AN ABSTRACT OF THE THESIS OF

KENJI KANBE for the MASTER OF SCIENCE
(Name) (Degree)

in INDUSTRIAL ENGINEERING presented on October 12, 1967
(Major) (Date)

Title: THE MAXIMUM PRINCIPLE AND ITS APPLICATION TO INDUSTRIAL ENGINEERING PROBLEMS

Abstract approved: Redacted for Privacy
Michael S. Inoue

The maximum principle developed by the Russian mathematician, L.S. Pontryagin is considered to be one of the most significant contributions to the recent advances in mathematical optimization techniques.

Unfortunately, most of the published literature on the application of the maximum principle is in the field of control system design, and very little has been published on the application of this principle to industrial engineering problems.

The purpose of this thesis is to apply the maximum principle to practical problems in industry and business. Examples from inventory control, production planning and investment problems are presented.

This thesis is also intended to critically compare the discrete version of Pontryagin's principle with other traditional optimization

techniques and to present an alternate derivation of the algorithm of the discrete maximum principle.

The dissertation is composed of two parts. Part I introduces briefly the basic theory of Pontryagin's maximum principle for time-optimizing continuous processes. It contains the algorithm, the derivation of the algorithm and the application to the inventory control problem. Part II discusses the discrete version of the maximum principle. It presents the statement of the algorithm, its derivation, the applications to the production planning and investment problems, and the analysis of the algorithm. Concluding remarks are presented in the last section of Part II.

A tabulating programming technique based on the maximum principle has been developed for industrial and business application. Its procedure and an example are also included in Part II.

The Maximum Principle and its Application to Industrial Engineering Problems

by

Kenji Kanbe

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

June 1968

APPROVED:

Redacted for Privacy

Associate Professor of Mechanical and Industrial Engineering

In charge of Major

Redacted for Privacy

Head of Department of Mechanical and Industrial Engineering

Redacted for Privacy

Dean of Graduate School

Date thesis is presented October 12, 1967

Typed by Clover Redfern for Kenji Kanbe

ACKNOWLEDGMENTS

I am greatly indebted to Dr. Michael S. Inoue for his unlimited academic assistance and personal encouragement. The innumerable hours spent with him in discussion and consultation have provided inspiration, not only for this thesis but for my graduate work.

Professor James L. Riggs' advice and recommendation presented at the early stage of my academic work enabled me to choose this project as a graduation thesis.

Dr. Donald A. Pierce of the Statistics Department provided me the opportunity to take the reading and conference on operations research and to initiate me into the study of optimization techniques. To Dr. George L. Nemhauser, I am grateful for his suggestion and recommendation for various materials in this research.

Other members of the Department of Mechanical and Industrial Engineering, especially Professor William F. Engesser, Mr. Wu Y. Chiu should be acknowledged.

Mrs. Clover Redfern did the greater part of typing on the final version of this thesis.

My greatest debt, however, is to my American family, the Lambeths. Miss Jo Anne Lambeth with her exceptional talent in English has undertaken willingly the tedious task of proofreading the manuscript. Mr. and Mrs. Raymond H. Lambeth's constant encouragement has been helpful not only through the completion of this ordeal but through my lifetime in the United States.

TABLE OF CONTENTS

LIST OF FIGURES

THE MAXIMUM PRINCIPLE AND ITS APPLICATION TO INDUSTRIAL ENGINEERING PROBLEMS

I. INTRODUCTION

There are many techniques of searching for the optimal value of a function. Some of the most frequently used techniques for solving problems of optimization are:

Direct method of calculation

Classical differential calculus method

Lagrange multiplier method

The calculus of variations

Experimental search method

Linear and nonlinear programming

Dynamic programming

The maximum principle.

Among the numerous attempts to find new mathematical methods, dynamic programming developed by Bellman (1957) and the maximum principle derived by a Russian mathematician, Pontryagin[1], are perhaps the most successful. The maximum principle was first proposed in 1956 by Pontryagin and his associates for individual types

[1] Pontryagin, L.S. Some mathematical problems arising in connection with the theory of optimum automatic control system. Session of the Academic of Science of the USSR on Scientific Problems of Automating Industry, October: 15-20. 1956. (In Russian)

of time-optimizing continuous processes[2].

Gamkrelidze (1958) extended the principle to a general case in which an arbitrary functional of an integral function is to be optimized. The first attempt to extend the maximum principle to the optimization of stage-wise processes was made by Rozonoer[3] for the processes with linear state variables.

Chang (1960) presented the discrete version of the maximum principle for non-linear, simple processes. An algorithm essentially identical to Chang's version, but different in notations, was independently obtained by Katz (1962). Following the procedure used by Katz in the derivation of the discrete maximum principle, Fan and Wang (1964) have extended the same algorithm with some modifications to solve optimization problems of a complex process.

---

[2]Less known perhaps is an independent development of a similar principle by M.R. Hestenes at the University of California, Los Angeles in 1958.

[3]Rozonoer, L. I. The maximum principle of L.S. Pontryagin in optimal system theory. Autmat, Telemech., Moscow, 20: 1320, 1441, 1561. 1960.

II. CONTINUOUS MAXIMUM PRINCIPLE

In this section, only the basic algorithm of the original version of Pontryagin's maximum principle is introduced. Its derivation is described from a dynamic programming point of view, and the continuous maximum principle is applied to an inventory control problem.

A. Continuous Processes

The maximum principle was originally presented by Pontryagin as a set of necessary conditions for the optimization of continuous simple processes. A simple process is defined as a process dynamically following a simple path as illustrated in Figure 1. Many other types of processes are found in actual industrial situations. Those are composed of several inter-connected branches, and are called continuous complex processes (see Figure 2). However, such a complex process can usually be solved by decomposing it into several simple processes and by exercising additonal special care in handling the state vector variables at the junction point of the sub-processes. This work extending simple process concept to complex processes has recently been published by Fan and his associates (1966). Only simple processes will be described here.

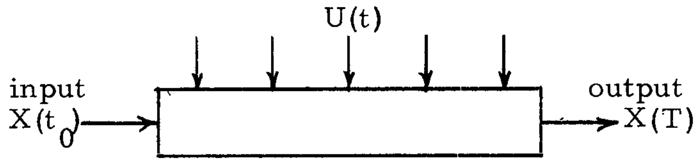


Figure 1. Simple process.

B. Algorithm for Simple Processes

A process refers to the dynamical change of a system. A process can be classified either as a time or a space process, depending on whether the change of the state is a function of a time or a space variable. Since these two types of processes are mathematically similar, the variable t will be used to denote either its time or space variable in the following discussion.

For a deterministic simple process, the state of a system at a certain time or position, t, is completely described by the state vector, $X(t)$. The change of the state is a result of the action of the decision vector, $U(t)$, which can be manipulated independently. In a simple continuous process (see Figure 1), the change of the state can be described by the following differential equations:

$$\frac{dx_i}{dt} = f_i[x_1(t), x_2(t), \ldots, x_s(t); \quad u_1(t), \ldots, u_r(t)] \tag{1a}$$

or

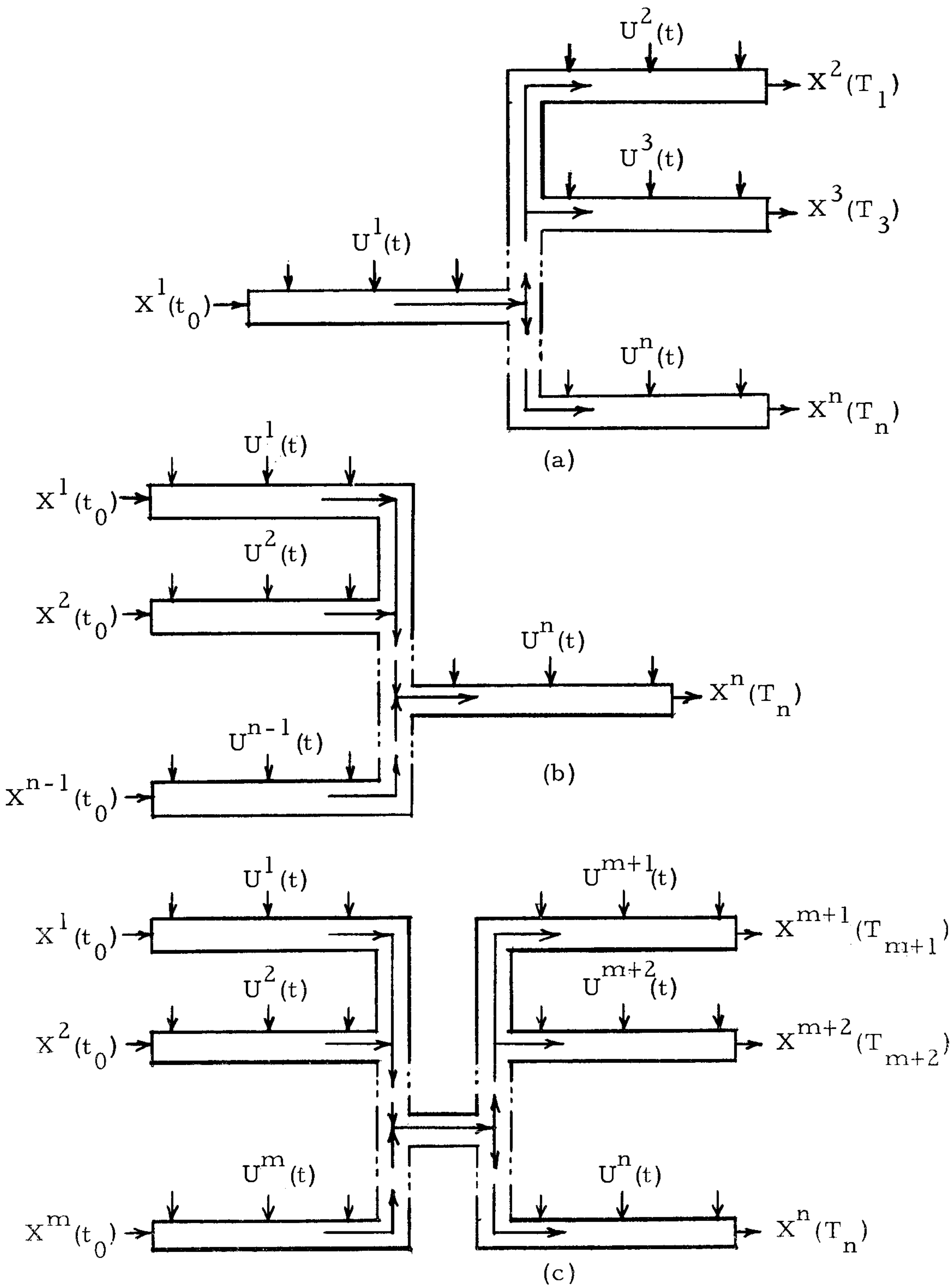


Figure 2. Complex processes with: (a) separating point, (b) combining point, (c) crossing point.

$$x_i(t) = \int_{t_0}^{t} f_i[x_1(t), \ldots, x_s(t);\ u_1(t), \ldots, u_r(t)]dt \tag{1b}$$

$$i = 1, 2, \ldots, s. \qquad t_0 \leqq t \leqq T$$

In its vector form, we may write (1a) as

$$\frac{dX(t)}{dt} = f(X(t),\ U(t)) \tag{1c}$$

s and r are the dimensions of X and U respectively. The length of the process is denoted by $(T-t_0)$ which is the distance between two end-points, t_0 and T, in the t-coordinate.

A basic optimization problem associated with such a process is to choose a piecewise continuous decision vector function, $U(t)$, subject to the constraints

$$g_i[u_1(t), \ldots,\ u_r(t)] \leqq 0 \qquad i = 1, 2, \ldots, p \tag{2}$$

so as to maximize (or minimize) a linear function of the final values of the state

$$J = \sum_{i=1}^{s} c_i x_i(T); \qquad c_i = \text{constant} \tag{3}$$

with the given initial condition $X(t_0) = K$.

The decision vector function so chosen is called an optimal decision function and denoted by $U^*(t)$.

Additional state variables may be added to Equations (1) in order to include non-linear and integral relationships in the return (or payoff) function J. For example, if the integral-square value of one of the state variables, x_m is to be maximized, then a new variable may be defined as

$$x_{s+1}(t) = \int_{t_0}^{t} x_m^2 dt$$

or,

$$\frac{dx_{s+1}}{dt} = x_m^2 ; \qquad x_{s+1}(t_0) = 0 .$$

The return function J under these conditions is

$$J = x_{s+1}(T).$$

We will assume in the development which follows that Equations (1) include any additional state variables required to specify the return function.

There are three different types of problems:

1. Fixed time problems with free right end;

   no given final conditions,

   known time interval.

2. Fixed time problems with fixed right end;

   given final conditions,

   known time interval.

3. Final time open problems;

unspecified final time.

Only the fixed time problem with free right end will be presented in this thesis. The other two types of problems and their conditions are discussed by Pontryagin et al. (1962), Kopp (1962), and Fan et al. (1966).

To solve this problem, we shall introduce an s-dimensional covariant vector function $Z(t)$ and a Hamiltonian function H satisfying the following recurrence relationships:

$$H[Z(t), X(t), U(t)] = \sum_{i=1}^{s} z_i f_i [X(t), U(t)] \tag{4}$$

$$\frac{dz_i(t)}{dt} = -\frac{\partial H(Z, X, U)}{\partial x_i(t)} \; ; \quad i = 1, 2, \ldots, s \tag{5}$$

$$z_i(T) = c_i \qquad i = 1, 2, \ldots, s \tag{6}$$

It can be seen that once the decision vector function $U(t)$ is chosen, the covariant vector $Z(t)$ is uniquely determined by Equations (5) and (6) and the initial condition $X(t_0) = K$. It may also be noted that the performance Equations (1) may be rewritten in terms of the Hamiltonian function as

$$\frac{dx_i(t)}{dt} = \frac{\partial H(Z, X, U)}{\partial z_i(t)} \; ; \quad i = 1, 2, \ldots, s \tag{7}$$

The optimal decision vector function $U^*(t)$ which makes the objective function J a maximum (or minimum), is the decision vector function $U(t)$ which renders the Hamiltonian function H a maximum for every t, $t_0 \leqq t \leqq T$. If the optimal decision vector is interior to the constraint set given by Equation (2), a necessary condition for J to be a maximum with respect to $U(t)$ is

$$\frac{\partial H}{\partial U} = 0 \tag{8}$$

If $U(t)$ is restricted, the optimal decision vector function $U^*(t)$ is determined either by solving Equation (8) for $U(t)$ or by seeking the boundary of the constraint set.

Thus, Pontryagin maximum principle can be summarized in the following theorem:

<u>Theorem</u>: Let $U(t)$, $t_0 \leqq t \leqq T$ be a piecewise continuous vector function satisfying the constraints given Equation (2). In order that the linear function J, (Equation 3), may be maximum (or minimum) for a process described by Equation (1) or (7), with initial condition $X(t_0) = K$ given, it is necessary that there exists a non-zero continuous vector function $Z(t)$ satisfying Equation (5) and (6), and that the decision vector function $U(t)$ be so chosen that

$$H[Z(t), X(t), U(t)] = \text{Maximum (or minimum)}$$

for every t, $t_0 \leq t \leq T$,

where the equations are identified by:

$$\frac{dX(t)}{dt} = f[X(t),\ U(t)] \tag{1}$$

$$g_i[u_1(t), \ldots, u_r(t)] \leq 0; \quad i = 1, 2, \ldots, p \tag{2}$$

$$J = \sum_{i=1}^{s} c_i x_i(T); \quad c_i = \text{constant} \tag{3}$$

$$H(Z, X, U) = \sum_{i=1}^{s} z_i f_i[X(t),\ U(t)] \tag{4}$$

$$\frac{dz_i(t)}{dt} = - \frac{\partial H(Z, X, U)}{\partial x_i(t)}; \quad i = 1, 2, \ldots, s \tag{5}$$

$$z_i(T) = c_i; \quad i = 1, 2, \ldots, s \tag{6}$$

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial z_i}; \quad i = 1, 2, \ldots, s \tag{7}$$

C. Derivation of the Algorithm

Originally, the continuous maximum principle was developed by Pontryagin, but the derivation by Pontryagin is unnecessarily cumbersome. Alternate derivations of the maximum principle have recently been presented by Kopp (1962) and Nemhauser (1966). In this thesis, a much simpler alternate derivation from a dynamic programming point of view will be presented.

