JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **136**, 325–333 (1988)

Duality and Sufficiency in Control Problems with Invexity

B. MOND AND I. SMART

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

Submitted by E. Stanley Lee

Received May 6, 1987

INTRODUCTION

Recently, Mond, Chandra, and Husain [3] established duality results for a variational problem under more general convexity conditions, called invex, than Mond and Hanson [4]. In a similar fashion, the results of Mond and Hanson [5] for duality in control problems are extended to more general convex, or invex, functions. It is also shown that for invex functions, the necessary conditions for optimality in the control problem are also sufficient.

CONTROL PROBLEM AND DUAL

Consider the real scalar function $f(t, x, u)$ where $t \in [t_0, t_f]$, $x \in \mathbf{R}^n$, and $u \in \mathbf{R}^m$. Here t is the independent variable, $u(t)$ is the control variable, and $x(t)$ is the state variable; u is related to x via the state equations $G(t, x, u) = x'$, where prime denotes derivative with respect to t.

If $x = (x^1, ..., x^n)^T$, the gradient vector of f with respect to x is denoted by

$$f_x = \left(\frac{\partial f}{\partial x^1}, ..., \frac{\partial f}{\partial x^n}\right)^T.$$

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

$$R_x = \begin{pmatrix} \frac{\partial R^1}{\partial x^1} & \cdots & \frac{\partial R^r}{\partial x^1} \\ \vdots & \ddots & \vdots \\ \frac{\partial R^1}{\partial x^n} & \cdots & \frac{\partial R^r}{\partial x^n} \end{pmatrix}.$$

0022-247X/88 $3.00

409/136/1-22

Gradients with respect to u are defined analogously. It is assumed that f, G, and R have continuous second derivatives.

The control problem is to transfer the state variable from an initial state x_0 at t_0 to a final state x_f at t_f so as to minimize a given functional, subject to constraints on the control and state variables.

Stated formally, this is:

Problem P (*Primal*). Minimize $\int_{t_0}^{t_f} f(t, x, u)\, dt$ subject to

$$x(t_0) = x_0, \qquad x(t_f) = x_f \tag{1}$$

$$G(t, x, u) = x' \tag{2}$$

$$R(t, x, u) \geqslant 0. \tag{3}$$

The following dual problem is given in [5].

Problem D (*Dual*). 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

$$x(t_0) = x_0, \qquad x(t_f) = x_f \tag{4}$$

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

$$f_u(t, x, u) - G_u(t, x, u)\, \lambda(t) - R_u(t, x, u)\, \mu(t) = 0 \tag{6}$$

$$\mu(t) \geqslant 0, \tag{7}$$

where λ: $[t_0, t_f] \mapsto \mathbf{R}^n$ and μ: $[t_0, t_f] \mapsto \mathbf{R}^r$.

$x(t)$ and $u(t)$ are required to be piecewise smooth functions on $[t_0, t_f]$; their derivatives are continuous except perhaps at points of discontinuity of $u(t)$, which has piecewise continuous first and second derivatives.

The constraints (2), (3), (5), and (6) may fail to hold at these points of discontinuity of u, but (2) and (5) must hold for left- and right-hand limits.

Remark. If f, G, and R are independent of t (without loss of generality, assume $t_f - t_0 = 1$), then the problems (P) and (D) reduce to a static primal and dual of mathematical programming.

Putting $z = \binom{x}{u}$, we have:

Problem PS. Minimize $f(z)$ subject to

$$G(z) = 0$$

$$R(z) \geqslant 0.$$

Problem DS. Maximize $f(z) - \lambda^T G(z) - \mu^T R(z)$ subject to

$$f_z(z) - G_z(z)\, \lambda - R_z(z)\, \mu = 0$$

$$\mu \geqslant 0,$$

where $\lambda \in \mathbf{R}^n$, $\mu \in \mathbf{R}^r$.

INVEXITY

In [5], (P) and (D) are shown to be a dual pair if f, $-\lambda^T G$, and $-R$ are all convex in x and u. Here we extend this duality by weakening the convexity requirement.

