CONSTRAINED MINIMIZATION UNDER VECTOR-VALUED CRITERIA IN LINEAR TOPOLOGICAL SPACES

by

N. O. Da Cunha

E. Polak

Memorandum No. ERL-M191

4 November 1966

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Manuscript submitted: 25 October 1966

The research reported herein was supported wholly by the National Aeronautics and Space Administration under Grant NsG-354, Supplement 3.

INTRODUCTION

Judging from the literature, a vector-valued criterion optimization problem was formulated for the first time by the economist V. Pareto in 1896 [1]. Since then, discussions of this problem have kept reappearing in the economics literature (see Karlin [2], Debreu [3]), in the operations research literature (see Kuhn and Tucker [4]) and, more recently, in the control engineering literature (see Zadeh [5], Chang [6]).

Basically, the vector-valued criterion optimization problem arises as follows. Suppose that we wish to minimize simultaneously q real valued functions h^i of a variable x subject to given constraints. Usually this cannot be done and we are therefore forced to reformulate the problem as that of finding the admissible values $\hat{x}$ of the variable x, which make the vector cost $h(\hat{x}) = (h^1(\hat{x}), \cdots, h^q(\hat{x}))$ noninferior to all other comparable and admissible vector costs $h(x)$, with respect to some partial ordering on the q dimensional Euclidean space E^q.

Although the vector-valued criterion formulation of an optimization problem is frequently much closer to reality than a formulation with a scalar-valued criterion, very few results have been obtained to date that shed light on the subject. Among the most important questions which remain largely unanswered is that of whether a problem with a vector-valued criterion can be "scalarized", i.e., converted into an equivalent family of optimization problems with real-valued criteria. This question is important for the following reasons. First, whenever scalarization can be performed, it is highly likely that solutions can be obtained by using standard algorithms. Second, if scalarization were always possible, there would be little reason for constructing a separate theory of necessary conditions for vector-criterion optimization problems. So far, there is no evidence to indicate that scalarization is or is not always possible by arbitrary means. However, there are examples that show that scalarization by linearly combining the components of h into a real valued cost does not produce an equivalent family of optimization problems. Thus, for the time being at least, we require a special theory for vector-valued optimization problems.

The present paper is devoted to developing a broad theory of necessary conditions which characterize noninferior points, and to establishing relations between the solutions of a vector-valued criterion problem and the solutions of certain families of optimization problems with scalar-valued criteria.

Finally, we show how the general conditions we obtain reduce to a Pontryagin type maximum principle for a class of optimal control problems.

I. Necessary Conditions for the Basic Problem[†]

Let E^s, where s is a positive integer, be the s-dimensional Euclidean space with the usual norm topology. Let $\mathcal{X}$ be a real, linear topological space; let $h : \mathcal{X} \to E^p$ and $r : \mathcal{X} \to E^m$ be continuous functions, and let Ω be a subset of $\mathcal{X}$.

Furthermore, suppose that we are given an ordering $\prec$ in E^p, with the following property:

1. For every y in E^p there exists an index set $J(y) \subset \{1, 2, \cdots, p\}$ and a ball $B(\epsilon_0, y)$ with center y and radius $\epsilon_0 > 0$ such that every $\tilde{y} \in B(\epsilon_0, y)$, with $\tilde{y}^i < y^i$ for all $i \in J(y)$, satisfies $\tilde{y} \prec y$ and $y \not\prec \tilde{y}$.

We shall call the index set $J(y)$, defined above, <u>the set of critical indices for the point y</u>.

2. <u>Examples</u>:

The following orderings $\prec$ satisfy (1):

(a) For $p = 1$, $y_1 \prec y_2$ if and only if $y_1 \leqq y_2$.

(b) For $p > 1$, $y_1 \prec y_2$ if and only if $y_1^i \leqq y_2^i$ for $i = 1, 2, \cdots, p$.

(c) For $p > 1$, $y_1 \prec y_2$ if and only if

$$\mathrm{Max}\{y_1^i \mid i = 1, 2, \cdots, p\} \leqq \mathrm{Max}\{y_2^i \mid i = 1, 2, \cdots, p\} .$$

---

† Our approach to necessary conditions is derived from the work of Neustadt[15], Cannon, Cullum and Polak [16] and Halkin and Neustadt [17].

An ordering $\prec$ which satisfies (1) may be partial as in Example (2)(b) or complete as in Examples (2)(a) and (c).

The problems we wish to consider can always be cast in the following standard form:

3. Basic Problem: Find a point $\hat{x}$ in $\mathfrak{X}$, such that:

4. (i) $\hat{x} \in \Omega$ and $r(\hat{x}) = 0$;

5. (ii) for every x in Ω with $r(x) = 0$, the relation $h(x) \prec h(\hat{x})$ implies that $h(\hat{x}) \prec h(x)$.

As a first step in obtaining necessary conditions for a point $\hat{x}$ in $\mathfrak{X}$ to be a solution to the Basic Problem (3), we introduce "linear" approximations to the set Ω and to the continuous functions h and r at $\hat{x}$.

6. Definition: We shall say that a convex cone $C(\bar{x}, \Omega)$ is a linearization of the constraint set Ω at the point $\bar{x} \in \Omega$, if there exist continuous linear functions $h'(\bar{x}) : \mathfrak{X} \to E^p$ and $r'(\bar{x}) : \mathfrak{X} \to E^m$ such that for any finite collection $\{x_1, x_2, \cdots, x_k\}$ of linearly independent vectors in $C(\bar{x}, \Omega)$, there exist a positive scalar ϵ_1, a continuous map ζ from $\epsilon S = co\{\epsilon x_1, \epsilon x_2, \cdots, \epsilon x_k\}$ into $\Omega - \{\bar{x}\}$, where $0 \leqq \epsilon \leqq \epsilon_1$, (possibly depending on ϵ and $\{x_1, x_2, \cdots, x_k\}$), and continuous functions $o_h : \mathfrak{X} \to E^p$ and $o_r : \mathfrak{X} \to E^m$ (possibly depending on ϵ and S), which satisfy (7), (8), (9) and (10) below.

7. $\lim_{\epsilon \to 0} \frac{\|o_h(\epsilon y)\|}{\epsilon} = 0$ uniformly for $y \in S$,

8. $\lim_{\epsilon \to 0} \frac{\|o_r(\epsilon y)\|}{\epsilon} = 0$ uniformly for $y \in S$,

9. $h(\bar{x} + \zeta(x)) = h(\bar{x}) + h'(\bar{x})(x) + o_h(x)$, for all $x \in S$, $0 \leqq \epsilon \leqq \epsilon_1$,

and

10. $r(\bar{x} + \zeta(x)) = r(\bar{x}) + r'(\bar{x})(x) + o_r(x)$, for all $x \in S$, $0 \leqq \epsilon \leqq \epsilon_1$.

