SUFFICIENCY AND DUALITY IN CONTROL PROBLEMS WITH GENERALIZED INVEXITY

I. HUSAIN, A. AHMED and B. AHMAD

Abstract. Sufficient optimality criteria are derived for a control problem under generalized invexity. A Mond-Weir type dual to the control problem is proposed and various duality theorems are validated under generalized invexity assumptions on functionals appearing in the problems. It is pointed out that these results can be applied to the control problem with free boundary conditions and have linkage with results for nonlinear programming problems in the presence of inequality and equality constraints already established in the literature.

1. Introduction

Optimal control models are very prominent amongst constrained optimization models because of their occurrences in a variety of popular contexts, notably, advertising investment, production and inventory, epidemic, control of a rocket etc. The planning of a river system, where it is required to make the best use of the water, can also be modelled as an optimal control problem. Optimal control models are also potentially applicable to

economic planning, and to the world models of the "Limits to Growth" kind.

Necessary optimality conditions for existence of extremal solution for a variational problem in the presence of inequality and equality constraints were obtained by Valentine [8]. Using Valentine's results, Berkovitz [2] obtained corresponding Fritz John type necessary optimality conditions for a control problem. Mond and Hanson [5] pointed out that if the optimal solution for the problem is normal, then the Fritz John type optimality conditions reduce to Karush-Kuhn-Tucker conditions. Using these Karush-Kuhn-Tucker optimality conditions, Mond and Hanson [5] presented Wolfe type dual and established weak, strong and converse duality theorems under convexity conditions. Abraham and Buie [1] studied duality for continuous programming and optimal control from a unified point of view. Later Mond and Smart [6] proved that for invex functions, the necessary conditions of Berkovitz [2] together with normality conditions, are sufficient for optimality and also derived some duality results under invexity.

In this paper, it is shown that for generalized invexity assumptions on functionals, the necessary conditions [2] in the control problems are also sufficient. As an application of Berkovitz's [2] optimality conditions with normality, a Mond-Weir [7] type dual to the control problem is constructed and under generalized invexity of functionals, various duality results are derived. It is indicated that these duality results are applicable to the control problem with free boundary conditions and also related to those for nonlinear programming problems already existing in the literature.

2. Control problem and related preliminaries

Let $\mathbb{R}^n$ denotes an n-dimensional Euclidean space, $I = [a, b]$ be a real interval and $f \colon I \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be a continuously differentiable with respect to each of its arguments. For the function $f(t, x, u)$, where $x \colon I \to \mathbb{R}^n$ is differentiable with its derivative $\dot{x}$ and $u \colon I \to \mathbb{R}^m$ is the smooth function, denote the partial derivatives of f by f_t, f_x and f_u, where

$$f_t \colon \frac{\partial f}{\partial t}, \ f_x = \left(\frac{\partial f}{\partial x^1}, \dots, \frac{\partial f}{\partial x^n} \right)^T, \ f_u = \left(\frac{\partial f}{\partial u^1}, \dots, \frac{\partial f}{\partial u^m} \right)^T,$$

$$x = \left(x^1, \dots, x^n \right)^T \quad \text{and} \quad u = \left(u^1, \dots, u^m \right)^T.$$

For an m-dimensional vector function $g(t, x, u)$, the gradient with respect to x is

$$g_x = \begin{pmatrix} \dfrac{\partial g^1}{\partial x^1}, & \cdots, & \dfrac{\partial g^p}{\partial x^n} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial g^1}{\partial x^n}, & \cdots, & \dfrac{\partial g^p}{\partial x^n} \end{pmatrix}, \quad \text{an } n \times p \text{ matrix of first order derivatives.}$$

Here $u(t)$ is the control variable and $x(t)$ is the state variable, u is related to x via the state equation $\dot{x} = h(t, x, u)$. Gradients with respect to u are defined analogously.

A control problem is to transfer the state vector from an initial state $x(a) = \alpha$ to a final state $x(b) = \beta$ so as to minimize a functional, subject to constraints on the control and state variables. A control problem can be stated formally as,

$$\text{Problem (CP) (Primal):} \quad \underset{x \in X,\, u \in U}{\text{Minimize}} \int_a^b f(t, x, u) dt,$$

subject to

$$x(a) = \alpha, \ x(b) = \beta, \tag{1}$$

$$h(t, x, u) = \dot{x}, \ t \in I, \tag{2}$$

$$g(t, x, u) \leq 0, \ t \in I, \tag{3}$$

(i) f is as before, $g \colon I \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ and $h \colon I \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are continuously differentiable functions with respect to each of its arguments.

(ii) X is the space of continuously differentiable state functions $x \colon I \to \mathbb{R}^n$ such that $x(a) = \alpha$, $x(b) = \beta$, equipped with the norm $\|x\| = \|x\|_\infty + \|Dx\|_\infty$, and u is the space of piecewise continuous control functions $u \colon I \to \mathbb{R}^m$ has the uniform norm $\|\cdot\|_\infty$, and

(iii) The differential equation (2) for x with the initial conditions expressed as $x(t) = x(a) + \int_a^t h(s, x(s), u(s)) ds$, $t \in I$, may be written as $Dx = H(x, u)$, where the map: $X \times U \to C(I, \mathbb{R}^n)$, $C(I, \mathbb{R}^n)$ being the space of continuous functions from $I \to \mathbb{R}^n$, defined by $H(x, u)(t) = h(t, x(t), u(t))$.

Following Craven [3], the control problem can be expressed as,

$$\text{(ECP):} \quad \underset{x \in X,\, u \in U}{\text{Minimize}} F(x, u)$$

subject to

$$Dx = H(x, u),$$

$$-G(x, u) \in S,$$

Where G is function from $X \times U$ into $C(I, \mathbb{R}^p)$ given by $G(x, u)(t) = g(t, x(t), u(t))$ from $x \in X$, $u \in U$, and $t \in I$; S is the convex cone of functions in $C(I, \mathbb{R}^p)$ whose components are non-negative; thus S has interior points.