In his original derivation, Pontryagin (1962) started with a minimization problem, demonstrating that the Hamiltonian must be maximized, named his finding "Maximum Principle." Kopp and Nemhauser, on the other hand, used a maximizing problem and concluded that the Hamiltonian should be minimized.

Our approach in this chapter, will be to examine a maximizing problem and to yield a maximizing Hamiltonian as the result. This technique was made possible by our judicial choice of multipliers in the definition of the Hamiltonian function.

We wish to find the decision vector function, $U^*(t)$ that maximizes the objective function, J, Equation (3):

$$\text{Maximum} \qquad \sum_{i=1}^{s} c_i x_i(T) = J$$

$$\text{subject to} \qquad \frac{dx_i}{dt} = f_i[X(t),\ U(t)] \quad t_0 \leqq t \leqq T$$

$$x_i(t_0) = k_i \qquad i = 1, 2, \ldots, s$$

Under the assumption that Equations (1) include any state variables required to specify the return (or payoff) function, J, this problem has already been transformed into a terminal optimization problem (i.e. a fixed final time problem).

The otpimum return function denoted by F is implicitly a function of initial state vector $X(t_0)$ and t_0. Then we can define

$$F[t_0, X(t_0)] = \max_{\substack{U(t)\\ [t_0,T]}} \sum_{i=1}^{s} c_i x_i(T).$$

Since initial time point, t_0, could occur at any time t, $t_0 \leqq t \leqq T$, the subscript of t_0 in the above equation can be dropped. Hence,

$$F[t, x(t)] = \max_{\substack{U(t)\\ [t,T]}} \sum_{i=1}^{s} c_i x_i(T) \tag{9a}$$

Let us now apply the principle of optimality[4] (Bellman, 1957). We can first rewrite Equaton (9a) as:

$$F[t, X(t)] = \max_{\substack{U(t)\\ [t,t+\Delta]}} \max_{\substack{U(t)\\ [t+\Delta,T]}} \sum_{i=1}^{s} c_i x_i(T) \tag{9b}$$

Noting that

$$F[t+\Delta, X(t+\Delta)] = \max_{\substack{U(t)\\ [t+\Delta,T]}} \sum_{i=1}^{s} c_i x_i(T),$$

---

[4] An optimal set of decisions has the property that whatever the first decision is, the remaining must be optimal with respect to the outcome which results from the first decision.

We obtain the functional equation

$$F[t,\ X(t)] = \max_{\substack{U(t)\\ [t,\, t+\Delta]}} \ F[t+\Delta,\ X(t+\Delta)] \tag{10}$$

Expanding $F[t+\Delta,\ X(t+\Delta)]$ in a Taylor series about $F(t,\ X(t))$ and neglecting second-order and higher terms yields

$$F[t+\Delta,\ X(t+\Delta)] = F[t,\ X(t)] + \Delta(F_t + \sum_{i=1}^{s} F_{x_i} f_i) \tag{11}$$

By substituting Equation (11) into Equation (10),

$$F[t,\ X(t)] = \max_{\substack{U(t)\\ [t,\, t+\Delta]}} \ \{F[t,\ X(t)] + \Delta(F_t + \sum_{i=1}^{s} F_{x_i} f_i)\} \tag{12a}$$

where,

$$F_t = \frac{\partial F}{\partial t} \quad \text{and} \quad F_{x_i} = \frac{\partial F}{\partial x_i}.$$

Subtracting F from both sides of Equation (12a) and then dividing by Δ yields in the limit as $\Delta \rightarrow 0$,

$$0 = \max_{U(t)} (F_t + \sum_{i=1}^{s} F_{x_i} f_i). \tag{12b}$$

Since F_t is not a function of U, the objective is to maximize

$$\sum_{i=1}^{s} F_{x_i} f_i .$$

We now define the covariant vector function as auxiliary variables which may be recognized as time dependent Lagrange multipliers:

$$z_i(t) = F_{x_i}(t)$$

From Equation (9a), $F_{x_i}(T) = c_i$, and the end condition becomes

$$z_i(T) = c_i .$$

Letting the Hamiltonian function be $\sum_{i=1}^{s} z_i(t) f_i$, that is,

$$H[Z(t),\ X(t),\ U(t)] = \sum_{i=1}^{s} z_i(t) f_i ,$$

then, the objective can be restated as

$$\text{Max. } H(Z, X, U).$$

Let f_i^* be the values of f_i evaluated at the optimum value of u_i.

From Equation (12b),

$$0 = F_t + \sum_{i=1}^{s} z_i f_i^*$$

$$F_t = - \sum_{i=1}^{s} z_i f_i^*$$

Partially differentiating the above equation with respect to x_i, we obtain

$$F_{tx_i} = F_{x_i t} = \frac{dz_i}{dt} = - \sum_{j=1}^{s} \frac{\partial z_j f_j^*}{\partial x_i}; \quad i = 1, 2, \ldots, s.$$

Summarizing the above results,

$$\text{Max. } H(Z, X, U) = \sum_{i=1}^{s} z_i f_i [X(t), U(t)]$$

subject to

$$\frac{dz_i}{dt} = - \frac{\partial H(Z, X, U)}{\partial x_i}$$

$$z_i(T) = c_i \qquad i = 1, 2, \ldots, s$$

$$\frac{dx_i}{dt} = f_i [X(t), U(t)]. \qquad t_0 \leq t \leq T$$

$$x_i(t_0) = k_i$$

These are the same conditions given by the maximum principle, and the derivation is therefore completed.

D. Application to Inventory Problem

An application of the continuous maximum principle to a problem of optimum production or purchasing planning under certainty will now be illustrated. The objective is to minimize the sum of the cost of holding inventories and stockout, and the cost of manufacturing or ordering. Pontryagin's maximum principle is employed to obtain an optimal production or procurement policy and the corresponding trajectory of the inventory for the manufacturing warehouse.

1. Description of the Problem

The total cost to be minimized is approximated by a quadratic form. The optimal control (decision) action and the corresponding trajectory of the state variables are respectively the optimal production plan and its inventory.

In a manufacturing company, forecasting is used in designing production rules which anticipate and prepare for sales fluctuations, and a buffer inventory is maintained so that errors in sales forecasts will not cause runouts or will not force rapid changes in the rate of plant operation.

Let $I(t)$, $P(t)$ and $S(t)$ be defined respectively as the inventory level, production rate and sales rate at time t. The functions I and P satisfy, in general, the condition that the rate of change in finished-good inventory is equal to the difference between the production and sales rates, that is,

$$\dot{I}(t) = P(t) - S(t) \tag{13}$$

where

$$\dot{I}(t) = \frac{dI(t)}{dt} .$$

Although the dynamic characteristics of an inventory system depend upon the relation between sales forecasts and actual sales, we assume in the present example that the sales are known with certainty, i. e. $S(t)$ is a known prescribed function of time (or constant).

The total cost of operation under these conditions may be composed of two parts, inventory costs and manufacturing cost. They are derived from three factors: holding inventories, stockouts, and the derivation of the production rate from that which is considered optimal for the plant. Let us define that $\bar{I}$ and $\bar{P}$ present the desired inventory and the desired production level of the plant, and that C_I and C_P are cost coefficients. The rate at which the holding cost or stockout cost is incurred at time t can be approximated by a quadratic function $C_I[I(t)-\bar{I}]^2$. The rate at which the manufacturing cost is incurred can be approximated by another quadratic function $C_P[P(t)-\bar{P}]^2$ so that the marginal cost function is linear and the unit cost function is U-shaped.[5] Sometimes $\bar{I}$ and $\bar{P}$ can be known

[5]Assuming, of course, that $C_I \geqq 0$; $C_P \geqq 0$.

functions of time t, but both of them are assumed to be constant in our example.

Therefore, the total cost incurred between time t_0 and T is the integral:

$$C_T = \int_{t_0}^{T} [C_I(I(t)-\bar{I})^2 + C_P(P(t)-\bar{P})^2]dt \qquad (14)$$

Here, T represents some future point of time, and $(T-t_0)$ is not necessarily the length of a season.

Now, the problem is briefly described as follows: find the optimum production and the corresponding inventory at time t which minimizes the cost function represented by Equation (14), subject to the constraint given by Equation (13).

This inventory model was also used by Holt <u>et al</u>. (1960).

2. Solution Procedure

Let us define $x_1(t) = I(t)$, $u(t) = P(t)$. Then, Equation (13) becomes

$$\frac{dx_1}{dt} = u(t) - S(t) \qquad (15)$$

$$x_1(t_0) = k_1$$

where $S(t)$ is a known fixed function.

Introducing an additional state variable, $x_2(t)$ such that

$$x_2(t) = \int_{t_0}^{t} [C_I(x_i(t)-\overline{I})^2 + C_P(u(t)-\overline{P})^2]dt$$

or

$$\frac{dx_2}{dt} = C_I[x_i(t)-\overline{I}]^2 + C_P[u(t)-\overline{P}]^2; \quad x_2(t_0) = 0 \tag{16}$$

the problem is thus transformed into:

$$\text{Min. } J = \sum_{i=1}^{s} c_i x_i(T) = x_2(T)$$

subject to Equations (15) and (16).

From Equations (3) and (6),

$$c_1 = z_1(T) = 0 \qquad\qquad c_2 = z_2(T) = 1$$

According to Equations (4) and (5), we have

$$H(Z, X, U) = z_1(u-S) + z_2[C_I(x_1-\overline{I})^2 + C_P(u-\overline{P})^2] \tag{17}$$

$$\frac{dz_1}{dt} = -\frac{\partial H}{\partial x_1} = -2z_2C_I(x_1-\overline{I}), \quad z_1(T) = 0 \tag{18}$$

$$\frac{dz_2}{dt} = -\frac{\partial H}{\partial x_2} = 0, \quad z_2(T) = 1 \tag{19}$$

Solve Equation (19) for z_2,

$$z_2(t) = \text{constant}$$

$$\therefore\ z_2(t) = z_2(T) = 1$$

The Hamiltonian function can now be written as

$$H(Z, X, U) = z_1(u-S) + C_I(x_1-\overline{I})^2 + C_P(u-\overline{P})^2 \tag{20}$$

In this problem no constraint is imposed on the decision variables, and we do not have to be concerned with the boundary of the constraint set. The optimal control action may be determined from Equation (8) as

$$\frac{\partial H}{\partial u} = 0 = z_1 + 2C_P(u-\overline{P})$$

or

$$z_1(t) = -\ 2C_P[u(t)-\overline{P}] \tag{21}$$

Combining Equation (18) with (21),

$$z_1(T) = 0 = -2C_P[u(T)-\overline{P}]$$

$$u(T) = \overline{P}$$

$$\frac{dz_1}{dt} = -2C_P\frac{du}{dt} = -2C_I(x_1-\overline{I})$$

$$\therefore\ C_I(x_1-\overline{I}) - C_P\frac{du}{dt} = 0 \quad u(T) = \overline{P} \tag{22}$$

Equation (22) together with Equation (15) constitutes a pair of linear differential equations in the two unknown functions $x_1(t)$ and $u(t)$. The solution of these two differential equations will yield the optimal functions of x_1 and u.

Solution of Equations (15) and (22):

We set a pair of differential equations as follows,

$$\frac{dx_1}{dt} - u(t) = -S(t) \tag{23}$$

$$C_I x_1 - C_P \frac{du}{dt} = C_I \bar{I} \tag{24}$$

differentiating Equation (23) with respect to t and (23) becomes

$$\frac{d^2x_1}{dt^2} C_P - \frac{du}{dt} C_P = -\frac{dS}{dt} C_P \tag{25}$$

and taking the difference of Equation (24) from Equation (25),

$$(C_P \frac{d^2}{dt^2} - C_I)x_1 = -C_P \frac{dS}{dt} - C_I \bar{I} \tag{26}$$

The characteristic equation of Equation (26) is

$$C_P m^2 - C_I = 0$$

$$\therefore m = \pm \sqrt{\frac{C_I}{C_P}} = \pm \lambda \ .$$

The solution of the pair of linear differential equations (23) and (24) is, therefore,

$$x_1 = A_1 e^{\lambda t} + A_2 e^{-\lambda t} + (x_1)_p$$

where $(x_1)_p$ is the particular solution of the equation to be decided by the form of the function $S(t)$ and by the value of $\bar{I}$. The substitution of Equation (27) into Equation (23) yields

$$\begin{aligned} u(t) &= \frac{dx_1}{dt} + S(t) \\ &= \frac{d}{dt}[A_1 e^{\lambda t} + A_2 e^{-\lambda t} + (x_1)_p] + S(t) \\ &= A_1 \lambda e^{\lambda t} - A_2 \lambda e^{-\lambda t} + \frac{d(x_1)_p}{dt} + S(t) \end{aligned}$$

where $\lambda = \sqrt{\frac{C_I}{C_P}}$, and A_1, A_2 are obtained from the initial conditions $x_1(t_0) = k_1$ and $x_2(t_0) = 0$.

III. DISCRETE MAXIMUM PRINCIPLE

The discrete version of the maximum principle is more useful for industrial and management systems than the continuous version of the maximum principle.

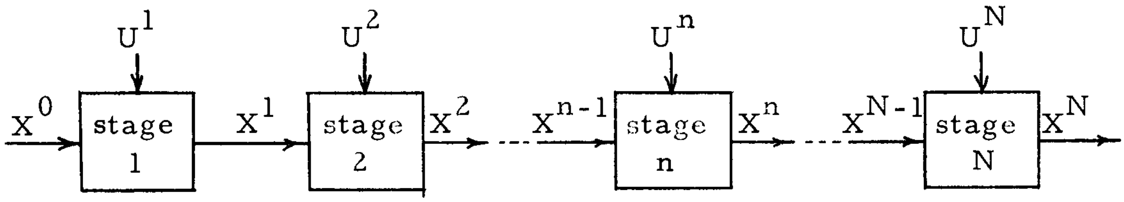
In Section B, the basic algorithm of the discrete maximum principle and its derivation will be presented. In Section C, applications to three problems in industry and business are illustrated: (1) production level planning, (2) capital allocation and (3) investment problems ((2) and (3) are the same things in the sense of investment, but different types of systems). In the section of the applications, a table is presented which contains the results optimized by the technique of the maximum principle for a particular system. Computational procedures to solve the problems are also explained in the application.

In Section D, the analysis of the principle algorithm will be presented by using other optimization techniques, and some criticisms of the discrete maximum principle are provided.

A. Discrete Processes

The discrete maximum principle is simply an extension of Pontragin's continuous maximum principle. This extension work has recently been achieved independently by Change (1960) and Katz (1962).

Their discrete version of the maximum principle is only for a simple process in which the output from one stage is the input to the next stage, as shown in Figure 3. Although other systems which may be more highly complex than the simple processes are employed in actual industry and business, these complex processes (see Figure 4) can be handled by decomposing a complex process into several simple processes with some appropriate modifications. We will, therefore, introduce the algorithm of the discrete maximum principle for the simple process. Wang and Fan (1964a) extended Katz's algorithm of the discrete maximum principle, for complex processes.



X: state vector variable
U: decision vector variable

Figure 3. Simple process.