11. <u>Theorem</u>: If $\hat{x}$ is a solution to the Basic Problem (3), if $C(\hat{x},\Omega)$ is a linearization of Ω at $\hat{x}$, and if $J(h(\hat{x}))$ is the set of critical indices for $h(\hat{x})$ (see (1)), then there exist a vector μ in E^p and a vector η in E^m such that

12. (i) $\mu^i \leqq 0$ for $i \in J(h(\hat{x}))$ and $\mu^i = 0$ for $i \in J^c(h(\hat{x}))$;

13. (ii) $(\mu, \eta) \neq 0$;

14. (iii) $\langle \mu, h'(\hat{x})(x)\rangle + \langle \eta, r'(\hat{x})(x)\rangle \leqq 0$ for all $x \in \overline{C(\hat{x}, \Omega)}$, where $h'(\hat{x})$, $r'(\hat{x})$ are the linear continuous maps appearing in the definition of $C(\hat{x}, \Omega)$, see (6).

<u>Proof</u>: Let $\hat{x}$ be a solution to the Basic Problem.

Let $J(h(\hat{x}))$ and $B(\epsilon_0, h(\hat{x}))$ be, respectively, the critical index set and the neighborhood of $h(\hat{x})$ in E^p which satisfy condition (1). Also, let q be the cardinality of $J(h(\hat{x}))$ and let f be the continuous function from $\mathcal{X}$ into E^q defined by:

15. $f(x) = (f^1(x), \cdots, f^q(x))$ where $f^j(x) = h^{i_j}(x)$ with $i_j \varepsilon J(h(\hat{x}))$ for $j = 1, 2, \cdots, q$, and $i_\alpha > i_\beta$ when $\alpha > \beta$. Also, let

16. $f'(\hat{x})(x) = (f'^1(\hat{x})(x), \cdots, f'^q(\hat{x})(x))$ where $f'^j(\hat{x}) = h'^{i_j}(\hat{x})$ with $i_j \varepsilon J(h(\hat{x}))$ for $j = 1, 2, \cdots, q$, and $i_\alpha > i_\beta$ when $\alpha > \beta$.

Now let

17. $A(\hat{x}) = \{y \varepsilon E^q | y = f'(\hat{x})(x), \; x \varepsilon C(\hat{x}, \Omega)\}$,

18. $B(\hat{x}) = \{z \varepsilon E^m | z = r'(\hat{x})(x), \; x \varepsilon C(\hat{x}, \Omega)\}$,

19. $K(\hat{x}) = \{u \varepsilon E^q \times E^m | u = (f'(\hat{x})(x), \; r'(\hat{x})(x)), \; x \varepsilon C(\hat{x}, \Omega)\}$.

Since, by definition, $f'(\hat{x})$ and $r'(\hat{x})$ are linear maps, $A(\hat{x})$, $B(\hat{x})$, and $K(\hat{x})$ are convex cones in E^q, E^m, and in $E^q \times E^m$, respectively. Clearly, $K(\hat{x}) \subset A(\hat{x}) \times B(\hat{x})$.

Let C and R be the convex cones in E^q and in $E^q \times E^m$, respectively, defined by

20. $C = \{y = (y^1, y^2, \cdots, y^q) \varepsilon E^q | y^i < 0, \; i = 1, 2, \cdots, q\}$,

21. $R = \{(y, 0) \varepsilon E^q \times E^m | y \varepsilon C, \; 0 \varepsilon E^m\}$.

Examining (12), (13), and (14), we observe that if we define $\mu^i = 0$ for $i \varepsilon J^c(h(\hat{x}))$, the complement of $J(h(\hat{x}))$ in $\{1, 2, \cdots, p\}$, then the claim of the theorem is that the convex sets $K(\hat{x})$ and R are separated in $E^q \times E^m$. We now construct a proof by contradiction.

Suppose that $K(\hat{x})$ and R are not separated in $E^q \times E^m$. We then find that the following two statements must be true.

22. (I) The convex sets $K(\hat{x})$ and R are not disjoint, i.e., $R \cap K(\hat{x}) \neq \emptyset$, the empty set.

23. (II) The convex cone $B(\hat{x})$ in E^m contains the origin as an interior point and hence $B(\hat{x}) = E^m$.

Statement (II) follows from the fact that if 0 is not an interior point of the convex set $B(\hat{x})$, then by the separation theorem [7], it can be separated from $B(\hat{x})$ by a hyperplane in E^m, i.e., there exists a nonzero vector η_0 in E^m such that

24. $\langle \eta_0, z \rangle \leqq 0$ for all $z \in B(\hat{x})$.

Clearly, the vector $(0, \eta_0)$ in $E^q \times E^m$ separates R from $A(\hat{x}) \times B(\hat{x})$ and hence from $K(\hat{x})$, contradicting our assumption that R and $K(\hat{x})$ are not separated.

We now proceed to utilize the facts (I) and (II). Since the origin in E^m belongs to the non-void interior of $B(\hat{x})$ ($B(\hat{x}) = E^m$, see (II)), we can construct a simplex Σ in $B(\hat{x})$, with vertices $z_1, z_2, \cdots, z_{m+1}$, such that

25. (i) 0 is in the interior of Σ;

26. (ii) there exists a set of vectors $\{x_1, x_2, \cdots, x_{m+1}\}$ in $C(\hat{x},\Omega)$ satisfying:

27. (a) $z_i = r'(\hat{x})(x_i)$ for $i = 1, 2, \cdots, m+1$;

28. (b) $\zeta(x) \varepsilon (\Omega - \{\hat{x}\}) \cap N$ for all $x \varepsilon \text{co}\{x_1, x_2, \cdots, x_{m+1}\}$,

where ζ is the map entering the definition of a linearization, see (6), and N is a neighborhood of 0 in $\mathfrak{X}$ such that $h(\{\hat{x}\} + N) \subset B(\epsilon_0, h(\hat{x}))$, where $B(\epsilon_0, h(\hat{x}))$ is the ball about $h(\hat{x})$ entering the definition of $J(h(\hat{x}))$, see (1). (Clearly, such an N exists since h is continuous).

29. (c) The points $y_i = f'(\hat{x})(x_i)$ are in C for $i = 1, 2, \cdots, m+1$.

