mu(t) = 0$$

$$\mu(t) \geq 0$$

where x and u are piecewise smooth functions on I, with continuous derivatives except perhaps at points of discontinuity of u, which has piecewise continuous first and second derivatives. Constraints in (CP) and (CD) may fail at these points of discontinuity, but (4) and (5) must hold for left and right hand limits.

The two subsequent results are proved in Mond and Smart [11].

<u>Theorem 4.3 (Weak Duality)</u>. If $\int_{t_0}^{t_f} f$, $\int_{t_0}^{t_f} - \lambda^T(G-x')$ and $\int_{t_0}^{t_f} -\mu^T R$ for any piecewise smooth $\lambda: I \to \mathbb{R}^n$ and $\mu: I \to \mathbb{R}^r$ with $\mu(t) \geq 0$, are invex with respect to the same functions η and ζ, then inf(CP) $\geq$ sup(CD).

<u>Theorem 4.4 (Strong Duality)</u>. Under the invexity conditions of Theorem 4.3, if (x^*,u^*) is an optimal solution of (CP) and a constraint qualification is satisfied, then there exist $\lambda: I \to \mathbb{R}^n$ and $\mu: I \to \mathbb{R}^r$ such that (x^*,u^*,λ,μ) is optimal for (CD), and the corresponding objective values are equal.

Strict invexity at $(\bar{x},\bar{u})$ occurs if there is strict inequality in the definition of invexity whenever $(x,u) \neq (\bar{x},\bar{u})$.

<u>Theorem 4.5 (Strict converse duality)</u> Let $(\bar{x},\bar{u},\bar{\lambda},\bar{\mu})$ be optimal for (CD), and (x^*,u^*) optimal for (CP). If a constraint qualification is satisfied, the invexity conditions of Theorem 4.3 hold, and $\int_{t_0}^{t_f} f$ is strictly invex at $(\bar{x},\bar{u})$, then $(x^*,u^*) = (\bar{x},\bar{u})$.

<u>Proof</u>. Assume $(x^*,u^*) \neq (\bar{x},\bar{u})$.

By Theorem 4.4, there exist $\lambda^*: I \to \mathbb{R}^n$ and $\mu^*: I \to \mathbb{R}^r$ such that

$$\int_{t_0}^{t_f} f(t,x^*,u^*)dt = \int_{t_0}^{t_f} (f(t,x^*,u^*) - \lambda^*(t)^T(G(t,x^*,u^*)-x^{*\prime})-\mu^*(t)^T R(t,x^*,u^*))dt$$

$$= \int_{t_0}^{t_f} (f(t,\bar{x},\bar{u}) - \bar{\lambda}^T(t)^T(G(t,\bar{x},\bar{u}) - \bar{\mu}(t)^T R(t,\bar{x},\bar{u}))dt \quad (6)$$

Write η for $\eta(t,x^*,\bar{x},x^{*\prime},\bar{x}',u^*,\bar{u})$ and ζ for $\zeta(t,x^*,\bar{x},x^{*\prime},\bar{x}',u^*,\bar{u})$.

Strict invexity of $\int_{t_0}^{t_f} f$ at $(\bar{x},\bar{u})$ gives

$$\int_{t_0}^{t_f} (f(t,x^*,u^*) - f(t,\bar{x},\bar{u}))dt > \int_{t_0}^{t_f} (\eta^T f_x(t,\bar{x},\bar{u}) + \zeta^T f_u(t,\bar{x},\bar{u}))dt \quad (7)$$

Invexity of $\int_{t_0}^{t_f} - \bar{\lambda}^T(G - x')$ implies that

$$\int_{t_0}^{t_f} -\bar{\lambda}(t)^T(G(t,x^*,u^*)-x^{*\prime})dt + \int_{t_0}^{t_f} \bar{\lambda}(t)^T(G(t,\bar{x},\bar{u})-\bar{x}')dt$$

$$\geq \int_{t_0}^{t_f} (-\eta^T G_x(t,\bar{x},\bar{u})\,\bar{\lambda}(t) - \eta^T\,\bar{\lambda}'(t) - \zeta^T G_u(t,\bar{x},\bar{u})\,\bar{\lambda}(t))dt \quad (8)$$

and invexity of $\int_{t_0}^{t_f} - \bar{\mu}^T R$ implies that

$$\int_{t_0}^{t_f} - \bar{\mu}(t)^T R(t,x^*,u^*)dt + \int_{t_0}^{t_f} \bar{\mu}(t)^T R(t,\bar{x},\bar{u})dt$$

$$\geq \int_{t_0}^{t_f} (- \eta^T R_x(t,\bar{x},\bar{u})\,\bar{\mu}(t) - \zeta^T G_u(t,\bar{x},\bar{u})\,\bar{\mu}(t))dt \quad (9)$$

