

NONLINEAR PROGRAMMING IN BANACH SPACES WITH APPLICATIONS TO OPTIMAL CONTROL PROBLEMS

Tagawa, Shojiro
International Institute for Advanced Study of Social Information Science

<https://doi.org/10.5109/13103>

出版情報：統計数理研究. 17 (1/2), pp.49-61, 1976-03. Research Association of Statistical Sciences

バージョン：

権利関係：



NONLINEAR PROGRAMMING IN BANACH SPACES WITH APPLICATIONS TO OPTIMAL CONTROL PROBLEMS

By

Shôjirô TAGAWA*

(Received March 29, 1975; revised August 31, 1975)

§ 1. Introduction.

This paper is devoted to present the necessary and sufficient conditions for the optimal solution to nonlinear programming in Banach spaces which are analogous to those given in the previous papers [8] and [10]. Indeed, we shall treat the following nonlinear programming problem (P_0) : Minimize $f(x)$ subject to $x \in A$ and $g(x) \in B$, where $f(x)$ is a real valued function defined over a real Banach space X , A is a quasi-convex subset of X , $g(x)$ is a function from X into a real Banach space Y , and B is a closed convex set with non-empty interior in Y . In contrast with the previous paper [8], the locally relatively convex set is replaced by the quasi-convex set, which is closely related to that introduced in [2], and the closed convex cone is replaced by the closed convex set. We also consider the case where the regularity assumption is set up. We shall apply these arguments to the optimal control problems with restricted phase coordinates. Then, the necessary and sufficient conditions for an trajectory to be an optimal solution to the control problem are obtained. There are many varieties of necessary conditions and sufficient conditions, separately, but there are only a few necessary and sufficient conditions. In the methods using the theory of functional analysis, there are necessary conditions, for example, see [5], [6] and [7], and sufficient conditions, for example, see [4].

In Section 2, we shall give basic definitions and preliminary results, in particular, the properties of the quasi-convex set. We shall present nonlinear programming problem (P_0) and describe the conditions which the optimal solution must satisfy, in Section 3. We also discuss about the regularity assumptions, under which necessary and sufficient conditions are derived in Theorem 3.4. In Section 4, we shall consider the optimal control problems with restricted phase coordinates, and show that these control problems fall under the category of nonlinear programming problem (P_0) and that, under some regularity assumptions, necessary and sufficient conditions are obtained.

* International Institute for Advanced Study of Social Information Science, Fujitsu Limited, 17-25, Shinkamata 1-Chome, Ohtaku, Tokyo, Japan.

§ 2. Preliminaries.

In this section, we shall introduce the notations, definitions and some preliminary results. Some notations and definitions are the same as those given in [8].

Let X be a real linear topological space, A a non-empty subset of X and $\bar{x} \in A$.

DEFINITION 2.1. (Varaiya [12]). The intersection of all closed cones containing the set $A - \bar{x} = \{a - \bar{x} | a \in A\}$ is called the *closed cone of A at \bar{x}* . We shall denote it by $C(A, \bar{x})$.

DEFINITION 2.2. The intersection of all closed convex cones containing the set $A - \bar{x}$ is called the *closed convex cone of A at \bar{x}* . We shall denote it by $CC(A, \bar{x})$.

DEFINITION 2.3. (Varaiya [12]). The set $LC(A, \bar{x})$ defined by

$$LC(A, \bar{x}) = \bigcap_{N \in \mathfrak{N}(\bar{x})} C(A \cap N, \bar{x}),$$

where $\mathfrak{N}(\bar{x})$ is the class of all neighborhoods of \bar{x} , is called the *local closed cone of A at \bar{x}* .

DEFINITION 2.4. The set $P(A, \bar{x})$ defined by

$$P(A, \bar{x}) = \bigcap_{N \in \mathfrak{N}(\bar{x})} CC(A \cap N, \bar{x})$$

is called the *local closed convex cone of A at \bar{x}* .

DEFINITION 2.5. The set A is called a *pseudo-cone with vertex at \bar{x}* if

$$x - \bar{x} \in LC(A, \bar{x}) \quad \text{for all } x \in A.$$

DEFINITION 2.6. The set A is called a *quasi-convex set* if $\text{co}(A) \subset \bar{A}$.

LEMMA 2.1. Let X be a topological space, N an open set in X and A an arbitrary set in X . Then, it is valid that $\overline{A \cap N} = \bar{A} \cap \bar{N}$.

PROOF. Immediate.

Q. E. D.

LEMMA 2.2. Let X be a linear topological space, A an arbitrary set in X which contains \bar{x} . Then, it is valid that

$$LC(A, \bar{x}) = LC(\bar{A}, \bar{x}).$$

PROOF. This follows from Lemma 2.1.

Q. E. D.

LEMMA 2.3. Let X be a locally convex linear topological space and let A be a set in X which contains \bar{x} . If A is convex, then it is true that

$$P(A, \bar{x}) = LC(A, \bar{x}) = C(A, \bar{x}).$$

PROOF. Since

$$\text{cone}(A \cap N - \bar{x}) = \text{cone}(A - \bar{x}),$$

for every neighborhood N of \bar{x} , the assertion of Lemma 2.3 is easily verified.

Q. E. D.

LEMMA 2.4. Let X be a locally convex linear topological space and let A be a quasi-convex set in X . Then, for each $\bar{x} \in A$, we have

$$P(A, \bar{x}) = LC(A, \bar{x}) = C(A, \bar{x}).$$

PROOF. This follows from Lemmas 2.2 and 2.3.

Q. E. D.

LEMMA 2.5. *Let X be a linear topological space. If A is a quasi-convex set in X , then it is a pseudo-cone with vertex at every point in A .*

PROOF. This follows from Lemma 2.4.

