

UPPER AND LOWER BOUNDS FOR A CERTAIN CLASS OF CONSTRAINED VARIATIONAL PROBLEMS

M. A. HANSON

(received 26 July 1962)

1. Introduction

Certain optimization problems involving inequality constraints, known as optimal control problems have been extensively studied during recent years especially in relation to the calculation of optimal rocket thrusts and trajectories. A summary of these works is given by Berkovitz [1] who also establishes necessary conditions for the existence of solutions for a wide class of such problems.

The analytical solution of such problems is in general quite intractable, and recourse is made to approximation methods, which makes it desirable to have some measure of the errors involved. In this paper a duality theorem will be established for a class of optimal control programmes which permits the calculation of upper and lower bounds for the solutions. A method of calculating bounds for a related class of variational problems without constraints has been given by Bellman [2].

2. The problem of optimal control

Let t , x , and u be elements of the sets T , X , and U contained in 1, n , and m dimensional euclidean spaces respectively. The sets T and X will be assumed to be bounded. The element x is called the *state vector* and u is called the *control vector*. Let $S = T \times X \times U$.

Broadly our problem will be to select, under given conditions, a control vector u , as a function of t , such that the state vector x , also a function of t , is brought from some specified initial state $x_0 = x(t_0)$ to some specified final state $x_1 = x(t_1)$.

Notationally we shall not distinguish between row and column vectors which if necessary can be identified from the context. Superscripts will denote components of vectors. Subscripts x or u will denote partial differentiation with respect to the vector mentioned; for example, if $Z(t, x, u)$ is a vector valued appropriately differentiable function then Z_x denotes the matrix of partial derivatives $[\partial Z^i / \partial x^j]$. A prime will signify differentiation

with respect to t . Let $f(t, x, u)$ be a scalar function convex on $X \times U$ and of class C'' on $T \times X$; let $G(t, x, u)$ be an n -dimensional vector linear in x and u , that is, of the form

$$G(t, x, u) = Mx + Nu + P$$

where M, N , and P are $n \times n, n \times m$, and $n \times 1$ matrices respectively, whose elements are functions of t ; and let $R(t, x, u)$ be an r -dimensional function concave on $X \times U$, of class C'' on S and satisfying Berkovitz' *constraint conditions*:

(i) If $r > m$ then at each point of S at most m components of R can vanish.

(ii) At each point of S the matrix $[\partial R^i / \partial u^j]$ where i ranges over those indices such that $R^i(t, x, u) = 0$, and $j = 1, \dots, m$, has maximum rank.

A set of *admissible controls* A is then defined as a subset of U such that corresponding to each $u \in A$ the function $u = u(t)$ is piecewise C'' on the closure $[t_0, t_1]$ of T and permits a continuous solution $x(t)$ of the differential equation

$$(1) \quad x' = G(t, x, u)$$

such that $x(t_0) = x_0$ and $x(t_1) = x_1$, and such that the curve K thus defined is interior to $T \times X$ and $x(t) \neq x_1$ for any $t_0 \leq t \leq t_1$; further, along K , the constraints

$$(2) \quad R(t, x(t), u(t)) \geq 0$$

are satisfied.

We shall also require that K is normal (Berkovitz Section VIII). The problem is then to find an admissible control, if it exists, which minimises

$$(3) \quad J(u) = \int_{t_0}^{t_1} f(t, x, u) dt.$$

Suppose u^* is an optimal admissible control, and $x^*(t)$ the corresponding function defining K^* . Berkovitz has proved that there exists inter alia an n -dimensional Lagrange multiplier $\lambda(t)$ defined and continuous on T and an r -dimensional Lagrange multiplier $\mu(t) \leq 0$ defined and continuous on T except at values of t corresponding to corners of K^* , where it possesses unique right and left hand limits, such that along K^* :

$$(4) \quad \lambda' = -(H_x + \mu R_x)$$

$$(5) \quad H_u + \mu R_u = 0$$

$$(6) \quad \text{and } \mu R = 0$$

where H is the Lagrangian function

$$(7) \quad H(t, x, u, \lambda) = f(t, x, u) + \lambda G(t, x, u).$$

We shall utilize these results in establishing a duality between Problems I and II below. The term “duality” is used by analogy with the corresponding duality theorems of mathematical programming. Although the symbols used in Problem I are also used in Problem II, it is to be understood, of course, that the two problems are quite distinct.

Problem I (Primal).

$$\begin{aligned}
 (8) \quad & \text{Minimise } J(u) = \int_{t_0}^{t_1} f(t, x, u) dt \\
 (9) \quad & \text{subject to } x' = G(t, x, u) \\
 (10) \quad & x(t_0) = x_0, x(t_1) = x_1 \\
 (11) \quad & \text{and } R(t, x, u) \geq 0.
 \end{aligned}$$

Problem II (Dual)

$$\begin{aligned}
 (12) \quad & \text{Maximise } I(u, \lambda, \mu) = \int_{t_0}^{t_1} \{f(t, x, u) + \mu(t)R(t, x, u)\} dt \\
 (13) \quad & \text{subject to } x' = G(t, x, u) \\
 (14) \quad & x(t_0) = x_0, x(t_1) = x_1 \\
 (15) \quad & f_x + \lambda G_x + \mu R_x = -\lambda' \\
 (16) \quad & f_u + \lambda G_u + \mu R_u = 0 \\
 (17) \quad & \mu \leq 0.
 \end{aligned}$$

In the objective functional of Problem II, λ is regarded as implicit.