DEFINITION. If there exist vector functions $\eta(t, x, x^*, x', x^{*\prime}, u, u^*) \in \mathbf{R}^n$, with $\eta = 0$ at t if $x(t) = x^*(t)$, and $\xi(t, x, x^*, x', x^{*\prime}, u, u^*) \in \mathbf{R}^m$ such that for the scalar function $h(t, x, x', u)$ the functional

$$H(x, x', u) \equiv \int_{t_0}^{t_f} h(t, x, x', u)\, dt$$

satisfies

$$\begin{aligned} &H(x, x', u) - H(x^*, x^{*\prime}, u^*) \\ &\quad \geqslant \int_{t_0}^{t_f} \Big(\eta^T h_x(t, x^*, x^{*\prime}, u^*) + \frac{d\eta^T}{dt} h_{x'}(t, x^*, x^{*\prime}, u^*) \\ &\qquad + \xi^T h_u(t, x^*, x^{*\prime}, u^*)\Big)\, dt \end{aligned}$$

then H is said to be invex in x, x', and u on $[t_0, t_f]$ with respect to η and ξ.

Remark. Invexity is defined here for functionals instead of functions, unlike the definition given in Mond, Chandra, and Husain [3]. This has been done so that invexity of a functional H is necessary and sufficient for its critical points to be global minima, which coincides with the original concept of an invex function being one for which critical points are also global minima (Craven and Glover [2]). We thus have the following characterization result.

LEMMA. *$H(x, x', u) \equiv \int_{t_0}^{t_f} h(t, x, x', u)\, dt$ is invex iff every critical point of H is a global minimum.*

(Note. (x^*, u^*) is a critical point of H if $h_x(t, x^*, x^{*\prime}, u^*) = (d/dt)\, h_{x'}(t, x^*, x^{*\prime}, u^*)$ and $h_u(t, x^*, x^{*\prime}, u^*) = 0$ almost everywhere in $[t_0, t_f]$. If $x(t_0)$ and $x(t_f)$ are free, the transversality conditions $h_{x'}(t, x^*, x^{*\prime}, u^*) = 0$ at t_0 and t_f are included.)

Proof. (a) Assume that there exist functions η and ξ such that H is invex with respect to η and ξ on $[t_0, t_f]$. Let (x^*, u^*) be a critical point of H. Then

$$
\begin{aligned}
&H(x, x', u) - H(x^*, x^{*\prime}, u^*) \\
&\quad \geqslant \int_{t_0}^{t_f} \Big(\eta^T h_x(t, x^*, x^{*\prime}, u^*) + \frac{d\eta^T}{dt} h_{x'}(t, x^*, x^{*\prime}, u^*) \\
&\qquad + \xi^T h_u(t, x^*, x^{*\prime}, u^*)\Big)\, dt \\
&\quad = \int_{t_0}^{t_f} \Big(\eta^T h_x(t, x^*, x^{*\prime}, u^*) - \eta^T \frac{d}{dt} h_{x'}(t, x^*, x^{*\prime}, u^*) \\
&\qquad + \xi^T h_u(t, x^*, x^{*\prime}, u^*)\Big)\, dt \\
&\qquad + \eta^T h_{x'}(t, x^*, x^{*\prime}, u^*)\big|_{t_0}^{t_f} \qquad \text{by integration by parts} \\
&\quad = 0
\end{aligned}
$$

as (x^*, u^*) is a critical point, and either fixed boundary conditions imply that $\eta = 0$ at t_0 and t_f, or free boundary conditions imply that $h_{x'} = 0$ at t_0 and t_f.

Therefore, (x^*, u^*) is a global minimum of H.

(b) Assume every critical point is a global minimum.

If (x^*, u^*) is a critical point, put $\eta = \xi = 0$.

If (x^*, u^*) is not a critical point, then if $h_x \neq (d/dt)\, h_{x'}$ at (x^*, u^*) put

$$\eta = \frac{h(t, x, x', u) - h(t, x^*, x^{*\prime}, u^*)}{2(h_x - (d/dt)\, h_{x'})^T (h_x - (d/dt)\, h_{x'})} (h_x - (d/dt)\, h_{x'})$$

or, if $h_x = (d/dt)\, h_{x'}$ put $\eta = 0$; and if $h_u \neq 0$, put

$$\xi = \frac{h(t, x, x', u) - h(t, x^*, x^{*\prime}, u^*)}{2 h_u^T h_u} h_u$$

or if $h_u = 0$ put $\xi = 0$.

