No. 132 (81-33)

Multistage Decision Process with Random Observations and its Applications*)

by

Seizo Ikuta**)

December 1981

Abstract

Consider a finite horizon, discrete-time stochastic decision process where a decision is made after drawing an observation from a known distribution. Assume that both an immediate reward and state transition law, associated with a present state and action taken, may depend on an observation drawn. The present paper studies mainly a case that both state and action spaces are discrete as well as finite. The objective to the process is to maximize the total expected reward or per-period expected reward over a given time horizon, possibly infinite. We provide an approach to the process which makes possible a very systematic treatment of a certain class of managerial, economic decision problems, say a commodity purchasing problem on a fluctuating market. Seven examples in application of the approach are demonstrated. An explicit relationship between the process and a Markovian decision process is revealed, by use of which it is shown how to apply a Horward's algorithm to an infinte time horizon version of the process. A slight extension is attempted to a case that both action and state spaces are given by finite intervals.

*) This is a revision of discussion papers, No.74 and No.94, by the auther.

**) Institute of Socio-Economic Planning, University of Tsukuba, Sakura-mura, Niihari-gun, Ibaraki-ken 305, Japan

0. INTRODUCTION

The model of a discrete-time stochastic decision process presented in the paper is formulated and developed with the intention of establishing a systematic approach to such problems as cited below: a sequential stochastic assignment probelm (Derman et al. (1972)), material purchasing problem on a fluctuating market (Fabian et al. (1959)), uniform game (Gilbert and Mosteller (1966)), strategy for divestiture (Hayes (1969)), stopping problem (Hockman (1973)), order selection problem (Ikuta (1975)), asset selling problem (Arrow (1962)), job search problem (Lippman and MaCall (1976)), target attacking problem (Mastran and Thomas (1973)), purchasing problem (Morris (1959)), sequential sampling problem (Sakaguchi (1961)), hause selling problem (Simon (1957)), timing problem (Taylor (1967)), warehouse problem (Wagner (1975)), and so on. Moreover it will be demonstrated that our approach to the model becomes a quite powerfull and efficient mean of treating not only more generalized versions of the problems cited above but also some topics related to them. In section 1 we shall define a model which is mainly discussed in the present paper. Here the model is assumed to have a finite time horizon and to have state and action spaces both of which are discrete as well as finite. A general form of a system of fundamental equations characterizing the model is given under some monotonicity conditions. Subsequently a group of necessary conditions are derived under which these monotonicity conditions are satisfied for all points in time over a given planning horizon. Section 2 reveals an explicit relationship between our model and a Markovian decision process. The use of the relationship provides a procedure for an application of a Howard's policy iteration algorithm to an infinite time horizon version of our model. In section 3 two specializations of the model are presented to such a case that an immediate reward is a linear function of an observation obtained. In section 4 a slight extension is attempted to a case that both state and action spaces are given by finite intervals. Section 5 demonstrates a wide applicability of our model to a certain type of managerial, economic decision problems, in particular, ones as being cited above. Some of them have been already investigated in detail by many authers, however readers will appreciate in the section that our model offers an even more systematic approach to these types of problems than conventional ones. In the last subsection of this section we shall present and examine an example of cases that the monotonicity conditions defined in section 1 are not always satisfied, just under which a successful treatment of our model can be made possible.

1. MODEL WITH FINITE TIME HORIZON

Consider the following discrete-time stochastic decision process. The process is assumed to have a given finite time horizon where for convenience let points in time be numbered backward as ...,2,1,0 from the terminating point in time of the process, denoted by $t = 0$. Let the process have a finite state space $I = \{i = \text{integer}; 0 \leqq i \leqq N\}$ with a fixed nonnegative integer N and finite state-dependent action spaces $A(i) = \{x = \text{integer}; a(i) \leqq x \leqq b(i)\}$ where both a(i) and b(i) are integers such as $a(i) \leqq b(i)$. Moreover we shall assume that A(i) is concave or linear where "concave" and "linear" are defined as follows;

DEFINITION 1 Let $[A(i)] = [a(i),b(i)]$, a closed interval, and define $\lambda[A(i)] + (1-\lambda)[A(i')] = \{z=\lambda x+(1-\lambda)x'; x \text{ in } [A(i)], x' \text{ in } [A(i')]\}$ for any λ in $[0,1]$. A(i) is said to be concave if $[A(\lambda i+(1-\lambda)i')] \supset \lambda[A(i)]+(1-\lambda)[A(i')]$ for any two integers i and i' in I and for any λ in $[0,1]$ such as $\lambda i+(1-\lambda)i'$ in I. A(i) is said to be linear if both a(i) and b(i) are linear on I.

