数理解析研究所講究録
第 16 巻 1966 年 53-75

ON AN OPTIMIZATION PROBLEM FOR DISCRETE-TIME CONTROL SYSTEMS

by Suguru Arimoto
Tokyo University

1. INTRODUCTION

In optimization problems of multi-stage processes or discrete-time control processes, some types of necessary conditions for optimality were proposed in several papers, like the maximum principle for continuous-time control processes. In the earliest paper [1], Chang proposed a necessary condition for optimality and called it "the digitized maximum principle." Similar results were also obtained by Katz [2]. Butkovsky [3], in addition to a counter-example to Katz's theorem, offered the local maximum principle which implies that the Hamiltonian attains the maximum value in a neighborhood of the optimal control. However, the local maximum principle for discrete-time processes does not hold in general as shown by a counter-example which will be given in the last section of this paper.

Recently, the two important papers by Halkin [4], Jordan and Polak [5] were published. Halkin showed geometric aspects of necessary conditions for optimality, and Jordan and Polak established the local maximum or stationary principle.

In the present paper we shall consider more general optimization problems for multi-stage processes than those in the above cited references.

The problem stated in Section 2 is a discrete version of

Berkovitz's problem [6] formulated for continuous-time control systems. It is also regarded as a generalization of nonlinear bottleneck-type programming problems in multi-stage production processes first discussed by Bellman [7]. In Section 4 a necessary condition for optimality will be proved.

Sections 5 and 6 treat a special case of the problem. In Section 5 a sufficient condition for local optimality will be given in Theorem 2. In Section 6, a global maximum principle will be proposed in Theorem 3 under an additional condition which is analogous to that given by Fillipov [8] for the proof of the existence of an optimal control for continuous-time systems.

In our previous work [9] we proposed the analogous theorem to Theorem 3 given in Section 6, but, in the proof, we falsely used the local maximum principle which does not hold in general.

2. PROBLEM STATEMENT

Let us consider a multi-stage process whose state at the t-th stage ($t = 0, 1, \ldots$) or time t is described by an n-vector x_t governed by the difference equation

$$x_{t+1} = f_t(x_t, u_t) \tag{2.1}$$

where u_t is an r-vector called a decision and f_t is an n-vector valued function which has continuous first derivatives with respect to all arguments of x_t and u_t. Given an initial state x_o and a sequence $u = \{u_t ;\ t\ 0, 1, \ldots, N-1\}$ of decisions, there exists a unique solution of (2.1) denoted by $x_t = x_t(x_o, u)$.

The problem to be considered is

Problem 1. Given an initial state x_o, find a sequence of decisions $u_o, u_1, \ldots, u_{N-1}$ which minimizes

$$J(N;\ x_o, u) = \sum_{t=o}^{N-1} \alpha_t(x_t(x_o, u), u_t) \tag{2.2}$$

subject to the process equation (2.1) and the inequality side constraints

$$g_t^{(i)}(x_t(x_o, u), u_t) \geqslant 0, \quad i = 1, 2, \ldots, m \tag{2.3}$$

for $t = 0, 1, \ldots, N-1$ where $g_t^{(i)}$ has continuous first derivatives with respect to all arguments.

Throughout this paper we assume that there exist at least one sequence of decisions and the corresponding sequence of states $x_t(x_o, u)$ along which (2.1) and (2.3) are satisfied. We shall call such a sequence of decisions to be admissible.