Q. E. D.

§ 3. Nonlinear programming in Banach spaces.

In this section, we shall present nonlinear programming in Banach spaces which is one of generalizations of Theorem 3.3 in [8].

DEFINITION 3.1. (Neustadt [5]). The function f from an open domain D in a linear topological space X into a linear topological space Y is called *differentiable at $\bar{x} \in D$ in the sense of Neustadt* if to every $x \in X$ there corresponds a vector $f_{\bar{x}}(x) \in Y$ such that

$$\frac{f(\bar{x} + \varepsilon y) - f(\bar{x})}{\varepsilon} \xrightarrow[\substack{\varepsilon \rightarrow 0+ \\ y \rightarrow x}]{} f_{\bar{x}}(x).$$

DEFINITION 3.2. Let X and Y be linear topological spaces, B a closed convex cone with vertex at the origin in Y , A a convex set in X , and h a function from A into Y . Then, h is called to be *B -convex* if

$$[h(\alpha x_1 + \beta x_2) - \alpha h(x_1) - \beta h(x_2)] \in B,$$

whenever $0 \leq \alpha, \beta \leq 1, \alpha + \beta = 1$, and $x_1, x_2 \in A$.

LEMMA 3.1. *Let Y be a Banach space and B a convex set having a non-empty interior in Y . Then, it is valid that*

$$(3.1) \quad C(B, \bar{y}) = \overline{\text{cone}(\text{int}(B - \bar{y}))},$$

where $\bar{y} \in B$. Furthermore, we have

$$(3.2) \quad \text{int } C(B, \bar{y}) = \text{cone}(\text{int}(B - \bar{y})).$$

PROOF. It is clear that

$$\overline{\text{cone}(\text{int}(B - \bar{y}))} \subset C(B, \bar{y}).$$

To show the contradiction, assume that there exists a vector $z \in C(B, \bar{y})$ such that

$$(3.3) \quad z \notin \overline{\text{cone}(\text{int}(B - \bar{y}))}.$$

By virtue of Lemma 2.1 in [8] and Lemma 2.3, there are a sequence of positive numbers $\{\lambda_n\}$ and a sequence of vectors $\{y_n\} \subset B$ such that

$$y_n \longrightarrow \bar{y}, \quad \lambda_n(y_n - \bar{y}) \longrightarrow z \quad \text{as } n \rightarrow \infty.$$

There is a vector $y_0 \in \text{int } B$. Since B is convex, and since $\lambda_n \geq 1$ for sufficiently large n , there is a positive integer M such that

$$w_n \equiv \left(1 - \frac{1}{n\lambda_n}\right)y_n + \frac{1}{n\lambda_n}y_0 \in \text{int } B \quad \text{for all } n \geq M.$$

We then have

$$\|w_n - \bar{y}\| \leq \|y_n - \bar{y}\| + \frac{1}{n\lambda_n}\|y_0 - y_n\| \longrightarrow 0 \quad \text{as } n \rightarrow \infty$$

and

$$\|\lambda_n(w_n - \bar{y}) - z\| \leq \|\lambda_n(y_n - \bar{y}) - z\| + \frac{1}{n} \|y_0 - y_n\| \longrightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Hence, we have

$$z \in \overline{\text{cone}(\text{int}(B - \bar{y}))},$$

which contradicts to (3.3). Consequently, the equality (3.1) holds.

Next, we shall show that the equality (3.2) holds. It follows from (3.1) that

$$\begin{aligned} \text{int } C(B, \bar{y}) &= \text{int}(\overline{\text{cone}(\text{int}(B - \bar{y}))}) \\ &= \text{int}(\text{cone}(\text{int}(B - \bar{y}))) \\ &= \text{cone}(\text{int}(B - \bar{y})). \end{aligned} \quad \text{Q. E. D.}$$

Now, we shall consider the following nonlinear programming problem in Banach spaces. Let X and Y be Banach spaces, A a quasi-convex set in X , B a closed convex set in Y having a non-empty interior, f a real valued function defined over X , and g a function from X into Y .

PROBLEM (P_0): Find a vector $\bar{x} \in A$ which minimizes $f(x)$ subject to $x \in A$ and $g(x) \in B$.

THEOREM 3.1. Let $\bar{x} \in A$ be an optimal solution to the problem (P_0). Assume that

- (a) A is a quasi-convex set,
- (b) f and g are differentiable at \bar{x} in the sense of Neustadt, the differential $f_{\bar{x}}(x)$ is convex continuous in $x \in X$ and the differential $g_{\bar{x}}(x)$ is $C(B, g(\bar{x}))$ -convex continuous in $x \in X$.

Then, there exist a real number $\bar{\eta}$ and a linear continuous functional $\bar{y}^* \in Y^*$, not both zero, such that

$$(3.4) \quad \bar{\eta} \geq 0,$$

$$(3.5) \quad \bar{y}^*(b) \leq \bar{y}^*(g(\bar{x})) \quad \text{for all } b \in B,$$

$$(3.6) \quad g(\bar{x}) \in B,$$

$$(3.7) \quad \bar{\eta} f_{\bar{x}}(x) + \bar{y}^*(g_{\bar{x}}(x)) \geq 0 \quad \text{for all } x \in P(A, \bar{x}).$$

PROOF. Let E be the set in $Y \times R^1$ defined by

$$E = \{(g_{\bar{x}}(x), f_{\bar{x}}(x)) - (b, r) \in Y \times R^1 \mid x \in P(A, \bar{x}), b \in C(B, g(\bar{x})), r \leq 0\}.$$