Then H is invex on $[t_0, t_f]$ with respect to η and ξ. ∎

Duality

We prove that problems (P) and (D) are a dual pair subject to invexity conditions on the objective and constraint functions.

THEOREM 1 (Weak Duality). *If* $\int_{t_0}^{t_f} f\, dt$, $\int_{t_0}^{t_f} -\lambda^T (G - x')\, dt$, *and* $\int_{t_0}^{t_f} -\mu^T R\, dt$, *for any* $\lambda(t) \in \mathbf{R}^n$ *and* $\mu(t) \in \mathbf{R}^r$ *with* $\mu(t) \geqslant 0$, *are all invex with respect to the same functions* η *and* ξ, *then* $\inf(\mathrm{P}) \geqslant \sup(\mathrm{D})$.

Proof. Let (x^*, u^*) be feasible for problem (P), and let (x, u, λ, μ) be feasible for problem (D). Then

$$\int_{t_0}^{t_f} (f(t, x^*, u^*) - f(t, x, u))\, dt$$

$$\geqslant \int_{t_0}^{t_f} (\eta^T f_x(t, x, u) + \xi^T f_u(t, x, u))\, dt \qquad \text{by invexity of } \int_{t_0}^{t_f} f(t, x, u)\, dt$$

$$= \int_{t_0}^{t_f} \eta^T [\lambda'(t) + G_x(t, x, u)\, \lambda(t) + R_x(t, x, u)\, \mu(t)]\, dt$$

$$+ \int_{t_0}^{t_f} \xi^T [G_u(t, x, u)\, \lambda(t) + R_u(t, x, u)\, \mu(t)]\, dt \qquad \text{by (5) and (6)}$$

$$= \eta^T \lambda(t) \Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \frac{d\eta^T}{dt} \lambda(t)\, dt + \int_{t_0}^{t_f} [\eta^T G_x(t, x, u)\, \lambda(t)$$

$$+ \xi^T G_u(t, x, u)\, \lambda(t)]\, dt$$

$$+ \int_{t_0}^{t_f} [\eta^T R_x(t, x, u)\, \mu(t) + \xi^T R_u(t, x, u)\, \mu(t)]\, dt$$

by integration by parts

$$= \int_{t_0}^{t_f} \left[\eta^T G_x(t, x, u) - \frac{d\eta^T}{dt} + \xi^T G_u(t, x, u) \right] \lambda(t)\, dt$$

$$+ \int_{t_0}^{t_f} [\eta^T R_x(t, x, u) + \xi^T R_u(t, x, u)]\, \mu(t)\, dt$$

as fixed boundary conditions give $\eta = 0$ at t_0 and t_f

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

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

by invexity of $\int_{t_0}^{t_f} -\lambda^T (G - x')\, dt$ and $\int_{t_0}^{t_f} -\mu^T R\, dt$

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

by (2), (3), and (7).

Therefore

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

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

As (x^*, u^*) and (x, u, λ, μ) are arbitrary feasible solutions of (P) and (D), respectively, we have $\inf(P) \geqslant \sup(D)$. ∎

Assuming the constraint conditions for the existence of multipliers $\lambda(t)$ and $\mu(t)$ at extrema of (P) hold, the necessary conditions for (x^*, u^*) to be optimal for (P) are (Berkovitz [1]):

There exist $\lambda_0 \in \mathbf{R}$, $\lambda(t)$, $\mu(t)$ such that

$$F \equiv \lambda_0 f - \lambda(t)^T [G - x'] - \mu(t)^T R$$

satisfies

$$F_x = \frac{d}{dt} F_{x'} \tag{8}$$

$$F_u = 0 \tag{9}$$

$$\mu^i R^i = 0, \qquad i = 1, ..., r \tag{10}$$

$$\mu \geqslant 0 \tag{11}$$

at almost all $t \in [t_0, t_f]$ (except that at t corresponding to discontinuities of $u^*(t)$, (8) holds for right-and left-hand limits.)

It is assumed from now on that the minimizing solution (x^*, u^*) of (P) is normal; that is, λ_0 is non-zero, so that without loss of generality, we can take $\lambda_0 = 1$.