Necessary optimality conditions for existence of extremal solution for a variational problem subject to both equality and inequality constraints was given by Valentine [8]. Invoking Valentine's [8] results, Berkovitz [2] obtained corresponding necessary optimality conditions for the above control problem (CP). Here we mention the Fritz John optimality conditions derived by Craven [3] in the form of the following proposition which will be required in the sequel.

Proposition 1 (Necessary optimality conditions). *If $(\bar{x}, \bar{u}) \in X \times U$ an optimal solution of* (CP) *and the Fréchet derivatives $Q' = (D - H_x(x, u), -H_u(x, u))$ is surjective, then there exist Lagrange multipliers $\lambda_0 \in \mathbb{R}$, and piecewise smooth functions $\lambda \colon I \to \mathbb{R}^p$ and $\mu \colon I \to \mathbb{R}^n$ satisfying, for all $t \in I$,*

$$\lambda_0 f_x(t, \bar{x}, \bar{u}) + \lambda(t)^T g_x(t, \bar{x}, \bar{u}) + \mu(t)^T h_x(t, \bar{x}, \bar{u}) + \dot{\mu}(t) = 0,$$

$$\lambda_0 f_u(t, \bar{x}, \bar{u}) + \lambda(t)^T g_u(t, \bar{x}, \bar{u}) + \mu(t)^T h_u(t, \bar{x}, \bar{u}) = 0,$$

$$\lambda(t)^T g(t, \bar{x}, \bar{u}) = 0,$$

$$(\lambda_0, \lambda_{(t)}) \geq 0,$$

$$(\lambda_0, \lambda_{(t)}, \mu_{(t)}) \neq 0.$$

The above conditions will become Karush-Kuhn-Tucker conditions if $\lambda_0 > 0$. Therefore, if we assume that the optimal solutions $(\bar{x}, \bar{u})$is normal, then without any loss of generality, we can set $\lambda_0 = 1$. Thus from the above we have the Karush-Kuhn- Tucker type optimality conditions

$$f_x(t, \bar{x}, \bar{u}) + \lambda(t)^T g_x(t, \bar{x}, \bar{u}) + \mu(t)^T h_x(t, \bar{x}, \bar{u}) + \dot{\mu}(t) = 0, \quad t \in I, \qquad (4)$$

$$f_u(t, \bar{x}, \bar{u}) + \lambda(t)^T g_u(t, \bar{x}, \bar{u}) + \mu(t)^T h_u(t, \bar{x}, \bar{u}) = 0, \quad t \in I, \qquad (5)$$

$$\lambda(t)^T g(t, \bar{x}, \bar{u}) = 0, \quad t \in I, \qquad (6)$$

$$\lambda(t) \geq 0, \quad t \in I. \qquad (7)$$

Using these optimality conditions, Mond and Hanson [5] constructed following Wolfe type dual.

Problem (CD) (Dual) :

Maximize $\int_a^b \left\{ f(t, x, u) + \lambda(t)^T g(t, x, u) + \mu(t)^T \left(h(t, x, u) - \dot{x} \right) \right\} dt$
subject to

$$f_x(t, x, u) + \lambda(t)^T g_x(t, \bar{x}, \bar{u}) + \mu(t)^T h_x(t, x, u) + \dot{\mu}(t) = 0, \quad t \in I,$$
$$f_u(t, x, u) + \lambda(t)^T g_u(t, \bar{x}, \bar{u}) + \mu(t)^T h_u(t, x, u) = 0, \quad t \in I,$$
$$\lambda(t) \geq 0, \quad t \in I.$$

In [5], [CP] and (CD) are shown to be a dual pair if f, g and h are all convex in x and u. Subsequently, Mond and Smart [6] extended this duality by introducing the following invexity requirement.

Definition 1 (Invex [6]). If there exists vector function $\eta(t, x, \bar{x}) \in \mathbb{R}^n$ with $\eta = 0$ at t if $x(t) = \bar{x}(t)$, and there exists vector function $\xi(t, u, \bar{u}) \in \mathbb{R}^m$ such that for scalar function $\Phi(t, x, \dot{x}, u)$, the functional

$$\Phi(x, \dot{x}, u) = \int_a^b \phi(t, x, \dot{x}, u)dt$$

satisfies

$$\Phi(x, \dot{x}, u) - \phi(\bar{x}, \dot{\bar{x}}, \bar{u})$$
$$\geq \int_a^b \left[\eta^T \phi_x \left(t, \dot{\bar{x}}, \bar{x}, \bar{u} \right) + \left(\frac{dn}{dt} \right)^T \phi_{\dot{x}}(t, \bar{x}, \dot{\bar{x}}, \bar{u}) + \xi^T \phi_u(t, \bar{x}, \dot{\bar{x}}, \bar{u}) \right] dt$$

then ϕ is said to be invex at x, $\dot{x}$ and u on I with respect to η and ξ.

In [6] Mond and Smart proved weak, strong and converse duality theorems under the invexity of $\int_a^b f dt$, $\int_a^b \lambda^T g dt$, for $\lambda(t) \in \mathbb{R}^p$ with $\lambda(t) \geq 0$, $t \in I$ and $\int_a^b \mu^T h dt$ for any $\mu(t) \in \mathbb{R}^n$, $t \in I$.

3. Generalized invexity

In this section, we extend the notion of invexity for a functional given in [6] to a large class of functionals, as these will be required for subsequent analysis.

