

SUFFICIENT CONDITIONS
IN THE CALCULUS OF VARIATIONS
AND IN THE THEORY
OF OPTIMAL CONTROL

by

Zeev Nehari

Report 72-16

June 1972

AUG 11 1972

WONT LIBRARY
CARNEGIE-MELLON UNIVERSITY

SUFFICIENT CONDITIONS IN THE CALCULUS OF VARIATIONS
AND IN THE THEORY OF OPTIMAL CONTROL

by

Zeev Nehari

Abstract

A method is described for obtaining criteria confirming the optimality of solutions found either by integrating the Euler-Lagrange equations, or by applying the Pontryagin necessary condition. The method is based on convexity considerations, and its success in a given problem depends on the judicious choice of certain arbitrary auxiliary functions.

SUFFICIENT CONDITIONS IN THE CALCULUS OF VARIATIONS
AND IN THE THEORY OF OPTIMAL CONTROL

by

Zeev Nehari

1. In the classical calculus of variations, a number of procedures are available which make it possible, at least in principle, to confirm the extremal character of a solution found by integrating the Euler-Lagrange equations. However, the practical usefulness of these procedures is rather limited, since they depend on auxiliary information -- such as the existence of a field of extremals, or the knowledge of a general integral of the associated Hamilton-Jacobi equation -- which often is not easy to come by. In the theory of optimal control, the situation is even less satisfactory, and verification of the optimality of the solutions provided by the Pontryagin principle is possible only in relatively simple cases.

We shall describe here a method which yields sufficient conditions for optimality and which does not require the knowledge of quantities which, in practice, are difficult to obtain. Instead, the success of the method depends on the judicious choice of certain arbitrary auxiliary functions.

Research supported by the National Science Foundation under Grant GP23112.

2. We begin by illustrating our procedure in the case of the variational problem

$$(1) \quad \#(x) = \int_{t_1}^{t_2} F(t, x, x', \dots, x^{(k)}) dt = \min.,$$

with the end conditions

$$(2) \quad x^{(r)}(t_1) = A_r, x^{(r)}(t_2) = B_r, \quad r = 0, 1, \dots, k-1,$$

where $x \in C^k$ is a vector (x^1, \dots, x^n) , $x^{(r)}$ denotes the r -th derivative of x , and A_r, B_r are given constant vectors. The function F is defined and has continuous derivatives of order up to $2k$ for $t \in [t_1, t_2]$ and all real values of its remaining (scalar) variables. We assume that this variational problem has a solution \bar{x} , and we simplify the writing by using overhead bars to indicate quantities which are to be evaluated along the extremal \bar{x} ; thus, we write \bar{F} for $F(t, \bar{x}, \bar{x}', \dots, \bar{x}^{(k)})$, \bar{F}_j for the gradient $F_j(x, x', \dots, x^{(k)})$, etc. With the abbreviation

$$(3) \quad L(F) = \bar{F} - \sum_{r=0}^{k-1} \bar{F}_j \bar{x}^{(r)},$$

it follows from (1) that

$$(4) \quad p(x) - p(\bar{x}) = \int_{t_1}^{t_2} L(F) dt + T(x) - T(\bar{x}),$$

where

$$T(x) = \int_{t_1}^{t_2} \sum_{r=0}^{k-1} \bar{F}_j x^{(r)} dt.$$

If $x \in C^k$ and if it is observed that both x and \bar{x} satisfy

the end conditions (2), a number of integrations by parts leads to the identity

$$T(x) - T(\bar{x}) = \int_{t_1}^{t_2} \psi(\bar{x})(x - \bar{x}) dt,$$

where

$$\psi(x) = F_x - \frac{d}{dt} F_x + \dots + (-1)^k \frac{d^k}{dt^k} F_x^{(k)}.$$

Hence, if \bar{x} is a solution of the Euler-Lagrange equation $\psi(\bar{x}) = 0$ we have $T(x) = T(\bar{x})$, and (4) reduces to

$$(5) \quad J(x) - J(\bar{x}) = \int_{t_1}^{t_2} L(F) dt.$$

Since the expression (3) will be nonnegative if, for every fixed $t \in [t_1, t_2]$, F is a convex function of the set of variables $x_s^{(r)}$ ($r = 0, \dots, k$; $s = 1, \dots, n$), (5) shows that in this case a solution of the Euler-Lagrange equation which satisfies the end conditions indeed solves our minimum problem.