The existence of such a simplex is easily established. First we construct any simplex Σ' in $B(\hat{x})$ (see (18)) with vertices $z'_1, z'_2, \cdots, z'_{m+1}$, which contains the origin in its interior. This is clearly possible since $B(\hat{x}) = E^m$ by (23). Let $x'_1, x'_2, \cdots, x'_{m+1}$ be any set of points in $C(\hat{x},\Omega)$ which satisfy (27), i.e., $z'_i = r'(\hat{x})(x'_i)$ for $i = 1, 2, \cdots, m+1$. If $f'(\hat{x})(x'_i) < 0$ for $i = 1, 2, \cdots, m+1$, then (29) is satisfied. It is easy to show that (25), (27), and (29), together with the fact that $r'(\hat{x})$ is a linear map, imply that the vectors $x'_1, x'_2, \cdots, x'_{m+1}$ are linearly independent. In order to satisfy (28) we first note that from the definition of linearization (6) there exists a positive scalar ϵ'_0 such that $\zeta(\text{co}\{\epsilon x'_1, \epsilon x'_2, \cdots, \epsilon x'_{m+1}\})$ is contained in $\Omega - \{\hat{x}\}$, for every real ϵ with $0 \leq \epsilon \leq \epsilon'_0$. Since ζ is a continuous function, there exist neighborhoods $N_1, \cdots, N_{m+1}$, of the origin in $\mathfrak{X}$ such that

$$\zeta(\{x'_i\} + N_i) \subset N \text{ for } i = 1, 2, \cdots, m+1.$$

Let $N' = N_1 \cap N_2 \cap \cdots \cap N_{m+1}$ and let N'' be a balanced† neighborhood of the origin in $\mathfrak{X}$ such that $\overbrace{N'' + N'' + \cdots + N''}^{m+1} \subset N'$ (N'' exists since $\mathfrak{X}$ is a topological linear space) and let ϵ'_1 be a positive real number such that $\epsilon'_1 x'_i \in N''$ for $i = 1, 2, \cdots, m+1$, (ϵ'_1 exists since in a topological linear space every neighborhood of the origin is absorbent). Clearly, $\zeta(co\{\epsilon'_1 x'_1, \epsilon'_1 x'_2, \cdots, \epsilon'_1 x'_{m+1}\}) \subset N$. Hence (28) is satisfied by letting $x_i = \epsilon' x'_i$, where $\epsilon' = \min\{\epsilon'_0, \epsilon'_1\}$.

But suppose, without loss of generality, that $f'(\hat{x})(x'_1) \geqq 0$ and $f'(\hat{x})(x'_i) < 0$ for $i = 2, \cdots, m+1$. Since by (22), $K(\hat{x}) \cap R \neq \emptyset$, there exists a point $u = (f'(\hat{x})(\tilde{x}), 0) \in K(\hat{x}) \cap R$, i.e., $f'(\hat{x})(\tilde{x}) < 0$ and $r'(\hat{x})(\tilde{x}) = 0$. Choose any scalar $\lambda > 0$ such that $f'(\hat{x})(\lambda x'_1 + (1-\lambda)\tilde{x}) < 0$, and let $x_1 = \lambda x'_1 + (1-\lambda)\tilde{x}$. Then, as above, the simplex Σ, with vertices $\epsilon\lambda z'_1, \cdots, \epsilon z'_{m+1}$, satisfies the conditions (25), (26), (27), (28) and (29), for the corresponding vectors $x_1, x'_2, x'_3, \cdots, x'_{m+1}$ and some $\epsilon > 0$.

It is easy to show that (25) implies that the vectors $(z_1 - z_{m+1}), \cdots, (z_m - z_{m+1})$ are linearly independent. Let $\ell_1, \ell_2, \cdots, \ell_m$ be any basis in E^m, let $Z : E^m \to E^m$ be a linear operator defined by $Z\ell_i = (z_i - z_{m+1})$ with

---

† <u>Definition:</u> A subset S of a linear vector space E will be called <u>balanced</u> if $\alpha x \in S$ whenever $x \in S$ and $-1 \leqq \alpha \leqq 1$. [7].

$i = 1, 2, \cdots, m$; and let $X : E^m \to \mathfrak{X}$ be a linear operator defined by $X\ell_i = (x_i - x_{m+1})$, with $i = 1, 2, \cdots, m$. Since the vectors $(z_i - z_{m+1})$, $i = 1, 2, \cdots, m$, are linearly independent, the operator Z is nonsingular. Let Z^{-1} denote the inverse of Z_1. Clearly the map $z \to XZ^{-1}(z - z_{m+1}) + x_{m+1}$ from Σ into $co\{x_1, x_2, \cdots, x_{m+1}\}$ is continuous.

Now, for $0 < \alpha \leqq 1$, let S_α be a sphere in E^m with radius $\alpha\rho$ (where $\rho > 0$) and center at the origin and contained in the interior of the simplex Σ.

We now define a continuous map G_α from the sphere S_α into E^m by

30. $$G_\alpha(\alpha z) = r(\hat{x} + \zeta(\alpha XZ^{-1}(z - z_{m+1}) + \alpha x_{m+1})),$$

where $||z|| \leqq \rho$, $\alpha z \in S_\alpha$, and ζ is the map specified by Definition (6).

From Definition (6),

31. $$G_\alpha(\alpha z) = r(\hat{x}) + r'(\hat{x})(\alpha XZ^{-1}(z - z_{m+1}) + \alpha x_{m+1}) + o_r(\alpha XZ^{-1}(z - z_{m+1}) + \alpha x_{m+1}).$$

But $r(\hat{x}) = 0$, $r'(\hat{x}) \circ X = Z$, and $r'(\hat{x})(x_{m+1}) = z_{m+1}$. Hence (31) becomes

32. $$G_\alpha(\alpha z) = \alpha z + o_r(\alpha XZ^{-1}(z - z_{m+1}) + \alpha x_{m+1}).$$

Now, since $\lim_{\alpha \to 0} \dfrac{||o_r(\alpha XZ^{-1}(z - z_{m+1}) + \alpha x_{m+1})||}{\alpha} = 0$, uniformly for $z \in \Sigma$ there exists for $||z|| = \rho$ an α_0, $0 < \alpha_0 \leqq 1$, such that

33. $\| o_r(\alpha X Z^{-1}(z - z_{m+1}) + \alpha x_{m+1}) \| < \alpha\rho$, for all $0 < \alpha \leqq \alpha_0$ and $\|z\| = \rho$.

Since h satisfies (9), it is clear that the components of f may be expanded as follows:

34. $f^i(\hat{x} + \zeta(\alpha X Z^{-1}(z - z_{m+1}) + \alpha x_{m+1}))$

$$= f^i(\hat{x}) + \alpha f^{i'}(\hat{x})(X Z^{-1}(z - z_{m+1}) + x_{m+1}) + o^i(\alpha X Z^{-1}(z - z_{m+1}) + \alpha x_{m+1}),$$

where $\| o^i(\epsilon z) \| / \epsilon \to 0$ as $\epsilon \to 0$ uniformly for $z \varepsilon \Sigma$, for $i = 1, 2, \cdots, q$.