Now suppose that at each point of time an observation θ is drawn from a known distribution F where succssive observations are assumed to be independent. Let an observation space be denoted by Ω. If an observation θ is obtained in state i and an action x in A(i) is chosen, then an immediate reward of $g(i,x,\theta)$, a real function, can be gained and the next state of the process becomes $j(i,x,s)$ with probability q_s, $s = 0,1,\ldots,$ $\sum_s p_s = 1$. From now on let the s be called a disturbance. Here it is a matter of course that j(i,x,s) must be always defined to be in I. If $j(i,x,s)$ depends only on $u \triangleq i+x$ and s, then we shall define

$$j(u,s) = j(i,x,s), \quad u = i+x \tag{1}$$

on $U = \{u = \text{integer}; \min_I a(i) \leqq u \leqq N+\max_I b(i)\}$. Throughout the paper if there exists only one kind of disturbance $s = s^*$, in other words, $q_{s^*} = 1$, then for simplicity let $j(i,x) = j(i,x,s^*)$ and $j(u) = j(u,s^*)$.

The objective to the above model is to find the optimal decision strategy $D[t] = (D(t), D(t-1), \ldots, D(1))$ for all $t \geqq 1$, or a sequence of optimal decision rules $D(t')$ at points in time $t' = t, t-1, \ldots, 1$, maximizing the total expected reward gained over times t to 0 under a given terminating condition of the process. Now we shall let $v(i,t,\theta)$ be the maximum of a total expected reward starting with state i at time $t \geqq 1$, given an observation θ. And let the expectation of $v(i,t,\theta)$ as to θ be denoted by

$$v_i(t) = E(v(i,t,\theta)) \tag{2}$$

which means the maximum total expected reward before an observation. Then the maximum total expected reward after an observation θ is drawn and an action x is chosen is provided by

$$v(i,x,t,\theta) = g(i,x,\theta) + \beta \sum_s q_s v_{j(i,x,s)}(t-1) \qquad t \geqq 1 \tag{3}$$

where $\beta (0 < \beta \leq 1)$ represents a discount factor and where $v_i(0)$ is assumed to be uniquly determined by the given terminating condition. Then clearly we have from the very definition of $v(i,t,\theta)$

$$v(i,t,\theta) = \max_{A(i)} v(i,x,t,\theta) \tag{4}$$

Now by the term k-partition of observation space Ω we shall mean a set of k mutually exclusive subsets of Ω such that the entire union of them equals Ω. It is easy to see that a set of the $k(i)(=b(i)-a(i)+1)$ subsets of Ω defined below provides a $k(i)$-partition of Ω.

$$\begin{aligned} D(i,a(i),t) &= \{\theta; v(i,a(i),t,\theta) \geq v(i,r',t,\theta) \text{ for } a(i) < r' \leq b(i)\} \\ D(i,x,t) &= \{\theta; v(i,x,t,\theta) > v(i,r,t,\theta),\ v(i,x,t,\theta) \geq v(i,r',t,\theta) \\ &\qquad \text{for } a(i) \leq r < x < r' \leq b(i)\} \qquad a(i) < x < b(i) \\ D(i,b(i),t) &= \{\theta; v(i,b(i),t,\theta) > v(i,r,t,\theta)\ a(i) \leq r < b(i)\} \end{aligned} \tag{5}$$

Here notice that for a given x^* in $A(i)$ the $D(i,x^*,t)$ provides a set of θ for which the maximum of $v(i,x,t,\theta)$ on $A(i)$ is attained at $x = x^*$. This implies that the $k(i)$ subsets $D(i,x,t)$'s completely characterize an optimal decision rule, attaining the maximum of the right hand of (4). In other words, an optimal decision rule is that if an observation θ obtained at time t when in state i is contained in $D(i,x,t)$, then take the action x in $A(i)$. From now on we shall refer to the above $k(i)$-partition as an optimal partition. Then clearly (4) can be written as, using the above optimal partition

$$v(i,t,\theta) = \sum_{A(i)} v(i,x,t,\theta) I(\theta; D(i,x,t)) \tag{6}$$

where $I(x;X)$ represents an indicator function in which X and x are a set and its element, respectively, that is, if x is in X, it takes value 1, otherwise 0.

Before proceeding to further discussions we shall provide some definitions as to a concavity and convexity of a discrete function, which are notions usually defined only when its domain is convex.