The discrete processes are defined as follows: A stage-wise process is a process consisting of a finite number of interconnected stages. A stage may present any real or abstract entity (for example, a space unit, or a time period or an economic activity) in which a certain transformation takes place. Those variables which are transformed in each stage are called state variables. The desired transformation for the state variables is achieved through

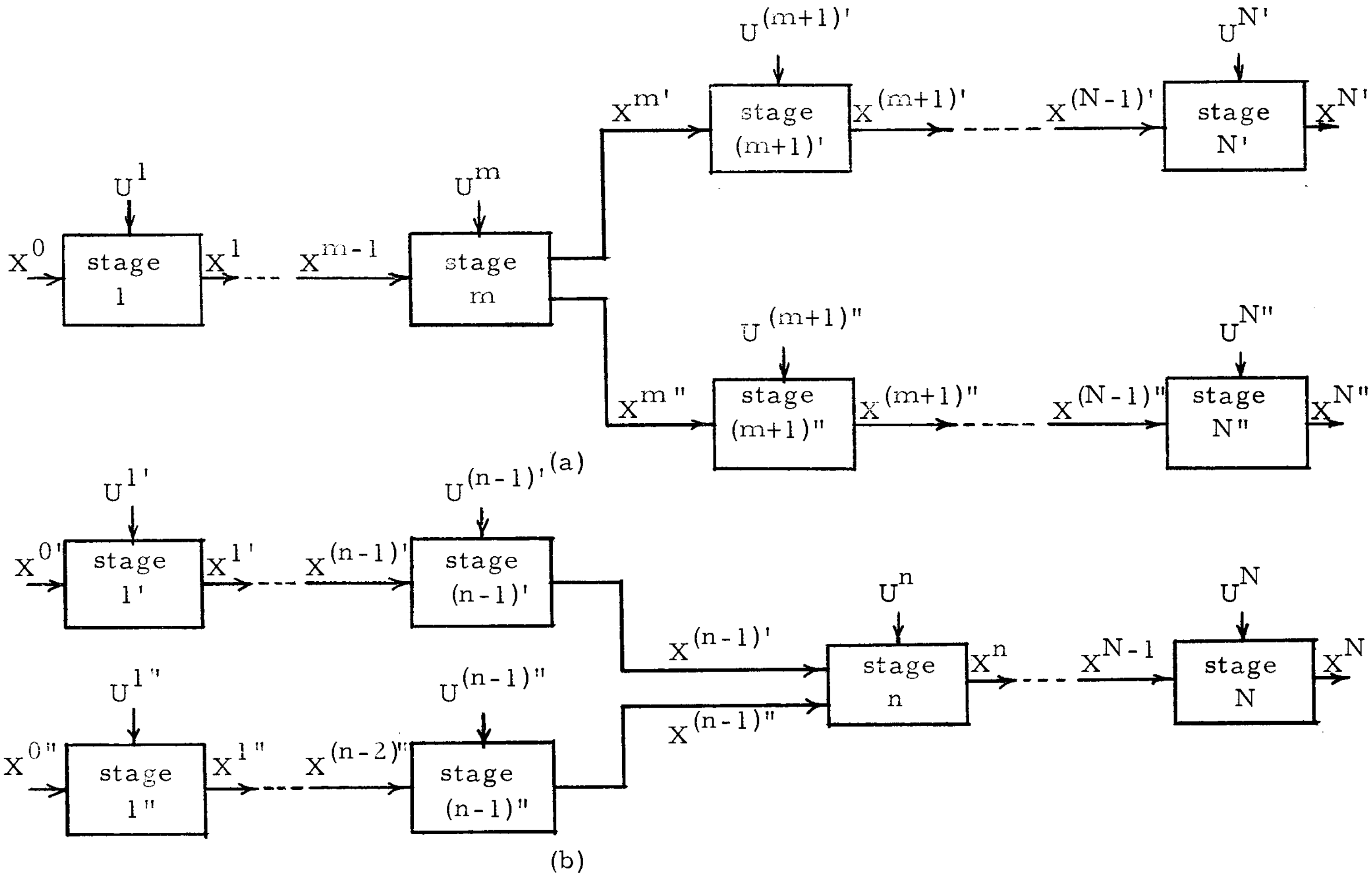


Figure 4. Complex processes: (a) process with separating branch, (b) process with combining branch.

manipulation of control (or decision) variables which remain constant within each stage of the process. The transformation at each stage is completely described by a set of performance equations.

Most processes in industry consist of one or more of the following three basic types of stages;

(a) Linking stage

(b) Separating stage

(c) Combining stage (see Figure 5).

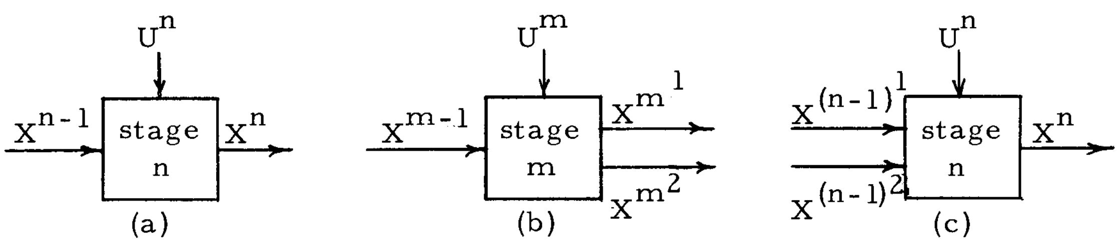


Figure 5. Three basic types of stages: (a) linking stage, (b) separating stage, (c) combining stage.

Whether a process is complex or simple depends upon the types of stages which the process includes. If a process consists entirely of linking stages, it is called a simple process. A complex process is a process containing at least one separating or combining stage or a stage more complex than either of these. Figure 4 shows two such typical complex processes.

A process can be further categorized either as a homogeneous or a heterogeneous process, depending on the form of the performance equations. A homogeneous process is a process in which the state vector and the decision vector are inter-related by the same set of performance equations throughout the process. A process is called heterogeneous if it is not a homogeneous process.

For a process with all the performance equations and the initial values of state variables given, a general optimization problem is to determine the decision variables at each stage, subject to certain constraints.

B. Discrete Maximum Principle

We will now present the discrete version of Pontryagin's maximum principle derived by Katz (1962) for simple homogeneous processes (see Figure 3).

1. Statement of the Algorithm

A N-multistage dynamic process is considered as an abstract notion by which a large number of human activities can be presented. A schematical representation of the simple multistage control process has been already illustrated in Figure 1. The process consists of N linking stages connected in series. The state of the process stream denoted by the s-dimensional vector, $X = (x_1, x_2, \ldots, x_s)$, is transformed by $T(\cdot)$ at each stage according to an r-dimensional decision vector, $U = (u_1, u_2, \ldots, u_r)$, which represents the decisions made at that stage.

Suppose the transformation of the state variables at the n^{th} stage is described by a set of difference equations (discrete performance equations),

$$x_i^n = T_i^n(x_1^{n-1}, x_2^{n-1}, \ldots, x_s^{n-1}; u_1^n, u_2^n, \ldots, u_r^n)$$

$$x_i^0 = k_i; \quad i = 1, 2, \ldots, s \quad n = 1, 2, \ldots, N,$$

or in vector form,

$$X^n = T^n(X^{n-1}, U^n)$$

$$X^0 = K \qquad (28)^6$$

An optimization problem associated with such a process is to find a sequence of control actions, U^n, $n = 1, \ldots, N$ which makes one of the final state variables, x_m^N, a maximum (or a minimum) when the initial condition $X^0 = K$ is given.

Therefore, the function x_m^N is the objective function of the process. The procedure for solving such an optimization problem by the discrete maximum principle calls for the introduction of s new variables (s-dimensional vector), $z_1^n, z_2^n, \ldots, z_s^n$, satisfying the following recurrence relations:

---

[6] The superscript n indicates the stage number. The exponents are written with parentheses or brackets such as $(X^n)^2$ or $[T^n(X^{n-1}, U^n)]^2$.

$$z_i^{n-1} = \sum_{j=1}^{s} \frac{\partial T_j^n(X^{n-1}, U^n)}{\partial x^{n-1}} z_j^n \; ; \quad i = 1, 2, \ldots, s \qquad (29)$$
$$n = 2, \ldots, N$$

and the final conditions,

$$z_i^N = \begin{cases} 1; \; i = m; \\ 0; \; i \neq m; \end{cases} \qquad i = 1, 2, \ldots, s \qquad (30)$$

To determine the optimal sequence of the decision, $\overline{U}^n$, the following condition must be satisfied;

$$\sum_{j=1}^{s} z_j^n T_j^n(X^{n-1}, U^n) = \text{min.} \; ; \quad n = 1, 2, \ldots, N \qquad (31)$$

The Equations (28) to (31) thus appear as a two-point boundary value problem in the variables, X^n and Z^n, whose solution carries with it the determination of best control action.

For the more general case, the problem can be stated as follows,

$$\text{Max. } J = \sum_{i=1}^{s} c_i x_i^N \qquad (32)$$

subject to the constraint,

$$g_i^n(u_1^n, \ldots, u_r^n) \leqq 0; \quad n = 1, \ldots, N \qquad (33)$$
$$i = 1, \ldots, p$$

Equations (28) to (31) which derive the solution of this problem can be compactly written in terms of Pontryagin's Hamiltonian formalism.

Let the Hamiltonian be

$$H^n = \sum_{j=1}^{s} z_j^n T_j^n (X^{n-1}, U^n); \quad n = 1, \ldots, N \tag{34}$$

Equations (28) to (31) may be written as

$$x_i^n = \frac{\partial H^n}{\partial z_i^n} = T_i^n (X^{n-1}, U^n); \quad x_i^0 = k_i \quad \begin{array}{l} n = 1, \ldots, N \\ i = 1, \ldots, s \end{array} \tag{35}$$

$$z_i^{n-1} = \frac{\partial H^n}{\partial x_i^{n-1}}; \quad z_i^N = c_i \quad \begin{array}{l} n = 2, \ldots, N \\ i = 1, \ldots, s \end{array} \tag{36}$$

$$H^n = \max \text{ (or min), or } \frac{\partial H^n}{\partial U^n} = 0 \quad n = 1, \ldots, N \tag{37}$$

If the optimal decision, $\overline{U}^n$, is interior to the set of the constraints, a necessary condition for J to be maximized is $\frac{\partial H^n}{\partial U^n} = 0$. If $\overline{U}^n$ is at a boundary of the constraint set, it can be determined from the condition that H^n = maximum (or minimum).

2. Derivation of the Algorithm

A simplified derivation of the algorithm stated for maximizing the objective function in the simple multistage process will be presented in this section. The algorithm for minimizing the objective function can be derived by reversing the direction of the inequality signs.

To derive the optimization algorithm, we first assume that the difference equation, $T^n(X^{n-1}, U^n)$, is continuous in its arguments and that its first partial derivatives exist and are piecewise continuous in its arguments. Furthermore, we assume that a set of optimal decisions denoted by $\overline{U}^n$ exists and can be found such that the objective function J attains its maximum. Then the corresponding optimal state vector is

$$\overline{X}^n = T^n(\overline{X}^{n-1}, \overline{U}^n); \quad n = 1, \ldots, N \tag{38}$$

If the decision vector is perturbed arbitrarily but slightly from the optimal value at each stage of the process, that is,

$$|\delta U^n| = |U^n - \overline{U}^n| < \epsilon\phi^n; \quad n = 1, \ldots, N \tag{39}$$

where ϕ^n is an r-dimensional vector.

The resulting perturbation of the state vector function is

$$\delta X^n = (\partial T^n / \partial X^{n-1})\delta X^{n-1} + (\partial T^n / \partial U^n)\delta U^n \tag{40}$$

$$\delta X^0 = 0 \tag{41}$$

where $\delta X^n = X^n - \overline{X}^n$, $n = 1, \ldots, N$ and ϵ is a small positive number.

Because of the continuity assumption, all neglected terms are of the order $(\epsilon)^2$ or higher, and the signs of both sides of Equation (40) are the same if ϵ is sufficiently small. The partial derivatives are evaluated at the upper barred quantities.

Multiplication of Equation (40) by Z^n, followed by a summation of n from $n = 1$ to $n = N$, yields the following results by virtue of Equation (36),

$$\sum_{n=1}^{N} (Z^n)^T \delta X^n = \sum_{n=1}^{N} (Z^n)^T (\partial T^n / \partial X^{n-1}) \delta X^{n-1} + \sum_{n=1}^{N} (Z^n)^T (\partial T^n / \partial U^n) \delta U^n$$

$$= \sum_{n=1}^{N} (Z^{n-1})^T \delta X^{n-1} + \sum_{n=1}^{N} (Z^n)^T (\partial T^n / \partial U^n) \delta U^n \qquad (42)^7$$

By making use of Equations (36) and (41), Equation (42) can be simplified to

$$(C)^T \delta X^N = \sum_{n=1}^{N} (Z^n)^T (\partial T^n / \partial U^n) \delta U^n$$

$$= \sum_{n=1}^{n} \sum_{j=1}^{s} \sum_{i=1}^{r} z_j^n (\partial T_j^n / \partial u_i^n) \delta u_i^n \qquad (43)$$

The variation of the objective function J from Equation (32) is

---

[7] The superscript T indicates a Transpose.

$$\delta J = (C)^T \delta X^N = \sum_{i=1}^{s} c_i \delta x_i^N \tag{44}$$

Since J is to be maximized, it is necessary that δJ be zero for all free variations δU^n, and that δJ be negative for all one-sided variations when the optimal decisions are at the boundaries of constraints as expressed by Equation (33).

Hence, it is necessary from Equations (43) and (44) that

$$\delta J = (C)^T \delta X^N \leqq 0$$

$$\sum_{j=1}^{s} \sum_{i=1}^{r} z_j^n (\partial T_j^n / \partial u_i^n) \delta u_i^n \leqq 0 \tag{45}$$

From the definition of the Hamiltonian function given by Equation (34), the necessary condition given by (45) can be written as follows

$$\sum_{j=1}^{s} z_j^n (\partial T_j^n / \partial u_i^n) \delta u_i^n \leqq 0 \qquad \begin{array}{l} i = 1, 2, \ldots, r \\ n = 1, 2, \ldots, N \end{array}$$

or in short

$$(\partial H^n / \partial u_i^n) \delta u_i^n \leqq 0 \tag{46}$$

because of the independence of the variation δU^n. However, for free variations of U^n a necessary and sufficient condition for J

to be zero is given by

$$\partial H^n / \partial U^n = 0 \tag{47}$$

Since $\delta J = 0$ is only a necessary condition for maximizing J, Equation (47) is a necessary but not a sufficient condition for maximizing J for free variations of U^n.

When any of the U^n, $n = 1, 2, \ldots, N$, lies at a boundary, it is necessary and sufficient that we have the following condition for all allowable variations in order to make $\delta J < 0$,

$$(\partial H^n / \partial U^n) \delta U^n \leqq 0.$$

This is equivalent to the condition that H^n be (locally) maximum at the boundary.

Thus, we are led to the conclusion that in order for the objective function J to attain its maximum (or minimum) value it is necessary to choose a set of decisions U^n, $n = 1, \ldots, N$ such that

a) Hamiltonian H^n is made stationary with respect to the optimal decision $\overline{U}^n$, when it is not constrained (or it lies in the interior of the admissible domain of U^n), that is, choose $\overline{U}^n$ such that

$$\partial H^n / \partial U^n = \sum_{j=1}^{s} z_j^n (\partial T_j^n / \partial U_i^n) = 0 \qquad \begin{array}{l} n = 1, \ldots, N \\ i = 1, \ldots, r \end{array}$$

b) Hamiltonian H^n is made an extremum with respect to $\overline{U}^n$ when it lies on a boundary of constraints, that is, choose $\overline{U}^n$ such that H^n = maximum (or minimum).

C. Application of Discrete Maximum Principle

In the preceding section, a simple multistage process with fixed terminal point and with given initial state values has been discussed. However, this basic algorithm presented can be extended to handle a variety of problems usually encountered in practice, for example, processes with fixed end point in state values, processes with choice of initial values, processes with different performance equations at each stage, processes with feed-back and so on.

Therefore, the discrete maximum principle presented in the previous section can now be applied to some more realistic problems of industry and business in this section. Note that we are still considering only the simple multistage process with given initial conditions.

The problems which will be presented later can be classified by constraints depending on whether some constraints are imposed or not. They can also be classified by the Hamiltonian function depending upon either where the stationary point of the Hamiltonian is located with respect to the constraints or whether the Hamiltonian function is linear or non-linear with respect to the decision variables. Thus, the classification of example problems may be summarized as follows:

Case 1. There exists no constraint on decision variables (or the optimal decisions are found to be interior to the set of constraints if some constraints are imposed), and the Hamiltonian function is non-linear.

Case 2. The optimal decisions are found to be at the lower bound or the upper bound of the constraints when the stationary point of the Hamiltonian function lies outside the constraints, and the Hamiltonian is non-linear.

Case 3. The optimal decisions are found to be at the boundary of the constraints when the Hamiltonian function is linear with respect to the decision variables.

Case 4. The Hamiltonian function is non-linear and its stationary point may be infinite with respect to the decision variables, and each stage has a different transformation form.

This classification is also illustrated in Figure 6. Each problem in the discrete maximum principle application will be identified by the above case number.

			OPTIMAL DECISION		
			Boundary	Interior	No constraint
HAMILTONIAN FUNCTION	HF linear		Case 3		
	HF non-linear	HF stationary point is interior		Case 1	Case 1
		HF stationary point is exterior	Case 2, 4		

Figure 6. Classification of problems.

1. Application to Production Level Optimization

a. Case 1 (Optimal Decisions are Interior to the Set of Constraints).