3. PRELIMINARY FORMULATIONS

Let v_t be an m-vector and define

$$\beta_{t+1} = \beta_t + \alpha_t(x_t, u_t) + v_t' g_t(x_t, u_t),$$
$$\beta_0 = 0, \tag{3.1}$$

where the prime denotes the transpose and $g_t(x_t, u_t)$ is the m-vector whose components are composed of $g_t^{(i)}(x_t, u_t)$ in (2.3). Let $u^* = \{u_t^*\}$ be a fixed admissible sequence of decisions and denote the corresponding state by $x_t^* = x_t(x_0, u^*)$ and the solution of (3.1) by $\beta^* = \beta(x_0, u^*)$. Then, after a lengthy but easy calculation, we obtain

$$x_{t+1} - x_{t+1}^*$$
$$= \frac{\partial f_t(x_t^*, u_t^*)}{\partial x_t^*}(x_t - x_t^*) + \frac{\partial f_t(x_t^*, u_t^*)}{\partial u_t^*}(u_t - u_t^*)$$
$$+ h_t(x_t - x_t^*) + k_t(u_t - u_t^*) \tag{3.2}$$

and

$$\beta_{t+1} - \beta_{t+1}^* - (\beta_t - \beta_t^*)$$
$$= \frac{\partial\left[\alpha_t(x_t^*, u_t^*) + v_t' g_t(x_t^*, u_t^*)\right]}{\partial x_t^*}(x_t - x_t^*)$$

$$+ \frac{\partial\left[\alpha_t(x_t^*, u_t^*) + v_t' g_t(x_t^*, u_t^*)\right]}{\partial u_t^*} (u_t - u_t^*)$$

$$+ h_t^o(x_t - x_t^*) + k_t^o(u_t - u_t^*) , \qquad (3.3)$$

where h_t, k_t are n-vector valued functions and h_t^o, k_t^o are scalar functions such that

$$\begin{cases} |h_t(x_t - x_t^*)| = o(|x_t - x_t^*|), \\ |k_t(u_t - u_t^*)| = o(|u_t - u_t^*|), \\ |h_t^o(x_t - x_t^*)| = o(|x_t - x_t^*|), \\ |k_t^o(u_t - u_t^*)| = o(|u_t - u_t^*|). \end{cases} \qquad (3.4)$$

Here the symbol $|x|$ implies the norm of the vector x, i.e.,

$$|x| = \max_j |x^{(j)}|$$

and the symbol $|h(x)| = o(|x|)$ implies that for an arbitrarily given $\varepsilon > 0$ there exists a positive number ρ such that $|h(x)| \leqslant \varepsilon |x|$ for every x satisfying $|x| \leqslant \rho$.

Next we introduce the following notations :

$$H_t(x_t^*, p_t^*, u_t) = - \alpha_t(x_t^*, u_t) + p_t^{*\prime} f_t(x_t^*, u_t), \qquad (3.5)$$

$$F_t(x_t^*, p_t^*, u_t) = H_t - v_t' g_t(x_t^*, u_t), \qquad (3.6)$$

where p_t^* is an n-vector determined by

$$\dot{p}_{t-1}^* = \partial F_t(x_t^*, p_t^*, u_t^*)/\partial x_t^* \tag{3.7}$$

together with the boundary condition

$$p_{N-1}^* = 0 . \tag{3.8}$$

Finally we define

$$d_t(u, u^*) = \beta_t - \beta_t^* - p_{t-1}^{*\prime}(x_t - x_t^*) . \tag{3.9}$$

Then it follows from (3.2) and (3.3) that

$$\begin{aligned} & d_{t+1}(u, u^*) - d_t(u, u^*) \\ & = - \frac{\partial F_t(x_t^*, p_t^*, u_t^*)}{\partial u_t^*} (u_t - u_t^*) \\ & \quad + h_t^o + k_t^o - p_t^{*\prime}(h_t + k_t) . \end{aligned} \tag{3.10}$$

Note that

$$\begin{cases} d_o(u, u^*) = 0, \\ d_N(u, u^*) = \beta_N - \beta_N^* . \end{cases} \tag{3.11}$$

4. NECESSARY CONDITION

Let

$$U_t(x) = \left\{ u \ ; \ g_t^{(i)}(x, u) \geqslant 0, \quad i=1, \ \dots \ , m \right\}$$

and define

Condition 1. For an arbitrarily fixed x, $U_t(x)$ is a convex subset of R^r .