DEFINITION 2 Consider a discrete function $f(x)$ defined on a discrete domain $K = \{0,1,\ldots,k\}$ with a nonnegative integer k, which is not convex. Then

(a) $f(x)$ is said to be concave(convex) if $f(\lambda x+(1-\lambda)x') \geqq (\leqq) \lambda f(x) + (1-\lambda)f(x')$ for any two integers x and x' in K and for any real number λ in $(0,1)$ such as $\lambda x + (1-\lambda)x'$ in K. The function is said to be linear if it is concave as well as convex.

(b) For any real number x in the interval $(K) = (0,k)$, define $\bar{f}(x) = (x-[x])f([x]+1) + (1-x+[x])f([x])$ where $[\ ]$ is Gauss's symbol. The $\bar{f}(x)$ is a continuous function generated from combining all two adjacent points $(x,f(x))$ and $(x+1,f(x+1))$ by a straight line. We shall call the continuous function a continuous form of the discrete function $f(x)$.

It is clear that the above definition 2a is equivalent to a decreasingness (increasingness) of a difference $f(x)-f(x-1)$ and moreover that if $f(i)$ is concave (convex), so also is $\bar{f}(i)$, and vice versa. However it will be easily understood that ordinary definitions of concavity and convexity cannot be applied for a discrete function of two or more variables. Thereupon we shall provide the following definition:

DEFINITION 3 Let $\bar{f}(i,x)$ be a continuous form of a discrete function $f(i,x)$ with respect to x (on $A(i)$) for each fixed i (on I). For any two integers i and i' in I, for any two real numbers x in $[A(i)]$ and x' in $[A(i')]$, and for any real number λ in $[0,1]$ such as $\lambda i+(1-\lambda)i'$ in I, if $\bar{f}(\lambda i+(1-\lambda)i', \lambda x+(1-\lambda)x') \geqq (\leqq)\ \lambda\bar{f}(i,x) + (1-\lambda)\bar{f}(i',x')$, then $f(i,x)$ is said to be concave(convex) with respect to i and x. Furthermore $f(i,x)$ is said to be strictly concave (strictly convex) if the above inequality holds strictly, or without equality sign.

The next two lemmas can be easily verified.

LEMMA 1 If $a(i)$ and $b(i)$ are, respectively, convex and concave, then $A(i)$ is concave, and vice versa.

LEMMA 2 Consider a discrete function $v(x)$ defined on $A = \{0,1,\ldots,k\}$ where k is a positive integer, and let $v = \max_A v(x)$.

(a) If $v(x)$ is concave, then v can be expressed as

$$v = v(0) + \sum_{A'}[\Delta v(x)]^+ = v(k) + \sum_{A'}[-\Delta v(x)]^+$$

where $A' = \{1,2,\ldots,k\}$, $\Delta v(x) = v(x) - v(x-1)$ on A', and $[y]^+ = \max\{y,0\}$ for any real number y. When $k = 0$, the terms of a summation in the above expressions vanish, that is, $v = v(0)$. That a given x in A is the smallest value attaining the v is equivalent to $\Delta v(x) > 0 \geqq \Delta v(x+1)$ where $\Delta v(0) = +\infty$ and $\Delta v(k+1) = -\infty$. If $v(x)$ is strictly concave, then the maximizing x is unique.

(b) If $v(x)$ is convex, then we have

$$v = \max\{v(0),\ v(k)\} = v(0) + [v(k)-v(0)]^+ = v(k) + [v(0)-v(k)]^+$$

where if $0 \geqq v(k) - v(0)$, $v = v(0)$, otherwise $v = v(k)$.

Now suppose that $v(i,x,t,\theta)$ is concave with respect to x for any i, any t, and all θ. Then by use of lemma 2a (2) can be transformed into

$$
\begin{aligned}
v_i(t) &= E(v(i,a(i),t,\theta)) + \sum_{A'(i)} E([\Delta v(i,x,t,\theta)]^+), \text{ or} \\
&= E(v(i,b(i),t,\theta)) + \sum_{A'(i)} E([-\Delta v(i,x,t,\theta)]^+)
\end{aligned}
\tag{7}
$$

where $A'(i) = \{x = \text{integer}; a(i) < x \leqq b(i)\}$ and $\Delta v(i,x,t,\theta) = v(i,x,t,\theta) - v(i,x-1,t,\theta)$ for x in $A'(i)$. Notice that if $a(i) = b(i)$, the second terms of (7) vanish. Then it follows from lemma 2a that a set of θ for which a specific x gives the smallest value maximizing the $v(i,x,t,\theta)$ over $A(i)$ is given by the subset of Ω, $\{\theta; \Delta v(i,x,t,\theta) > 0 \geqq \Delta v(i,x+1,t,\theta)\}$, where $\Delta v(i,a(i),t,\theta) = +\infty$ and $\Delta v(i,b(i)+1,t,\theta) = -\infty$. By making use of a concavity of $v(i,x,t,\theta)$ with respect to x, it is easy to check that the above k(i) subsets becomes equal to $D(i,x,t)$, that is,

$$
D(i,x,t) = \{\theta; \Delta v(i,x,t,\theta) > 0 \geqq \Delta v(i,x+1,t,\theta)\}, \quad x \text{ in } A(i) \tag{8}
$$