Since by construction, (see (29)), $f^{i'}(\hat{x})(x_j) < 0$, for $i = 1, 2, \cdots, q$ and $j = 1, 2, \cdots, m+1$, and the point $X Z^{-1}(z - z_{m+1}) + x_{m+1}$ is in $co\{x_1, x_2, \cdots, x_{m+1}\}$, we have $f^{i'}(\hat{x})(X Z^{-1}(z - z_{m+1}) + x_{m+1}) < 0$, for $i = 1, 2, \cdots, q$. Hence there exist positive real numbers α_i, $i = 1, 2, \cdots, q$, such that for $i = 1, 2, \cdots, q$, and $\|z\| = \rho$

35. $f^i(\hat{x} + \zeta(\alpha X Z^{-1}(z - z_{m+1}) + \alpha x_{m+1})) < f^i(\hat{x})$ for all $0 < \alpha \leqq \alpha_i$.

Let α^* be the minimum of $\{\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_q\}$. It now follows from Brouwer's Fixed Point Theorem [8] that there exists a point $\alpha^* z^*$ in S_{α^*} such that $G_{\alpha^*}(\alpha^* z^*) = 0$.

Now, let $x^* = \hat{x} + \zeta(\alpha^* X Z^{-1}(z^* - z_{m+1}) + \alpha^* x_{m+1})$, then

36. (a) $r(x^*) = 0$ (since $r(x^*) = G_{\alpha^*}(\alpha^* z^*) = 0$),

37. (b) $x^* \varepsilon \Omega$, since $(x^* - \hat{x}) \varepsilon \zeta(co\{x_1, x_2, \cdots, x_{m+1}\}) \subset \Omega - \{\hat{x}\}$.

But (35), (28), (15), and (1) imply that

38. $h(x^*) < h(\hat{x})$ and that $h(\hat{x}) \nless h(x^*)$.

Now, (36), (37) and (38) contradict the assumption that $\hat{x}$ is a solution to the Basic Problem. Therefore the convex cones $K(\hat{x})$ and R are separated in $E^q \times E^m$, i.e., there exists a nonzero vector $(\bar{\mu}, \eta)$ in $E^q \times E^m$ such that

39. (i) $\langle \bar{\mu}, f'(\hat{x})(x)\rangle + \langle \eta, r'(\hat{x})(x)\rangle \leqq 0$ for all $x \in C(\hat{x}, \Omega)$,

40. (ii) $\langle \bar{\mu}, y\rangle + \langle \eta, 0\rangle \geqq 0$ for all $y \in C$.

But (40) implies that $\bar{\mu}^i \leqq 0$ for $i = 1, 2, \cdots, q$. Let $\mu = (\mu^1, \cdots, \mu^p)$ be the vector in E^p defined by $\mu^{i_j} = \bar{\mu}^j$, $i_j \in J(h(\hat{x}))$ for $j = 1, 2, \cdots, q$ and $i_\alpha > i_\beta$, when $\alpha > \beta$ and $\mu^k = 0$ for $k \in J^c(h(\hat{x}))$. Hence,

(i) $\mu^i \leqq 0$ for $i \in J(h(\hat{x}))$ and $\mu^i = 0$ for $i \in J^c(h(\hat{x}))$,

(ii) $(\mu, \eta) \neq 0$,

and (39) together with the continuity of $f'(\hat{x})$ and $r'(\hat{x})$ yield

(iii) $\langle \mu, f'(\hat{x})(x)\rangle + \langle \eta, r'(\hat{x})(x)\rangle \leqq 0$ for all $x \in \overline{C(\hat{x}, \Omega)}$.

II. Reduction of a vector-valued criterion to a family of scalar-valued criteria

In this section we restrict ourselves to the partial ordering defined in Example (2)(c), i.e., given y_1, y_2 in E^p, $y_1 \prec y_2$ if and only if $y_1^i \leqq y_2^i$ for $i = 1, 2, \cdots, p$. It follows from (1) that for any vector y in E^p, the set of critical indices $J(y)$ is the set $\{1, 2, \cdots, p\}$.

Let us denote by Π a particular Basic Problem, characterized by the vector valued criterion h, the constraint function r, and the constraint set Ω. Suppose that $\hat{x}$ is a solution to Π and that $\mu \in E^p$ is a vector satisfying the conditions of Theorem (11) for $\hat{x}$. Now consider the problem $\Pi(\mu)$, which is characterized by the same constraints as Π, but which has the scalar valued criterion $h_\mu(x) = -\langle \mu, h(x) \rangle$. It is clear that if $\hat{x}$ is also a solution to $\Pi(\mu)$, then Theorem (11) yields identical necessary conditions for $\Pi(\mu)$ and Π. This observation leads us to the question: can the solutions to the Basic Problem (3) be obtained by solving a family of scalar-valued criterion problems? This question is partially answered below by Theorems (48), (49), (50), (57), and (60).

In order to simplify our exposition, we combine the constraint set Ω with the set $\{x \in \mathfrak{X} \mid r(x) = 0\}$ into a set $A = \Omega \cap \{x \in \mathfrak{X} \mid r(x) = 0\}$. We therefore consider a subset A of $\mathfrak{X}$, a continuous mapping h from $\mathfrak{X}$ into E^p and introduce the following definitions.

41. Definition: We shall denote by P the problem of finding a point $\hat{x}$ in A such that for every x in A, the relation $h(x) \leqq h(\hat{x})$ (componentwise) implies that $h(x) = h(\hat{x})$.

42. Definition: Let Λ be the set of all vectors $\lambda = (\lambda^1, \cdots, \lambda^p)$ in E^p such that $\Sigma_{i=1}^p \lambda^i = 1$ and $\lambda^i > 0$ for $i = 1, 2, \cdots, p$; let $\bar{\Lambda}$ be the closure of Λ in E^p.

43. Definition: Given any vector λ in E^p, we shall denote by $P(\lambda)$ the problem of finding a point $\bar{x}$ in A such that $\langle \lambda, h(\bar{x}) \rangle \leqq \langle \lambda, h(x) \rangle$ for all x in A.

We shall consider the following subsets of $\mathfrak{X}$:

44. $L = \{x \in A \mid x \text{ solves } P\}$,

45. $M = \{x \in A \mid x \text{ solves } P(\lambda) \text{ for some } \lambda \in \Lambda\}$,

46. $N = \{x \in A \mid x \text{ solves } P(\lambda) \text{ for some } \lambda \in \bar{\Lambda}\}$.

47. Remark: Clearly the set M is contained in the set N; furthermore, it is easy to show that if h is a continuous function, then the closure of the set M is contained in the set N and by very simple examples we can show that this last inclusion may be proper (see [9]).