Definition 2. For a scalar function $\phi(t, x, \dot{x}, u)$ the functional $\Phi(x, \dot{x}, u) = \int_a^b \phi(t, x, \dot{x}, u)dt$ is said to be pseudoinvex at x, $\dot{x}$ and u if there exist vector function $\eta(t, x, \bar{x}) \in \mathbb{R}^n$ with $\eta = 0$ at t if $x(t) = \bar{x}(t)$ and $\xi(t, u, \bar{u}) \in \mathbb{R}^m$ such that for all $(x, \dot{x}, u) \neq (\bar{x}, \dot{\bar{x}}, \bar{u})$

$$\int_a^b \left[\eta^T \phi_x(t, \bar{x}, \dot{\bar{x}}, \bar{u}) + \left(\frac{d\eta}{dt} \right)^T \phi_{\dot{x}}(t, \bar{x}, \dot{\bar{x}}, \bar{u}) + \xi^T \phi_u(t, \bar{x}, \dot{\bar{x}}, \bar{u}) \right] dt \geq 0$$
$$\Rightarrow \Phi(x, \dot{x}, u) \geq \Phi(\bar{x}, \dot{\bar{x}}, \bar{u}).$$

Definition 3 (Strictly Pseudoinvex)**.** The functional Φ is said to be strictly pseudoinvex, if there exist vector functions $\eta(t, x, \bar{x}) \in \mathbb{R}^n$ with $\eta = 0$ at t if $x(t) = \bar{x}(t)$ and $\xi(t, u, \bar{u}) \in \mathbb{R}^m$ such that

$$\int_a^b \left[\eta^T \phi_x(t, \bar{x}, \dot{\bar{x}}, \bar{u}) + \left(\frac{d\eta}{dt} \right)^T \phi_{\dot{x}}(t, \bar{x}, \dot{\bar{x}}, \bar{u}) + \xi^T \phi_u(t, \bar{x}, \dot{\bar{x}}, \bar{u}) \right] dt \geq 0$$

$$\Rightarrow \Phi(x, \dot{x}, u) > \Phi(\bar{x}, \dot{\bar{x}}, \bar{u}).$$

Definition 4 (Quasi-invex)**.** The functional Φ is said to be quasi-invex, if there exist vector functions $\eta(t, x, \bar{x}) \in \mathbb{R}^n$ with $\eta = 0$ at t if $x(t) = \bar{x}(t)$ and $\xi(t, u, \bar{u}) \in \mathbb{R}^m$ such that

$$\Phi(x, \dot{x}, u) \leq \Phi(\bar{x}, \dot{\bar{x}}, \bar{u})$$

$$\Rightarrow \int_a^b \left[\eta^T \phi_x(t, \bar{x}, \dot{\bar{x}}, \bar{u}) + \left(\frac{d\eta}{dt} \right)^T \phi_{\dot{x}}(t, \bar{x}, \dot{\bar{x}}, \bar{u}) + \xi^T \phi_u(t, \bar{x}, \dot{\bar{x}}, \bar{u}) \right] dt \leq 0.$$

4. Sufficiency of optimality conditions

It can be proved that for generalized invex functionals, the Karush-Kuhn-Tucker optimality conditions given in Section 2 are sufficient for optimality.

Theorem 1. *If there exists $(\bar{x}, \bar{u}, \bar{\lambda}, \bar{\mu})$ such that the conditions (4)–(7) hold with $(\bar{x}, \bar{u})$ feasible for* (CP) *and $\int_a^b f dt$ is pseudoinvex and $\int_a^b \left(\bar{\lambda}^T g + \bar{\mu}^T(h - \dot{x}) \right) dt$ is quasi-invex with respect to the same η and ξ, then $(\bar{x}, \bar{u})$ is an optimal solution of* (CP).

Proof. Assume that $(\bar{x}, \bar{u})$ is not optimal for (CP). Then there exists $(x, u) \neq (\bar{x}, \bar{u})$, i.e., (x, u) feasible for (CP), such that $\int_a^b f(t, x, u) dt < \int_a^b f(t, \bar{x}, \bar{u}) dt$.

This, because of pseudoinvexity of $\int_a^b f dt$ with respect to the same η and ξ, it follows that

$$\int_a^b \left[\eta^T f_x(t, \bar{x}, \bar{u}) + \xi^T f_u(t, \bar{x}, \bar{u}) \right] dt < 0.$$

Using (4) and (5), this yields

$$0 < \int_a^b \left[\eta^T \left(\bar{\lambda}^T(t) g_x(t, \bar{x}, \bar{u}) + \bar{\mu}(t)^T h_x(t, \bar{x}, \bar{u}) + \dot{\bar{\mu}}(t) \right) \right.$$

$$\left. + \xi^T \left(\bar{\lambda}^T(t)^T g_u(t, \bar{x}, \bar{u}) + \bar{\mu}(t)^T h_u(t, \bar{x}, \bar{u}) \right) \right] dt$$

$$= \int_a^b \eta^T \left(\bar{\lambda}(t)^T g_x(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T h_x(t,\bar{x},\bar{u}) \right) dt + \int_a^b \eta^T \dot{\bar{\mu}}(t) dt$$

$$+ \int_a^b \xi^T \left(\bar{\lambda}(t)^T g_u(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T h_u(t,\bar{x},\bar{u}) \right) dt$$

$$= \int_a^b \left[\eta^T \left(\bar{\lambda}^T(t) g_x(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T h_x(t,\bar{x},\bar{u}) \right) - \left(\frac{d\eta}{dt} \right)^T \bar{\mu}(t) \right.$$

$$\left. + \xi^T \left(\bar{\lambda}^T(t) g_u(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T h_u(t,\bar{x},\bar{u}) \right) \right] dt + \eta\mu(t)|_{t=b}^{t=a}$$

(by integrating by parts)

$$= \int_a^b \left[\eta^T \left(\bar{\lambda}^T(t) g_x(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T h_x(t,\bar{x},\bar{u}) \right) - \left(\frac{d\eta}{dt} \right)^T \bar{\mu}(t) \right.$$

$$\left. + \xi^T \left(\bar{\lambda}^T(t) g_u(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T h_u(t,\bar{x},\bar{u}) \right) \right] dt$$

(using $\eta = 0$ at t if $x(t) = \bar{x}(t)$).