If $v(i,x,t,\theta)$ is strictly concave with respect to x for any i, any t, and any θ, then the maximizing x, or optimal action x is unique for the i, t, and θ.

On the other hand if $v(i,x,t,\theta)$ is convex with respect to x for any i, any t, and all θ, then from lemma 2b (2) becomes

$$v_i(t) = E(v(i,a(i),t,\theta)) + E([V(i,t,\theta)]^+), \text{ or}$$
$$= E(v(i,b(i),t,\theta)) + E([-V(i,t,\theta)]^+) \qquad (9)$$

$$V(i,t,\theta) = v(i,b(i),t,\theta) - v(i,a(i),t,\theta) \qquad (10)$$

,and the optimal partition is provided by

$$D(i,a(i),t) = \{\theta; 0 \geqq V(i,t,\theta)\}, \quad D(i,b(i),t) = \{\theta; V(i,t,\theta) > 0\} \quad (11)$$

, which are, in other words, the sets of θ for which optimal are the choices of actions a(i) and b(i), respectively. This implies that an intermediate action, or an action x such as $a(i) < x < b(i)$, is not taken at all under the optimal decision rule. Mathematically this may be treated by setting D(i,x,t) to be an empty set for $a(i) < x < b(i)$. From now on we shall refer to (7) and (9) as fundamental equations.

REMARK 1 Inequality signs $\geqq$ and $>$ in the above optimal partitions (8) and (11) may be always exchanged in an arbitraly way as far as a mutual exclusiveness of D(i,x,t)'s is not broken.

Now in general let v(i,x), j(i,x), and v_i be functions defined on I and A(i) where j(i,x) is assumed to take only integers on I. Then define $v(i) = \max_{A(i)} v(i,x)$ and $vj(i,x) = v_{j(i,x)}$. Below by symbols $\nearrow$, $\swarrow$, $\cap$, $\cup$, and L we shall mean, respectively, increasingness, decreasingness, concavity, convexity, or linearity which a function in question has. For a function of two variables i and x, we shall attach subscripts i and/or x to the above symboles as, for instance, $\nearrow_i$, $\cup_{i,x}$, ... in order to show with respect to which of the two variables the structural forms of the function are specified. Then we have the next lemma where $A(i)$ is said to be increasing (decreasing) if $A(i') \subset (\supset) A(i)$ for any $i' \leqq i$ in I.

LEMMA 3 The statements below, (a) to (j), are all true.

	$v(i,x)$	$A(i)$	$\rightarrow$	$v(i)$
(a)	$\nearrow_i$	$\nearrow$	$\rightarrow$	$\nearrow$
(b)	$\swarrow_i$	$\swarrow$	$\rightarrow$	$\swarrow$
(c)	$\cap_{i,x}$	$\cap$	$\rightarrow$	$\cap$
(d)	$\cup_{i,x}$	L	$\rightarrow$	$\cup$

	$j(i,x)$	v_i	$\rightarrow$	$vj(i,x)$
(e)	$\cap_{i,x}$	$\cap, \nearrow$	$\rightarrow$	$\cap_{i,x}$
(f)	$\cup_{i,x}$	$\cap, \swarrow$	$\rightarrow$	$\cap_{i,x}$
(g)	$L_{i,x}$	$\cap$	$\rightarrow$	$\cap_{i,x}$
(h)	$\cap_{i,x}$	$\cup, \swarrow$	$\rightarrow$	$\cup_{i,x}$
(i)	$\cup_{i,x}$	$\cup, \nearrow$	$\rightarrow$	$\cup_{i,x}$
(j)	$L_{i,x}$	$\cup$	$\rightarrow$	$\cup_{i,x}$