Let us consider the production level of a particular perishable commodity. The following information is given:

- The excess production over the sales forecast is wasted at a cost of $10 per unit.
- The cost of changing the level is three times the square of the difference between two production levels.

- The last quarter production level was $136\frac{2}{3}$ units.
- The following sales forecast must be met (no shortage is allowed):

Quarter (n)	1	2	3	4
Sales forecast (Q^n)	115	125	100	95 units

The problem is to determine the production level at each period which minimizes the total cost subject to the sales forecast for all periods.

We can define each quarter as a stage as shown in Figure 7.

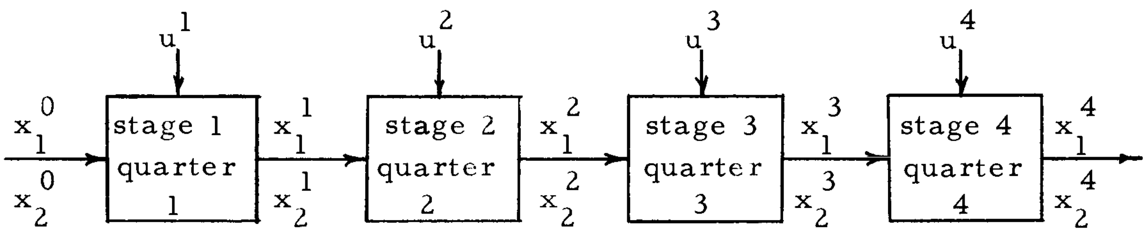


Figure 7. Production planning process.

Let

x_1^n = Production level at the n^{th} stage,

x_2^n = The sum of the cost up to and including the n^{th} stage,

u^n = The change in production level from the $(n-1)^{th}$ stage to the n^{th} stage,

Q^n = The sales forecast for the n^{th} stage,

$G^n(x_1^{n-1}, u^n)$ = The cost at the n^{th} stage.

According to Equation (35),

$$x_1^n = T_1^n(x_1^{n-1}, u^n) = x_1^{n-1} + u^n \quad n = 1, \ldots, 4 \tag{48}$$

$$x_2^n = x_2^{n-1} + G^n(x_1^{n-1}, u^n)$$

$$= x_2^{n-1} + 3(u^n)^2 + 10(x_1^n - Q^n) \tag{49}$$

Substitution of Equation (48) into (49),

$$x_2^n = T_2^n(x_1^{n-1}, x_2^{n-1}, u^n) \quad n = 1, 2, 3, 4$$

$$= x_2^{n-1} + 3(u^n)^2 + 10(x_1^{n-1} + u^n - Q^n) \tag{50}$$

where

$$x_1^0 = 136\frac{2}{3}, \; x_2^0 = 0.$$

The problem in which the objective function J to be minimized can be written as

Minimum

$$J = \sum_{n=1}^{4} G^n(x_1^{n-1}, u^n) = \sum_{i=1}^{2} c_i x_i^4 = x_2^4 \tag{51}$$

subject to

$$x_1^n = x_1^{n-1} + u^n$$

$$x_2^n = G^n(x_1^{n-1}, u^n) + x_2^{n-1}$$

$$x_1^n \geqq Q^n.$$

Let us start solving this problem by using the discrete maximum principle. As the first step, we obtain from Equation (51),

$$c_1 = z_1^4 = 0 \qquad c_2 = z_2^4 = 1 \tag{52}$$

According to the recurrence Equations (34), (36), (48) and (50);

$$H^n = \sum_{j=1}^{2} z_j^n T_j^n (X^{n-1}, u^n)$$

$$= z_1^n (x_1^{n-1} + u^n) + z_2^n [x_2^{n-1} + 3(u^n)^2 + 10(x_1^{n-1} + u^n - Q^n)] \tag{53}$$

$$z_2^{n-1} = \frac{\partial H^n}{\partial x_2^{n-1}} = z_2^n \tag{54a}$$

$$z_1^{n-1} = \frac{\partial H^n}{\partial x_1^{n-1}} = z_1^n + 10 z_2^n \tag{54b}$$

Equations (54a) and (52) show that $z_2^4 = z_2^3 = z_2^2 = z_2^1 = 1$. By substituting $z_2^4 = 0$ and $z_2^4 = 1$ into Equation (54b), z_1^n can be obtained,

$$z_1^4 = 0 \quad z_1^3 = 10 \quad z_1^2 = 20 \quad z_1^1 = 30.$$

Now, we have all values of the multiplier vector Z^n. We are therefore able to determine the sequence of optimal decisions from Equations (37) and (53),

$$\frac{\partial H^1}{\partial u^1} = \frac{\partial}{\partial u^1}\{z_1^1(x_1^0+u^1)+z_2^1[x_2^0+3(u^1)^2+10(x_1^0+u^1-Q^1)]\}$$

$$= \frac{\partial}{\partial u^1}[30(x_1^0+u^1)+x_2^0+3(u^1)^2+10(x_1^0+u^1-Q^1)]$$

$$= 30 + 6u^1 + 10 = 0 \quad \therefore u^1 = -\frac{20}{3}$$

$$\frac{\partial H^2}{\partial u^2} = \frac{\partial}{\partial u^2}[20(x_1^1+u^2)+x_2^1+3(u^2)^2+10(x_1^1+u^2-Q^2)]$$

$$= 20 + 6u^2 + 10 = 0 \quad \therefore u^2 = -5$$

$$\frac{\partial H^3}{\partial u^3} = \frac{\partial}{\partial u^3}[10(x_1^2+u^3)+x_2^2+3(u^3)^2+10(x_1^2+u^3-Q^3)]$$

$$= 10 + 6u^3 + 10 = 0 \quad \therefore u^3 = -\frac{10}{3}$$

$$\frac{\partial H^4}{\partial u^4} = \frac{\partial}{\partial u^4}[0(x_1^3+u^4)+x_2^3+3(u^4)^2+10(x_1^3+u^4-Q^4)]$$

$$= 6u^4 + 10 = 0 \quad \therefore u^4 = -\frac{5}{3}.$$

Then, the production level x_1^n will be obtained by substituting these decisions into Equation (48),

$$x_1^1 = 136\frac{2}{3} - \frac{20}{3} = 130 \qquad x_1^2 = 130 - 5 = 125$$

$$x_1^3 = 125 - \frac{10}{3} = 121\frac{2}{3} \qquad x_1^4 = 121\frac{2}{3} - \frac{5}{3} = 120.$$

Fortunately, any stationary point of H^n is interior to the constraints. Hence, these results are the solution of the problem and summarized as

	Initial state	1st quarter	2nd quarter	3rd quarter	4th quarter
Sales forecast Q^n		115	125	100	95
Production level x_1^n	$136\frac{2}{3}$	130	125	$121\frac{2}{3}$	120
Change of level u^n		$-\frac{20}{3}$	-5	$-\frac{10}{3}$	$-\frac{5}{3}$

If one of the optimal decisions lies at the boundary of the constraints when the stationary point of its Hamiltonian function is outside the constraints, the solution is not so easy and much more complicated. In order to compare the discrete maximum principle with dynamic programming, this problem has also been solved by dynamic programming technique in Appendix I.

b. Case 2 (One of the Optimal Decisions Lies at the Boundary). Let us now consider the application of the maximum principle in which one of the optimal decisions is at the boundary of the set of constraints when the stationary point of the Hamiltonian function H^n lies outside the constraints.

The same example is again considered. In order for the derivation of such a problem with an optimal decision at the boundary of the constraints, let us change one of the values of the preceding sales forecast in Section 1-a, as follows

Quarter	1	2	3	4
Sales forecast	115	125	200	95 units .

The estimated sales units of the third quarter has been changed from 100 to 200 units.

If we use the same procedure in the previous section, we can obtain the same sequence of the optimal decisions U^n regardless of the change of the value in sales forecast:

Quarter	Decision u^n	State x_1^n	Sales forecast Q^n
1	-20/3	130	115
2	-5	125	125
3	-10/3	$121\frac{2}{3}$	200
4	-5/3	120	95

Looking at the 3rd stage, we note:

$$x_1^3 = 121\frac{2}{3} \quad Q^3 = 200.$$

This does not satisfy the constraint that $x_1^3 \geqq Q^3$. Obviously, the stationary point of the Hamiltonian function H^3 is outside the constraint. Hence, the procedure in Section 1-a that makes $\frac{\partial H^n}{\partial u^n}$ zero without consideration of the constraints is no longer applicable for this problem.

New solution procedure:

The problem is now solved by using the backward approach with explicit consideration of the constraints.

Step 1 (Stage 4)

Since $J = x_2^4$, $z_1^4 = 0$ and $z_2^4 = 1$.

$$H^4 = z_2^4 x_2^4 = x_2^3 + 3(u^4)^2 + 10(x_1^3 + u^4 - Q^4)$$

$$\frac{\partial H^4}{\partial u^4} = 6u^4 + 10 = 0 \quad \therefore\ u^4 = -\frac{5}{3} \tag{55}$$

However, there exists a constraint at the 4^{th} stage such that

$$x_1^4 \geqq Q^4.$$

According to Equation (48),

$$u^4 = x_1^4 - x_1^3$$

$$x_1^4 = (x_1^3 + u^4) \geqq Q^4 = 95 \tag{56}$$

Substituting $u^4 = -\frac{5}{3}$ (55) into (56),

$$x_1^3 \geqq (Q^4 + \frac{5}{3}) = 96\frac{2}{3} \tag{57}$$

Equation (57) is satisfied because

$$x_1^3 \geqq 200 = Q^3.$$

Step 2 (Stage 3)

$$z_1^3 = \frac{\partial H^4}{\partial x_1^3} = 10 \qquad z_2^3 = \frac{\partial H^4}{\partial x_2^3} = 1$$

$$H^3 = 10(x_1^2+u^3) + x_2^3 + 3(u^3)^2 + 10(x_1^2+u^3-Q^3)$$

$$\frac{\partial H^3}{\partial u^3} = 10 + 6u^3 + 10 = 0$$

$$u^3 = -\frac{10}{3} \qquad \text{if } x_1^2 \geqq (Q^3+\frac{10}{3}) = 203\frac{1}{3} \tag{58}$$

$$x_1^3 = Q^3 = 200 \quad \text{if } x_1^2 < 203\frac{1}{3} \tag{59}$$

To choose either Equation (58) or (59), we must consider Q^2. Since Q^2 is equal to 125 units, x_1^2 will definitely be less than $203\frac{1}{3}$. Hence Equation (59) is chosen:

$$\therefore x_1^3 = Q^3 = 200 \tag{59}$$

Step 3 (Stage 2)

From Equation (48),

$$u^3 = x_1^3 - x_1^2 = 200 - x_1^2 \tag{60}$$

By substitution of Equation (60) into H^3,

$$H^3 = x_2^2 + 3(200-x_1^2)^2 + 2000$$

then,

$$\left.\begin{aligned} z_1^2 &= \frac{\partial H^3}{\partial x_1^2} = 6x_1^2 - 1200 \\ z_2^2 &= \frac{\partial H^3}{\partial x_2^2} = 1 \end{aligned}\right\} \quad (61)$$

The Hamiltonian function at the Stage 2 is

$$H^2 = z_1^2(x_1^1+u^2) + z_2^2[x_2^1+3(u^2)^2+10(x_1^1+u^2-Q^2)].$$

Substituting Equation (61) into H^2,

$$H^2 = (6x_1^2-1200)(x_1^1+u^2) + x_2^1 + 3(u^2)^2 + 10(x_1^1+u^2-Q^2)$$

$$\frac{\partial H^2}{\partial u^2} = 6x_1^2 - 1200 + 6u^2 + 10 = 0$$

$$u^2 = \frac{1190}{6} - x_1^2.$$

From Equation (48), we obtain $u^2 = x_1^2 - x_1^1$.

Since

$$x_1^2 - x_1^1 = \frac{1190}{6} - x_1^2,$$

$$\therefore \begin{cases} x_1^2 = \frac{1}{2}x_1^1 + \frac{595}{6} & \text{if } x_1^1 \geqq 2(Q^2-99\frac{1}{6}) = 51\frac{2}{3} \\ x_1^2 = Q^2 = 125 & \text{if } x_1^1 < 51\frac{2}{3}. \end{cases}$$

x_1^1 should be more than $51\frac{2}{3}$, because Q^1 equals to 115 units. So, we choose that

$$x_1^2 = \frac{1}{2}x_1^1 + \frac{595}{6} \tag{62}$$

Step 4 (Stage 1)

$$z_1^1 = \frac{\partial H^2}{\partial x_1^1} = 6x_1^2 - 1200 + 10$$

Substituting Equation (62) into the above equation, z_1^1,

$$z_1^1 = 3x_1^1 - 595 \quad \text{and} \quad z_2^1 = \frac{\partial H^2}{\partial x_2^1} = 1.$$

$$\therefore\ H^1 = (3x_1^1 - 595)(x_1^0 + u^1) + x_2^0 + 3(u^1)^2 + 10(x_1^0 + u^1 - Q^1)$$

where

$$x_1^0 = 136\frac{2}{3} \quad \text{and} \quad x_2^0 = 0$$

$$\frac{\partial H^1}{\partial u^1} = 3x_1^1 - 595 + 6u^1 + 10 = 0 \tag{63}$$

By the substitution of $x_1^1 = x_1^0 + u^1$ into Equation (63), we obtain

$$u^1 = 19\frac{4}{9} \tag{64}$$

which satisfies the constraint at the first stage. Through the Equations (48), (55), (59), (62) and (64), one can obtain the following results:

	Initial state	1st quarter	2nd quarter	3rd quarter	4th quarter
Sales forecast Q^n		115	125	200	95
Production level x_1^n	$136\frac{2}{3}$	$156\frac{1}{9}$	$177\frac{2}{9}$	200	$198\frac{1}{3}$
Change of level u^n		$19\frac{4}{9}$	$21\frac{1}{9}$	$22\frac{7}{9}$	$-1\frac{2}{3}$

This model of the production level optimization was first presented by Fan and Hwang (1966).

The reader can recognize that this particular procedure of the discrete maximum principle is very similar to that of dynamic programming (Appendix II).

c. Table of Optimal Decisions for Production Level Problem.

Based on the foregoing discussions, a tabulating method has been devised to minimize computational difficulties for production level optimization problems of type 1-a.

Following a simple step-by-step procedure, the table is aimed to give us the sequence of optimal decisions u^n with very little computational work.

We assume, based on the production level problem given in the Section 1-a, that the cost function in general takes the form:

the cost of changing production level;

$$A(x_1^n - x_1^{n-1})^m = A(u^n)^m$$

the cost of over-production;

$$B(x_1^n - Q^n) = B(x_1^{n-1} + u^n - Q^n)$$

where

A = the cost coefficient with respect to the change of production level,

B = the cost of producing one excess unit,

u^n = the change of production level from $(n-1)^{th}$ stage to nth stage,

x_1^n = production level at the n^{th} stage,

x_2^n = the cumulative cost up to and including n^{th} stage,

Q^n = sales forecast at n^{th} stage,

m = the exponent of the changing level cost function.

According to Equation (53),

$$H^n = z_1^n(x_1^{n-1} + u^n) + z_2^n[x_2^{n-1} + A(u^n)^m + B(x_1^{n-1} + u^n - Q^n)].$$

From Equation (36), we obtain

$$\left.\begin{aligned} z_1^{n-1} &= \frac{\partial H^n}{\partial x_1^{n-1}} = z_1^n + Bz_2^n \\ z_2^{n-1} &= \frac{\partial H^n}{\partial x_2^{n-1}} = z_2^n \end{aligned}\right\} \qquad (65)$$

From the objective function $J = x_2^N$ and Equation (36), we obtain

$$z_1^N = 0, \quad z_2^N = 1 \tag{66}$$

By combining Equation (65) and (66),

$$z_2^{n-1} = z_2^n = 1 \tag{67}$$

$$z_1^N = 0$$

$$z_1^{N-1} = B$$

$$z_1^{N-2} = 2B$$

$$\vdots$$

$$z_1^n = (N-n)B \tag{68}$$

(Note that N is the number of the final stage.)

We also assume that no constraint is imposed on the decision or state variables (or, if some constraints exist, we assume the stationary point of H^n is interior to the constraints for all n).