48. Theorem: The set L contains the set M.

Proof: Suppose $\bar{x} \in M$ and $\bar{x} \notin L$. Then, there must exist a point x' in A such that $h(x') \leq h(\bar{x})$. But for any $\lambda \in \Lambda$, this implies

$\langle \lambda, h(x') \rangle < \langle \lambda, h(\bar{x}) \rangle$, and hence $\bar{x}$ is not in M, a contradiction.

49. Theorem: If for each $\lambda \varepsilon \bar{\Lambda}$ either $P(\lambda)$ has a unique solution or else it has no solution, then the set L contains the set N.

Proof: Suppose that $\bar{x} \varepsilon N$ for some $\bar{\lambda} \varepsilon \bar{\Lambda}$ and that $\bar{x} \notin L$. Then there must exist a point $x' \neq \bar{x}$ in A such that $h(x') \leq h(\bar{x})$. But for any $\lambda \varepsilon \bar{\Lambda}$, this implies that $\langle \lambda, h(x') \rangle \leqq \langle \lambda, h(\bar{x}) \rangle$, and hence x' is also a solution to $P(\bar{\lambda})$, which contradicts the assumption that $\bar{x}$ is the unique solution to $P(\bar{\lambda})$.

50. Theorem: Suppose that h is a convex function (component-wise) and that A is a convex set. Then the set N contains the set L.

Proof: Let $\hat{x}$ be a point in L, i.e., $\hat{x}$ is a solution to the problem P (41). Let

51. $$\Delta = \{\alpha = (\alpha^1, \alpha^2, \cdots, \alpha^p) \mid h^i(x) - h^i(\hat{x}) < \alpha^i,\ i = 1, 2, \cdots, p, \text{ for some } x \varepsilon A\}.$$

Since $\hat{x}$ is a solution to P, Δ does not contain the origin. Furthermore, since h is convex, Δ is a convex set in E^p. By the separation theorem [7] there exists a hyperplane in E^p separating Δ from the origin, i.e., there exists a vector $\bar{\alpha}$ in E^p, $\bar{\alpha} \neq 0$ such that

52. $\langle \bar{\alpha}, \alpha \rangle \geqq 0$ for all $\alpha \varepsilon \Delta$.

Since each α^i can be made as large as we wish, we must have $\bar{\alpha}^i \geqq 0$ and hence $\bar{\alpha} \geq 0$. For any positive scalar $\epsilon > 0$, let $\alpha = h(x) - h(\hat{x}) + \epsilon e$ for some x in A and $e = (1, 1, \cdots, 1)$. The vector α is in Δ by definition, and hence, from (52),

53. $\langle \bar{\alpha}, h(x) - h(\hat{x}) \rangle \geqq -\epsilon \langle \bar{\alpha}, e \rangle$.

Relation (53) holds for every x in A, and since ϵ is arbitrary,

54. $\langle \bar{\alpha}, h(x) - h(\hat{x}) \rangle \geqq 0$ for all x in A.

55. If we define $\bar{\lambda} = \bar{\alpha} / \Sigma_{i=1}^{p} \bar{\alpha}^i$, then $\bar{\lambda} \varepsilon \bar{\Lambda}$ and

56. $\langle \bar{\lambda}, h(\hat{x}) \rangle \leqq \langle \bar{\lambda}, h(x) \rangle$ for all x in A.

But (55) and (56) implies that $\hat{x} \varepsilon N$.

57. Corollary: If A is convex and h is strictly convex (component-wise), then $L = N$.

Proof: This follows from (49) and (50).

58. Definition: We shall say that a solution $\hat{x}$ of the problem P defined in (41), is regular if there exists a closed convex neighborhood U of $\hat{x}$ such that for any $y \varepsilon A \cap U$ the relation $h(\hat{x}) = h(y)$ implies that $\hat{x} = y$.

59. Definition: We shall say that the problem P is regular if every solution of P is a regular solution.

Remark: It is easy to verify that if h is convex and one if its components is strictly convex then P is regular.

60. Theorem: Suppose that the problem P is regular, that h is continuous and convex, and that the constraint set A is a closed convex subset of a Hausdorff, locally convex, linear topological space $\mathfrak{X}$, with the property that for some closed convex neighborhood V of the origin, the set $(A - \{x\}) \cap V$ is compact for every x in A. Then the set L is contained in the closure of the set M.

Proof: We shall show that for every $\hat{x} \in L$ there exists a sequence of points in M which converges to $\hat{x}$.

We begin by constructing a sequence which converges to an arbitrary, but fixed, $\hat{x}$ in L. We shall then show that this sequence is in M.

Let $\hat{x}$ be any point in L. Since we can translate the origins of $\mathfrak{X}$ and E^p, we may suppose, without loss of generality, that $\hat{x} = 0$ and that $h(\hat{x}) = 0$.

Let U be a closed convex neighborhood of $\hat{x}$ satisfying the conditions of definition (58) with respect to $\hat{x}$, and let $V \subset U$ be a closed convex neighborhood of $\hat{x}$ such that $A \cap V$ is compact. For any positive scalar ϵ, $0 < \epsilon \leqq 1/p$, (where p is the dimension of the space containing the range of $h(\cdot)$), let

61. $\Lambda(\epsilon) = \{\lambda = (\lambda^1, \lambda^2, \cdots, \lambda^p) \in E^p \mid \sum_{i=1}^{p} \lambda^i = 1, \ \lambda^i \geqq \epsilon \text{ for } i = 1, 2, \cdots, p\}$.

Let g be the real-valued function with domain $A \cap V \times \Lambda(\epsilon)$, defined by

62. $g(\lambda, x) = \langle \lambda, h(x) \rangle$.

Clearly, since h is continuous and convex, g is continuous in $A \cap V \times \Lambda(\epsilon)$. Furthermore, g is convex in x for fixed λ and linear in λ for fixed x. Since the sets $A \cap V$ and $\Lambda(\epsilon)$ are compact, the sets

63. $\{x \in A \cap V \mid g(\bar{\lambda}, x) = \min_{\eta \in A \cap V} g(\bar{\lambda}, \eta)\}$,

64. $\{\lambda \in \Lambda(\epsilon) \mid g(\lambda, \bar{x}) = \operatorname{Max}_{\xi \in \Lambda(\epsilon)} g(\xi, \bar{x})\}$,

are well defined for every $\bar{\lambda} \in \Lambda(\epsilon)$ and every $\bar{x} \in A \cap V$, respectively. Obviously, because of the form of g and because the sets $A \cap V$ and $\Lambda(\epsilon)$ are convex the sets defined in (63) and (64) are also convex.