Adding (7), (8) and (9), and using feasibility of $(\bar{x},\bar{u},\bar{\lambda},\bar{\mu})$ in (CD),

$$\int_{t_0}^{t_f} (f(t,x^*,u^*) - \bar{\lambda}(t)^T(G(t,x^*,u^*) - x^{*\prime}) - \bar{\mu}(t)^T R(t,x^*,u^*))dt$$

$$- \int_{t_0}^{t_f} (f(t,\bar{x},\bar{u}) - \bar{\lambda}(t)^T(G(t,\bar{x},\bar{u}) - \bar{x}') - \bar{\mu}(t)^T R(t,\bar{x},\bar{u}))dt > 0.$$

This implies, by (6), that

$$\int_{t_0}^{t_f} (\bar{\lambda}(t)^T(G(t,x^*,u^*) - x^{*\prime}) + \bar{\mu}(t)^T R(t,x^*,u^*))dt > 0$$

But this is a contradiction since $G(t,x^*,u^*) = x^{*\prime}$, $\bar{\mu}(t) \geq 0$ and $R(t,x^*,u^*) \geq 0$ on I.

Hence $(x^*,u^*) = (\bar{x},\bar{u})$. ∎

A Mond-Weir dual may be formulated by shifting the $\mu(t)^T R(t,x,u)$ term from the objective of (CD), and inserting a new constraint:

$$\mu(t)^T \; R(t,x(t),u(t)) \le 0.$$

The corresponding weak, strong, and strict converse duality results may then be established assuming pseudo-invexity of $\int_{t_0}^{t_f} (f - \lambda^T(G - x'))$ and quasi-invexity of $\int_{t_0}^{t_f} - \mu^T R$.

It is also possible to deal with control problems with free boundary conditions; that is, $x(t_0)$ and $x(t_f)$ unrestricted. The sufficiency and duality theorems hold using a supplementary constraint from the transversality conditions, namely $\lambda(t_0) = \lambda(t_f) = 0$. (See Mond and Smart [11]).

REFERENCES

[1] Ben-Israel, A. and B. Mond, "What is Invexity?", *J. Austral. Math. Soc.* Series B, 28 (1986), 1-9.

[2] Berkovitz, L.D., "Variational methods in Problems of Control and Programming", *J. Math. Anal. Appl.*3 (1961), 145-149.

[3] Craven, B.D., "Duality for generalized convex fractional programs", in *Generalized Concavity in Optimization and Economics*, (eds. S. Schaible and W.T. Ziemba), (Academic Press, New York, 1981), 473-489.

[4] Craven, B.D. and B.M. Glover, "Invex functions and duality", *J. Austral. Math. Soc.* 24 (1981), 1-20.

[5] Crouzeix, J-P. and J.A. Ferland, "Criteria for quasi-convexity and pseudo-convexity: Relationships and comparisons", *Math. Prog.* 23 (1982), 193-205.

[6] Hanson, M.A., "On sufficiency of the Kuhn-Tucker Conditions", *J. Math. Anal. Appl.* 80 (1981), 545-550.

[7] Kuhn, H.W. and A.W. Tucker, "Nonlinear programming" in *Proceedings of the Second Symposium on Mathematical Statistics and Probability*, (ed. J.Neyman), (University of California Press, Berkeley, Calif., 1951), 481-492.

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

[9] Martin, D.H., "The essence of invexity", *J. Optim. Th. Appl.* 47 (1985), 65-76.

[10] Mond, B. "Generalized convexity in mathematical programming", *Bull. Austral. Math. Soc.* 27 (1983), 185-202.

[11] Mond, B. and I. Smart, "Duality and sufficiency in control problems with invexity", *J. Math. Anal. Appl.* (to appear)

[12] Mond, B. and T. Weir, "Generalized concavity and duality" in *Generalized Concavity in Optimization and Economics*, (eds. S. Schaible and W.T. Ziemba), (Academic Press, New York, 1981), 263-279.

[13] Zang, I., E.U. Choo and M. Avriel, "On functions whose stationary points are global minima", *J. Optim. Th. Appl.* 22 (1977), 195-208.

Department of Mathematics,
La Trobe University,
Bundoora, Victoria, 3083,
Australia.