Under the above assumption, the approach which sets $\frac{\partial H^n}{\partial u^n}$ equal to zero is available to obtain an optimal policy of the problem. Hence,

$$\frac{\partial H^n}{\partial u^n} = \frac{\partial}{\partial u^n}\{z_1^n(x_1^{n-1}+u^n)+z_2^n[x_2^{n-1}+A(u^n)^m+B(x_1^{n-1}+u^n-Q^n)]\}$$

$$= z_1^n + z_2^n mA(u^n)^{m-1} + B = 0 \tag{69}$$

By substituting Equations (67) and (68) into (69),

$$\frac{\partial H^n}{\partial u^n} = (N-n)B + mA(u^n)^{m-1} + B = 0$$

$$u^n = \left[\frac{(n-N-1)B}{mA}\right]^{\frac{1}{m-1}}; \quad n = 1, 2, \ldots, N \tag{70}$$

Thus, we have obtained the N-sequence of the optimal decisions of u^n from Equation (70) as a function of the number of stages (N), cost coefficients $(A$ and $B)$ and the exponent of changing level cost (m).

For N, A, B and m, let us assume that

$$N = 1, 2, \ldots, 12.$$

$$A = 1, 2, \ldots, 10.$$

$$B = 1, 2, \ldots, 10.$$

$$m = 2, 4, 6.$$

Thus, 3600 (12 x 10 x 10 x 3) possible problems are presented here. It is almost impossible to solve all these problems by hand-calculation. A FORTRAN II computer program shown below was used to accomplish this calculation.

```
      DIMENSION U(10, 10, 6)
      N=1
    1 AN=N
      PUNCH 100, N
  100 FORMAT (3HAN=, I2)
      DO 4 I=1, N
  200 FORMAT (5H  BN=, I2)
      PUNCH 200, I
      BN=I
      DO 4 IA=1, 10
      A=IA
      DO3 IB=1, 10
      B=IB
      DO 2 M=2, 6, 2
      XM=M
    2 U(IA, IB, M)=((BN-AN-1.)*B/(XM*A))**(1./(XM-1.))
  300 FORMAT (2HA=, I3, 4H  B=, I3, 5X, 3E14. 8)
    3 PUNCH 300, IA, IB, U(IA, IB, 2), U(IA, IB, 4), U(IA, IB, 6)
    4 CONTINUE
      N=N+1
      IF(N-12)1, 1, 5
    5 STOP
      END
```

AN, BN and XM are respectively N, n and m of Equation (70).

Table: For $m = 2$, following results were obtained:

exponent $m = 2$

parameters A, B, N varying as

$A = 1, 2, \ldots, 5$

$B = 5, 6, \ldots, 10$

$N = 1, 2, \ldots, 6$

that is, cost of changing production level

$= A(u^n)^2$ for all n, $n = 1, 2, \ldots, N$

cost of over-production

$$= B(x_1^n - Q^n) \qquad \text{for all } n,\ n = 1, 2, \ldots, N$$

Table 1. Table of decisions (u^n) for optimal policy.

Number of stages, N											
1	2	3	4	5	6	B=5	B=6	B=7	B=8	B=9	B=10
								A=1			
					1	15	18	21	24	27	30
				1	2	12.5	15	17.5	20	22.5	25
			1	2	3	10	10	14	16	18	20
		1	2	3	4	7.5	9	10.5	12	13.5	15
	1	2	3	4	5	5	6	7	8	9	10
1	2	3	4	5	6	2.5	3	3.5	4	4.5	5
								A=2			
					1	7.5	9	10.5	12	13.5	15
				1	2	6.25	7.5	8.75	10	11.25	12.5
			1	2	3	5	6	7	8	9	10
		1	2	3	4	3.75	4.5	5.25	6	6.75	7.5
	1	2	3	4	5	2.5	3	3.5	4	4.5	5
1	2	3	4	5	6	1.25	1.5	1.75	2	2.25	2.5
								A=3			
					1	5	6	7	8	9	10
				1	2	4.17	5	5.83	6.67	7.5	8.33
			1	2	3	3.33	4	4.67	5.33	6	6.67
		1	2	3	4	2.5	3	3.5	4	4.5	5
	1	2	3	4	5	1.67	2	2.33	2.67	3	3.33
1	2	3	4	5	6	0.83	1	1.17	1.33	1.5	1.67
								A=4			
					1	3.75	4.5	5.25	6	6.75	7.5
				1	2	3.13	3.75	4.38	5	5.63	6.25
			1	2	3	2.5	3	3.5	4	4.5	5
		1	2	3	4	1.88	2.25	2.63	3	3.38	3.75
	1	2	3	4	5	1.25	1.5	1.75	2	2.25	2.5
1	2	3	4	5	6	0.63	0.75	0.88	1	1.13	1.25
								A=5			
					1	3	3.6	4.2	4.8	5.4	6
				1	2	2.5	3	3.5	4	4.5	5
			1	2	3	2	2.4	2.8	3.2	3.6	4
		1	2	3	4	1.5	1.8	2.1	2.4	2.7	3
	1	2	3	4	5	1	1.2	1.4	1.6	1.8	2
1	2	3	4	5	6	0.5	0.6	0.7	0.8	0.9	1

(Note: all values of u^n are negative)

The state variables with respect to its sequence of the optimal decisions u^n can be easily obtained from the difference equations:

$$x_1^n = x_1^{n-1} + u^n$$

$$x_2^n = x_2^{n-1} + A(u^n)^2 + B(x_1^{n-1} + u^n - Q^n)$$

with initial conditions

$$x_1^0 = k_1, \quad x_2^0 = 0.$$

As we can recognize from the table or from Equation (70), the value of the final decision, u^N, is fixed regardless of the variation of the stage number, N.[8]

A numerical example of the production level problem which shows how to use this table is illustrated in the shaded area in Table 1.

We may reasonably respect that many authors who are working on the maximum principle may be currently preparing monographs or tables including results obtained by the discrete maximum principle such as our example.

2. Application to Capital Allocation Problem (Case 3)

Consider a five-stage allocation process for which each stage

[8] The extension of this program to the case of the optimal decisions at the boundary is possible but rather tedious.

represents one year. The first state variable is the available capital in dollars to be invested. The decision variable is the amount of capital at each year that will be invested in stock A, with the remaining capital being invested in stock B. Let stock A be that of a growing company with respected stock appreciation of ten percent per year and dividend of ten percent per year, and stock B be that of a gold mining company whose mine is being depeleted and the stock is expected to depreciate 20 percent, but whose dividend is 30 percent per year.

The problem is to find the optimal policy that will maximize the total dividend for five years. We assume the initial capital to be \$1, 000.

Let us define:

x_1^{n-1} = amount of capital available for investment at the n^{th} stage.

x_2^n = sum of the dividend up to and including the n^{th} stage.

u^n = amount of capital invested in stock A at the n^{th} stage.

$x_1^{n-1}-u^n$ = amount of capital invested in stock B at the n^{th} stage.

The difference equations can be written as:

$$x_1^n = T_1^n(x_1^{n-1}, u^n) = (1+0.1)u^n + (1-0.2)(x_1^{n-1}-u^n)$$

$$x_1^0 = 1,000 \qquad n = 1, 2, \ldots, 5 \tag{71}$$

$$x_2^n = T^n(x_1^{n-1}, x_2^{n-1}, u^n) = x_2^{n-1} + 0.1u^n + 0.3(x_1^{n-1} - u^n)$$

$$x_2^0 = 0 \qquad\qquad n = 1, 2, 3, 4, 5 \tag{72}$$

where the dividend earned at the n^{th} stage is $[0.1u^n+0.3(x_1^{n-1}-u^n)]$ and x_2^5 is the total amount of dividend earned in five years.

The objective function, J, is to maximize the total dividend x_2^5, that is,

$$\text{Max} \quad J = \sum_{i=1}^{2} c_i x_i^5 = c_1 x_1^5 + c_2 x_2^5 = x_2^5 \tag{73}$$

Hence,

$$c_1 = 0 \qquad\qquad c_2 = 1 \tag{74}$$

According to Equations (34) and (36),

$$H^n = z_1^n[1.1u^n+0.8(x_1^{n-1}-u^n)] + z_2^n[x_2^{n-1}+0.1u^n+0.3(x_1^{n-1}-u^n)]$$

$$n = 1, \ldots, 5 \tag{75}$$

$$z_1^{n-1} = \frac{\partial H^n}{\partial x_1^{n-1}} = 0.8z_1^n + 0.3z_2^n; \quad z_1^5 = c_1 = 0; \quad n = 2, \ldots, 5 \tag{76}$$

$$z_2^{n-1} = \frac{\partial H^n}{\partial x_2^{n-1}} = z_2^n = c_2 = 1; \quad n = 2, \ldots, 5 \tag{77}$$

Hence the Hamiltonian function can be rewritten as

$$H^n = z_1^n[1.1u^n+0.8(x_1^{n-1}-u^n)] + x_2^{n-1} + 0.1u^n + 0.3(x_1^{n-1}-u^n)$$

$$= (0.8z_1^n x_1^{n-1}+x_2^{n-1}+0.3x_1^{n-1}) + (0.3z_1^n-0.2)u^n \tag{78}$$

Since the values of z_1^n, x_1^{n-1}, and x_2^{n-1} at the n^{th} stage are considered as constants in extremizing the Hamiltonian function H^n, the variable portion of H^n as given by Equation (78) is

$$H_v^n = (0.3z_1^n-0.2)u^n \tag{79}$$

The function H_v^n obviously becomes a linear function of u^n. The optimal value of u^n that makes H_v^n maximum should, therefore, occur at a boundary of the admissible region of u^n, $(0 \leqq u^n \leqq x_1^{n-1})$.

The sign of q^n, given by

$$q^n = 0.3z_1^n - 0.2 \tag{80}$$

decides in which one of the boundaries $\overline{u}^n$ lies. For a positive value of qn, $\overline{u}^n$ is x_1^{n-1} (it is equivalent to investing all the money in stock A), and for a negative q^n, $\overline{u}^n$ is zero (it is equivalent to investing all the money in stock B).

Summarizing the above description,

$$\bar{u}^n = x_1^{n-1} \qquad \text{if } q^n > 0$$
$$\bar{u}^n = 0 \qquad \text{if } q^n < 0 \tag{81}$$
$$0 < \bar{u}^n < x_1^{n-1} \qquad \text{if } q^n = 0$$

In order to solve the problem, we first calculate z_1^n from Equations (76) and (77),

$$z_1^5 = 0$$
$$z_1^4 = 0.8z_1^5 + 0.3z_2^5 = 0.3$$
$$z_1^3 = 0.8(0.3) + 0.3 = 0.54$$
$$z_1^2 = 0.8(0.54) + 0.3 = 0.732$$
$$z_1^1 = 0.8(0.732) + 0.3 = 0.886$$

According to Equations (80) and (81),

$$q^5 = 0.3z_1^5 - 0.2 = -0.2 < 0 \qquad \bar{u}^5 = 0$$
$$q^4 = 0.3z_1^4 - 0.2 = (0.09-0.2) < 0 \qquad \bar{u}^4 = 0$$
$$q^3 = 0.3z_1^3 - 0.2 = [0.3(0.54)-0.2] < 0 \qquad \bar{u}^3 = 0$$
$$q^2 = 0.3z_1^2 - 0.2 = [0.3(0.732)-0.2] > 0 \qquad \bar{u}^2 = x_1^1$$
$$q^1 = 0.3z_1^1 - 0.2 = [0.3(0.886)-0.2] > 0 \qquad \bar{u}^1 = x_1^0 = 1000$$

We just obtained optimal investment policy, and substituting these values into Equations (71) and (72),

$$x_1^1 - 1.1u^1 + 0.8(x_1^0 - u^1) = 1,100$$

$$x_1^2 = 1.1u^2 + 0.8(x_1^1 - u^2) = 1.1x_1^1 = 1,210$$

$$x_1^3 = 1.1u^3 + 0.8(x_1^2 - u^3) = 0.8x_1^2 = 968$$

$$x_1^4 = 1.1u^4 + 0.8(x_1^3 - u^4) = 0.8x_1^3 = 774$$

$$x_1^5 = 1.1u^5 + 0.8(x_1^4 - u^5) = 0.8x_1^4 = 619$$

$$x_2^1 = x_2^0 + 0.1u^1 + 0.3(x_1^0 - u^1) = 0.1u^1 = 100$$

$$x_2^2 = x_2^1 + 0.1u^2 + 0.3(x_1^1 - u^2) = x_2^1 + 0.1x^1 = 210$$

$$x_2^3 = x_2^2 + 0.3x_1^2 = 210 + 0.3(1210) = 573$$

$$x_2^4 = x_2^3 + 0.3x_1^3 = 573 + 0.3(968) = 863$$

$$x_2^5 = x_2^4 + 0.3x_1^4 = 863 + 0.3(774) = 1,095$$

Hence, the total dividend for five years according to the optimal policy ($u^1 = \$1,000$, $u^2 = \$1,100$, $u^3 = 0$, $u^4 = 0$, $u^5 = 0$) is \$1,095.

A similar model was used by Fan and Hwang (1967).

3. Application to Investment Problem (Case 4)

Let us now attempt to apply the discrete form of the maximum principle to an investment problem in which there are several alternatives (three alternatives in our example).

For the optimization problem, three methods will be used to demonstrate unique characteristics of each method and to find some correlations among these mathematical techniques. Let us assume the following investment problem is given to us.

The management of a company is faced to decide a policy for investment. There are three items to be invested, and the expected returns for item 1, 2, and 3 are functions $R_1(x_1)$, $R_2(x_2)$ and $R_3(x_3)$, respectively. However, the capital available is limited as $\sum_{i=1}^{3} x_i \leqq \$15M$. Decide an optimal policy to maximize the total expected return, where

$$R_1(x_1) = 2x_1 \tag{82}$$

$$R_2(x_2) = 10(x_2)^{\frac{1}{2}} \tag{83}$$

$$R_3(x_3) = 10(x_3)^{\frac{1}{3}} \tag{84}$$

and

$$x_i \text{ is integer.}$$

The graph of each function is shown in Figure 8.

a. <u>Solution by Dynamic Programming.</u> We set the problem as

$$\text{Max} \quad [R_1(x_1) + R_2(x_2) + R_3(x_3)]$$

subject to

$$0 \leqq \sum_{i=1}^{3} x_i \leqq 15M.$$

Letting

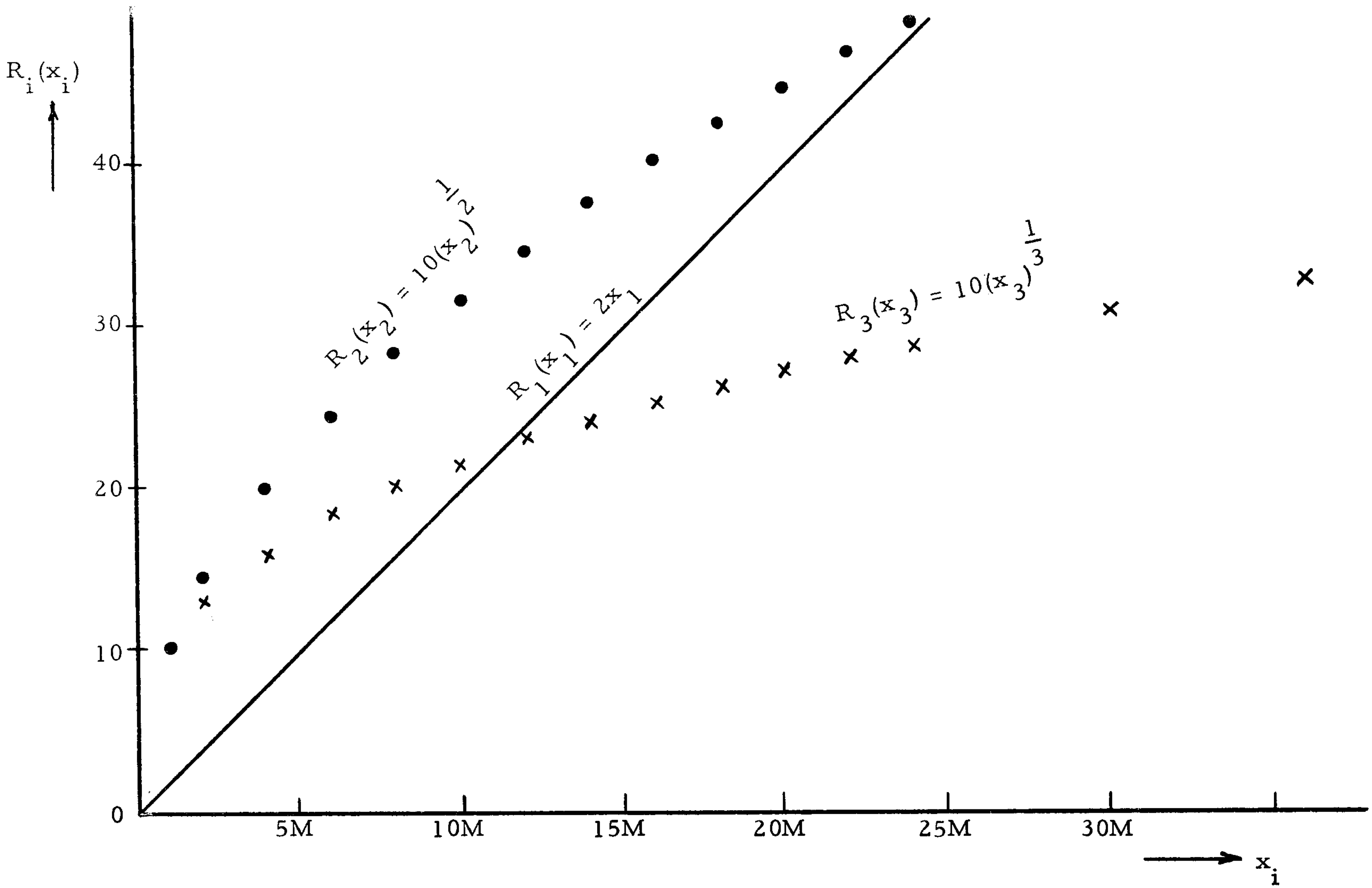


Figure 8. Return functions.