By Ky Fan's Theorem [10], there exist a point $\lambda(\epsilon)$ in $\Lambda(\epsilon)$ and a point $x(\epsilon)$ in $A \cap V$ such that

65. $\langle \lambda, h(x(\epsilon))\rangle \leqq \langle \lambda(\epsilon), h(x(\epsilon))\rangle \leqq \langle \lambda(\epsilon), h(x)\rangle$

for every x in $A \cap V$ and λ in $\Lambda(\epsilon)$.

Since $\hat{x} = 0$ is in $A \cap V$ and $h(\hat{x}) = 0$, we have from (65)

66. $\langle \lambda(\epsilon), h(x(\epsilon))\rangle \leqq 0$,

And, from (65) and (66),

67. $\langle \lambda, h(x(\epsilon))\rangle \leqq 0$ for every λ in $\Lambda(\epsilon)$.

Since $A \cap V$ is compact, we can choose a sequence ϵ_n, $n = 1, 2, \cdots$, with $0 < \epsilon_n \leqq 1/p$, converging to zero in such a way that the resulting sequence of points $x(\epsilon_n)$, satisfying (65), converges, i.e.,

68. $\lim_{n\to\infty} x(\epsilon_n) = x^*, \; x^* \varepsilon A \cap V$.

Since $g(\lambda, x)$ is continuous, it follows from (67) and (68) that

69. $\langle \lambda, h(x^*)\rangle \leqq 0$ for all $\lambda \varepsilon \Lambda$,

which implies that $h(x^*) \leqq 0$. But $\hat{x}$ is a solution to P; hence, $h(x^*) \leqq 0 = h(\hat{x})$ implies that $h(x^*) = h(\hat{x})$. Consequently, since P is regular, $x^* = \hat{x} = 0$. Thus, we have constructed a sequence $\{x(\epsilon_n)\}$, which converges to $\hat{x}$.

We shall now show that the sequence $\{x(\epsilon_n)\}$ contains a subsequence, $\{x(\epsilon_n)\}$ also converging to $\hat{x}$, which is contained in M.

Since $\hat{x}$ is in the interior of V, there exists a positive integer n_0 such that the points $x(\epsilon_n) \varepsilon A \cap V$ belong to the interior of V for $n \geqq n_0$.

We will show that for $n \geqq n_0$, $x(\epsilon_n)$ is a solution to $P(\lambda(\epsilon_n))$, i.e., that for $n \geqq n_0$, $x(\epsilon_n) \varepsilon M$. By way of contradiction, suppose that for $n \geqq n_0$, $x(\epsilon_n)$ is not a solution to $P(\lambda(\epsilon_n))$. Then there must exist a point x' in A such that

70. $\langle \lambda(\epsilon_n), h(x') \rangle < \langle \lambda(\epsilon_n), h(x(\epsilon_n)) \rangle$.

Let $x''(\alpha) = (1-\alpha)x(\epsilon_n) + \alpha x'$, with $0 < \alpha < 1$. Since A is convex, $x''(\alpha)$ is in A for $0 < \alpha < 1$. But for $n \geqq n_0$, $x(\epsilon_n)$ is in the interior of V, and hence there exists an α^*, $0 < \alpha^* < 1$, such that $x''(\alpha^*)$ belongs to V.

Now,

71. $\langle \lambda(\epsilon_n), h(x''(\alpha^*)) \rangle = \langle \lambda(\epsilon_n), h((1-\alpha^*)x(\epsilon_n) + \alpha x')\rangle$.

But for $\lambda(\epsilon_n) \epsilon \Lambda(\epsilon_n)$, $\langle \lambda(\epsilon_n), h(x) \rangle$ is convex in x. Hence (70) and (71) imply that

72. $\langle \lambda(\epsilon_n), h(x''(\alpha^*)) \rangle < \langle \lambda(\epsilon_n), h(x(\epsilon_n)) \rangle$,

which contradicts (65).

Therefore, for $n \geqq n_0$, $x(\epsilon_n)$ is a solution to $P(\lambda(\epsilon_n))$, i.e., $x(\epsilon_n)$ is in M.

Thus, for any given $\hat{x} \varepsilon L$, there exists a sequence $\{x(\epsilon_n)\}$ contained in M such that $x(\epsilon_n) \to \hat{x}$ as $n \to \infty$. This completes our proof.

III - Applications to optimal control

To illustrate the applicability of the theory just developed, we shall use it to obtain a maximum principle for an optimal control problem with a vector-valued cost function. It will be observed that when the vector cost function degenerates into a scalar cost function, our maximum principle becomes identical with the Pontryagin Maximum Principle.

Consider a dynamical system described by the differential equation:

73. $$\frac{dx}{dt} = f(x, u)$$

for all t in the compact interval $I = [t_1, t_2]$, where $x(t) \varepsilon E^n$ is the state of the system at time t, $u(t) \varepsilon E^m$ is the input of the system at time t, and f is a function defined in $E^n \times E^m$ with range in E^n.

The Optimal Control Problem is that of finding a control $\hat{u}(t)$, $t \varepsilon I$, and a corresponding trajectory $\hat{x}(t)$, determined by (73), such that

74. (i) for $t \varepsilon I$, $\hat{u}(t)$ is a measurable, essentially bounded function

whose range is contained in an arbitrary but fixed subset U of E^m;

75. (ii) $\hat{x}(t_1) = x_0$, where x_0, a fixed vector in E^n, is the given initial condition;

76. (iii) $\hat{x}(t_2) \varepsilon X_2$, where $X_2 = \{x \varepsilon E^n | g(x) = 0\}$, and g maps E^n into E^ℓ (X_2 is the fixed target set);

77. (iv) for every control $u(t)$, $t \varepsilon I$, and corresponding trajectory $x(t)$, satisfying the conditions (74), (75), and (76), the relation

$$\int_{t_1}^{t_2} c(x(t), u(t))\, dt \leqq \int_{t_1}^{t_2} c(\hat{x}(t), \hat{u}(t))\, dt \text{ implies that}$$

$$\int_{t_1}^{t_2} c(x(t), u(t))\, dt = \int_{t_1}^{t_2} c(\hat{x}(t), \hat{u}(t))\, dt,$$ where $c(x,u)$ maps $E^n \times E^m$ into E^p.

We make the following assumptions:

78. (i) the functions $f(x,u)$ and $c(x,u)$ are continuous in both x and u, and are continuously differentiable in x;

79. (ii) the function $g(x)$ is continuously differentiable and the corresponding Jacobian matrix $\frac{\partial g(x)}{\partial x}$ is of maximum rank for every x in X_2.

To transcribe the control problem into the form of the Basic Problem (3), we require the following definitions:

Let I_α denote the $\alpha \times \alpha$ identity matrix and let $0_{\alpha,\beta}$ denote the $\alpha \times \beta$ zero matrix. We define the projection matrices P_1 and P_2 as

80. $P_1 = (I_p, 0_{p,n})$,

and

81. $P_2 = (0_{n,p}, I_n)$.