By quasi-invexity of $\int_a^b (\bar{\lambda} g + \bar{\mu}(h - \dot{x})) dt$, this implies

$$\int_a^b \left\{ \bar{\lambda}(t)^T g(t,x,u) + \bar{\mu}(t)^T \left(h(t,x,u) - \dot{x} \right) \right\} dt$$

$$> \int_a^b \left\{ \bar{\lambda}(t)^T g(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T \left(h(t,\bar{x},\bar{u}) - \dot{\bar{x}} \right) \right\} dt.$$

Using (6) and also $\bar{\mu}(t)^T (h(t,\bar{x},\bar{u}) - \dot{\bar{x}}) = 0$, the above inequality gives

$$\int_a^b \left\{ \bar{\lambda}(t)^T g(t,x,u) + \bar{\mu}(t)^T \left(h(t,x,u) - \dot{x} \right) \right\} dt > 0. \tag{8}$$

Since (x,u) is feasible for (CP), $g(t,x,u) \leq 0$, $t \in I$ and $h(t,x,u) - \dot{x} = 0$. Hence for $\bar{\lambda}(t) \geq 0$, $t \in T$ and $\bar{\mu}(t) \in \mathbb{R}^n$, we have

$$\int_a^b \left\{ \bar{\lambda}(t)^T g(t,x,u) + \bar{\mu}(t)^T \left(h(t,x,u) - \dot{x} \right) \right\} dt \leq 0. \tag{9}$$

Consequently (8) contradicts (9). Thus $(\bar{x},\bar{u})$ is, indeed, an optimal solution of the control problem (CP). $\qquad\square$

5. DUALITY

We formulate the following dual (CD) to the primal problem (CP) in the spirit of Mond and Weir [7].

Problem (CD) (Dual) : Maximize $\int_a^b f(t, x, u)dt$

subject to

$$x(a) = \alpha, \ x(b) = \beta, \tag{10}$$

$$f_x(t, x, u) + \lambda(t)^T g_x(t, x, u) + \mu(t)^T h_x(t, x, u) + \dot{\mu}(t) = 0, \quad t \in I, \tag{11}$$

$$f_u(t, x, u) + \lambda(t)^T g_u(t, x, u) + \mu(t)^T h_u(t, x, u) = 0, \quad t \in I, \tag{12}$$

$$\int_a^b \left(\lambda(t)^T g(t, x, u) + \mu(t)^T \left(h(t, x, u) - \dot{x} \right) \right) dt \geq 0, \tag{13}$$

$$\lambda(t) \geq 0, \quad t \in I. \tag{14}$$

Theorem 2 (Weak Duality). *Let $(\bar{x}, \bar{u})$ and (x, u, λ, μ) be feasible solution for (CP) and (CD) respectively. If for all feasible $(\bar{x}, \bar{u}, x, u, \lambda, \mu)$, $\int_a^b f dt$ is pseudoinvex and $\int \left(\lambda^T g + \mu^T (h - \dot{x}) \right) dt$ for $\lambda(t) \in \mathbb{R}^n$, $\lambda(t) \geq 0$, $t \in I$ and $\mu(t) \in \mathbb{R}^n$ is quasi-invex with respect to the same η and ξ, then*

$$\max(\text{CP}) \geq \min(\text{CD}).$$

Proof. Since $(\bar{x}, \bar{u})$ is feasible for the problem (CP) and (x, u, λ, μ) feasible for the problem (CD), it implies that

$$\int_a^b \left(\lambda(t)^T g(t, \bar{x}, \bar{u}) + \mu(t)^T (h(t, \bar{x}, \bar{u}) - \dot{\bar{x}}) \right) dt$$

$$\leq \int_a^b \left(\lambda(t)^T g(t, x, u) + \mu(t)^T \left(h(t, x, u) - \dot{x} \right) \right) dt.$$

This, because of quasi-invexity of $\int_a^b (\lambda^T g + \mu^T(h - \dot{x}))dt$, implies

$$0 \geq \int_a^b \left\{ \eta^T \left(\lambda(t)^T g_x(t, x, u) + \mu(t)^T h_x(t, x, u) \right) \right.$$

$$\left. + \xi^T \left(\lambda(t)^T g_u(t, x, u) + \mu(t)^T h_u(t, x, u) \right) \right\} dt - \int_a^b \left(\frac{d\eta}{dt} \right)^T \mu(t)dt$$

$$= \int_a^b \left\{ \eta^T \left(\lambda(t)^T g_x(t, x, u) + \mu(t)^T h_x(t, x, u) \right) \right.$$

$$\left. + \xi^T \left(\lambda(t)^T g_u(t, x, u) + \mu(t)^T h_u(t, x, u) \right) \right\} dt - \mu(t)\eta|_{t=a}^{t=b} + \int_a^b \dot{\mu}(t)^T \eta dt$$

(by integration by parts)

$$= \int_a^b \left\{ \eta^T \left(\lambda(t)^T g_x(t,x,u) + \mu(t)^T h_x(t,x,\dot{u}) + \dot{\mu}(t) \right) \right.$$
$$\left. + \xi^T \left(\lambda(t)^T g_u(t,x,u) + \mu(t)^T h_u(t,x,u) \right) \right\} dt$$

(as fixed boundary conditions give $\eta = 0$ at $t = a$ and $t = b$).

Using (11) and (12), we have

$$\int_a^b \left\{ \eta^T f_x(t,x,u) + \xi^T f_x(t,x,u) \right\} dt \geq 0.$$

By pseudoinvexity $\int_a^b f dt$, this gives

$$\int_a^b f(t,\bar{x},\bar{u}) dt \geq \int_a^b f(t,x,u) dt.$$

That is,

$$\inf(\mathrm{CP}) \geq \sup(\mathrm{CD}).$$

$\square$

Theorem 3 (Strong Duality). *Under generalized invexity conditions of Theorem 2, if $(\bar{x}, \bar{u})$ is an optimal solution of the problem (CP) and is also normal, then there exist piecewise smooth functions $\bar{\lambda} \colon I \to \mathbb{R}^p$ and $\bar{\mu} \colon I \to \mathbb{R}^n$ such that $(\bar{x}, \bar{u}, \bar{\lambda}, \bar{\mu})$ is an optimal solution of (CP) and the corresponding objective values are equal.*