PROOF (a) For any two integers i' and i in I such as $i' \leqq i$, we get $v(i) \geqq \max_{A(i')} v(i,x) \geqq \max_{A(i')} v(i',x) = v(i')$. Similarly for (b). (c) First notice that we have $v(i) = \max_{x \in [A(i)]} \bar{v}(i,x)$. Then for any two integers i and i' in I and for any real number λ in $[0,1]$ such as $\lambda i+(1-\lambda)i'$ in I, we have $v(\lambda i+(1-\lambda)i')$

$$= \max_{x \in [A(\lambda i+(1-\lambda)i')]} \bar{v}(\lambda i+(1-\lambda)i',x)$$

$$\geqq \max_{x \in \lambda[A(i)]+(1-\lambda)[A(i')]} \bar{v}(\lambda i+(1-\lambda)i',x)$$

$$= \max_{y \in [A(i)],\, y' \in [A(i')]} \bar{v}(\lambda i+(1-\lambda)i',\lambda y+(1-\lambda)y')$$

$$\geqq \max_{y \in [A(i)],\, y' \in [A(i')]} \{\lambda \bar{v}(i,y) + (1-\lambda)\bar{v}(i',y')\}.$$

$= \lambda \max_{y \in [A(i)]} \bar{v}(i,y) + (1-\lambda) \max_{y' \in [A(i')]} \bar{v}(i',y')$

$= \lambda v(i) + (1-\lambda) v(i')$ (refer to Bellman(1975))

(d) We have $v(i) = \max\{v(i,a(i)), v(i,b(i))\} = \max\{\bar{v}(i,a(i)), \bar{v}(i,b(i))\}$. Hence for any i, i', and λ such as defined in the proof of (c) above, we have $v(\lambda i+(1-\lambda)i')$

$= \max\{\bar{v}(\lambda i+(1-\lambda)i', \lambda a(i)+(1-\lambda)a(i')), \bar{v}(\lambda i+(1-\lambda)i', \lambda b(i)+(1-\lambda)b(i'))\}$

$\leqq \max\{\lambda\bar{v}(i,a(i))+(1-\lambda)\bar{v}(i',a(i')), \lambda\bar{v}(i,b(i))+(1-\lambda)\bar{v}(i',b(i'))\}$

$\leqq \lambda\max\{\bar{v}(i,a(i)), \bar{v}(i,b(i))\} + (1-\lambda)\max\{\bar{v}(i',a(i')), \bar{v}(i',b(i'))\}$

$= \lambda v(i) + (1-\lambda)v(i')$, noticing that $\max\{A+A',B+B'\} \leqq \max\{A,B\} + \max\{A',B'\}$ for any real numbers A,A',B, and B'.

(e) Notice that $vj(i,x) = \bar{v}_{j(i,x)}$ for any integers i in I and x in A(i) where $\bar{v}_i$ is a continuous form of v_i with respect to i. Hence for any i, i', and λ such as defined in the above (c) and for any real numbers x and x' in the intervals [A(i)] and [A(i')], respectively, we have $vj(\lambda i+(1-\lambda)i', \lambda x+(1-\lambda)x') = \bar{v}_{j(\lambda i+(1-\lambda)i', \lambda x+(1-\lambda)x')}$

$\geqq \bar{v}_{\lambda j(i,x)+(1-\lambda)j(i',x')} \geqq \lambda\bar{v}_{j(i,x)} + (1-\lambda)\bar{v}_{j(i',x')}$

$= \lambda vj(i,x) + (1-\lambda)vj(i',x')$. Similarly for (f) to (j). Q.E.D.

Now define $vj(i,x,s)(t-1) = v_{j(i,x,s)}(t-1)$, and let $\bar{v}(i,x,t,\theta)$, $\bar{g}(i,x,\theta)$, and $\overline{vj}(i,x,s)(t-1)$ be continuous forms with respect to i and x. Then clearly we get

$$\bar{v}(i,x,t,\theta) = \bar{g}(i,x,\theta) + \sum_s q_s \overline{vj}(i,x,s)(t-1) \tag{12}$$

Applying lemma 3 to (2), (3), and (4) with attention to the just above things yields immediately the next theorem.

THEOREM 1 The six statements below are all true for <u>any</u> $t \geq 1$.