Let $F: E^{p+n} \times E^m \to E^{p+n}$ be the function defined by

82. $F(z,u) = (c(P_2 z, u), \; f(P_2 z, u)), \; z \varepsilon E^{p+n}, \; u \varepsilon E^m$.

Now consider the differential equation

83. $\frac{dz}{dt} = F(z,u)$

for some $u(t) \varepsilon E^m$ for $t \varepsilon I$.

It is clear that the optimal control problem is equivalent to the problem of finding a control $\hat{u}(t)$, $t \varepsilon I$ and a corresponding trajectory $\hat{z}(t)$, determined by (83), such that

84. (i) for $t \varepsilon I$, $\hat{u}(t)$ is a measurable, essentially bounded function, whose range is contained in an arbitrary but fixed subset U of E^m;

85. (ii) $\hat{z}(t_1) = (0, x_0) = z_0$; where x_0, a fixed vector in E^n, is the given initial condition;

86. (iii) $\hat{z}(t_2) \varepsilon X_2'$, where $X_2' = \{z \varepsilon E^{p+n} \mid g(P_2 z) = 0\}$, where g maps E^n into E^ℓ ;

87. (iv) for every control u(t), with $t \varepsilon I$, and corresponding trajectory z(t), satisfying (83) and the conditions (i), (ii), and (iii) above, the relation $P_1 z(t_2) \leqq P_1 \hat{z}(t_2)$ implies that $P_1 z(t_2) = P_1 \hat{z}(t_2)$.

Finally, we define

88. $h(z) = P_1 z(t_2)$,

89. $r(z) = g(P_2 z(t_2))$,

90. and we let Ω be the set of all absolutely continuous functions z from I into E^{p+n} which, for some measurable, essentially bounded function u from I into $U \subset E^m$, satisfy the differential equation (83) for almost all t in I, with $z(t_1) = (0, x_0)$.

91. Remark: It is clear that with h, r, and Ω defined as in (88), (89), and (90), respectively, we have transcribed the optimal control problem into the form of the Basic Problem (3). We shall call the transcribed optimal control "the optimal control problem in standard form."

We still have not defined the linear topological space $\mathfrak{X}$. From (90) it is clear that Ω is a subset of the linear space of all absolutely continuous functions from I into E^{p+n}. However, since we wish to use a linearization constructed first by Pontryagin et al. [11], we find it necessary to imbed Ω into a larger topological linear vector space which we define below.

Let $\mathcal{U}$ be the set of all upper semi-continuous real valued functions* defined on I, and let $\mathcal{S} = \mathcal{U} - \mathcal{U}$. From the properties of upper and lower semi-continuous functions (see [12]), it follows that $\mathcal{S}$ is a linear vector space. We then define $\mathfrak{X}$ to be the Cartesian product $\mathcal{S}^{n+p} = \mathcal{S} \times \mathcal{S} \times \cdots \times \mathcal{S}$, with the pointwise topology, [13] i.e., the topology which is constructed from the sub-base consisting of the family of all subsets of the form $\{f \in \mathfrak{X} : f(t) \in N\}$, where t is a point in I and N is an open set in E^{p+n}.

It is easy to show that h and r, as respectively defined by (88) and (89), are continuous.

Let $\hat{z}(t)$, corresponding to the control $\hat{u}(t)$, be a solution to the optimal control problem in standard form (91). We now proceed to construct a linearization for the constraint set Ω at $\hat{z}$.

Let $I_1 \subset I$ be the set of all points t at which $\hat{u}(t)$ is regular, i.e.,

92. $$I_1 = \{\underline{t} \mid t_1 < \underline{t} < t_2, \lim_{\text{meas}(T)\to 0} \frac{\text{meas}(\hat{u}^{-1}(N) \cap T)}{\text{meas}(T)} = 1, \text{ for every neighborhood } N \text{ of } \hat{u}(t), \ \underline{t} \in T \subset I\}.$$

---

* Definition: A real valued function $f: E^1 \to E^1$ is called upper semi-continuous at a point t_0 in E^1, if $\limsup_{t\to t_0} f(t) \leqq f(t_0)$. And it is called lower semi-continuous if $-f$ is upper semi-continuous [12].

Let $\Phi(t,\tau)$ be the $(p+n) \times (p+n)$ matrix which satisfies the linear differential equation

93. $$\frac{d}{dt}\Phi(t,\tau) = \frac{\partial F}{\partial z}(\hat{z}(t), \hat{u}(t))\Phi(t,\tau)$$

for almost all $t \in I$, with $\Phi(\tau,\tau) = I_{p+n}$, the $(p+n)$ identity matrix.

For any $s \in I_1$ and $v \in U$ we define

94. $$\delta z_{s,v}(t) = \begin{cases} 0 \text{ for } t_1 \leqq t < s \\ \Phi(t,s)[F(\hat{z}(s),v) - F(\hat{z}(s), \hat{u}(s))], \quad s \leqq t \leqq t_2, \end{cases}$$

and

95. $$C(\hat{z},\Omega) = \left\{ \delta z \in \mathfrak{X} \;\middle|\; \delta z(t) = \sum_{i=1}^{k} \alpha_i \delta z_{s_i,v_i}(t),\ \{s_1, s_2, \cdots, s_k\} \subset I_1, \right.$$

$$\left. \{v_1, v_2, \cdots, v_k\} \subset U,\ \alpha_i \geqq 0, \text{ for } i=1,2,\cdots,k,\ k \text{ arbitrary finite} \right\}.$$

The work by Pontryagin et al. [11] provides a proof that the set $C(\hat{z},\Omega)$ defined in (95), is a linearization for the set Ω at $\hat{z}$. The linear maps $h'(\hat{z})$ and $r'(\hat{z})$ which one uses with this linearization are defined as follows. For every $\delta z \in \mathfrak{X}$,

96. $$h'(\hat{z})(\delta z) = P_1 \delta z(t_2)$$

and

97. $r'(\hat{z})(\delta z) = \dfrac{\partial g(P_2 \hat{z}(t_2))}{\partial x} P_2 \delta z(t_2)$.

Therefore, from Theorem (11), there exist a vector μ in E^p and a vector η in E^ℓ such that

98. (i) $\mu^i \leqq 0$ for $i = 1, 2, \cdots, p$;

99. (ii) $(\mu, \eta) \neq 0$;

100. (iii) $\langle \mu, P_1 \delta z(t_2)\rangle + \langle \eta, \dfrac{\partial g(P_2 \hat{z}(t_2))}{\partial x} P_2 \delta z(t_2)\rangle \leqq 0$ for all $\delta z \in \overline{C(\hat{z}, \Omega)}$.