Proof. Since $(\bar{x}, \bar{u})$ is optimal solution for (CP) and is normal, by Proposition 1, there exist piecewise smooth functions $\bar{\lambda} \colon I \to \mathbb{R}^p$ and $\bar{\mu} \colon I \to \mathbb{R}^n$ such that the condition (4)–(7) are satisfied. Since $\bar{\lambda}(t)^T g(t,\bar{x},\bar{u}) = 0$ and $\bar{\mu}(t)^T (h(t,\bar{x},\bar{u}) - \dot{\bar{x}}) = 0$, $\int_a^b (\bar{\lambda}(t)^T g(t,\bar{x},\bar{u}) + \bar{\mu}(t)^T g(t,\bar{x},\bar{u})) dt = 0$. Thus, this together with (4), (5) and (7) implies that $(\bar{x}, \bar{u}, \bar{\lambda}, \bar{\mu})$ is feasible for (CD) and the corresponding objective values are the same as it is evident from the formulation of the primal and dual problems. So by Theorem 2, $(\bar{x}, \bar{u}, \bar{\lambda}, \bar{\mu})$ is an optimal solution for (CD). $\square$

Theorem 4 (Strict Converse Duality). *Let $(\bar{x}, \bar{u})$ be an optimal solution of (CP) and also normal. If $(\hat{x}, \hat{u}, \hat{\lambda}, \hat{\mu})$ is an optimal solution; and $\int_a^b f dt$ is strictly pseudoinvex and $\int_a^b (\hat{\lambda}^T g + \hat{\mu}^T (h - \dot{x})) dt$ is quasi-invex at $(\hat{x}, \hat{u})$ with respect to the same η and ξ, then $(\bar{x}, \bar{u}) = (\hat{x}, \hat{u})$, i.e., $(\hat{x}, \hat{u})$ is an optimal solution of (CP).*

Proof. Assume that $(\bar{x}, \bar{u}) \neq (\hat{x}, \hat{u})$.

Since $(\bar{x}, \bar{u})$ is an optimal solution of (CP) at which normality condition is met, and since conditions of Theorem 1 are satisfied, then, by Theorem 3, there exist piecewise smooth $\bar{\lambda} \colon I \to \mathbb{R}^p$ and $\bar{\mu} \colon I \to \mathbb{R}^n$ such that $(\bar{x}, \bar{u}, \bar{\lambda}, \bar{\mu})$ is an optimal solution of (CD) and

$$\int_a^b f(t, \bar{x}, \bar{u})dt = \int_a^b f(t, \hat{x}, \hat{u})dt. \tag{15}$$

By the feasibility of $(\bar{x}, \bar{u})$ for (CP) and $(\hat{x}, \hat{u}, \hat{\lambda}, \hat{\mu})$ for (CD), it implies,

$$\int_a^b \left(\hat{\lambda}(t)^T g(t, \bar{x}, \bar{u}) + \hat{\mu}(t)^T \left(h(t, \bar{x}, \bar{u}) - \dot{\bar{x}} \right) \right) dt \leq 0,$$

and

$$\int_a^b \left(\hat{\lambda}(t)^T g(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T \left(h(t, \hat{x}, \hat{u}) - \dot{\hat{x}} \right) \right) dt \geq 0.$$

Combining these inequalities we have

$$\int_a^b \left(\hat{\lambda}(t)^T g(t, \bar{x}, \bar{u}) + \hat{\mu}(t)^T \left(h(t, \bar{x}, \bar{u}) \right) \right) dt$$

$$\leq \int_a^b \left(\hat{\lambda}(t)^T g(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T \left(h(t, \hat{x}, \hat{u}) \right) \right) dt.$$

Because of the quasi-invexity of $\int_a^b \left(\hat{\lambda}^T g + \hat{\mu}^T h(x - \dot{x}) \right) dt$ at $(\hat{x}, \hat{u})$, this yields

$$0 \geq \int_a^b \left\{ \eta^T \left(\hat{\lambda}(t)^T g_x(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T h_x(t, \hat{x}, \hat{u}) \right) - \left(\frac{d\eta}{dt} \right)^T \hat{\mu}(t) \right.$$

$$\left. + \xi^T \left(\hat{\lambda}(t)^T g_u(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T h_u(t, \hat{x}, \hat{u}) \right) \right\} dt$$

$$= \int_a^b \left\{ \eta^T \left(\hat{\lambda}(t)^T g_x(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T h_x(t, \hat{x}, \hat{u}) + \hat{\mu}(t) \right) \right.$$

$$\left. + \xi^T \left(\hat{\lambda}(t)^T g_u(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T h_u(t, \hat{x}, \hat{u}) \right) \right\} dt - \eta^T \hat{\mu}(t)|_{t=a}^{t=0}$$

(by integration by parts)

$$0 \geq \int_a^b \left\{ \eta^T \left(\hat{\lambda}(t)^T g_x(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T h_x(t, \hat{x}, \hat{u}) + \hat{\mu}(t) \right) \right.$$

$$\left. + \xi^T \left(\hat{\lambda}(t)^T g_u(t, \hat{x}, \hat{u}) + \hat{\mu}(t)^T h_u(t, \hat{x}, \hat{u}) \right) \right\} dt. \tag{16}$$

Because $(\hat{x}, \hat{u})$ is feasible for (CD), we have that

$$f_x(t, \hat{x}, \hat{u}) + \hat{\lambda}(t)^T g_x(t, \hat{x}, \hat{u}) + \hat{\mu}(t) h_x(t, \hat{x}, \hat{u}) + \dot{\hat{\mu}}(t) = 0, \quad t \in I,$$

$$f_u(t, \hat{x}, \hat{u}) + \hat{\lambda}(t) g_u(t, \hat{x}, \hat{u}) + \hat{\mu}(t) h_u(t, \hat{x}, \hat{u}) = 0, \quad t \in I.$$

Using these equations in (16), we have

$$\int_a^b \left(\eta^T f_x(t, \hat{x}, \hat{u}) + \xi^T f_u(t, \hat{x}, \hat{u}) \right) dt \geq 0.$$

Thus, by strict pseudoinvexity of $\int_a^b f \, dt$ yield,

$$\int_a^b f(t, \bar{x}, \bar{u}) dt > \int_a^b f(t, \hat{x}, \hat{u}) dt.$$

This contradicts (15). Hence $(\hat{x}, \hat{u}) = (\bar{x}, \bar{u})$, i.e., $(\hat{x}, \hat{u})$ is an optimal solution of (CP). $\qquad\square$