	$g(i,x,\theta)$	$j(i,x,s)$	$v_i(t-1)$	$A(i)$	$\rightarrow$	$v(i,x,t,\theta)$	$v_i(t)$
(a)	$\cap_{i,x}$	$\cap_{i,x}$	$\cap, \nearrow$	$\cap$	$\rightarrow$	$\cap_{i,x}$	$\cap$
(b)	$\cap_{i,x}$	$\cup_{i,x}$	$\cap, \swarrow$	$\cap$	$\rightarrow$	$\cap_{i,x}$	$\cap$
(c)	$\cap_{i,x}$	$L_{i,x}$	$\cap$	$\cap$	$\rightarrow$	$\cap_{i,x}$	$\cap$
(d)	$\cup_{i,x}$	$\cap_{i,x}$	$\cup, \swarrow$	L	$\rightarrow$	$\cup_{i,x}$	$\cup$
(e)	$\cup_{i,x}$	$\cup_{i,x}$	$\cup, \nearrow$	L	$\rightarrow$	$\cup_{i,x}$	$\cup$
(f)	$\cup_{i,x}$	$L_{i,x}$	$\cup$	L	$\rightarrow$	$\cup_{i,x}$	$\cup$
(g)	$\nearrow_i$	$\nearrow_i$	$\nearrow$	$\nearrow$	$\rightarrow$	$\nearrow_i$	$\nearrow$
(h)	$\swarrow_i$	$\nearrow_i$	$\swarrow$	$\swarrow$	$\rightarrow$	$\swarrow_i$	$\swarrow$

From the theorem the next corollary can be easily derived by induction.

COROLLARY 1 . The eight statements below are all true for <u>all</u> $t \geq 0$.

	$g(i,x,\theta)$	$j(i,x,s)$	$v_i(0)$	$A(i)$	$\rightarrow$	$v(i,x,t,\theta)$	$v_i(t)$
(a')	$\cap_{i,x}, \nearrow_i$	$\cap_{i,x}, \nearrow_i$	$\cap, \nearrow$	$\cap, \nearrow$	$\rightarrow$	$\cap_{i,x}, \nearrow_i$	$\cap, \nearrow$
(b')	$\cap_{i,x}, \swarrow_i$	$\cup_{i,x}, \nearrow_i$	$\cap, \swarrow$	$\cap, \swarrow$	$\rightarrow$	$\cap_{i,x}, \swarrow_i$	$\cap, \swarrow$
(c')	$\cap_{i,x}$	$L_{i,x}$	$\cap$	$\cap$	$\rightarrow$	$\cap_{i,x}$	$\cap$
(d')	$\cup_{i,x}, \swarrow_i$	$\cap_{i,x}, \nearrow_i$	$\cup, \swarrow$	L, $\swarrow$	$\rightarrow$	$\cup_{i,x}, \swarrow_i$	$\cup, \swarrow$
(e')	$\cup_{i,x}, \nearrow_i$	$\cup_{i,x}, \nearrow_i$	$\cup, \nearrow$	L, $\nearrow$	$\rightarrow$	$\cup_{i,x}, \nearrow_i$	$\cup, \nearrow$
(f')	$\cup_{i,x}$	$L_{i,x}$	$\cup$	L	$\rightarrow$	$\cup_{i,x}$	$\cup$
(g')	$\nearrow_i$	$\nearrow_i$	$\nearrow$	$\nearrow$	$\rightarrow$	$\nearrow_i$	$\nearrow$
(h')	$\swarrow_i$	$\nearrow_i$	$\swarrow$	$\swarrow$	$\rightarrow$	$\swarrow_i$	$\swarrow$

The next things should be noted in identifying the concavity or convexty of the $g(i,x,\theta)$ and $j(i,x,s)$ for each example in application of section 5.

1. If any two functions of i and x are concave with respect to i and x, then so also is the sum of them.

2. Let $f(i)$ and $g(x)$ are concave, respectively, in i on I and x on $A(i)$. Then both the functions may be regarded as being concave with respect to i and x.

3. A function of z, $[z]^+$, is concave.

Consequently it follows that $i+x$, $i+nx$, $[i+nx-s]^+$, $[i-s]^+ + x$ and so on, where n and s are integers, are all concave with respect to i and x. It is clear that the above 1. and 2. are also true for a convexty.

In case that θ is an one-dimensional vector where let its element be denoted by θ, the corollary below can be frequently applied if readers wish to represent a subset $D(i,x,t)$ of R^1 in a form of interval. Now define

$$\begin{aligned} d(i,x,t) &= \sup\{\theta;\Delta v(i,x,t,\theta) \geq 0\} \quad \text{if } v(i,x,t,\theta)\cap_x \text{ for all } \theta \\ d(i,x) &= \sup\{\theta; V(i,t,\theta) \geq 0\} \quad \text{if } v(i,x,t,\theta)\cup_x \text{ for all } \theta \end{aligned} \tag{14}$$

COROLLARY 2 Suppose that an observation is an one-dimensional vector. Then

(a) Suppose that $v(i,x,t,\theta)$ is concave with respect to x for all θ. If $\Delta v(i,x,t,\theta)$ is increasing (decreasing) in θ, then $d(i,x,t)$ is also increasing (decreasing) in x and we have $D(i,x,t) = (d(i,x,t), d(i,x+1,t)]$ $([d(i,x+1,t), d(i,x,t)))$ where $d(i,a(i),t) = -\infty (+\infty)$ and $d(i,b(i)+1,t) = +\infty (-\infty)$. Furthermore with respect to i or t, $d(i,x,t)$ has a monotonicity in the opposite direction to (in the same direction as) a direction of the monotonicity of $\Delta v(i,x,t,\theta)$ with respect to i or t.