Condition 2. If $g_t^{(j)}(x, u)=0$ for $j=i_1, \ \dots \ , i_k$ and for arbitrarily fixed x, u, then it holds

$$\operatorname{rank} \left(\frac{\partial g_t^{(j)}(x, u)}{\partial u^{(i)}} \right) = k .$$

Theorem 1. Assume Conditions 1 and 2. Let $u^* = \left\{ u_t^* \right\}$ be an optimal admissible sequence of decisions which minimizes (2.2). Then for every $t=0, 1, \ \dots \ , N-1$ it holds that

i) there exists an m-vector v_t which satisfies

$$v_t^{(i)} g_t^{(i)}(x_t^*, u_t^*) = 0 \qquad i=1, \ \dots \ , m, \tag{4.1}$$

$$\partial F_t(x_t^*, p_t^*, u_t^*)/\partial u_t^* = 0 , \tag{4.2}$$

ii) $v_t^{(i)} \leqslant 0 \qquad i=1, \ \dots \ , m,$ (4.3)

iii) $$\frac{\partial H_t(x_t^*, p_t^*, u_t^*)}{\partial u_t^*}(u_t - u_t^*) \leqslant 0 \qquad (4.4)$$

for all $u_t \in U_t(x_t^*)$.

Proof. At first we assume that the part i) holds for every t. Assume that for $t = s+1, \ldots, N-1$ the conclusions ii) and iii) hold but for $t = s$ do not. Then, by Lemma 1 stated later, there exists a decision $\bar{u}_s \in U_s(x_s^*)$ such that

$$\frac{\partial H_s(x_s^*, p_s^*, u_s^*)}{\partial u_s^*}(\bar{u}_s - u_s^*) = \rho > 0 . \qquad (4.5)$$

Let λ be a small positive number and

$$u_s(\lambda) = \lambda \bar{u}_s + (1-\lambda)u_s^* .$$

Then we have $u_s(\lambda) \in U_s(x_s^*)$ by Condition 1 and

$$\frac{\partial H_s(x_s^*, p_s^*, u_s^*)}{\partial u_s^*}(u_s(\lambda) - u_s^*) = \lambda\rho > 0$$

by using (4.5). Consider now the following process :

$$\begin{cases} x_{t+1}(\lambda) = f_t(x_t(\lambda), u_t(\lambda)) & t = s, s+1, \ldots, N-1, \\ x_s(\lambda) = x_s^* . & \end{cases} \qquad (4.6)$$

Here, $u_t(\lambda)$ is determined successively such that

$$v_t' g_t(x_t(\lambda), u_t(\lambda)) = 0 \qquad t = s+1, \ldots, N-1. \quad (4.7)$$

We note that, from Lemma 2 stated later, it is possible to choose $u_t(\lambda)$ such that, in addition to (4.7),

$$u_t(\lambda) \in U_t(x_t(\lambda)), \qquad |u_t(\lambda) - u_t^*| \leq c\lambda \quad (4.8)$$

for $t = s+1, \ldots, N-1$, where c is a positive constant independent of λ. Thus, if λ is sufficiently small, we have from (3.1), (3.8) to (3.11) that

$$J(N-s \; ; x_s^*, u^*) = \sum_{t=s}^{N-1} \alpha_t(x_t^*, u_t^*)$$

$$= \sum_{t=s}^{N-1} \left[\alpha_t(x_t^*, u_t^*) + v_t' g_t(x_t^*, u_t^*)\right]$$

$$= \beta_N(\lambda) - d_N(u(\lambda), u^*)$$

$$= \sum_{t=s}^{N-1} \left[\alpha_t(x_t(\lambda), u_t(\lambda)) + v_t' g_t(x_t(\lambda), u_t(\lambda))\right] + o(\lambda)$$

$$= J(N-s \; ; x_s^*, u(\lambda)) + v_s' g_s(x_s^*, u_s(\lambda)) + o(\lambda) .$$