THEOREM 2 (Strong Duality). *Under the invexity conditions of Theorem* 1, *if* (x^*, u^*) *is an optimal solution of* (P), *then there exist* $\lambda(t)$ *and* $\mu(t)$ *such that* (x^*, u^*, λ, μ) *is optimal for* (D), *and the corresponding objective values are equal.*

Proof. The conditions (8), (9), and (11) imply that there exist $\lambda(t)$ and $\mu(t)$ such that (x^*, u^*, λ, μ) is feasible for (D).

Since $G(t, x^*, u^*) = x^{*\prime}$, and $\mu(t)^T R(t, x^*, u^*) = 0$ by (10), the dual objective has the same value as the primal objective, so by Theorem 1, (x^*, u^*, λ, μ) is optimal for (D). ∎

Similarly, converse duality holds if we further assume [Mond and Hanson [5]] that f, G, and R have continuous third derivatives, and, writing (6) as $P(t, x, u, \lambda, \mu)=0$, the matrix (dP^i/dz^j), $i=1, ..., m, j=1, ..., n+m$ where $z=\binom{x}{u}$, has rank m.

THEOREM 3 (Converse Duality). *If* $(x^*, u^*, \lambda^*, \mu^*)$ *is optimal for* (D), *and if*

$$\begin{pmatrix} f_{xx}-(G_x\lambda)_x-(R_x\mu)_x & f_{ux}-(G_x\lambda)_u-(R_x\mu)_u \\ f_{xu}-(G_u\lambda)_x-(R_u\mu)_x & f_{uu}-(G_u\lambda)_u-(R_u\mu)_u \end{pmatrix}$$

is non-singular for all $t \in [t_0, t_f]$, *then* (x^*, u^*) *is optimal for* (P), *and the corresponding objective values are equal.*

Proof. See Mond and Hanson [5]. ∎

SUFFICIENCY

It can be shown that, for invex functions, the necessary conditions of Berkovitz [1], together with normality of the constraints, are sufficient for optimality.

THEOREM 4. *If there exists* $(x^*, u^*, \lambda^*, \mu^*)$ *such that conditions* (8), (9), (10), *and* (11) *hold, with* (x^*, u^*) *feasible for* (P), *and* $\int_{t_0}^{t_f} f\,dt$, $\int_{t_0}^{t_f} -\lambda^{*T}(G-x')\,dt$, *and* $\int_{t_0}^{t_f} -\mu^{*T} R\,dt$ *are all invex with respect to the same functions* η *and* ξ, *then* (x^*, u^*) *is optimal for* (P).

Proof. Assume (x^*, u^*) is not optimal for (P). Then there exists $(x, u) \neq (x^*, u^*)$, (x, u) feasible for (P), such that

$$\int_{t_0}^{t_f} f(t, x, u)\,dt < \int_{t_0}^{t_f} f(t, x^*, u^*)\,dt.$$

As $\int_{t_0}^{t_f} f\,dt$ is invex with respect to η and ξ, it follows that

$$\int_{t_0}^{t_f} (\eta^T f_x(t, x^*, u^*) + \xi^T f_u(t, x^*, u^*))\,dt < 0. \tag{12}$$

Now,

$$\lambda^*(t)^T[G(t, x, u) - x'] = 0 = \lambda^*(t)^T[G(t, x^*, u^*) - x^{*\prime}]$$

implies

$$\int_{t_0}^{t_f} -\lambda^*(t)^T[G(t, x, u) - x' - G(t, x^*, u^*) + x^{*\prime}]\,dt = 0 \leqslant 0.$$

Thus, by invexity of $\int_{t_0}^{t_f} -\lambda^{*T}[G-x']\,dt$,

$$-\int_{t_0}^{t_f}\left[\eta^T G_x(t,x^*,u^*)\,\lambda^*(t)-\frac{d\eta^T}{dt}\lambda^*(t)+\xi^T G_u(t,x^*,u^*)\,\lambda^*(t)\right]dt\leqslant 0. \tag{13}$$

Also,

$$\mu^*(t)^T R(t,x,u)\geqslant 0=\mu^*(t)^T R(t,x^*,u^*)$$

implies

$$\int_{t_0}^{t_f} -\mu^*(t)^T[R(t,x,u)-R(t,x^*,u^*)]\,dt\leqslant 0.$$

By invexity of $\int_{t_0}^{t_f} -\mu^{*T} R\,dt$,

$$-\int_{t_0}^{t_f}[\eta^T R_x(t,x^*,u^*)\,\mu^*(t)+\xi^T R_u(t,x^*,u^*)]\,dt\leqslant 0. \tag{14}$$