Then, the set E is convex. Let \tilde{B} be the set in $Y \times R^1$ defined by

$$\tilde{B} = \text{int } C(B, g(\bar{x})) \times \{r \in R^1 \mid r < 0\}.$$

It is then valid that \tilde{B} is an open convex cone in $Y \times R^1$. To show the contradiction, assume that there is a vector $(y, \alpha) \in Y \times R^1$ such that $(y, \alpha) \in E \cap \tilde{B}$. Then, there exist a vector $b_1 \in \text{int } C(B, g(\bar{x}))$, a vector $b_2 \in C(B, g(\bar{x}))$ and a vector $x \in P(A, \bar{x})$ such that

$$(3.8) \quad g_{\bar{x}}(x) = b_1 + b_2,$$

$$(3.9) \quad f_{\bar{x}}(x) \leq f_{\bar{x}}(x) - r = \alpha < 0.$$

It follows from Lemma 3.1 in [6] and (3.8) that

$$g_{\bar{x}}(x) \in \text{int } C(B, g(\bar{x})).$$

It is true, by virtue of Lemma 3.1, that there exist a positive real number μ and a vector $b_3 \in \text{int } B$ such that

$$g_{\bar{x}}(x) = \mu(b_3 - g(\bar{x})).$$

Since $x \in P(A, \bar{x})$, it follows from Lemma 2.2 in [8] that there exist a sequence of positive numbers $\{\lambda_n\}$, a sequence of positive integers $\{m_n\}$, a sequence of vectors $(\alpha_{n1}, \dots, \alpha_{nm_n}) \in P^{m_n}$ and a sequence of vectors $\{x_{ni}\}_{i=1, \dots, m_n} \subset A$ such that

$$x_{ni} \longrightarrow \bar{x}, \quad \lambda_n \left(\sum_{i=1}^{m_n} \alpha_{ni} x_{ni} - \bar{x} \right) \longrightarrow x \quad \text{as } n \rightarrow \infty.$$

For a neighborhood $U(b_3)$ of b_3 satisfying $U(b_3) \subset \text{int } B$, there is a positive integer M_1 such that

$$g\left(\sum_{i=1}^{m_n} \alpha_{ni} x_{ni}\right) \in \frac{\mu}{\lambda_n} U(b_3) + \left(1 - \frac{\mu}{\lambda_n}\right) g(\bar{x}).$$

Moreover, there is a positive integer $M_2 \geq M_1$ such that $\lambda_n > \mu$ for all $n \geq M_2$. Since B is convex, we have

$$\sum_{i=1}^{m_n} \alpha_{ni} x_{ni} \in \text{co}(A) \cap \text{int } g^{-1}(B) \quad \text{for all } n \geq M_2.$$

Therefore, it is valid that

$$x \in \text{LC}(\text{co}(A) \cap \text{int } g^{-1}(B), \bar{x}).$$

Since A is quasi-convex, it follows from Lemma 2.1 that we have

$$\text{LC}(\text{co}(A) \cap \text{int } g^{-1}(B), \bar{x}) = \text{LC}(A \cap \text{int } g^{-1}(B), \bar{x}).$$

It is then true, on the basis of Theorem 3.2 in [8], that $f_{\bar{x}}(x) \geq 0$, which contradicts to (3.9).

Consequently, we conclude that the sets E and \tilde{B} are disjoint, and hence it follows from the separation theorem (see [1]) that there exist a real number $\bar{\eta}$ and a linear continuous functional $\bar{y}^* \in Y^*$, not both zero, such that

$$(3.10) \quad \bar{y}^*(y_1) + \bar{\eta} \alpha_1 \leq 0 \leq \bar{y}^*(g_{\bar{x}}(x) - b) + \bar{\eta}(f_{\bar{x}}(x) - r) \\ \text{for all } y_1 \in \text{int } C(B, g(\bar{x})), \quad \text{all } \alpha_1 < 0, \quad \text{all } x \in P(A, \bar{x}), \\ \text{all } b \in C(B, g(\bar{x})) \quad \text{and all } r \leq 0.$$

If we set $x=0$ and $b=0$ in (3.10), it is clear that (3.4) holds. By setting $x=0$ and $r=0$, we have

$$\bar{y}^*(b) \leq 0 \quad \text{for all } b \in C(B, g(\bar{x})),$$

which immediately implies (3.5). At last, by setting $b=0$ and $r=0$, we obtain (3.7).

Q. E. D.

DEFINITION 3.3. Let f be a real valued function defined on a linear topological space X . Then, for a subset A of X , f is called to be *pseudo-convex over A at \bar{x}* ($\in A$)

if f is differentiable at \bar{x} in the sense of Neustadt and if

$$x \in A \quad \text{and} \quad f_{\bar{x}}(x - \bar{x}) \geq 0 \quad \text{imply} \quad f(x) - f(\bar{x}) \geq 0.$$

If f is pseudo-convex over A at each point in A , the f is called to be *pseudo-convex* on A .

REGULARITY ASSUMPTION: Let X and Y be linear topological spaces, A a subset of X which contains a vector \bar{x} , B a closed convex set with non-empty interior and g a function from X into Y which is differentiable at \bar{x} in the sense of Neustadt. Then, there is a vector $x_0 \in P(A, \bar{x})$ such that

$$g(\bar{x}) + g_{\bar{x}}(x_0) \in \text{int } B.$$

THEOREM 3.2. *Let all notations and hypotheses be the same as those in Theorem 3.1. Suppose that Regularity Assumption holds. If $\bar{x} \in A$ is an optimal solution to the problem (P_0) , then there is a linear continuous functional $\bar{y}^* \in Y^*$ such that*

$$(3.11) \quad \bar{y}^*(b) \leq \bar{y}^*(g(\bar{x})) \quad \text{for all } b \in B,$$

$$(3.12) \quad g(\bar{x}) \in B,$$

$$(3.13) \quad f_{\bar{x}}(x) + \bar{y}^*(g_{\bar{x}}(x)) \geq 0 \quad \text{for all } x \in P(A, \bar{x}).$$