Noting that

$$v_s' g_s(x_s^*, u_s(\lambda)) = v_s' g_s(x_s^*, u_s^*)$$

$$+ \frac{\partial v_s' g_s(x_s^*, u_s^*)}{\partial u_s^*} (u_s(\lambda) - u_s^*) + o(\lambda)$$

$$= \frac{\partial H_s(x_s^*, u_s^*)}{\partial u_s^*} (u_s(\lambda) - u_s^*) + o(\lambda)$$

$$= \lambda \rho + o(\lambda) ,$$

we get

$$J(N-s; x_s^*, u(\lambda)) = J(N-s; x_s^*, u^*) - \lambda \rho + o(\lambda)$$

$$< J(N-s; x_s^*, u^*)$$

by choosing λ small enough. This contradicts the optimality of the sequence u^*. By an analogous method to the above we can prove the part i) using Lemma 3.

Now, it remains only to prove the following lemmata.

Lemma 1. If the inequality (4.4) holds for all $u_t \in U_t(x_t^*)$, then (4.3) holds.

Proof. For simplicity we omit the subscript and asterisk. We assume, without loss of generality, that for $j = 1, \ldots , m_o$, $g^{(j)}(x, u) = 0$ and for $j = m_o+1, \ldots , m$, $g^{(j)}(x, u) > 0$. Note that by Condition 2 there exist the vectors $u^1, u^2, \ldots , u^{m_o}$ such that

$$\frac{\partial g^{(j)}(x, u)}{\partial u} (u^k - u) = \begin{cases} 1 & \text{for } j = k , \\ 0 & \text{for } j \neq k , \end{cases}$$

$$k, j = 1, 2, \dots , m_o .$$

To prove this lemma by contradiction we assume that $v^{(k)} > 0$ for some k. Let

$$u^k(\lambda) = \lambda u^k + (1-\lambda)u,$$

$$\bar{u}^k(\lambda) = (1-\gamma)u^k(\lambda) + \delta \sum_{j \neq k} u^j(\lambda)$$

where γ and δ are small positive numbers such that $\gamma = (m_o - 1)\delta$. If λ is taken small enough, then

$$\frac{\partial g^{(j)}(x, u)}{\partial u} (\bar{u}^k(\lambda) - u) = \begin{cases} (1-\gamma)\lambda & \text{for } j = k, \\ \delta \lambda & \text{for } j \neq k. \end{cases}$$

This implies $\bar{u}^k(\lambda) \in U(x)$. Hence

$$\lambda(1-\gamma)v^{(k)} + \lambda \delta \sum_{j \neq k} v^{(j)}$$

$$= \frac{\partial H(x, u)}{\partial u} (\bar{u}^k(\lambda) - u) \leq 0$$

by using (4.2). The last inequality follows from the assumption of the lemma. On the other hand, the left hand side of the above equation becomes positive by choosing δ sufficiently small. Thus the contradiction has been derived.

Lemma 2. Assume that

$$v^{(j)} g^{(j)}(x, u) = 0 \quad \text{for} \quad j = 1, \ldots, m \tag{4.9}$$

and

$$|x(\lambda) - x| \leqslant M_0 \lambda, \qquad u \in U(x) \tag{4.10}$$

for all λ such that $0 \leqslant \lambda \leqslant \rho$, where M_0 and ρ are positive constants. Then for any sufficiently small λ there is a decision $u(\lambda)$ such that

$$v' g(x(\lambda), u(\lambda)) = 0 \tag{4.11}$$

and

$$|u(\lambda) - u| \leqslant M_1 \lambda, \qquad u(\lambda) \in U(x(\lambda)), \tag{4.12}$$

where M_1 is a positive constant independent of λ.