Since every $\delta z_{s,v}(t)$, as defined in (94), is in $\overline{C(\hat{z}, \Omega)}$, (100) implies that

101. $\langle \mu, P_1 \Phi(t_2, s)[F(\hat{z}(s), v) - F(\hat{z}(s), \hat{u}(s))]\rangle +$

$$+ \langle \eta, \frac{\partial g(P_2 \hat{z}(t_2))}{\partial x} P_2 \Phi(t_2, s)[F(\hat{z}(s), v) - F(\hat{z}(s), \hat{u}(s))]\rangle \leqq 0$$

for every $s \in I$, and $v \in U$.

Hence,

102. $\langle \Phi^T(t_2, t)\left[P_1^T \mu + P_2^T \dfrac{\partial g^T(P_2 \hat{z}(t_2))}{\partial x} \eta\right], F(\hat{z}(t), v) - F(\hat{z}(t), \hat{u}(t))\rangle \leqq 0$

for every $t \in I_1$, and $v \in U$.

103. Let $\psi(t) = \Phi^T(t_2, t)(P_1^T \mu + P_2^T \frac{\partial g^T(P_2\hat{z}(t_2))}{\partial x} \eta)$, i.e.,

for almost all t in I, $\psi(t)$ satisfies the differential equation:

104. $$\frac{d}{dt} \psi^T(t) = -\psi^T(t) \frac{\partial F(\hat{z}(t), \hat{u}(t))}{\partial z}; \quad \psi^T(t_2) = \mu^T P_1 + \eta^T \frac{\partial g(P_2 \hat{z}(t_2))}{\partial x} P_2 .$$

Combining (102) and (103), we obtain

105. $$\langle \psi(t), F(\hat{z}(t), \hat{u}(t)) \rangle = \text{Maximum} \{ \langle \psi(t), F(\hat{z}(t), v) \rangle \mid v \in U \} \text{ for } t \in I_1 .$$

Since meas (I_1) = meas (I), (105) holds for almost all t in I.

Remark: By assumption (see(79)), $\frac{\partial g(P_2\hat{z}(t_2))}{\partial x}$ is of maximum rank, and since $(\mu, \eta) \neq 0$, $\psi(t)$ as a solution to (103) is not identically zero.

Thus, we have proved the following theorem, which we state in terms of the original quantities defining the optimal control problem.

105. Theorem: If the control $\hat{u}(t)$ and the corresponding trajectory $\hat{x}(t)$, $t \in I$, solve the optimal control problem, then there exist a vector $\psi_1 \in E^p$, $\psi_1 \leqq 0$, and a vector-valued function $\psi_2(t) \in E^n$, with $(\psi_1, \psi_2(t)) \not\equiv 0$, such that

(i) $$\frac{d\psi_2^T(t)}{dt} = -\psi_1^T \frac{\partial c(\hat{x}(t), \hat{u}(t))}{\partial x} - \psi_2^T(t) \frac{\partial f(\hat{x}(t), \hat{u}(t))}{\partial x},$$

(ii) $$\psi_2(t_2) = \left(\frac{\partial g(\hat{x}(t_2))}{\partial x} \right)^T \eta, \text{ for some } \eta \in E^\ell,$$

(iii) for every $v \in U$ and almost all $t \in I$,

$$\langle \psi_1, c(\hat{x}(t), \hat{u}(t)) \rangle + \langle \psi_2(t), f(\hat{x}(t), \hat{u}(t)) \rangle \geqq \langle \psi_1, c(\hat{x}(t), v) \rangle + \langle \psi_2(t), f(\hat{x}(t), v) \rangle .$$

CONCLUSION

In this paper we have presented a theory of necessary conditions for a canonical vector-valued criterion optimization problem. To demonstrate that many complex optimization problems can be transcribed into our canonical form, we have used our necessary conditions to construct a Pontryagin type maximum principle for an optimal control problem with vector cost. It is not difficult to show that most nonlinear programming problems of interest can also be treated within our framework (see [14]).

We have also considered the possibility of "scalarizing" a vector-valued criterion problem by using convex combinations of the components of the vector cost. Our results indicate that the solution sets of the vector and scalar criterion problems do not necessarily coincide.

Since the conditions presented in this paper are considerably more general than hitherto available in the literature, it is hoped that they will open up important classes of optimization problems.

REFERENCES

1. Pareto, V., Cours d'Economie Politique, Lausanne, Rouge, 1896.
2. Karlin, S., Mathematical Methods and Theory in Games, Programming and Economics, 1, Addison-Wesley, Massachusetts, 1959.
3. Debreu, G., Theory of Value, John Wiley, New York, 1959.
4. Kuhn, H. W. and Tucker, A. W., "Nonlinear Programming," Proc. of the Second Berkeley Symposium on Mathematic Statistics and Probability, University of California Press, Berkeley, California, 1951, pp. 481-492.
5. Zadeh, L. A., "Optimality and Non-Scalar-Valued Performance Criteria," IEEE Transactions on Automatic Control, vol. AC-8, number 1, pp. 59-60; January, 1963.
6. Chang, S. S. L., "General Theory of Optimal Processes," J. SIAM Control, Vol. 4, No. 1, 1966, pp. 46-55.
7. Edwards, R. E., Functional Analysis Theory and Applications, Holt Rinehart and Winston, New York, 1965.
8. Dieudonne', J., Foundations of Modern Analysis, Academic Press, New York, 1960.
9. Klinger, A., "Vector-valued Performance Criteria," IEEE Transactions on Automatic Control, vol. AC-9, number 2, January 1964, pp. 117-118.
10. Fan, K., "Fixed Point and Minimax Theorems in Locally Convex Topological Linear Spaces," Proc. Natl. Acad. Sci., U. S. 38 (1952), pp. 121-126.

11. Pontryagin, L. S. et al., The Mathematical Theory of Optimal Processes, John Wiley, New York 1962.

12. McShane, E. J., Integration, Princeton University Press, Princeton, 1964.

13. Kelly, J. L., General Topology, Van Nostrand, New York, 1955.

14. Da Cunha, N. O. and Polak, E., "Constrained Minimization under Vector-Valued Criteria in Finite Dimensional Spaces," Electronics Research Laboratory, University of California, Berkeley, California, Memorandum No. ERL-M188, October 1966.

15. Neustadt, L. W., "An Abstract Variational Theory with Applications to a Broad Class of Optimization Problems II," USCEE Report 169, 1966.

16. Halkin, H. and Neustadt, L. W., "General Necessary Conditions for Optimization Problems," USCEE Report 173, 1966.

17. Canon, M., Cullum, C. and Polak, E., "Constrained Minimization Problems in Finite Dimensional Spaces," J. SIAM Control, Vol. 4, No. 3, 1966, pp. 528-547.