PROOF. It follows from Theorem 3.1 that there are a real number $\bar{\eta}$ and a linear continuous functional $\bar{y}^* \in Y^*$, not both zero, such that conditions (3.4)–(3.7) hold. It suffices to show that $\bar{\eta} \neq 0$. To show the contradiction, suppose that $\bar{\eta} = 0$. Let the vector x_0 satisfy Regularity Assumption. By virtue of (3.5), we have

$$\bar{y}^*(g_{\bar{x}}(x_0)) \leq 0.$$

On the other hand, (3.7) with $\bar{\eta} = 0$ implies

$$\bar{y}^*(g_{\bar{x}}(x_0)) \geq 0.$$

Hence, we have $\bar{y}^*(g_{\bar{x}}(x_0)) = 0$ and so

$$(3.14) \quad \bar{y}^*(g(\bar{x}) + g_{\bar{x}}(x_0)) = \bar{y}^*(g(\bar{x})).$$

Let $y_0 = g(\bar{x}) + g_{\bar{x}}(x_0)$ and note that $y_0 \in \text{int } B$. Then, for every $y \in Y$ there exists a small positive number ε such that

$$(3.15) \quad y_0 + \varepsilon y \in B \quad \text{and} \quad y_0 - \varepsilon y \in B.$$

By virtue of (3.5), (3.14) and (3.15), we have

$$\bar{y}^*(y) = 0 \quad \text{for all } y \in Y.$$

This is a contradiction.

Q. E. D.

LEMMA 3.2. *Let X be a linear topological space, G a closed convex set with non-empty interior in X and A a quasi-convex set in X . If $\bar{A} \cap \text{int } G \neq \emptyset$, then the set $A \cap G$ is non-empty and quasi-convex.*

PROOF. We first show that

$$\bar{A} \cap G = \overline{\bar{A} \cap \text{int } G}.$$

It is clear that $\overline{\bar{A} \cap \text{int } G} \subset \bar{A} \cap G$. By the assumption, there is a vector $y \in \bar{A} \cap \text{int } G$.

Then, for each $x \in \bar{A} \cap G$, we have

$$(1-\lambda)y + \lambda x \in \bar{A} \cap \text{int } G \quad \text{for } 0 \leq \lambda < 1,$$

so that $x \in \overline{\bar{A} \cap \text{int } G}$.

Therefore, we obtain

$$\begin{aligned} \text{co}(A \cap G) &\subset \text{co}(A) \cap \text{co}(G) \subset \bar{A} \cap G = \overline{\bar{A} \cap \text{int } G} \\ &= \overline{\bar{A} \cap \text{int } G} \quad (\text{by Lemma 2.1}) \\ &\subset \overline{A \cap \text{int } G} = \bar{A} \cap G. \end{aligned}$$

Q. E. D.

THEOREM 3.3. Consider the problem (P_0) . Suppose that A is a quasi-convex set, that g is differentiable on X in the sense of Neustadt, that f is pseudo-convex on A , that the set $G \equiv \{x \in X \mid g(x) \in B\}$ is convex and that $\bar{A} \cap \text{int } G \neq \emptyset$. Then, if the vector $\bar{x} \in A$ satisfies the conditions (3.4)–(3.7) for some $\bar{\eta} \neq 0$ and some linear continuous functional $\bar{y}^* \in Y^*$, then the vector \bar{x} is an optimal solution to (P_0) .

PROOF. It follows from Lemma 3.2 that the set $A \cap G$ is quasi-convex and hence $A \cap G$ is a pseudo-cone with vertex at \bar{x} (by Lemma 2.5). Then, by virtue of Theorem 3.4 in [8], the vector \bar{x} is an optimal solution to (P_0) . Q. E. D.

THEOREM 3.4. Under the same circumstances as in Theorem 3.1, suppose that the set $G \equiv \{x \in X \mid g(x) \in B\}$ is convex, that f is pseudo-convex on $A \cap G$ and that $\bar{A} \cap \text{int } G \neq \emptyset$. If Regularity Assumption holds, then for the vector $\bar{x} \in A$ to be an optimal solution to the problem (P_0) , it is necessary and sufficient that there exists a linear continuous functional $\bar{y}^* \in Y^*$ which satisfies the conditions (3.11)–(3.13).

PROOF. It is true, on the basis of Theorems 3.2 and 3.3, that the assertion of Theorem 3.4 holds. Q. E. D.

LEMMA 3.3. Let X and Y be Banach spaces, A a quasi-convex set in X , B a closed convex cone with non-empty interior in Y and g a B -convex function from X into Y . Then, the condition $(*)$: There is a vector $x_0 \in A$ such that $g(x_0) \in \text{int } B$, implies Regularity Assumption.