Proof. Without loss of generality, we assume that for $j = m_0+1, \ldots, m$, $g^{(j)}(x, u) > 0$, and for $j = 1, \ldots, m_0$,

$$g^{(j)}(x, u) = 0 . \tag{4.13}$$

Keeping in mind of Condition 2 and the property (4.10), and applying the theory of implicit functions to equation (4.13), we find that there exists a $u(\lambda)$ such that

$$g^{(j)}(x(\lambda),\ u(\lambda)) = 0 \qquad j = 1, \ldots, m_o$$

and

$$|u(\lambda) - u| \leqslant M_1 \lambda \ .$$

On the other hand, if λ is sufficiently small, we have

$$g^{(j)}(x(\lambda),\ u(\lambda)) > 0 \qquad j = m_o+1, \ldots, m.$$

These imply (4.11) and (4.12).

<u>Lemma 3.</u> <u>Assume that for a fixed</u> v, <u>given</u> x <u>and</u> u,

$$g^{(j)}(x,\ u) = 0 \qquad j = 1, \ldots, k$$

<u>where</u> $k < m$, <u>and</u>

$$\frac{\partial F(x,\ p,\ u)}{\partial u} \neq 0 \ .$$

<u>Then, for any sufficiently small</u> λ, <u>there exists a</u> $u(\lambda)$ <u>such that</u>

$$g^{(j)}(x,\ u(\lambda)) = 0 \qquad j = 1, \ldots, k,$$

$$\frac{\partial F(x,\ p,\ u)}{\partial u}\,(\ u(\lambda) - u\ ) > \rho\,\lambda > 0$$

<u>where</u> ρ <u>is a positive constant independent of</u> ρ .

<u>Proof.</u> The proof of this lemma is almost clear and will not be given.

5. SUFFICIENT CONDITION

In this section and the subsequent we consider the special case of Problem 1 when $U_t(x)$ is independent of x, namely, x_t does not enter the constraint inequality (2.3). Hence we say that a sequence $u = \{ u_t \}$ of decisions is admissible if every u_t belongs to the set U_t which is a subset of R^r.

Corollary. Assume that U_t is a convex set for all t. Let $u^* = \{ u_t^* \}$ be an optimal admissible sequence of decisions minimizing (2.2). Then it holds for all $t = 0, 1, \ldots, N-1$ that

$$\frac{\partial H_t(x_t^*, p_t^*, u_t^*)}{\partial u_t^*} (u_t - u_t^*) \leqslant 0$$

for all $u_t \in U_t$.

Now we introduce the following notions.

Definition. If for an admissible sequence $u^* = \{ u_t^* \}$ of decisions there is a positive number ε such that it holds

$$J(N; x_o, u^*) \leqslant J(N; x_o, u) \tag{5.1}$$

for all admissible sequences $u = \{ u_t \}$ satisfying $| u_t - u_t^* | \leqslant \varepsilon$ for $t = 0, 1, \ldots, N-1$, we say that the sequence u^* is locally optimal.

Definition. In addition to the local optimality, if the equality symbol in (5.1) occurs only when $u = u^*$, we say that u^* is locally strict-optimal.

Theorem 2. Let $u^* = \{u_t^*\}$ be an admissible sequence of decisions and assume that for all $t = 0, 1, \dots, N-1$ it holds

$$\frac{\partial H_t(x_t^*, p_t^*, u_t^*)}{\partial u_t^*} (u_t - u_t^*) < 0 \tag{5.2}$$

for any u_t such that $u_t \in U_t$, $u_t \neq u_t^*$ and $|u_t - u_t^*| \leq \varrho$, where ϱ is some positive constant. Then the sequence u^* is locally strict-optimal.