Combining (12), (13), and (14),

$$\begin{aligned}\int_{t_0}^{t_f}\Bigg(&\eta^T f_x(t,x^*,u^*)+\xi^T f_u(t,x^*,u^*)\\ &-\left[\eta^T G_x(t,x^*,u^*)\,\lambda^*(t)-\frac{d\eta^T}{dt}\lambda^*(t)+\xi^T G_u(t,x^*,u^*)\,\lambda^*(t)\right]\\ &-[\eta^T R_x(t,x^*,u^*)\,\mu^*(t)+\xi^T R_u(t,x^*,u^*)\,\mu^*(t)]\Bigg)\,dt<0.\end{aligned} \tag{15}$$

Now, premultiplying (8) by η^T and (9) by ξ^T, adding and integrating gives

$$\begin{aligned}\int_{t_0}^{t_f}(&\eta^T f_x(t,x^*,u^*)+\xi^T f_u(t,x^*,u^*)\\ &-[\eta^T G_x(t,x^*,u^*)\,\lambda^*(t)+\xi^T G_u(t,x^*,u^*)\,\lambda^*(t)]\\ &-[\eta^T R_x(t,x^*,u^*)\,\mu^*(t)+\xi^T R_u(t,x^*,u^*)\,\mu^*(t)]-\eta^T\lambda^{*\prime}(t))\,dt=0.\end{aligned}$$

But

$$\begin{aligned}\int_{t_0}^{t_f}\eta^T\lambda^{*\prime}(t)\,dt&=\eta^T\lambda^*(t)\Big|_{t_0}^{t_f}-\int_{t_0}^{t_f}\frac{d\eta^T}{dt}\lambda^*(t)\,dt\\ &=-\int_{t_0}^{t_f}\frac{d\eta^T}{dt}\lambda^*(t)\,dt\end{aligned}$$

as fixed boundary conditions give $\eta=0$ at t_0 and t_f.

This contradicts Eq. (15). Hence (x^*,u^*) is optimal for (P). ∎

Free Boundary Conditions

The above results may also be applied to the control problem with free boundary conditions. If the "targets" $x(t_0)$ and $x(t_f)$ are not restricted, we have

Problem PF (*Primal*). Minimize $\int_{t_0}^{t_f} f(t, x, u)\, dt$ subject to

$$G(t, x, u) = x'$$

$$R(t, x, u) \geqslant 0.$$

The dual now includes the transversality conditions $F_{x'} = 0$ at t_0 and t_f as new constraints. That is, $F_{x'} = \lambda(t) = 0$ at t_0 and t_f. This gives:

Problem DF (*Dual*). Maximixe $\int_{t_0}^{t_f} [f(t, x, u) - \lambda(t)^T [G(t, x, u) - x'] - \mu(t)^T R(t, x, u)]\, dt$ subject to

$$\lambda(t_0) = 0, \qquad \lambda(t_f) = 0$$

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

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

$$\mu(t) \geqslant 0.$$

In order to prove the results corresponding to Theorems 1 and 4, it is necessary only to alter the reason for discarding the term $\eta^T \lambda(t)|_{t_0}^{t_f}$; instead of having $x(t_0) = x_0$ and $x(t_f) = x_f$ so that $\eta = 0$ at t_0 and t_f, we have $\lambda(t_0) = \lambda(t_f) = 0$.

For problems with mixed boundary conditions, the transversality condition associated with the free end is maintained in the dual, and the fact that $\eta = 0$ at the fixed end is utilized.

References

1. L. D. Berkovitz, Variational methods in problems of control and programming, *J. Math. Anal. Appl.* **3** (1961), 145–169.
2. B. D. Craven and B. M. Glover, Invex functions and duality, *J. Austral. Math. Soc. Ser. A.* **39** (1985), 1–20.
3. B. Mond, S. Chandra, and I. Husain, Duality for variational problems with invexity, *J. Math. Anal. Appl.* **134** (1988), 322–328.
4. B. Mond and M. A. Hanson, Duality for variational problems, *J. Math. Anal. Appl.* **18** (1967), 355–364.
5. B. Mond and M. A. Hanson, Duality for control problems, *SIAM J. Control* **6** (1968), 114–120.

Printed in Belgium