PROOF. We first show that

$$(3.16) \quad \text{int } C(B, g(\bar{x})) + B = \text{int } C(B, g(\bar{x})).$$

It is evident that $\text{int } C(B, g(\bar{x})) + B \supset \text{int } C(B, g(\bar{x}))$. For each vector $y \in \text{int } C(B, g(\bar{x})) + B$, there exist vectors $y_1 \in \text{int } C(B, g(\bar{x}))$ and $y_2 \in B$ such that $y = y_1 + y_2$. It follows from (3.2) in Lemma 3.1 that there are a positive number η_1 and a vector $y_{11} \in \text{int } B$ such that $y_1 = \eta_1(y_{11} - g(\bar{x}))$. Then, we have

$$y = \eta_1 \left(\left(y_{11} + \frac{1}{\eta_1} y_2 \right) - g(\bar{x}) \right).$$

Since (see Lemma 3.1 in [6])

$$y_{11} + \frac{1}{\eta_1} y_2 \in \text{int } B + B = \text{int } B,$$

it is valid that

$$y \in \text{cone}(\text{int } (B - g(\bar{x}))) = \text{int } C(B, g(\bar{x})).$$

Now, it is true that

$$g_{\bar{x}}(x_0 - \bar{x}) - [g(x_0) - g(\bar{x})] \in B,$$

since g is a B -convex function. By virtue of (3.16), we have

$$g_{\bar{x}}(x_0 - \bar{x}) = g(x_0) - g(\bar{x}) + \{g_{\bar{x}}(x_0 - \bar{x}) - [g(x_0) - g(\bar{x})]\} \in \text{int } C(B, g(\bar{x})).$$

Then, there are a positive number ζ and a vector $b_0 \in \text{int } B$ such that $g_{\bar{x}}(x_0 - \bar{x}) = \zeta(b_0 - g(\bar{x}))$. If we set $x'_0 = \frac{1}{\zeta}(x_0 - \bar{x})$, then $x'_0 \in P(A, \bar{x})$ and $g(\bar{x}) + g_{\bar{x}}(x'_0) = b_0 \in \text{int } B$.
Q. E. D.

REMARK: The number ζ in the proof of Lemma 3.3 can be taken as $\zeta > 1$.

§ 4. Applications to optimal control problems.

In this section, we shall describe optimal control problems as applications of Theorem 3.4. We shall derive necessary and sufficient conditions for the optimal solution.

We first consider the following optimal control problem with restricted phase coordinates in case of continuous time. Let I be a compact interval $[t_0, t_1]$, U an arbitrary set in Euclidian space R^m , and \mathcal{U} an admissible control class consisting of all m -vector valued functions defined on I which are measurable and essentially bounded on I , and satisfy $u(t) \in U$ for all $t \in I$. The dynamical system is described by

$$(4.1) \quad \dot{x}(t) = A(t)x(t) + b(u(t), t) \quad \text{for almost all } t \in I,$$

where $x(t)$ is an absolutely continuous n -vector valued function on I , $A(t)$ is a continuous $n \times n$ matrix valued function on I , $b(u, t)$ is a continuous n -vector valued function on $U \times I$, and $u \in \mathcal{U}$. Let $f^0(x, t)$ and $g^0(x, t)$ be continuously differentiable functions on $R^n \times I$ which are convex in x , and moreover, g^0 is of class C^2 in x . Let B_0 and B_1 be closed convex set with non-empty interior in R^n . Then, consider the following problem (CP): Minimize $\int_{t_0}^{t_1} f^0(x(t), t) dt$ subject to x satisfies (4.1) for some $u \in \mathcal{U}$, $x(t_0) \in B_0$, $x(t_1) \in B_1$, and $g^0(x(t), t) \leq 0$ for all $t \in I$.

Let X be a Banach space of all continuous n -vector valued functions on I with the uniform norm. Define a subset H of X by

$$(4.2) \quad H = \{x \in X \mid x \text{ satisfies (4.1) for some } u \in \mathcal{U}, x(t_0) \in R^n\}.$$

Let $f(x) = \int_{t_0}^{t_1} f^0(x(t), t) dt$ for all $x \in X$, then f is a continuous convex function from X into R^1 . Let Z be a Banach space of all continuous real valued functions on I with the uniform norm and D a closed convex cone in Z defined by

$$D = \{z \in Z \mid z(t) \leq 0 \text{ for all } t \in I\}.$$

Let $Y = R^n \times R^n \times Z$, and set

$$g(x) = (x(t_0), x(t_1), g^0(x(t), t)),$$

then g is a continuous convex function from X into Y . The control problem (CP) is now reformulated as follows: Minimize $f(x)$ subject to $x \in H$ and $g(x) \in B$, where $B = B_0 \times B_1 \times D$.

In order to apply Theorem 3.4 to the problem (CP), we must prove that the set H is quasi-convex.

LEMMA 4.1. *The set H defined in (4.2) is quasi-convex.*

PROOF. It suffices to show that for arbitrary two vectors x and y in H and for any λ , $0 \leq \lambda \leq 1$,

$$(4.3) \quad \lambda x + (1-\lambda)y \in \bar{H}.$$

Let the corresponding controllers to x and y be u and v in \mathbb{U} , respectively. To see (4.3), we shall show that for any positive number ε , there exists a vector $z \in H$ with the corresponding controller w such that

$$|\lambda x(t) + (1-\lambda)y(t) - z(t)| < \varepsilon \quad \text{for all } t \in I.$$

If we set $z(t_0) = \lambda x(t_0) + (1-\lambda)y(t_0)$, then we have

$$\begin{aligned} |\lambda x(t) + (1-\lambda)y(t) - z(t)| &\leq \int_{t_0}^t |A(s)| |\lambda x(s) + (1-\lambda)y(s) - z(s)| ds \\ &\quad + \left| \int_{t_0}^t (\lambda b(u(s), s) + (1-\lambda)b(v(s), s) - b(w(s), s)) ds \right|. \end{aligned}$$

It follows from Gronwall's inequality that

$$|\lambda x(t) + (1-\lambda)y(t) - z(t)| \leq \varphi(t) + \int_{t_0}^t |A(s)| \varphi(s) \exp \left(\int_s^t |A(\sigma)| d\sigma \right) ds,$$

where

$$\varphi(t) \equiv \left| \int_{t_0}^t (\lambda b(u(s), s) + (1-\lambda)b(v(s), s) - b(w(s), s)) ds \right|.$$