Proof. We prove this theorem by contradiction. Assume that there are infinite admissible sequences of decisions $u(k) = \{u_t(k)\}$, $k = 1, 2, \dots$, such that

$$0 < |u(k) - u^*| = \max_t |u_t(k) - u^*| \leq 1/k \tag{5.3}$$

and

$$J(N;\, x_0,\, u(k)) \leq J(N;\, x_0,\, u^*), \tag{5.4}$$

and denote the corresponding states by $x_t(k) = x_t(x_0, u(k))$.

Let

$$\frac{\partial H_t(x_t^*, p_t^*\, u_t^*)}{\partial u_t^*} (u_t(k) - u_t^*) = - c_t(k) < 0.$$

At first we note that it follows from the assumptions (5.2) and (5.3) that

$$\left| u_t(k) - u_t^* \right| \leqslant M_1 c_t(k) \leqslant M_2/k$$

where M_1 and M_2 are positive constants independent of t and k. Hence, by the same proceeding as in the proof of Theorem 1, we obtain

$$J(N;\ x_o,\ u^*) = J(N;\ x_o,\ u(k)) - \sum_{t=o}^{N-1} c_t(k) + o\left(\sum_{t=o}^{N-1} c_t(k) \right).$$

Thus we have

$$J(N;\ x_o,\ u(k)) > J(N;\ x_o,\ u^*)$$

if k is sufficiently large. This contradicts (5.4).

Remark. It should be noted that the conclusions in Theorem 1 and Corollary are valid even if u^* is locally optimal.

6. GLOBAL MAXIMUM PRINCIPLE

We require the following property.

Condition 3. For all x, any given u^1, $u^2 \in U_t$, and any positive number $0 \leq \lambda \leq 1$, there exists at least a decision $u^3 \in U_t$ such that

$$\lambda f_t(x, u^1) + (1-\lambda) f_t(x, u^2) = f_t(x, u^3),$$
$$\lambda \alpha_t(x, u^1) + (1-\lambda) \alpha_t(x, u^2) \geqslant \alpha_t(x, u^3) \quad (6.1)$$

for every $t = 0, 1, \ldots, N-1$.

Now we prove the global maximum principle.

Theorem 3. Assume Condition 3. Let $u^* = \{u_t^*\}$ be an optimal admissible sequence of decisions. Then it holds for all $t = 0, 1, \ldots, N-1$ that

$$H_t(x_t^*, p_t^*, u_t^*) \geqslant H_t(x_t^*, p_t^*, u_t) \quad (6.2)$$

for all $u_t \in U_t$.

Proof. Let $u = \{u_t\}$ be an arbitrarily fixed admissible sequence of decisions and $\lambda = \{\lambda_t\}$ be a sequence such that $0 \leqslant \lambda_t \leqslant 1$ for $t = 0, 1, \ldots, N-1$. Let

$$x_{t+1} = \lambda_t f_t(x_t, u_t) + (1-\lambda_t) f_t(x_t, u_t^*) \quad (6.3)$$

and denote the state vector for given x_o and $\lambda = \{\lambda_t\}$ by $x_t(\lambda) = x_t(x_o, \lambda)$. **We now consider the new optimization** problem of choosing an optimal sequence $\lambda = \lambda^* = \{\lambda_t^*\}$ such that it minimizes

$$J(N;\, x_o, \lambda) = \sum_{t=o}^{N-1} \alpha_t(x_t(\lambda), \lambda_t) \; . \tag{6.4}$$

subject to $0 \leqslant \lambda_t \leqslant 1$. Of course, it follows immediately from the meaning of this problem that

$$\min_{\lambda} J(N;\, x_o, \lambda) \leqslant J(N;\, x_o, u^*) \; . \tag{6.5}$$

On the other hand, we have from Condition 3 that

$$\min_{\lambda} J(N;\, x_o, \lambda) = J(N;\, x_o, \lambda^*) \geqslant J(N;\, x_o, u^*) \; . \tag{6.6}$$