(b) Assume that $v(i,x,t,\theta)$ is convex with respect to x for all θ. Then if $V(i,t,\theta)$ is increasing (decreasing) in θ, we get $D(i,a(i),t) = (-\infty, d(i,t)]$ $([d(i,t), +\infty))$, and the same thing as the above (a) holds as to a direction of monotonicity of $d(i,t)$ and $V(i,t,\theta)$.

2. INFINITE TIME HORIZON VERSION

Here an infinite time horizon version of the model in the previous sections is discussed. For this at first we shall reveal the relationship between the model and a Markovian decision process under the assumption that $v(i,x,t,\theta)$ is concave with respect to x for all i, all t, and all θ. For any i and j in I and for any x in A(i) define $S(i,x,j) = \{s; j=j(i,x,s)\}$, which is a set of s for which the next state becomes j if an action x is taken when in state i. Then the probability of the next state being j can be expressed as $p_{ij}(x) = \sum_{S(i,x,j)} q_s$ if the present state is i and an action x is taken. When using this, the second term of the right-hand of (3) can be converted into $\beta \sum_{j \in I} p_{ij}(x) v_j(t-1)$. Now (2) can be expressed using (6)

$$v_i(t) = \sum_{A(i)} E(v(i,x,t,\theta) I(\theta; D(i,x,t))) \tag{15}$$

Let $\mathcal{B}(i)$ be a space of all k(i)-partitions of Ω and $\mathcal{B}$ be a product of them on I, that is, $\mathcal{B} = \times_{i=0}^{N} \mathcal{B}(i)$. Furthermore let an element of the product be denoted by $B = (B(0), B(1), \ldots, B(N))$ where B(i) is a k(i)-partition, in other words, an element of $\mathcal{B}(i)$. If elements of B(i) are denoted by B(i,x)'s, (subsets of Ω) that is, $B(i) = \{B(i,x); a(i) \leqq x \leqq b(i)\}$, a decision rule for each state i may be prescribed as follows: If an observation drawn in state i, θ, belongs to a subset B(i,x), then take the action x. Noticing in general $I(x;X) = I(x; X \cap Y) + I(x; X \cap Y^c)$ for any sets X and Y, we have for any k(i)-partition B(i)

$\sum_{A(i)}E(v(i,x,t,\theta)I(\theta;D(i,x,t))) - \sum_{A(i)}E(v(i,x,t,\theta)I(\theta;B(i,x))) =$
$\sum_{x,r\in A(i)}E((v(i,x,t,\theta) - v(i,r,t,\theta))I(\theta;D(i,x,t)\cap B(i,r))) \ldots (*)$.

Now for any fixed x in A(i) and for any specific θ in D(i,x,t) (see (8)) we have $\Delta v(i,x',t,\theta) > 0 \geqq \Delta v(i,x'',t,\theta)$ for any two integers x' and x'' in A(i) such as $x' \leqq x < x''$ because $\Delta v(i,x,t,\theta)$ is decreasing in x. The inequality yields $v(i,x,t,\theta) \geqq v(i,r,t,\theta)$ for any given θ in D(i,x,t) and all r in A(i). Consequently we have $(*) \geqq 0$. This means that the optimal partition which was defined by (8), that is, $D(i,x) = \{D(i,x,t);a(i) \leqq x \leqq b(i)\}$ exists within (i). Therefore it follows that (15) may be rewritten as

$$v_i(t) = \max_{\mathcal{B}(i)}\sum_{A(i)}E(v(i,x,t,\theta)I(\theta;B(i,x))) \qquad (16)$$