It is true, on the basis of Lemma 4.1 in [2], that we can take $w \in \mathbb{U}$ such as

$$\varphi(t) < \varepsilon / \exp \left(\int_{t_0}^{t_1} |A(\sigma)| d\sigma \right) \quad \text{for all } t \in I.$$

Then, we can conclude that

$$\begin{aligned} |\lambda x(t) + (1-\lambda)y(t) - z(t)| &\leq \left(\varepsilon / \exp \left(\int_{t_0}^{t_1} |A(\sigma)| d\sigma \right) \right) \left(1 + \int_{t_0}^t |A(s)| \exp \left(\int_s^t |A(\sigma)| d\sigma \right) ds \right) \leq \varepsilon \\ &\quad \text{for all } t \in I. \quad \text{Q. E. D.} \end{aligned}$$

We shall suppose the following assumption:

ASSUMPTION (A): There is at least one trajectory $x_0 \in H$ such that $x_0(t_0) \in \text{int } B_0$, $x_0(t_1) \in \text{int } B_1$, and $g^0(x_0(t), t) < 0$ for all $t \in I$.

By virtue of Lemma 3.3, it is valid that Regularity Assumption holds if we set up Assumption (A). Now, we can apply Theorem 3.4 to this control problem.

Let $\bar{x} \in H$ be a solution to the problem (CP), then there exists a linear continuous functional $\bar{y}^* \in Y^*$ such that the conditions (3.11)-(3.13) hold. Since $\bar{y}^* = (\bar{\alpha}, \bar{\beta}, \bar{z}^*)$, where $\bar{\alpha}, \bar{\beta} \in R^n$, and $\bar{z}^* \in Z^*$, and since

$$\begin{aligned} f_{\bar{x}}(x) &= \int_{t_0}^{t_1} \frac{\partial f^0(\bar{x}(t), t)}{\partial x} x(t) dt, \\ g_{\bar{x}}(x) &= \left(x(t_0), x(t_1), \frac{\partial g^0(\bar{x}(t), t)}{\partial x} x(t) \right), \end{aligned}$$

it is valid, by (3.11) and (3.13), that

$$(4.4) \quad \bar{\alpha}\bar{\tau} + \bar{\beta}\mu + \bar{z}^*(z) \leq \bar{\alpha}\bar{x}(t_0) + \bar{\beta}\bar{x}(t_1) + \bar{z}^*(g^0(\bar{x}(\cdot), \cdot))$$

for all $\tau \in B_0$, all $\mu \in B_1$ and all $z \in D$,

$$(4.5) \quad \int_{t_0}^{t_1} \frac{\partial f^0(\bar{x}(t), t)}{\partial x} x(t) dt + \bar{\alpha}x(t_0) + \bar{\beta}x(t_1) + \bar{z}^*\left(\frac{\partial g^0(\bar{x}(\cdot), \cdot)}{\partial x} x(\cdot)\right) \geq 0,$$

for all $x \in C(H, \bar{x})$.

If we set $z = g^0(\bar{x}(\cdot), \cdot)$ in (4.4), then we have

$$(4.6) \quad \bar{\alpha}\tau \leq \bar{\alpha}\bar{x}(t_0) \quad \text{for all } \tau \in B_0,$$

$$(4.7) \quad \bar{\beta}\mu \leq \bar{\beta}\bar{x}(t_1) \quad \text{for all } \mu \in B_1.$$

It follows from the representation theorem for the space Z^* that there is a real valued function $v(t)$ of bounded variation on I which satisfies

$$\bar{z}^*(z) = \int_{t_0}^{t_1} z(t) dv(t) \quad \text{for all } z \in Z.$$

Moreover, we can suppose that

$$(4.8) \quad v(t) \text{ is continuous from right in } (t_0, t_1) \text{ and } v(t_0) = 0.$$

Since $0 \in D$ and $2g^0(x(\cdot), \cdot) \in D$, we have (by (4.4))

$$\int_{t_0}^{t_1} g^0(\bar{x}(t), t) dv(t) = 0,$$

and

$$\int_{t_0}^{t_1} z(t) dv(t) \leq 0 \quad \text{whenever } z(t) \leq 0 \text{ for all } t \in I.$$

It is then valid that

$$(4.9) \quad \frac{dv(t)}{dt} \geq 0 \quad \text{for almost all } t \in I,$$

and

$$(4.10) \quad v(t) \text{ is constant on subinterval of } I \text{ on which } g^0(\bar{x}(t), t) < 0.$$

It is easy to verify that

$$C(H, \bar{x}) = \{\bar{\lambda}x(t; \bar{\xi}, u) \mid \bar{\lambda} > 0, \bar{\xi} \in R^n, u \in \mathbb{U}\},$$

where

$$x(t; \bar{\xi}, u) = \Phi(t)\bar{\xi} + \Phi(t) \int_{t_0}^t \Phi(s)^{-1} [b(u(s), s) - b(\bar{u}(s), s)] ds,$$

for all $t \in I$, $\bar{\xi} \in R^n$, $u \in \mathbb{U}$,

and $\Phi(t)$ is the $n \times n$ matrix valued function which satisfies

$$\frac{d\Phi(t)}{dt} = A(t)\Phi(t) \quad \text{for almost all } t \in I,$$

$$\Phi(t_0) = \text{identity}.$$

Therefore, by setting $u = \bar{u}$ in (4.5), we have

$$(4.11) \quad \int_{t_0}^{t_1} \frac{\partial f_0(\bar{x}(t), t)}{\partial x} \Phi(t) dt + \bar{\alpha} + \bar{\beta} \Phi(t_1) + \int_{t_0}^{t_1} \frac{\partial g^0(\bar{x}(t), t)}{\partial x} \Phi(t) dv(t) = 0.$$