$f_n(X)$ = the return from the optimal investment of XM dollars divided between the first n investments.

$$f_1(X) = R_1(X) = R_1(x_1) \tag{85}$$

$$f_n(X) = \max_{0 \leq x_n \leq X} [R_n(x_n) + f_{n-1}(X - x_n)]; \quad n = 2, 3 \tag{86}$$

for all X, $0 \leq X \leq 15M$.

An optimal investment policy can be obtained by the above recurrence equations, and the solution procedure is as follows:

a) From the Equations (82), (83) and (84), let us construct Table 2 which indicates the values of the expected return function, $R_n(x_n)$.

b) Find $f_n(X)$ from $n = 1$ to $n = 3$, in turn, by using Equations (85) and (86). Table 3 illustrates those results.

Table 2. Expected return.

x_n	$R_1(x_1)$	$R_2(x_2)$	$R_3(x_3)$
		(M dollars)	
1	2	10.0	10.0
2	4	14.1	12.6
3	6	17.3	14.4
4	8	20.0	15.9
5	10	22.4	17.1
6	12	24.5	18.2
7	14	26.5	19.1
8	16	28.3	20.0
9	18	30.0	20.8
10	20	31.6	21.5
11	22	33.2	22.2
12	24	34.6	22.9
13	26	36.1	23.5
14	28	37.4	24.1
15	30	38.7	24.7

Table 3. $F_n(X)$ and x_n.

X	$f_1(X)$	x_1	$f_2(X)$	x_2	$f_3(X)$	x_3
					(M dollars)	
0	0	0	0	0	0	0
1	2	1	10	1	10	0, 1
2	4	2	14.1	2	20	1
3	6	3	17.3	3	24.4	1
4	8	4	20.0	4	27.3	1
5	10	5	22.4	5	30.0	1
6	12	6	24.5	6	32.6	2
7	14	7	26.5	6	35.0	2
8	16	8	28.5	6	37.1	2
9	18	9	30.5	6	.	.
10	20	10	32.5	6	.	.
11	22	11	34.5	6	.	.
12	24	12	36.5	6	.	.
13	26	13	38.5	6	.	.
14	28	14	40.5	6	.	.
15	30	15	42.5	6	51.1	2

As we can see in Table 3, $f_3(15)$ is the maximum return of the total investment, \$15M for investing to three items, and x_3, the amount of investment to item 3 is \$2M. Therefore we can obtain by Equations (85) and (86) that

$$f_3(15) = 51.1M \qquad x_3 = 2M$$

$$f_2(15-x_3) = f(13) = 38.5M \qquad x_2 = 6M$$

$$f_1(13-x_2) = f_1(7) = 14M \qquad x_1 = 7M$$

Hence, the optimal policy is $(x_1 = 7M,\ x_2 = 6M,\ x_3 = 2M)$ and the maximum return is \$51.1M.

<u>b. Solution by Lagrange Multiplier Method.</u> To make use of "Kuhn=Tucker conditions" (Kuhn and Tucker, 1951), let us introduce a Lagrangian function in this section.

$$\text{Max} \quad f(X) = R_1(x_1) + R_2(x_2) + R_3(x_3)$$

subject to

$$g(X) = (15-x_1-x_2-x_3) \geqq 0$$

$$x_i \geqq 0$$

The investment problem can be formulated as above, and the Lagrangian function corresponding to the above problem becomes

$$L(X, \lambda) = f(X) + \lambda g(X) = 2x_1 + 10(x_2)^{\frac{1}{2}} + 10(x_3)^{\frac{1}{3}} + \lambda(15-x_1-x_2-x_3) \quad (87)$$

Since the objective function $f(X)$ and the constraint function $g(X)$ are concave, the Kuhn-Tucker conditions

$$\frac{\partial L(X, \lambda)}{\partial x_i} \leqq 0 \qquad \text{if } < \text{ holds, } x_i = 0$$

$$\frac{\partial L(X, \lambda)}{\partial \lambda} \geqq 0 \qquad \text{if } > \text{ holds, } \lambda = 0$$

are necessary and sufficient to maximize $f(X)$. Let us solve the problem by using these conditions.

From Equation (87), we obtain

$$\frac{\partial L}{\partial x_1} = (2-\lambda) \leqq 0 \tag{88}$$

$$\frac{\partial L}{\partial x_2} = [5(x_2)^{-\frac{1}{2}}-\lambda] \leqq 0 \tag{89}$$

$$\frac{\partial L}{\partial x_3} = [\frac{10}{3}(x_3)^{-\frac{2}{3}}-\lambda] \leqq 0 \tag{90}$$

$$\frac{\partial L}{\partial \lambda} = (15-x_1-x_2-x_3) \geqq 0 \tag{91}$$

We know that x_i is not zero from the solution of dynamic programming in the previous section.[9]

Then, from Equation (88), λ can be obtained as

$$\lambda = 2.$$

Substituting $\lambda = 2$ into Equations (89) and (90),

$$x_2 = 6.25 \qquad x_3 = 2.15.$$

With these values we can obtain from Equation (91), $x_1 = 6.60$.

According to the initial assumption, x_i should be integer as

$$x_1 = 7 \quad x_2 = 6 \quad x_3 = 2 \quad \text{(M dollars)}$$

This optimal policy is identical to the policy derived in Section 3-a.

[9] If we do not know whether λ is zero or not, a value of λ can be assumed.

c. Solution by the Discrete Maximum Principle. We define each investment as a stage, and the state variables, the decision variables as:

u^n = amount of money invested at the n^{th} stage.

x_1^{n-1} = amount of money available for the investment at the n^{th} stage.

x_2^n = sum of the return up to and including the n^{th} stage.

The problem is, then, associated with three-stage discrete process. Difference equations of state variables can be written as

$$x_1^n = T_1^n(x_1^{n-1}, u^n) = x_1^{n-1} - u^n$$

$$x_2^n = T_2^n(x_2^{n-1}, u^n) = x_2^{n-1} + R_n(u^n); \quad n = 1, 2, 3$$

where

$$R_1(u^1) = 2u^1 \qquad \text{initial conditions}$$

$$R_2(u^2) = 10(u^2)^{\frac{1}{2}} \qquad x_1^0 = 15 \text{ (M dollars)}$$

$$R_3(u^3) = 10(u^3)^{\frac{1}{3}} \qquad x_2^0 = 0$$

Solution:

The objective function to be maximized is

$$J = x_2^3 = \sum_{i=1}^{2} c_i x_i^3 .$$

$$z_1^3 = c_1 = 0 \qquad z_2^3 = c_2 = 1 \tag{92}$$

Formulate the Hamiltonian function, H^n, and z_i^n:

$$H^n = \sum_{j=1}^{2} z_j^n T_j^n (X^{n-1}, u^n)$$

$$= z_1^n (x_1^{n-1} - u^n) + z_2^n [x_2^{n-1} + R(u^n)]; \quad n = 1, 2, 3 \tag{93}$$

$$z_i^{n-1} = \frac{\partial H^n}{\partial x_i^{n-1}} \qquad ; \quad n = 2, 3$$

By Equations (92) and (93), z_i^n is calculated as

$$z_1^{n-1} = \frac{\partial H^n}{\partial x_1^{n-1}} = z_1^n = 0 \tag{94}$$

$$z_2^{n-1} = \frac{\partial H^n}{\partial x_2^{n-1}} = z_2^n = 1 \tag{95}$$

However, Equation (94) is not valid because the difference equation of x_1^n actually exists. Hence, at least one of z_1^2 and z_1^1 should have non-zero value.

Let us begin from the last stage,

$$H^n = z_2^3[x_2^2+R_3(u^3)] = x_2^2 + 10(u^3)^{\frac{1}{3}} \tag{96}$$

According to Equation (96), the ordinary procedure which makes the derivative of H^3 zero $(\frac{\partial H^3}{\partial u^3} \rightarrow 0)$ is not applicable any more, since the function H^3 is strictly increasing. The stationary point of H^3 lies at u^3 = infinite, and the optimal decision, $\bar{u}^3$ should be at a boundary of the admissible region of u^3.

The admissible region of u^3 is, obviously, $0 \leq u^3 \leq x_1^2$. Since the function H^3 is increasing with respect to u^3, we can choose one of these boundaries in which $\bar{u}^3$ exists,

$$\bar{u}^3 = x_1^2 \tag{97}$$

Then, substitute Equation (97) into Equation (96):

$$H^3 = x^2 + 10(x_1^2)^{\frac{1}{3}}$$

$$\therefore \; z_1^2 = \frac{\partial H^3}{\partial x_1^2} = \frac{10}{3}(x_1^2)^{-\frac{2}{3}} \tag{98}$$

Going into the second stage, the function H^2 is written from Equations (95) and (98):

$$H^2 = z_1^2(x_1^1 - u^2) + x_2^1 + 10(u^2)^{\frac{1}{2}}$$

$$= \frac{10}{3}(x_1^2)^{-\frac{2}{3}}(x_1^1 - u^2) + x_2^1 + 10(u^2)^{\frac{1}{2}} \tag{99}$$

Consider x_1^1, x_1^2 and x_2^1 as constants in H^2 denoted by Equation (99), the curve of H^2 then seems to be Figure 9. The stationary points of H^2 may be within the constraint.

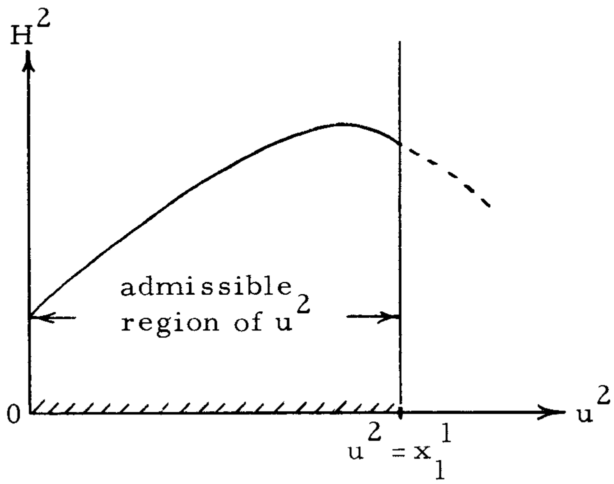


Figure 9. Function H^2.

Therefore, we can obtain the stationary point of H^2 with respect to u^2 in this way:

$$\frac{\partial H^2}{\partial u^2} = \frac{5}{\sqrt{u^2}} - \frac{10}{3}(x_1^2)^{-\frac{2}{3}} = 0$$

$$u^2 = \left[\frac{5}{\frac{10}{3}(x_1^2)^{-\frac{2}{3}}}\right]^2 \tag{100}$$

Then, go to the first stage,

$$H^1 = z_1^1(x_1^0 - u^1) + z_2^1[x_2^0 + R_1(u^1)]$$

where

$$x_1^0 = 15, \quad x_1^0 = 0 \text{ (M dollars)}; \quad z_2^1 = 1$$

$$\therefore H^1 = z_1^1(15-u^1) + 2u^1 = (2-z_1^1)u^1 + 15z_1^1 \tag{101}$$

Since the Hamiltonian function H^1 is linear with respect to u^1, the optimal decision $\overline{u}^1$ should be at a boundary of the admissible region which is $0 \leqq u^1 \leqq x_1^0$. Equation (101) can be analyzed as follows:

a) If $z_1^1 > 2$, maximum H^1 occurs ar $u^1 = 0$.

b) If $z_1^1 < 2$, maximum H^1 occurs at $u^1 = x_1^0 = 15$.

c) If $z_1^1 = 2$, maximum H^1 occurs at $0 < u^1 < 15$.

We should follow the case c), because we know that u^2 is neither zero nor 15. In other words, u^1 is between 0 and 15. Therefore,

$$z_1^1 = 2$$

According to the recurrence equation,

$$z_1^1 = \frac{\partial H^2}{\partial x_1^1} = z_1^2 = \frac{10}{3}(x_1^2)^{-\frac{2}{3}}$$

$$\frac{10}{3}(x_1^2)^{-\frac{2}{3}} = 2$$

$$x_1^2 = 2.15$$

From Equation (97),

$$u^3 = x_1^2 = 2.15$$

Substituting $\frac{10}{3}(x_1^2)^{-\frac{2}{3}} = 2$ into Equation (11), we have

$$u^2 = \left[\frac{5}{\frac{10}{3}(x_1^2)^{-\frac{2}{3}}}\right]^2 = (\frac{5}{2})^2 = 6.25$$

u^1 is obtained easily from the constraint that

$$(15 - \sum_{n=1}^{3} u^n) \geqq 0, \quad \text{and} \quad u^1 \text{ lies at this upper boundary.}$$

$$\therefore u^1 = 15 - u^2 - u^3 = 6.60$$

From the initial assumption that u^n is integer, the optimal policy is (u^1 = \$7M, u^2 = \$6M, u^3 = \$2M).

The optimal policy given by the technique of the discrete maximum principle is exactly the same as that given by the Langrange multiplier method and dynamic programming.

The most interesting fact in the relation between the maximum principle and the Lagrange multiplier method is that the value of the

Langrange multiplier λ is identical to the values of z_1^1 and z_1^2.

Although the application of the discrete maximum principle to several problems have been discussed, we must realize that it is usually very difficult to find an extremum of the Hamiltonian function H^n when the optimal decision $\overline{u}^n$ lies at a boundary of the constraint. Fortunately we have found a unique procedure in each case. However, such a unique procedure does work only for that problem, and there is no generalized solution to all problems. An attempt will now be made to present one possible general procedure using the discrete maximum principle where the optimal decison lies at a boundary of the constraint which is imposed on the decision or state variables, or where the stationary point of the Hamiltonian function is outside the constraint.

General procedure: This general procedure is applicable only when all initial conditions x_i^0 are given.

Step 1. Formulate difference equations for each stage. A proper choice of the state and decision variables is required to formulate the difference equations.

Step 2. Formulate the recurrence relation for the multiplier vector Z^n, and the Hamiltonian function H^n.

Step 3. Guess a sequence U of U^n, $n = 1, 2, \ldots, N$.

Step 4. Solve the difference equations for X^n, forward from $n = 1$ to $n = N$.

Step 5. With these X^n, obtain Z^n, backward from $n = N$ to $n = 1$ by their recurrence equations.

Step 6. With those X^n and Z^n, compute a new sequence U' of U^n from the condition, H^n = extremum.

Step 7. Return to step 4 until the new sequence U' is sufficiently close to the preceding sequence U.

This seven-step procedure can also be represented by a flow diagram illustrated in Figure 10.