It will be shown that the value of u that minimises the primal is the same as the value of u that maximises the dual. Further the extreme values of the objective functionals in both problems are identical. Thus any admissible solution of the primal is an upper bound to the optimal solution and any admissible solution of the dual is a lower bound. In particular if (x^0, u^0) is some admissible solution of the primal, that is, is an upper bound for the problem, and is sufficiently close to optimum, then a lower bound can easily be established by putting the values x^0 and u^0 in the dual and solving the resulting simple linear equations in λ and μ . The validity of such a procedure is shown by the following considerations: Since (x^0, u^0) satisfies the primal constraints it also satisfies the dual constraints (13) and (14). It remains to be seen whether values of λ and $\mu \leq 0$ can be found satisfying (15) and (16) for the particular values $x = x^0, u = u^0$. Let Z be any matrix such that $\mu R_u Z = \mu R_x$. Then from (16) we have

$$(18) \quad f_u Z + \lambda G_u Z + \mu R_x = 0$$

which, on substitution in (15) gives

$$(19) \quad (f_x - f_u Z) + \lambda(G_x - G_u Z) = -\lambda'.$$

Suppose (x^*, u^*) is an optimal solution of the primal. Then Berkovitz' theorem ensures the existence of λ^* and $\mu^* \leq 0$ satisfying (15) and (16), and consequently (19), for the values $x = x^*, u = u^*$. It follows (see, for example, [3]) that if (x^0, u^0) is sufficiently close to (x^*, u^*) , there exist λ^0 and $\mu^0 \leq 0$ arbitrarily close to λ^* and μ^* respectively, satisfying (19) for the values $x = x^0, u = u^0$.

3. The duality theorem

THEOREM. *If for problem I, $u^*(t) \in A$ is an optimal control and $x^*(t)$ is the corresponding optimal state defining the curve K^* then there exists a vector $\lambda^*(t)$ defined and continuous on T and a vector $\mu^*(t) \leq 0$ defined and continuous on T except perhaps at corners of K^* , where it possesses unique right and left hand limits, such that u^*, x^*, μ^* , and λ^* are optimal in problem II and the extreme values of the objective functions in both problems are equal.*

PROOF. Since (x^*, u^*) is an optimal solution of problem I then by Berkovitz' theorem λ^* and μ^* exist and clearly satisfy the constraints of problem II.

Let (x, u, λ, μ) be any admissible solution of problem II.

Put

$$f^* \equiv f(t, x^*, u^*)$$

$$R^* \equiv R(t, x^*, u^*)$$

and

$$G^* \equiv G(t, x^*, u^*).$$

Then

$$\begin{aligned} & I(u^*, \lambda^*, \mu^*) - I(u, \lambda, \mu) \\ &= \int_{t_0}^{t_1} \{(f^* + \mu^* R^*) - (f + \mu R)\} dt \\ &= \int_{t_0}^{t_1} \{(f^* - f) + \mu(R^* - R) - \mu R^*\} dt && \text{by (6)} \\ &\geq \int_{t_0}^{t_1} \{(x^* - x)f_x + (u^* - u)f_u \\ &\quad + (x^* - x)\mu R_x + (u^* - u)\mu R_u - \mu R^*\} dt \\ &\quad \text{since } f \text{ is convex and } R \text{ is concave, and } \mu \leq 0, \\ &= \int_{t_0}^{t_1} \{(x^* - x)(f_x + \mu R_x) + (u^* - u)(f_u + \mu R_u) - \mu R^*\} dt \\ &= \int_{t_0}^{t_1} \{(x^* - x)(-\lambda' - \lambda G_x) + (u^* - u)(-\lambda G_u) - \mu R^*\} dt, \\ &\quad \text{by (15) and (16),} \\ &= \int_{t_0}^{t_1} \{-\lambda[(x^* - x)G_x + (u^* - u)G_u + x_i^* - x_i] - \mu R^*\} dt \\ &\quad - [(x^* - x)]_{t_0}^{t_1}, && \text{by integrating by parts,} \end{aligned}$$

$$\begin{aligned}
 &= \int_{t_0}^{t_1} \{-\lambda[(x^* - x)G_x + (u^* - u)G_u + G^* - G] - \mu R^*\} dt, \\
 &\qquad\qquad\qquad \text{using (13) and the fact that } x^*(t_0) = x(t_0) \\
 &\qquad\qquad\qquad \text{and } x^*(t_1) = x(t_1), \\
 &= - \int_{t_0}^{t_1} \mu R^* dt \\
 &\geq 0 \text{ by (17) and (11).}
 \end{aligned}$$

Hence $(x^*, u^*, \lambda^*, \mu^*)$ is the optimal solution of problem II. Further,

$$\begin{aligned}
 \text{Max } I(u, \lambda, \mu) &= \int_{t_0}^{t_1} \{f^* + \mu^* R^*\} dt \\
 &= \int_{t_0}^{t_1} f^* dt \qquad\qquad\qquad \text{by (6)} \\
 &= \min J(u).
 \end{aligned}$$

The theorem is thus proved.

References

[1] Berkovitz, L. D., Variational Methods in Problems of Control and Programming, *Journal of Mathematical Analysis and Applications*, **3** (1961), p. 145.
 [2] Bellman, R., Quasi-linearization and Upper and Lower Bounds for Variational Problems, *Quarterly of Appl. Maths.* **15** (1962), p. 349.
 [3] Nemytskii, V. V. and Stepanov, V. V., *Qualitative Theory of Differential Equations*, Chap. I Sect. 2-4 (Princeton University Press, 1960).

Department of Statistics,
 University of New South Wales.