In order to make this convexity argument applicable to a wider class of integrands F , we employ the time-honored artifice -- first used by Legendre and much later made the basis of Carathéodory's treatment of the calculus of variations [1] -- of replacing F by suitable "equivalent" integrands. If $G(t, x, x', \dots, x^{(k-1)})$ is a function of class C^2 in all its variables and otherwise arbitrary, and if it is observed that $dG/dt = G_x x' + \dots + G_{x^{(k-1)}} x^{(k)} + G_t$, a computation shows that

$$(6) \quad L\left(\frac{dG}{dt}\right) = \frac{d}{dt} L(G),$$

where L is the operation defined in (3). Since, by (2),

$$\int_{t_1}^{t_2} \frac{d}{dt} L(G) dt = [L(G)]_{t_1}^{t_2} = 0,$$

we find that (5) may be replaced by

$$J(x) - J(\bar{x}) = \int_{t_1}^{t_2} L[F - \frac{dG}{dt}] dt.$$

This leads to the following sufficiency criterion.

Let \bar{x} be a solution of the Euler-Lagrange equation associated with the problem (1) which satisfies the end conditions (2). If there exists a function $G = G(t, x, x', \dots, x^{(k-1)})$ such that, for all fixed $t \in [t_1, t_2]$, $F - G_{x'} x' - \dots - G_{x^{(k-1)}} x^{(k-1)} - G_t$ is a convex function of the variables $x_s^{(r)}$ ($r = 0, \dots, k$; $s = 1, \dots, n$), then \bar{x} is a solution of the minimum problem (1).

We illustrate the use of this condition in the case $k = 1$. If we write P, Q, R, φ for the symmetric $n \times n$ matrices F_{xx} , $\frac{1}{2}[F_{xx'} + F_{x'x}]$, $F_{x'x'}$, G_{xx} , respectively, the convexity of $F - G_{x'} x' - G_t$ is equivalent to the positive-semidefiniteness of the quadratic form $\alpha R \alpha + 2\alpha(Q - \varphi)\beta + \beta(P - d\varphi/dt)\beta$ (where α, β are n -vectors). Since this quadratic form will have the property in question if R is positive-definite and the matrix $P - (Q - \varphi)R^{-1}(Q - \varphi) - d\varphi/dt$ is positive-semidefinite, we have the following result.

Set $P = F_{xx}$, $Q = \frac{1}{2}[F_{xx'} + F_{x'x}]$, $R = F_{x'x'}$, where F is the integrand of the variational problem (1) (with $k = 1$). If, for all fixed $t \in [t_1, t_2]$ and all $x, x' \in E^n$, the matrix R is

positive-definite, and if there exists a matrix $\langle p = \langle p(x,t)$ such that $P - (Q - \langle p) R^{-1} (Q - \langle p) - ckp/dt$ is positive-semidefinite, then a solution \bar{x} of the Euler-Lagrange equation associated with (1) which satisfies the prescribed end conditions is a solution of the minimum problem (1) .

It is easy to modify this statement so as to make it applicable to cases in which x and x^T are restricted to subsets X, X^f of E^n . If the conditions of the statement are satisfied for $x \in X, x^f \in X^f$ and if the convex hull of the Cartesian products $x \times x^f$ and $\bar{x}^n \times \bar{x}^f$ is contained in $X \times X^f$, the same argument as before shows that $\#(x) \wedge J(\bar{x})$.

In particular, if we restrict ourselves to the immediate neighborhood of $x \times x^T$ and we set $\langle p = Q + RU^1U$, we obtain the

classical result that x provides a weak local minimum for (1) if R is positive-definite and the Jacobi equation $(RU^T)^1 + (Q-P)U = 0$ has a matrix solution which is non-singular on $[t_1, t_2]$.

3. We now consider the fixed-time optimal control problem

$$(7) \quad \hat{\wedge}(x) = \int_t^2 f_0(x, u, t) dt = \min. , \quad u \in U, \quad x(t_1) = A, \quad x(t_2) = B,$$

where x, u are vectors of dimension n and m , respectively, $x \in C^1, U \in D^1, U$ is a closed convex set in E^m and x is subject to the side conditions

$$(8) \quad \dot{x} = f(x, u, t).$$

The scalar function f and the vector function f are required to have two continuous derivatives with respect to the components of u . We assume that an optimal trajectory x associated with an optimal control $u \in D$ exists and, as before, we will denote quantities associated with x, T by overhead bars.