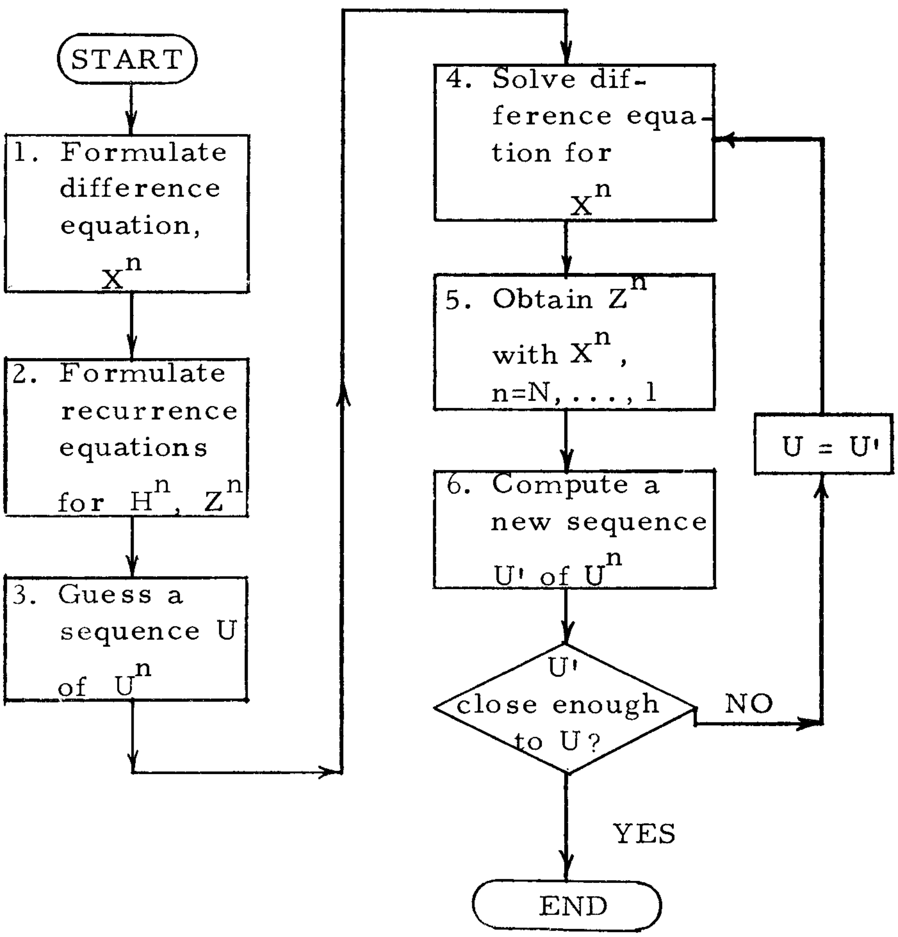


Figure 10. Flow diagram of general procedure.

D. Analysis of the Algorithm

In Section B-2, the algorithm of the discrete maximum principle has already been derived in the sense of the "variational approach" which is to make arbitrary perturbations in the decision variables. However, this procedure to derive the algorithm is unnecessarily cumbersome, and it is desirable to have a more direct derivation with intuitively logical proof.

To satisfy the above requirement we intend to analyze "why the maximum principle works" by studying its algorithm with the production level optimization problem presented in Section C-a. The discrete maximum principle is also developed from the view point of the Langrange multiplier method.

1. Analysis of the Algorithm through Production Level Problem

We consider the production level problem (Case 1).

Since the function $J = \sum_{i=1}^{2} c_i x_i^n$ is the objective function of the problem, one of the state variables of the final stage should be the total cost of the production system. Therefore, the Hamiltonian function of the last stage, H^4, corresponds to the total cost to be minimized.

Let TC be the total cost, then

$$TC = H^4 = x_2^3 + 3(u^4)^2 + 10(x_1^3+u^4-Q^4)$$

Since x_2^3 is the sum of cost up to the 3^{rd} stage,

$$[3(u^4)^2+10(x_1^3+ u^4-Q^4)]$$

is the cost at the 4^{th} stage. In general, if

$$G^n(x_1^{n-1}, u^n)$$

is the cost at the n^{th} stage,

$$G^n(x_1^{n-1}, u^n) = 3(u^n)^2 + 10(x_1^{n-1}+u^n-Q^n) \tag{102}$$

$$\therefore J = TC = H^4 = \sum_{n=1}^{4} G^n(x_1^{n-1}, u^n) \tag{103}$$

By substituting Equation (102) into Equation (103),

$$TC = 3(u^1)^2 + 10(x_1^0+u^1-Q^1) + 3(u^2)^2 + 10(x_1^1+u^2-Q^2) + 3(u^3)^2$$
$$+ 10(x_1^2+u^3-Q^3) + 3(u^4)^2 + 10(x_1^3+u^4-Q^4) \tag{104}$$

From the definition of the state variable x_1^n in the maximum principle, x_1^n can be represented in terms of u^n as follows:

$$x_1^1 = x_1^0 + u^1$$

$$x_1^2 = x_1^0 + u^1 + u^2$$

$$x_1^3 = x_1^0 + u^1 + u^2 + u^3$$

where x_1^0 = constant.

By substitution of the above equations into Equation (104), the total cost is described in terms of only the decision variables.

$$TC = 3(u^1)^2 + 10(x_1^0+u^1-Q^1) + 3(u^2)^2 + 10(x_1^0+u^1+u^2-Q^2) + 3(u^3)^2 + 10(x_1^0+u^1+u^2+u^3-Q^3) + 3(u^4)^2 + 10(x_1^0+u^1+u^2+u^3+u^4-Q^4) \quad (105)$$

Now, let us make an assumption that no constraint is imposed on the decision variables, or that the stationary point of the function TC with respect to u^n lies interior to the constraint if it is imposed.

Then, the problem is to minimize the total cost function TC. One can solve this problem by using the classical differential calculus method as

$$\frac{\partial TC}{\partial u^n} = 0$$

$$\frac{\partial TC}{\partial u^1} = 0 = 6u^1 + 40 \qquad u^1 = -\frac{20}{3}$$

$$\frac{\partial TC}{\partial u^2} = 0 = 6u^2 + 30 \qquad u^2 = -5$$

$$\frac{\partial TC}{\partial u^3} = 0 = 6u^3 + 20 \qquad u^3 = -\frac{10}{3}$$

$$\frac{\partial TC}{\partial u^4} = 0 = 6u^4 + 10 \qquad u^4 = -\frac{5}{3}$$

We can now recognize that this result is exactly the same as the sequence of u^n solved in Section C-a. We can, therefore, say that the discrete maximum principle is identical to the classical differential calculus method when no constraint is imposed.

Let us next assume that the constraint is imposed on the decision variable u^n or the state variable x_1^n. According to the production level problem, x_1^n must satisfy the condition:

$$x_1^n \geqq Q^n$$

Then, the problem can be stated as follows:

$$\text{Min} \quad TC(u^n)$$

subject to

$$x_1^n \geqq Q^n; \quad n = 1, 2, 3, 4$$

However the constraint, $x_1^n \geqq Q^n$, can be represented in terms of u^n. Hence, the problem may be rewritten as

$$\text{Min} \quad TC(u^1, u^2, u^3, u^4)$$

subject to

$$(Q^n - x_1^0 - \sum_{i=1}^{n} u^i) \leqq 0; \qquad n = 1, 2, 3, 4$$

This problem may be solved by introducing a Lagrangian function $L(u^n, \lambda^n)$ with Kuhn-Tucker conditions, that is,

$$L(u^n, \lambda^n) = TC(u^1, u^2, u^3, u^4) + \lambda^1(Q^1 - x_1^0 - u^1) + \lambda^2(Q^2 - x_1^0 - u^1 - u^2)$$

$$+\lambda^3(Q^3 - x_1^0 - u^1 - u^2 - u^3) + \lambda^4(Q^4 - x_1^0 - u^1 - u^2 - u^3 - u^4)$$

$$\frac{\partial L}{\partial u^n} \geqq 0, \qquad \frac{\partial L}{\partial \lambda^n} \leqq 0.$$

With Equation (105),

$$\frac{\partial L}{\partial u^1} = (6u^1 + 40 - \lambda^1 - \lambda^2 - \lambda^3 - \lambda^4) \geqq 0$$

$$\frac{\partial L}{\partial u^2} = (6u^2 + 30 - \lambda^2 - \lambda^3 - \lambda^4) \geqq 0$$

$$\frac{\partial L}{\partial u^3} = (6u^3 + 20 - \lambda^3 - \lambda^4) \geqq 0$$

$$\frac{\partial L}{\partial u^4} = (6u^4 + 10 - \lambda^4) \geqq 0$$

$$\frac{\partial L}{\partial \lambda^1} = (Q^1 - x_1^0 - u^1) \leqq 0$$

$$\frac{\partial L}{\partial \lambda^2} = (Q^2 - x_1^0 - u^1 - u^2) \leqq 0$$

$$\frac{\partial L}{\partial \lambda^3} = (Q^3 - x_1^0 - u^1 - u^2 - u^3) \leqq 0$$

$$\frac{\partial L}{\partial \lambda^4} = (Q^4 - x_1^0 - u^1 - u^2 - u^3 - u^4) \leqq 0$$

Then, guess λ^n which minimizes the function TC, based on these conditions.

For the production level problem in Section C-a, the optimal λ^n is $\lambda^1 = \lambda^2 = \lambda^3 = \lambda^4 = 0$. For the production level problem in Section C-b, the optimal λ^n is $\lambda^1 = \lambda^2 = \lambda^4 = 0$, $\lambda^3 = 156\frac{2}{3}$.

2. The Maximum Principle from the Lagrange Multiplier Point of View

It is possible to derive the algorithm of the discrete maximum principle by the Lagrange multiplier technique with some assumptions. This application of the Lagrange multiplier approach to the maximum principle has also been developed by Thomas (1967), independently. However, the derivation from the Lagrange multiplier method is useful only if the difference equation of the state variable is non-linear. This is the main assumption on this approach.

Following the original concept of the maximum principle that one of the state variables in the process represents the return function, the problem can be, in general, stated as

$$\text{Max}^{10} \quad J = x_m^N = \sum_{n=1}^{N} G^n(X^{n-1}, U^n)$$

subject to

$$X^n = T^n(X^{n-1}, U^n) \qquad n = 1, \ldots, N$$

$$X^0 = K$$

where x_m^N is the m^{th} state variable at the final stage and describes the sum of the return functions up to the N^{th} stage, and $G^n(X^{n-1}, U^n)$ is the return function at stage n resulting from the state variable X^{n-1} and the decision variable U^n. Initial condition $X^0 = K$ is given.

Then, we assume that each of these functions, $G^n(\cdot)$ has continuous second partial derivatives with respect to each of its arguments. The Lagrangian multiplier technique requires the Lagrangian function, $L(\cdot)$.

$$L(X^0, \ldots, X^N; U^1, \ldots, U^N; \lambda^0, \ldots, \lambda^N)$$

$$= \sum_{n=1}^{N} G^n(X^{n-1}, U^n) - \sum_{n=1}^{N} \lambda^n[X^n - T^n(X^{n-1}, U^n)] - \lambda^0(X^0 - K) \tag{106}$$

A necessary condition for a maximum of the objective, J, is that the partial derivative of the lagrangian function with respect to each

---

[10] The superscript indicates the stage number, the subscript indicates the variable's dimension. If a variable has only the superscript, it is represented in vector form.

of its arguments vanishes. Hence,

$$\left.\begin{aligned}
&\frac{\partial L}{\partial X^{n-1}} = 0 = \frac{\partial G^n(X^{n-1}, U^n)}{\partial X^{n-1}} - \lambda^{n-1} + \lambda^n \frac{\partial T^n(X^{n-1}, U^n)}{\partial X^{n-1}}; \\
&\qquad\qquad\qquad\qquad\qquad\qquad n = 1, 2, \ldots, N \\
&\frac{\partial L}{\partial x_i^N} = 0 = \lambda_i^N; \quad i = 1, \ldots, s \quad i \neq m \\
&\frac{\partial L}{\partial x_m^N} = 0 = 1 - \lambda_m^N
\end{aligned}\right\} \tag{107}$$

$$\frac{\partial L}{\partial U^n} = 0 = \frac{\partial G^n(X^{n-1}, U^n)}{\partial U^n} + \lambda^n \frac{\partial T^n(X^{n-1}, U^n)}{\partial U^n}; \quad n = 1, 2, \ldots, N \tag{108}$$

$$\left.\begin{aligned}
&\frac{\partial L}{\partial \lambda^0} = 0 = -X^0 + K \\
&\frac{\partial L}{\partial \lambda^n} = 0 = -X^n + T^n(X^{n-1}, U^n); \quad n = 1, \ldots, N
\end{aligned}\right\} \tag{109}$$

These are the same conditions we have obtained from the discrete maximum principle and λ^n is identical to Z^n, the multiplier vector of the maximum principle. Thus, by application of the well-known Langrange multiplier method we obtain the necessary conditions for an optimal solution. However, we dispense with the sufficient conditions because of the complexity.[11]

---

[11] The sufficiency conditions are described in references, (Thomas, 1967) and (Hancock, 1917).

Constraint: We have discussed the maximum principle from the point of view of the Lagrange multiplier technique with no constraints on the decision or state variables. We will outline, next, an alternate derivation of the Lagrange multiplier method under the assumption that some constraints are imposed on the state or decision variables. It is easily handled by introducing an additional Lagrange multiplier for each constraint. No additional conceptual difficulties are encountered.

Recall the definition of the state variables in the discrete maximum principle,

$$x_i^n = T_i^n(x_1^{n-1}, \ldots, x_s^{n-1};\ u_1^u, \ldots, u_r^n) \qquad n = 1, \ldots, N$$

$$i = 1, \ldots, s$$

$$x_{s+1}^n = T_{s+1}^n(x_1^{n-1}, \ldots, x_{s+1}^{n-1};\ u_1^n, \ldots, u_r^n)$$

$$x_i^0 = k_i \qquad i = 1, \ldots, s$$

where x_i^n is the state variable at n^{th} stage, and has s-dimension; similarly x_{s+1}^n is the $(s+1)^{th}$ state variable and represents the return functions up to and including n^{th} stage. Then, the problem can be written with the objective function, as

$$\text{Max} \quad J = x_{s+1}^N = \sum_{n=1}^{N} G^n(X^{n-1}, U^n)$$

subject to

$$x_i^n = T_i^n(x_1^{n-1}, \ldots, x_s^{n-1};\ u_1^n, \ldots, u_r^n)$$

$$x_i^0 = k_i \qquad\qquad i = 1, \ldots, s \qquad n = 1, \ldots, N$$

bounded by the constraint

$$g_i^n(x_1^{n-1}, \ldots, x_s^{n-1};\ u_1^n, \ldots, u_r^n) \leqq 0 \qquad i = 1, \ldots, s \qquad n = 1, \ldots, N$$

The Lagrangian function with respect to the above problem is

$$L(X^0, \ldots, X^N : U^1, \ldots, U^N : \lambda^0, \ldots, \lambda^N : \alpha^1, \ldots, \alpha^N)$$

$$= \sum_{n=1}^{N} \{G^n(X^{n-1}, U^n) - \lambda^n[X^n - T^n(X^{n-1}, U^n)] + \alpha^n g^n(X^{n-1}, U^n)\} + \lambda^0(X^0 - K)$$

Then, the necessary conditions become

$$\frac{\partial L}{\partial X^n} = 0 \quad n = 0, \ldots, N \qquad \frac{\partial L}{\partial \alpha^n} \leqq 0 \quad n = 1, \ldots, N$$

$$\frac{\partial L}{\partial U^n} = 0 \quad n = 1, \ldots, N \qquad \alpha^n g^n(\cdot) = 0 \quad n = 1, \ldots, N$$

$$\frac{\partial L}{\partial \lambda^n} = 0 \quad n = 0, \ldots, N \qquad \alpha^n \geqq 0 \quad n = 1, \ldots, N$$

We have seen that the discrete maximum principle may be obtained by application of the Lagrange multiplier technique with assumption that the objective function has continuous second partial derivatives. The constraints imposed on the state or decision variables bring additional multipliers into the Lagrangian function and introduce no new conceptual difficulty.

DISCUSSION AND CONCLUDING REMARKS

The discrete maximum principle is applicable to the processes with well-defined performance equations which must be continuously differentiable with respect to the state variables. In other words, the maximum principle enables us to solve the problem whose process can be either linear or non-linear. However, this optimization technique may be most useful for the quadratic programming.