Now, we shall prove converse duality under the assumption that f, g and h are twice continuously differentiable. The problem (CD) may be written in minimization form as follows:

$$\text{Minimize} -\psi(x, u, \lambda, \mu)$$

$$\text{subject to}$$

$$x(a) = \alpha, x(b) = \beta$$

$$\theta_1(t, x(t), u(t), \lambda(t), \mu(t), \dot{\mu}(t))$$

$$= f_x + \lambda(t)^T g_x + \mu(t)^T h_x + \dot{\mu}(t) = 0, \ t \in I,$$

$$\theta_2(t, x(t), u(t), \lambda(t), \mu(t)) = f_u$$

$$\lambda(t)^T g_u + \mu(t)^T h_u = 0, \ t \in I,$$

with $f_x \equiv f_x(t, x(t), u(t))$, $g_x \equiv g_x(t, x(t), u(t))$, $h_x \equiv h_x(t, x(t), u(t))$, etc.

Consider $\theta_1(\cdot, x(\cdot), u(\cdot), \lambda(\cdot), \mu(\cdot), \dot{\mu}(\cdot))$ as defining a mapping $Q^1 \colon X \times U \times V \times \Lambda \to B_1$, where V is the space of piecewise smooth functions λ, Λ is the space of differentiable functions μ and B_1, is a Banach Space; and also consider $\theta_2(\cdot, x(\cdot), u(\cdot), \lambda(\cdot), \mu(\cdot))$ as defining a mapping $Q^2 \colon X \times U \times V \times \Lambda \to B_2$, where B_2 is another Banach Space. In order to apply Proposition 1 or results of Valentine [8], some restrictions are needed on the equality constraints

$$\theta_1(\cdot) = 0 \text{ and } \theta_2(\cdot) = 0.$$

It suffices if Fréchet derivatives

$$Q^{1'} = \left[Q_x^1, Q_u^1, Q_\lambda^1, Q_\mu^1 \right] \text{ and } Q^{2'} = \left[Q_x^2, Q_u^2, Q_\lambda^2, Q_\mu^2 \right]$$

have weak* closed range. Denote $\bar{f} \equiv f(t, \bar{x}(t), \bar{u}(t))$, $\bar{f}_x \equiv f_x(t, \bar{x}(t), \bar{u}(t))$, etc.

Theorem 5 (Converse Duality). *Let $(\bar{x}, \bar{u}, \bar{\lambda}, \bar{\mu})$ be an optimal solution of* (CD). *Assume that*

(i) *the Fréchet derivative Q^1 and Q^2 have weak closed range.*

(ii) *Corresponding to (5), there exists a piecewise smooth Lagrange multiplier $\beta \colon I \to \mathbb{R}^n$*

with its derivative $\dot{\beta}(t) \geq 0, \quad t \in I$ and $\beta(a) = 0 = \beta(b)$.

(iii) *$\int_a^b \sigma(t)^T M(t) \sigma(t) dt = 0 \Rightarrow \sigma(t) = 0$, where $\sigma(t) \in \mathbb{R}^{n+m}$ and*

$$M(t) = \begin{pmatrix} \bar{f}_{xx} + \bar{\lambda}(t)^T \bar{g}_{xx} + \bar{\mu}(t)^T \bar{h}_{xx}, & \bar{f}_{ux} + \bar{\lambda}(t)^T g_{ux} + \bar{\mu}(t)^T \bar{h}_{ux} \\ \bar{f}_{xu} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T h_{xu}, & f_{xu} + \bar{\lambda}(t)^T g_{uu} + \bar{\mu}(t)^T h_{uu} \end{pmatrix},$$

is a positive definite and

(iv) *$\bar{f}_x + \bar{\lambda}(t)^T g_x + \bar{\mu}(t)^T \bar{h}_x + \dot{\bar{\mu}}(t) \neq 0, \ \bar{f}_u + \bar{\lambda}(t)^T g_u + \bar{\mu}(t)^T h_u \neq 0, \ t \in I.$*

If the hypotheses of Theorem 2 are satisfied, $(\bar{x}, \bar{u})$ is an optimal solution of (CP) *and the objective values of* (CP) *and* (CD) *are equal.*

Proof. Since $(\bar{x}, \bar{y}, \bar{\lambda}, \bar{\mu})$ is an optimal solution of (CD), an application of Proposition 1 shows that there exist Lagrange multipliers $\alpha \in \mathbb{R}$, piecewise smooth functions $\beta \colon I \to \mathbb{R}^n$ with $\beta(a) = 0 = \beta(b)$ and its derivative $\dot{\beta}(t) \geq 0, t \in I, \theta \colon I \to \mathbb{R}^m$, and $\zeta \colon I \to \mathbb{R}^m$ and $\gamma \in \mathbb{R}$ such that

$$\alpha f_x + \beta(t)^T \left(f_{xx} + \bar{\lambda}(t)^T g_{xx} + \bar{\mu}(t)^T h_{xx} \right) + \theta(t)^T \left(\bar{f}_{ux} + \bar{\lambda}(t)^T g_{ux} \right.$$
$$\left. + \bar{\mu}(t)^T h_{ux} \right) + \gamma \left(\bar{\lambda}(t)^T g_x + \bar{\mu}(t)^T h_x \right) = 0, \quad t \in I, \tag{17}$$

$$\alpha f_u + \beta(t)^T \left(f_{xu} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T h_{xu} \right) + \theta(t)^T \left(\bar{f}_{uu} + \bar{\lambda}(t)^T g_{uu} \right.$$
$$\left. + \bar{\mu}(t)^T h_{uu} \right) + \gamma \left(\bar{\lambda}(t)^T g_u + \bar{\mu}(t)^T h_u + \dot{\bar{\mu}}(t) \right) = 0, \quad t \in I, \tag{18}$$

$$\beta(t)^T g_x + \theta(t)^T g_u + \gamma g + \xi(t) = 0, \quad t \in I, \tag{19}$$

$$\beta(t)^T h_x - \dot{\beta}(t) + \theta(t)^T h_u + \gamma(h - \dot{x}) = 0, \quad t \in I, \tag{20}$$

$$\gamma \int_a^b (\bar{\lambda}(t)^T g + \bar{\mu}(t)^T (\bar{h} - \dot{x})) dt = 0, \tag{21}$$

$$\bar{\lambda}(t)^T \zeta(t) = 0, \quad t \in I, \tag{22}$$

$$(\alpha, \zeta(t), \gamma) \geq 0, \quad t \in I, \tag{23}$$

$$(\alpha, \beta(t), \theta(t), \zeta(t), \gamma) \neq 0, \quad t \in I. \tag{24}$$