Introducing an arbitrary vector $A = A(t) = (A^1, \dots, A^n)$, $A \in C$, we have, by (7) and (8),

$$\int_0^T [f - A(x-f)] dt = \int_0^T [f + Af + Ax] dt - \int_0^T [Ax] dt, /$$

H $f c_i$ 1

$(\dot{x} = dx/dt),$

and thus, using the end conditions and the abbreviation

$$0 = f_0 + Af + \dot{A}x,$$

$$\phi(x) - p(x) = \int_0^{t_2} (\varphi - \bar{\varphi}) dt.$$

Defining the operator L by

$$(9) \quad L(\bar{c}p) = \bar{\varphi} - \bar{c}p - \bar{c}p_x(x-\bar{x}) - \bar{\sigma}_u(u-\bar{u}),$$

we may replace this by

$$P(x) - \phi(\bar{x}) = \int_{t_1}^{t_2} L(p) dt + \int_{t_1}^{t_2} [\bar{p}_x(x-\bar{x}) + \bar{p}_u(u-\bar{u})] dt.$$

If we choose A so that $\bar{c}p_x = 0$, i.e., if $\dot{A} = -\bar{H}$, where

$H = H(x, u, t) = f_0 + Af$ is the Hamiltonian of the problem, and observe that because of $\bar{c}p = H + Ax$, $L(\bar{c}p) = L(H)$ and $\bar{c}p^u = H^u$, we obtain

$$(10) \quad \mathcal{J}(x) - \mathcal{J}(\bar{x}) = \int_{t_1}^{t_2} L(H) dt + \int_{t_1}^{t_2} \bar{H}_u(u-u) dt.$$

By Pontryagin's necessary condition [5], $H(\bar{x}, \bar{u}, t) \leq H(\bar{x}, v, t)$ for any $v \in U$ and all $t \in [t_1, t_2]$. Since U is convex, we may take $v = (1 - \epsilon) \bar{u} + \epsilon u = \bar{u} + \epsilon(u - \bar{u})$, where $\epsilon \in [0, 1]$ and is otherwise arbitrary. By Taylor's theorem,

$$0 \leq H(\bar{x}, v, t) - H(\bar{x}, \bar{u}, t) = \epsilon H_{uu}[\bar{x}, \bar{u} + \epsilon^*(u - \bar{u}), t] (u - \bar{u}),$$

where $0 \leq \epsilon^* \leq \epsilon$. Letting $\epsilon \rightarrow 0$, we obtain $\bar{H}_u(u - \bar{u}) \geq 0$, and (10) leads to the inequality

$$(11) \quad \mathcal{J}(x) - \mathcal{J}(\bar{x}) \geq \int_{t_1}^{t_2} L(H) dt.$$

Accordingly, \bar{x} will be optimal if the Hamiltonian is, for $u \in U$ and all fixed $t \in [t_1, t_2]$, a convex function of $(x_1, \dots, x_n, u_1, \dots, u_m)$. (For other sufficient conditions depending on convexity properties of the Hamiltonian, see [2,3,4].)

This is of course a very special situation. To make this convexity argument applicable to more general cases, we again modify the integrand by means of a suitable arbitrary function $G(x, t)$. A computation -- using (8) and the definition (9) of L -- shows that $L[(G_x - \bar{G}_x) f + G_t] = \frac{d}{dt} L(G)$. Since both x and \bar{x} are subject to the same end conditions, we have

$$\int_{t_1}^{t_2} \frac{d}{dt} L(G) dt = [L(G)]_{t_1}^{t_2} = 0$$

and (11) may therefore be replaced by

$$(12) \quad J(x) - \bar{J}(x) \geq \int_{t_1}^{t_2} L[H - (G_x - \bar{G}_x) f - G_t] dt.$$

This establishes the following result.

Let $H = H(x, u, t)$ be the Hamiltonian of the optimal control problem (7)-(8), and let \bar{x} satisfy Pontryagin's necessary condition for this problem. If there exists a function $G = G(x, t)$ such that, for $u \in U$ and every fixed $t \in [t_1, t_2]$, the function $H - (G_x - \bar{G}_x) f - G_t$ is convex in $(x_1, \dots, x_n, u_1, \dots, u_m)$, then \bar{x} is an optimal trajectory.

References

- [1] Caratheodory, C., Variationsrechnung und partielle Differentialgleichungen erster Ordnung, Leipzig, B. G. Teubner, 1935.
- [2] Funk, J. E. and E. G. Gilbert, "Some sufficient conditions for optimality in control problems with state space constraints", SIAM J. Control 8(1970), 498-504.
- [3] Kamien, M. I. and N. L. Schwartz, "Sufficient conditions in optimal control theory", J. Econ. Theory 3(1971), 207-214.
- [4] Mangasarian, O. L., "Sufficient conditions for the optimal control of nonlinear systems", SIAM J. Control 4(1966), 139-152.
- [5] Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, New York, Interscience, 1962.

Department of Mathematics
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

/ps -- 6/19/72