There is no difficulty in obtaining a sequence of optimal decisions by the maximum principle, when no constraint is imposed on the state and/or decision variables, or when the optimal policy is interior to the admissible region of the constraint. But, if the constraint is imposed, and at least one of the optimal decisions must lie at the boundary of its constraint, it is quite complicated and cumbersome in the computational procedure to obtain the solution. This is because it is not easy to calculate the multiplier, Z^n, in the latter case, and we must guess a sequence of the decisions so as to obtain Z^n. According to the characteristics of transformation function, we can compute Z^n with special care even if the stationary point of the Hamiltonian function H^n lies outside the set of constraint. In such a case, the computational procedure by the discrete maximum principle is very similar to that of dynamic programming techniques.

The discrete maximum principle may also be obtained by the application of the Lagrange multiplier technique which has been found useful in non-linear programming problems. This derivation is more direct than that of the variational approach and it is useful regardless of whether or not there exists a constraint on decision or state variables. Denn and Aris (1965) derived the algorithm of the discrete maximum principle through the "Green's Function."

Some doubts have been thrown on the correctness of the discrete maximum principle by Horn and Jackson (1965) using two simple counter-examples. According to them, the discrete version of the maximum principle is invalid unless the objective function or one of the Hamiltonian functions takes the stationary value with respect to the decision variable. They have listed some circumstances in which each H^n should take the stationary point in order that the objective, J, has its stationary point. Horn and Jackson seem to be at least partly correct. However, if some constraints are imposed on the decision variables, it is not necessary that the objective function or each of the Hamiltonian functions takes the stationary value.

We have presented the discrete maximum principle for simple multistage processes with given initial conditions. However, this basic concept can be extended to handle a variety of problems in practice. For example, complex multistage processes, processes with feed-back loop, processes with fixed-end point and so on can be solved by extending the basic concept. The problems solved in this thesis are limited, but other applications to the problems in business and industry such as equipment replacement and transportation problems have been developed by Fan and Wang (1964).

Inasmuch as an optimization problem with a stagewise nature has only one global (or local) optimal policy, the maximum principle may be one of the most powerful techniques at present in solving such a problem. We may say that this is the greatest contribution made by the maximum principle. It is widely recognized that there is no single mathematical optimization technique superior to all other techniques in handling all types of problems. Each method has its own merits and shortcomings. It is well known that, among many optimization techniques, dynamic programming is also very powerful in solving problems of a discrete process. The numerical calculation involved in solving a problem by dynamic programming is usually carried out by a digital computer. The calculation of dynamic programming is limited by the memory capacity of a computer. However, it must be admitted that, in solving a problem by the maximum principle, the difficulty in storage is avoided by introducing the multiplier vector z^n. In this point, we can say that the maximum principle is a better technique than dynamic programming. The best way to solve a stagewise optimization problem may be that the method of dynamic programming be employed first to locate approximately the position of global maximum and then the maximum principle be applied to pinpoint the maximum point.

All the processes which we have discussed are deterministic, but many industrial and management systems are also stochastic in nature. The application to stochastic processes is now being developed by industrial engineers.

BIBLIOGRAPHY

Bellman, Richard E. 1953. An introduction to the theory of dynamic programming. Santa Monica, Calif., Rand. 99 p.

________________ 1957. Dynamic programming. Princeton, N.J., Princeton University. 342 p.

Bellman, Richard E. and Stuart E. Dreyfus. 1962. Applied dynamic programming. Princeton, N.J., Princeton University. 363 p.

Carr, Charles R. and Charles W. Howe. 1964. Quantitative decision procedures in management and economics. New York, McGraw-Hill. 383 p.

Chang, S.S.L. 1960. Digitized maximum principle. Proceedings of the Institute of Radio Engineers 48:2030-2031.

Chen, S.K. et al. 1966. The stochastic discrete maximum principle for simple processes. Manhattan, Kansas. 21 p. (Kansas. Engineering Experiment Station. Special report no. 64)

Denn, M.M. and Rutherford Aris. 1965a. Green's functions and optimal systems. Industrial and Engineering Chemistry, Fundamentals 4:7-16.

________________ 1965b. Second order variational equations and the strong maximum principle. Chemical Engineering Science 20: 373-384.

Fan, L.T. and C.S. Wang. 1964. The discrete maximum principle. New York, Wiley. 158 p.

Fan, L.T. et al. 1966. The continuous maximum principle: a study of multistage systems optimization. New York, Wiley. 411 p.

Flagle, Charles D. et al. (eds.) 1960. Operations research and systems engineering. Baltimore, Johns Hopkins. 889 p.

Freeman, Herbert. 1965. Discrete time systems. New York, Wiley. 241 p.

Gamkrelidze, R. V. 1958. On the general theory of optimum processes. Doklady Akademiia Nauk SSSR 123:223-226. (Translated in Automation Express 1:37-39. 1959)

Hancock, Harris. 1917. Theory of maxima and minima. New York, Dover. 137 p.

Hadley, G. 1962. Linear programming. Reading, Mass., Addison-Wesley. 520 p.

_______________ 1964. Nonlinear and dynamic programming. Reading, Mass., Addison-Wesley. 484 p.

Hohn, F. and R. Jackson. 1965. Correspondence on the discrete maximum principle. Industrial and Engineering Chemistry, Fundamentals 4:110-112.

Holt, C. C. et al. 1960. Planning production, inventories and work force. Englewood Cliffs, N. J., Prentice-Hall. 419 p.

Hwang, C. L. and L. T. Fan. 1966. The application of the maximum principle to industrial and management systems. The Journal of Industrial Engineering 17:589-593.

_______________ 1967. A discrete version of Pontryagin's maximum principle. Operations Research 15:139-146.

Hwang, C. L., L.E. Erickson and L. T. Fan. 1967. Optimum production planning by the maximum principle. Management Science 13:751-755.

Inoue, Michael S. 1967. Systems tableau: an integrated approach to systems theory. Ph. D. thesis. Corvallis, Oregon State University. 315 numb. leaves (Microfilm)

Katz, Stanley. 1962. A discrete version of Pontryagin's maximum principle. Journal of Electronics and Control 13:179-184.

Kopp, Richard E. 1962. Pontryagin maximum principle. In: Optimization techniques with applications to aerospace systems, ed. by George Leitmann. New York, Academic. p. 255-279.

Kuhn, W. H. and A. W. Tucker. 1951. Nonlinear programming. In: Proceedings of the symposium on mathematical statistics and probability, ed. by Jerry Neyman. Berkeley, Calif., University of California.

Lee, E.S. 1963. Optimum design and operation of chemical processes. Industrial and Engineering Chemistry 55(8): 30-39.

Magee, John F. 1958. Production planning and inventory control. New York, McGraw-Hill. 333 p.

Matsuda, S., M. Sunouchi and S. Sugiyama. 1963. Fundamental mathematics for operations research 3. Tokyo, Maruzen. 320 p. (In Japanese)

Naha, Seiichi. 1960. Differential equations. Tokyo, Nihon University, Department of Science and Engineering. 232 p. (In Japanese)

Nemhauser, George L. 1966. Introduction to dynamic programming. New York, Wiley. 256 p.

Pontryagin, L. S. et al. 1962. The mathematical theory of optimal processes, tr. by K. N. Trirogoff. New York, Interscience. 360 p.

Thomas, Michael E. 1967. An alternate derivation of the discrete maximum principle. Paper presented at the Thirty-first National Meeting of ORSA, Sess. FP2. 2 on Allocation Problems, Control Theory, and Systems, National Institute of Health, Bethesda, Maryland, June 2. 11 numb. leaves. (Mimeographed preprint)

Wald, Abraham. 1950. Statistical decision functions. New York, Wiley. 179 p.

Wang, C. S. and L. T. Fan. 1964a. Multistage optimization by the generalized discrete maximum principle. Journal of Electronics and Control 16: 441-450.

________________ 1964b. The optimization of continuous complex processes by the maximum principle. Journal of Electronics and Control 17: 199-209.

APPENDIX

APPENDIX I

Solution for the Problem in Section C-1-a (by Dynamic Programming)

Problem: Determine the production level at each period, for a perishable commodity, under the following conditions.

- The excess production over the sales forecast is wasted at a cost of $10 per unit.
- The cost of changing the level is three times the square of the difference between two production levels.
- The following sales forecast must be met.

Quarter	1	2	3	4
Sales forecast	115	125	100	95 units

- The last quarter production level was $136\frac{2}{3}$ units.

Solution: Backward approach.

Period	1	2	3	4
Stage	4	3	2	1
	f_4	f_3	f_2	f_1

Let

$f_n(x_{n+1})$ = the cost of an optimal policy for the last n stages.

x_{n+1} is the production level at the previous stage.

where

$$f_0 = 0$$

$$f_n(x_{n+1}) = \min_{x_n \geqq Q_n} \{3(x_n - x_{n+1})^2 + 10(x_n - Q_n) + f_{n-1}(x_n)\}$$

$n = 1, 2, 3, 4$ Q_n: sales forecast at the n^{th} stage.

$$\therefore f_1(x_2) = \min_{x_1 \geqq Q_1} \{3x_1^2 - 6x_1x_2 + 3x_2^2 + 10x_1 - 10Q_1\}$$

By letting $\frac{\partial f_1}{\partial x_1}$ be zero, we can find the absolute minimum of function $f_1(x_2)$. However, the stationary point (absolute min) may exist outside the constraint $x_1 \geqq Q_1$. If the absolute min lies outside the constraint, we must take thelocal min at the boundary.

Stage 1.

$$\frac{\partial f_1}{\partial x_1} = 6x_1 - 6x_2 + 10 = 0$$

(1a) $\underline{x_1 = x_2 - \frac{5}{3}}$ if $x_2 \geqq Q_1 + \frac{5}{3}$ (since $x_1 \geqq Q_1$)

(1b) $x_1 = Q_1 = 95$ if $x_2 < Q_1 + \frac{5}{3} = 96\frac{2}{3}$

We choose $x_1 = x_2 - \frac{5}{3}$, because $x_2 \geqq 100$. Then, substituting $x_1 = x_2 - \frac{5}{3}$ into $f_1(x_2)$,

$$f_1(x_2) = 10x_2 - C_1 \quad C_1 = \text{constant}$$

Stage 2.

$$f_2(x_3) = \min_{x_2 \geqq Q_2} \{3(x_2-x_3)^2 + 10(x_2-Q_2) + f_1(x_2)\}$$

$$= \min \{3(x_2-x_3)^2 + 10(x_2-Q_2) + 10x_2 - C_1\}$$

$$\frac{\partial f_2(x_3)}{\partial x_2} = 6(x_2-x_3) + 20 = 0$$

(2a) $x_2 = x_3 - \frac{10}{3}$ if $x_3 \geqq (Q_2+\frac{10}{3}) = 103\frac{1}{3}$

(2b) $x_2 = Q_2 = 100$ if $x_3 < 103\frac{1}{3}$

We choose $x_2 = x_3 - \frac{10}{3}$, since $x_3 \geqq 125$. Then substituting $x_2 = x_3 - \frac{10}{3}$ into $f_2(x_3)$,

$$f_2(x_3) = 20x_3 - C_2$$

Stage 3.

$$f_3(x_4) = \min_{x_3 \geqq Q_3} \{3(x_3-x_4)^2 + 10(x_3-Q_3) + 20x_3 - C_2\}$$

$$\frac{\partial f_3}{\partial x_3} = 6(x_3-x_4) + 30 = 0$$

(3a) $x_3 = x_4 - 5$ if $x_4 \geqq (Q_3+5) = 130$

(3b) $x_3 = Q_3 = 125$ if $x_4 < 130$

We can choose either $x_3 = x_4 - 5$ or $x_3 = 125$, since $x_4 \geqq 115$.

For example if we take $x_3 = 125$, then, substituting $x_3 = 125$ into $f_3(x_4)$,

$$f_3(x_4) = 3(125-x_4)^2 + 10(125-Q_3) + (20)(125) - C_2$$

$$= 3x_4^2 - 750x_4 + C_3$$

Stage 4.

$$f_4(x_5) = \min_{x_4 \geqq Q_4} \{3(x_4-x_4)^2 + 10(x_4-Q_4) + 3x_4^2 - 750x_4 + C_3\}$$

where x_5 = initial value = $136\frac{2}{3}$

$$\frac{\partial f_4}{\partial x_4} = 6(x_4-x_5) + 6x_4 - 740 = 0$$

$\therefore x_4 = 130$ It satisfies the constraint $x_4 \geqq Q_4$.

Then, through Equations (3b), (2a) and (1a) in this order, we can obtain

$$x_4 = 130 \quad x_3 = 125 \quad x_2 = 121\frac{2}{3} \quad x_1 = 120$$

If we indicate subscript in terms of periods,

$$x_1 = 130 \quad x_2 = 125 \quad x_3 = 121\frac{2}{3} \quad x_4 = 120$$

APPENDIX II

Solution for the Problem in Section C-1-b (by Dynamic Programming)

Problem: Determine the production level, under the assumption that the sales forecast at 3rd quarter is changed from 100 to 200 units. The other conditions remain the same.

Quarter	1	2	3	4
Sales forecast	115	125	200	95 units
Let stage	4	3	2	1

Solution: Backward approach. All notations are the same as those in Appendix I.

Stage 1.

$$f_1(x_2) = \min_{x_1 \geqq 95} \{3(x_1 - x_2)^2 + 10(x_1 - Q_1) + f_0(x_1)\}$$

$$f_0 = 0$$

$$\frac{\partial f_1}{\partial x_1} = 6x_1 - 6x_2 + 10 = 0$$

(1a) $x_1 = x_2 - \frac{5}{3}$ if $x_2 \geqq 96\frac{2}{3}$

(1b) $x_1 = Q_1 = 95$ if $x_2 < 96\frac{2}{3}$

We choose Equation (1a), because of $x_2 \geqq 200$, substituting Equation (1a) into $f_1(x_2)$ to obtain

$$f_1(x_2) = 10x_2 - C_1 \qquad C_1 = \text{constant}$$

Stage 2.

$$f_2(x_3) = \min_{x_2 \geqq 200} \{3(x_2-x_3)^2 + 10(x_2-200) + 10x_2 - C_1\}$$

$$\frac{\partial f_2}{\partial x_2} = 6(x_2-x_3) + 20 = 0$$

(2a) $x_2 = x_3 - \frac{10}{3}$ if $x_3 \geqq (Q_2+\frac{10}{3}) = 203\frac{1}{3}$

(2b) $x_2 = Q_2 = 200$ if $x_3 < 203\frac{1}{3}$

We take Equation (2b), since $x_3 \geqq 125$. Substituting $x_2 = 200$ into $f_2(x_3)$,

$$f_2(x_3) = 3x_3^2 = 1200x_3 + C_2$$

Stage 3.

$$f_3(x_4) = \min_{x_3 \geqq 125} \{3(x_3-x_4)^2 + 10(x_3-125) + 3x_3^2 - 1200x_3 + C_2\}$$

$$\frac{\partial f_3}{\partial x_3} = 12x_3 - 6x_4 - 1190 = 0$$

(3a) $x_3 = \frac{1}{2}x_4 + 99\frac{1}{6}$ if $x_4 \geqq 2(Q_3-99\frac{1}{6}) = 25\frac{5}{6}$

(3b) $x_3 = Q_3 = 125$ if $x_4 < 25\frac{5}{6}$

We choose Equation (3a) because of $x_4 \geqq 115$. Substitute Equation

(3a) into $f_3(x_4)$,

$$f_3(x_4) = \frac{3}{2}x_4^2 - 595x_4 + C_3$$

Stage 4.

$$f_4(x_5) = \min_{x_4 \geqq 115} \{3(x_4 - 136\tfrac{2}{3})^2 + 10(x_4 - 115) + \tfrac{3}{2}x_4^2 - 595x_4 + C_3\}$$

where $x_5 = 136\frac{2}{3}$

$$\frac{\partial f_4}{\partial x_4} = 6x_4 - 820 + 10 + 3x_4 - 595 = 0$$

$$\therefore\ x_4 = 156\tfrac{1}{9} \qquad \text{if} \quad x_4 \geqq Q_4 = 115$$

Then, substituting $x_4 = 156\frac{1}{9}$ into Equations (3a), (2b) and (1a), one can obtain

$$x_4 = 156\tfrac{1}{9} \quad x_3 = 177\tfrac{2}{9} \quad x_2 = 200 \quad x_1 = 198\tfrac{1}{3}$$

In terms of period's subscripts,

$$x_1 = 156\tfrac{1}{9} \quad x_2 = 177\tfrac{2}{9} \quad x_3 = 200 \quad x_4 = 198\tfrac{1}{3}$$