Using (11) and (12) in (17) and (18) respectively, we have

$$(\gamma - \alpha) \left(\bar{\lambda}(t)^T g_x + \bar{\mu}(t)^T h_x + \dot{\bar{\mu}}(t) \right) + \beta(t)^T \left(\bar{f}_{xx} + \bar{\lambda}(t)^T \bar{g}_{xx} \right.$$
$$\left. + \bar{\mu}(t)^T h_{xx} \right) + \theta(t)^T \left(f_{xu} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T \bar{h}_{xu} \right) = 0, \ t \in I, \tag{25}$$

$$(\gamma - \alpha) \left(\bar{\lambda}(t)^T g_u + \bar{\mu}(t)^T \bar{h}_u \right) + \beta^T(t)^T \left(\bar{f}_{xx} + \bar{\lambda}(t) \bar{g}_{xx} + \bar{\mu}(t)^T h_{xx} \right)$$

$$+\theta(t)^T \left(\bar{f}_{uu} + \bar{\lambda}(t)^T g_{uu} + \bar{\mu}(t)^T h_{uu}\right) = 0, \quad t \in I. \tag{26}$$

Multiplying (19) and (20) by $\bar{\lambda}(t)^T$ respectively and then adding the resulting equations, we have

$$\int_a^b \left\{ \beta(t)^T \left(\lambda(t)^T g_x + \bar{\mu}(t)^T h_x\right) + \theta(t)^T \left(\lambda(t)^T g_u + \mu(t)^T h_u\right) \right\} dt$$

$$+\gamma \int_a^b \left(\lambda(t)^T g + \mu(t)^T h\right) dt + \int_a^b \lambda(t)^T \zeta(t) dt$$

$$= \int_a^b \beta(t)\mu(t) dt = \mu(t)^T \beta(t)|_{t=a}^{t=b} - \int_a^b \beta(t)^T \dot{\mu}(t) dt$$

(by integration by parts).

Using $\beta(a) = 0 = \beta(b)$, (21) and (22), this implies,

$$\int_a^b \left\{ \beta(t)^T \left(\bar{\lambda}(t)^T g_x + \bar{\mu}(t)^T h_x + \dot{\mu}(t)\right) + \theta(t)^T \left(\bar{\lambda}(t)^T g_u + \bar{\mu}(t)^T h_u\right) \right\} dt = 0.$$

Equivalently, this can be written as,

$$\int_a^b (\beta(t), \theta(t))^T \begin{pmatrix} \bar{\lambda}(t)^T \bar{g}_x + \bar{\mu}(t)^T h_x + \dot{\mu}(t) \\ \bar{\lambda}(t)^T g_u + \bar{\mu}(t)^T h_u \end{pmatrix} dt = 0. \tag{27}$$

The equation (25) and (26) can be combined to be written in the following matrix form,

$$(\gamma - \alpha) \begin{pmatrix} \bar{\lambda}(t)^T + \bar{\mu}(t)^T h_x + \dot{\mu}(t) \\ \bar{\lambda}(t)^T g_u + \bar{\mu}(t)^T h_u \end{pmatrix}$$

$$+ \begin{pmatrix} f_{xx} + \bar{\lambda}(t)^T + g_{xx}\bar{\mu}(t)^T h_{xx}, & f_{xu} + \bar{\lambda}(t)g_{xu} + \bar{\mu}(t)^T h_{xu} \\ f_{xu} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T h_{xu}, & f_{uu} + \bar{\lambda}(t)^T g_{uu} + \bar{\mu}(t)^T h_{uu} \end{pmatrix}$$

$$\begin{pmatrix} \beta(t) \\ \theta(t) \end{pmatrix} = 0, \ t \in I. \tag{28}$$

Multiplying this by $(\beta(t), \theta(t))^T$, and then integrating we obtain

$$(\gamma - \alpha) \int_a^b (\beta(t), \theta(t))^T \begin{pmatrix} \bar{\lambda}(t)^T + \bar{g}_x + \bar{\mu}(t)^T h_x + \dot{\mu}(t) \\ \bar{\lambda}(t)^T \bar{g}_u + \bar{\mu}(t)^T h_u \end{pmatrix} dt$$

$$+ \int_a^b (\beta(t), \theta(t))^T \begin{pmatrix} f_{xx} + \bar{\lambda}(t)^T + g_{xx}\bar{\mu}(t)^T h_{xx}, & f_{xx} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T h_{xu} \\ f_{xu} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T h_{xu}, & f_{uu} + \bar{\lambda}(t)^T g_{uu} + \bar{\mu}(t)^T h_{uu} \end{pmatrix}$$

$$\begin{pmatrix} \beta(t) \\ \theta(t) \end{pmatrix} dt = 0. \tag{29}$$

Using (27) in (29), we have

$$\int_a^b (\beta(t), \theta(t))^T \begin{pmatrix} \bar{f}_{xx} + \bar{\lambda}(t)^T g_{xx} + \bar{\mu}(t)^T h_{xx}, f_{xu} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T h_{xu} \\ f_{xu} + \bar{\lambda}(t)^T g_{xu} + \bar{\mu}(t)^T h_{xu}, f_{uu} + \bar{\lambda}(t)^T g_{uu} + \bar{\mu}(t)^T h_{uu} \end{pmatrix}$$
$$\begin{pmatrix} \beta(t) \\ \theta(t) \end{pmatrix} dt = 0.$$

In view of the hypothesis (iii), this implies

$$\sigma(t) = (\beta(t), \theta(t)) = 0 \Rightarrow \quad \beta(t) = \theta(t) = 0, \ t \in I. \tag{30}$$

The relation (28) together with (30) yields

$$(\gamma - \alpha) \begin{pmatrix} \bar{\lambda}(t)^T g_x + \bar{\mu}(t)^T h_x + \dot{\bar{\mu}}(t) \\ \bar{\lambda}(t)^T g_u + \bar{\mu}(t)^T h_u \end{pmatrix} = 0, \quad t \in I.$$

Because of the hypothesis (iv), this gives

$$\alpha = \gamma. \tag{31}$$

If $\alpha = 0$, then $\gamma = 0$. Consequently using (30), (19) implies that $\zeta(t) = 0$, $t \in I$.