To prove this, we assume that $J(N;\, x_o, \lambda)$ attains the minimum at $\lambda = \lambda^*$. Then there exists another admissible sequence $\bar{u} = \{\bar{u}_t\}$ satisfying

$$\lambda_t^* f_t(x_t(\lambda^*), u_t) + (1-\lambda_t^*) f_t(x_t(\lambda^*), u_t^*) = f_t(x_t(\lambda^*), \bar{u}_t) \; , \tag{6.7}$$

$$\lambda_t^* \alpha_t(x_t(\lambda^*), u_t) + (1-\lambda_t^*) \alpha_t(x_t(\lambda^*), u_t^*) \geqslant \alpha_t(x_t(\lambda^*), \bar{u}_t) \; . \tag{6.8}$$

Noting that $x_t(\lambda^*) = x_t(x_o, \lambda^*) = x_t(x_o, \bar{u})$ and taking into account of (6.8), we find

$$J(N;\ x_o, \lambda^*) \geqslant J(N;\ x_o, \bar{u}) . \tag{6.9}$$

Consequently,

$$J(N;\ x_o, \lambda^*) = J(N;\ x_o, u^*) . \tag{6.10}$$

Now we take λ^* such that $\lambda_t^* = 0$ for $t = 0, 1, \ldots, N-1$ and apply Corollary to the above mentioned problem. Then we have

$$\frac{\partial\left[\lambda_t^* H_t(x_t^*, p_t^*, u_t) + (1-\lambda_t^*)H_t(x_t^*, p_t^*, u_t^*)\right]}{\partial \lambda_t^*} (\lambda_t - \lambda_t^*) \leqslant 0.$$

This implies (6.2).

7. COUNTER-EXAMPLE

Consider the problem of minimizing $J(2;\, x_o,\, u)$ subject to

$$x_{t+1} = x_t + u_t \,,$$

$$J(2;\, x_o,\, u) = \sum_{t=o}^{1} \left[2x_t^{\,2} - u_t^{\,2} \right],$$

$$x_o = 0 \,, \qquad |u_t| \leqslant 1 \,.$$

By the easy calculation, the optimal decisions are

$$u_o^* = 0 \,, \qquad u_1^* = 1 \;\text{ or }\; -1 \,.$$

Along these decisions, the Hamiltonian becomes

$$H_o(x_o^*,\, p_o^*,\, u_o) = u_o^{\,2} \,.$$

This implies that H_o does not attain the local maximum at the optimal decision $u_o^* = 0$.

ACKNOWLEDGMENT

The author wishes to express his sincere appreciation to Professor Jin-ichi Nagumo of the University of Tokyo for his helpful suggestions, direction and encouragement.

REFERENCES

1. S.S.L. Chang. Digitized maximum principle. Proc. IRE. 48(1960), 2030-2031.
2. S. Katz. Discrete version of Pontryagin's maximum principle. J. Electronics and Control. 13(1962), 179-184.
3. A.G. Butkovskii. The necessary and sufficient conditions for optimality of discrete control systems. Automation and Control. 24(1964), 963-970.
4. H. Halkin. Optimal control for systems described by difference equations. In "Advances in control systems : Theory and applications. Vol. 1," pp. 173-196, Academic Press, N.Y., (1964).
5. B.W. Jordan and E. Polak. Theory of a class of discrete control systems. J. Electronics and Control. 17(1964), 697-711.
6. L.D. Berkovitz. Variational methods in problems of control and programming. J. Math. Anal. Appl. 5(1961), 145-169.

7. R. Bellman. " Dynamic programming," Princeton University Press, Princeton, N.J., (1957).

8. A.F. Fillipov. On certain questions in the theory of optimal control. J. SIAM, Control, Ser. A. 1(1962), 76-84.

9. S. Arimoto. On the optimization problem for sampled-data control systems. Proc. of 13-th Japan National Congr. for Appl. Mech., pp. 244-247 (University of Tokyo, Tokyo, Japan, 1963).