Furthermore, if we set $\xi=0$ in (4.5), we obtain

$$(4.12) \quad \int_{t_0}^{t_1} \left(\int_s^{t_1} \frac{\partial f^0(\bar{x}(t), t)}{\partial x} \Phi(t) dt + \bar{\beta} \Phi(t_1) + v(t_1) \frac{\partial g^0(\bar{x}(t_1), t_1)}{\partial x} \Phi(t_1) \right) \Phi(s)^{-1} [b(u(s), s) \\ - b(\bar{u}(s), s)] ds - \int_{t_0}^{t_1} \int_s^{t_1} v(t) \left\{ \frac{\partial^2 g^0(\bar{x}(t), t)}{\partial x^2} (A(t) \bar{x}(t) + b(\bar{u}(t), t)) \right. \\ \left. + \frac{\partial^2 g^0(\bar{x}(t), t)}{\partial x \partial t} \right\} \Phi(t) dt \Phi(s)^{-1} [b(u(s), s) - b(\bar{u}(s), s)] ds \\ - \int_{t_0}^{t_1} \int_s^{t_1} v(t) \frac{\partial g^0(\bar{x}(t), t)}{\partial x} A(t) \Phi(t) dt \Phi(s)^{-1} [b(u(s), s) - b(\bar{u}(s), s)] ds \\ - \int_{t_0}^{t_1} v(t) \frac{\partial g^0(\bar{x}(t), t)}{\partial x} [b(u(t), t) - b(\bar{u}(t), t)] dt \geq 0 \quad \text{for all } u \in \mathcal{U}.$$

Let us define

$$\psi(s) \equiv \left(\int_s^{t_1} \left[\frac{\partial f^0(\bar{x}(t), t)}{\partial x} \Phi(t) - v(t) \right] \left\{ \frac{\partial^2 g^0(\bar{x}(t), t)}{\partial x^2} (A(t) \bar{x}(t) + b(\bar{u}(t), t)) \right. \right. \\ \left. \left. + \frac{\partial^2 g^0(\bar{x}(t), t)}{\partial x \partial t} \right\} \Phi(t) - v(t) \frac{\partial g^0(\bar{x}(t), t)}{\partial x} A(t) \Phi(t) \right) dt + \bar{\beta} \Phi(t_1) \\ \left. + v(t_1) \frac{\partial g^0(\bar{x}(t_1), t_1)}{\partial x} \Phi(t_1) \right) \Phi(s)^{-1} \quad \text{for all } s \in I.$$

Then, the (row) n -vector valued absolutely continuous function $\psi(s)$ satisfies

$$(4.13) \quad \psi(t_0) = -\bar{\alpha},$$

$$(4.14) \quad \psi(t_1) = \bar{\beta} + v(t_1) \frac{\partial g^0(\bar{x}(t_1), t_1)}{\partial x},$$

$$(4.15) \quad \dot{\psi}(s) = -\psi(s) A(s) - \frac{\partial f^0(\bar{x}(s), s)}{\partial x} - v(s) \left\{ \frac{\partial^2 g^0(\bar{x}(s), s)}{\partial x^2} (A(s) \bar{x}(s) \right. \\ \left. + b(\bar{u}(s), s)) + \frac{\partial^2 g^0(\bar{x}(s), s)}{\partial x \partial t} \right\} - v(s) \frac{\partial g^0(\bar{x}(s), s)}{\partial x} A(s) \\ \text{for almost all } s \in I.$$

Now, the inequality (4.12) is restated as follows:

$$\int_{t_0}^{t_1} \left[\psi(t) - v(t) \frac{\partial g^0(\bar{x}(t), t)}{\partial x} \right] [b(u(t), t) - b(\bar{u}(t), t)] \geq 0 \quad \text{for all } u \in \mathcal{U},$$

which is equivalent to the pointwise minimum principle (see [10]):

$$(4.16) \quad \left[\psi(t) - v(t) \frac{\partial g^0(\bar{x}(t), t)}{\partial x} \right] b(\bar{u}(t), t) = \min_{u \in \mathcal{U}} \left[\psi(t) - v(t) \frac{\partial g^0(\bar{x}(t), t)}{\partial x} \right] b(u, t) \\ \text{for almost all } t \in I.$$

Thus, we can conclude the following result:

THEOREM 4.1. *Consider the linear optimal control problem (CP) with restricted phase coordinates. Suppose that the Assumption (A) holds. Then, for the trajectory*

$\bar{x}(t)$ with the corresponding controller $\bar{u}(t)$ to be an optimal solution to (CP), it is necessary and sufficient that there exist (row) n -vectors $\bar{\alpha}$ and $\bar{\beta}$, which satisfy (4.6) and (4.7), a real valued function $v(t)$ of bounded variation on I , which satisfies (4.8)–(4.10), and a (row) n -vector valued, absolutely continuous function $\phi(s)$, which satisfies (4.13)–(4.15) and the pointwise minimum principle (4.16).

We next consider the optimal control problem with restricted phase coordinates where the dynamical system is described by the difference equation. Let U be an arbitrary set in R^n , J a set of integers $\{1, 2, \dots, k-1\}$, $b(u, i)$ a function from $U \times J$ into R^n , U_1, \dots, U_{k-1} subsets of U such that each set $b(U_i, i)$, $i \in J$, is convex in R^n and \mathcal{U}_J a set of functions $u(i)$ from J into R^n such that $u(i) \in U_i$ for all $i \in J$. The dynamical system is given by

$$(4.17) \quad x(i+1) = A(i)x(i) + b(u(i), i) \quad \text{for all } i \in J,$$

where $A(i)$ and $x(i)$ are $n \times n$ matrix and n -vector, respectively. Let B_0 and B_1 be closed convex sets with non-empty interior, $f^0(x)$ a function from R^n into R^1 , which is convex and of class C^1 , and $g^0(x, i)$ a function from $R^n \times \{1, \dots, k\}$ into R^1 , which is convex and of class C^1 in x . Then, consider the control problem (CP'): Minimize $f^0(x(k))$ subject to $x(i)$ satisfies (4.17) for some $u \in \mathcal{U}_J$, $x(1) \in B_0$, $x(k) \in B_1$ and $g^0(x(i), i) \leq 0$ for $i=1, 2, \dots, k$. We also suppose the following regularity assumption (A').