Thus $(\alpha, \beta(t), \theta(t), \zeta(t), \gamma) = 0$, $t \in I$. This contradicts the Fritz John condition (24). Hence $\gamma = \alpha > 0$.

Using (30) and $\gamma > 0$ in (19), we have $g(t, \bar{x}, \bar{u}) = -\zeta(t)/\gamma \leq 0$, $t \in I$. Also from (20), we have $h(t, \bar{x}, \bar{u}) - \dot{\bar{x}} = \dot{\beta}(t)/\gamma \geq 0$. Thus, it shows that $(\bar{x}, \bar{u})$ is feasible for (CP) and the objective values of (CP) and (CD) are equal. In view of the hypotheses of Theorem 1, the optimality of $(\bar{x}, \bar{u})$ for (CP) follows. $\qquad\square$

6. Control problem with free boundary conditions

The results validated in the preceding sections may be applied to the control problems with free boundary conditions. If the "targets" $x(a)$ and $x(b)$ are not restricted, we have

Problem (CPF) (Primal): Minimize $\int_a^b f(t, x, u) dt$

subject to

$$h(t, x, u) = \dot{x}, t \in I,$$
$$g(t, x, u) \leq 0, t \in I.$$

The dual control problem now includes the transversality conditions $\mu(t) = 0$, $t = a$ and $t = b$ as the new constraints. This yields

Problem CDF (Dual): Maximize $\int_a^b f(t, x, u)dt$

subject to

$\mu(a) = 0, \ \mu(b) = 0,$

$f_x(t, \bar{x}, \bar{u}) + \lambda(t)^T g_x(t, \bar{x}, \bar{u}) + \mu(t)^T h_x(t, \bar{x}, \bar{u}) + \dot{\mu}(t) = 0, \ t \in I,$

$f_u(t, \bar{x}, \bar{u}) + \lambda(t)^T g_u(t, \bar{x}, \bar{u}) + \mu(t)^T h_u(t, \bar{x}, \bar{u}) = \quad 0, \ t \in I,$

$\int_a^b \left(\lambda(t)^T g(t, x, u) + \mu(t)^T \left(h(t, x, u) - \dot{x}(t) \right) \right) dt \geq 0,$

$\lambda(t)^T \geq 0, \ t \in I.$

In order to prove the results, corresponding to Theorem 1 to Theorem 4, we will have the term $\eta^T \mu(t) \mid_{t=a}^{t=b}$ vanished by using $\mu(a) = 0$ and $\mu(b) = 0$ instead of having $x(a) = \alpha$ and $x(b) = \beta$ so that $\eta = 0$ at $t = a$ and $t = b$.

7. Mathematical programming problems

If f, g and h are independent of t (without any loss of generality $b-a = 1$) then the problems (CP) and (CD) reduce to the static primal and dual of mathematical programming problems treated by Mond and Weir [7] under generalized convexity and also under invexity by Craven and Glover [4].

Put $z = \begin{pmatrix} x \\ u \end{pmatrix}$, we have

Problem (PS): Minimize $f(z)$

subject to

$h(z) = 0,$

$g(z) \leq 0.$

Problem (DS): Maximize $f(z)$

subject to

$f_z(z) + \lambda^T g_z(z) + \mu^T h_z(z) = 0,$

$\lambda^T g(z) + \mu^T h(z) \geq 0,$

$\lambda \geq 0.$

References

[1] Abraham, J., Buie, R. N., *Duality in continuous programming and optimal control*, Utilitas Math. **17** (1980), 103–119.
[2] Berkovitz, L. D., *Variational methods in problems of control and programming*, J. Math. Anal. Appl. **3** (1961), 145–169.

[3] Craven, B. D., *Mathematical Programming and Control Theory*, Chapman & Hall, London, 1978.

[4] Craven, B. D., Glover, B. M., *Invex functions and duality*, J. Austral. Math. Soc. Ser. A **39** (1985), 1–20.

[5] Mond, B., Hanson, M. A., *Duality for control problems*, SIAM J. Control **6** (1968), 114–120.

[6] Mond, B., Smart, I., *Duality and sufficiency in control problems with invexity*, J. Math. Anal. Appl. **136**(15) (1988), 325–333.

[7] Mond, B., Weir, T., *Generalized concavity and duality*, in "Generalized Concavity in Optimization and Economics", S. Schaible and W. T. Ziemba (eds.), Academic Press, New York, 1981, 263–271.

[8] Valentine, F. A., *The problem of Lagrange with differentiable inequalities as added side constraints*, in "Contribution to the Calculus of Variation", 1933–1937, Univ. Chicago Press, Chicago, Illinois, 407–448, 1937.

[9] Wolfe, P., *A duality theorem for nonlinear programming*, Quart. Appl. Math. **19** (1961), 239–244.

Iqbal Husain
Department of Mathematics
Jaypee Institute of Engineering
and Technology
AB-Road, Raghogarh
Guna (M.P.) – 473 226, India
e-mail: iqbal.husain@jiet.ac.in
e-mail: ihusain11@yahoo.com

Aquil Ahmed
Department of Statistics
University of Kashmir
Hazratbal, Srinagar 190 006
Kashmir, India
e-mail: aqlstat@yahoo.com

Bashir Ahmad
Department of Statistics
University of Kashmir
Hazratbal, Srinagar 190 006
Kashmir, India
e-mail: bashir_qau@yahoo.com