ASSUMPTION (A'): There is at least one trajectory $x(i)$, $i \in J$, which satisfies (4.17) for some $u \in \mathcal{U}_J$, $x(1) \in \text{int } B_0$, $x(k) \in \text{int } B_1$, and $g^0(x(i), i) < 0$, for $i=1, \dots, k$.

Let us define

$$\begin{aligned} F &= \{x = (x(1), \dots, x(k)) \in R^{nk} \mid x \text{ satisfies (4.17) for some } u \in \mathcal{U}_J \\ &\quad \text{and some } x(1) \in B_0, x(k) \in B_1\}, \\ g(x) &= (x(1), x(k), g^0(x(1), 1), \dots, g^0(x(k), k)), \\ f(x) &= f^0(x(k)) \quad \text{for all } x = (x(1), \dots, x(k)) \in R^{nk}, \\ G &= \{(r_1, \dots, r_k) \in R^k \mid r_i \leq 0 \text{ for } i=1, 2, \dots, k\}. \end{aligned}$$

Then, the problem (CP') is restated as follows: Minimize $f(x)$ subject to $x \in F$ and $g(x) \in B_0 \times B_1 \times G$.

Since the set F is convex, we can apply Theorem 3.4 to the problem (CP'), and we obtain the following results.

THEOREM 4.2. Consider the optimal control problem (CP') with restricted phase coordinates where the dynamical system is described by the difference equation (4.17). Suppose that the assumption (A'). Then, for the trajectory \bar{x} to be an optimal solution to (CP'), it is necessary and sufficient that there exist real numbers $\bar{\gamma}_i$, $i=1, \dots, k$, (row) n -vectors $\bar{\alpha}$ and $\bar{\beta}$ such that

$$\begin{aligned} \bar{\gamma}_i &\geq 0 && \text{for } i=1, \dots, k, \text{ and } \bar{\gamma}_i = 0 \text{ if } g^0(\bar{x}(i), i) < 0, \\ \bar{\alpha}b_0 &\leq \bar{\alpha}\bar{x}(1) && \text{for all } b_0 \in B_0, \\ \bar{\beta}b_1 &\leq \bar{\beta}\bar{x}(k) && \text{for all } b_1 \in B_1, \end{aligned}$$

and (row) n -vector valued function $\phi(i)$ defined on $\{1, 2, \dots, k\}$ such that

$$\phi(1) = -\bar{\alpha},$$

$$\phi(k) = \bar{\beta} + \frac{\partial f^0(\bar{x}(k))}{\partial x} + \bar{\gamma}_k \frac{\partial g^0(\bar{x}(k), k)}{\partial x},$$

$$\phi(j) = \phi(j+1)A(j) + \bar{\gamma}_j \frac{\partial g^0(\bar{x}(j), j)}{\partial x} \quad \text{for } j=1, 2, \dots, k-1,$$

and

$$\phi(j+1)b(\bar{u}(j), j) = \min_{v \in U_j} \phi(j+1)b(v, j) \quad \text{for } j=1, \dots, k-1.$$

Acknowledgement. The author wishes to his gratitude to Director Tosio Kitagawa for his helpful suggestions. He is also indebted to Professor Nagata Furukawa of Kyushu University for his valuable comments and corrections to original version of this paper.

References

- [1] N. DUNFORD and J.T. SCHWARTZ, *Linear Operators, Part I*, Interscience, New York, 1964.
- [2] R.V. GAMKRELIDZE, *On some extremal problems in the theory of differential equations with applications to the theory of optimal control*, SIAM J. Control, 3 (1965), 106-128.
- [3] E.B. LEE and L. MARKUS, *Foundations of Optimal Control Theory*, John Wiley & Sons, 1967.
- [4] L.W. NEUSTADT, *Sufficiency conditions and a duality theory for mathematical programming problems in arbitrary linear spaces*, in *Nonlinear Programming*, edited by J.B. Rosen, O.L. Mangasarian and K. Ritter, Academic Press, 1970.
- [5] L.W. NEUSTADT, *An abstract variational theory with applications to a broad class of optimization problems. I. General theory*, SIAM J. Control, 4 (1966), 505-527.
- [6] L.W. NEUSTADT, *A general theory of extremals*, J. Computer and System Sciences, 3 (1969), 57-92.
- [7] B.N. PSHENICHNYI, *Necessary Conditions for an Extremum*, Marcel Dekker, 1971.
- [8] S. TAGAWA, *On a class of nonlinear programming in a Banach space*, Bull. Math. Statist., 16 (1974), 95-114.
- [9] S. TAGAWA, *Generalized convexities of continuous functions and their applications to mathematical programming*, Ibid., 16 (1974), 115-125.
- [10] S. TAGAWA, *On a class of nonlinear programming with equality constraints in Banach spaces*, Bull. Math. Statist., 17 (1976), 17-32.
- [11] S. TAGAWA, *On a convergence of minimizing sequences for uniformly convex functions*, Bull. Math. Statist., 17 (1976), 1-16.
- [12] P. VARAIYA, *Nonlinear programming in Banach space*, SIAM J. Appl. Math., 15 (1967), 284-293.