Now since the right hand of (16) is maximized by the D(i,t) and D(i,x,t) takes the form of (8), the above maximization problem over $\mathcal{B}(i)$ may be reduced to a maximization problem over a subclass $\mathcal{H}(i)$ of $\mathcal{B}(i)$ which is a space of all k(i)-partitions of such a form that

$$H(x) = \{\theta;h(x,\theta) > 0 \geqq h(x+1,\theta)\} \quad x \text{ in } A(i) \qquad (17)$$

where $h(x,\theta)$ is a decreasing function of x with $h(a(i),\theta) = +\infty$ and $h(b(i)+1,\theta) = -\infty$. Let $H = \{H(x);x \text{ in } A(i)\}$. Then (16) with $\mathcal{H}(i)$ instead of $\mathcal{B}(i)$ can be transformed into by substituting (3)

$$v_i(t) = \max_{\mathcal{H}(i)}\{R_i(H) + \beta\sum_j Q_{ij}(H)v_j(t-1)\} \qquad (18)$$
$$R_i(H) = \sum_{A(i)}E(g(i,x,\theta)I(\theta;H(x))) \qquad (19)$$
$$Q_{ij}(H) = \sum_{A(i)}p_{ij}(x)E(I(\theta;H(x))) \qquad (20)$$

Here notice that the above (19) and (20) corresponds to, respectively, an immediate reward and transition probability which are usually definded in a Markovian decision process. Arranging them by substituting (17) into yields

$$R_i(H) = E(g(i,a(i),\theta)) + \sum_{A'(i)} E(\Delta g(i,x,\theta)I(h(x,\theta)>0)) \quad (21)$$

$$Q_{ij}(H) = p_{ij}(a(i)) + \sum_{A'(i)}(p_{ij}(x) - p_{ij}(x-1))E(I(h(x,\theta)>0)) \quad (22)$$

where $I(h(x,\theta)>0) \triangleq I(\theta;\{\theta;h(x,\theta)>0\})$.

Next we shall discuss an infinite time horizon version of our model under the assumption that the product $\mathcal{H} = \times_{i=0}^{N}\mathcal{H}(i)$ is finite. For example, if the distribution F of θ is of discrete type such that the number of θ with positive probability is finite, then clearly $\mathcal{H}(i)$, hence $\mathcal{H}$ can be set to be finite. First let us discuss the case of $\beta < 1$. Now let v_i be the total expected reward gained over an infinite time horizon associated with any stationary strategy H^{∞} where $H = (H(0),H(1),\ldots,H(N))$, an element of $\mathcal{H}$. And let the maximums of v_i over $\mathcal{H}$ for all i be attained at $H^* = (H^*(0),H^*(1),\ldots,H^*(N))$, where $H^*(i)$ is called a limiting optimal partition for state i. Below we shall show the procedure for determining the H^* using Howard's policy iteration algorithm. It will be easily understood that the v_0, v_1, ..., v_N associated with H^{∞} satisfy a system of equations for all i

$$v_i = R_i(H) + \beta\sum_j Q_{ij}(H)v_j \quad (23)$$

Arranging this by substituting (21) and (22) into yields

$$v_i = E(v(i,a(i),\theta)) + \sum_{A'(i)} E(\Delta v(i,x,\theta) I(h(x,\theta)>0)) \tag{24}$$

$$v(i,x,\theta) = g(i,x,\theta) + \beta \sum_j p_{ij}(x) v_j \tag{25}$$

where $\Delta v(i,x,\theta) = v(i,x,\theta) - v(i,x-1,\theta)$. Then a value determination operation in Howard's algorithm can be carried out by solving the system of equations (23) for any given H, and the subsequent policy improvement routine can be done by finding the $h(x,\theta)$ maximizing the right-hand of (24) for any given v_i under the condition that $h(x,\theta)$ is decreasing in x for all θ. Let v_i from subsequent two value determination operations be denoted by v_i' and v_i''. Then it has been already proved (Howard's (1960)) that we have always $v_i' \leqq v_i''$ for all i and that $v_i^* = v_i' = v_i''$ for all i are attained in a finite number of subsequent value determination operations where v_i^* is associated with the H^* .

Here it should be noted that in any policy improvement routine $v(i,x,\theta)$ is not always guaranteed to be concave with respect to x.

(a) Suppose that the $v(i,x,\theta)$ is concave with respect to x for all θ, given i. Then it is immediately seen from the lemma below (Its proof is quite easy) that the maximum of the right-hand of (24) is attained by

$$h(x,\theta) = \Delta v(i,x,t) \qquad x \text{ in } A'(i), \text{ all } \theta \tag{26}$$

LEMMA 4 $gI(h>0) \leqq gI(g>0) = [g]^+$ for any real functions g and h of θ.

(b) If $v(i,x,\theta)$ cannot be assured to be concave with respect to x, the maximum can be obtained by applying a dynamic programming to the maximization problem below. For each θ

$$\max \sum_{A'(i)} E(\Delta v(i,x,\theta) I(h(x,\theta)>0)) \tag{27}$$

subject to $h(x,\theta)$ decreasing in x for all θ

For the model with $\beta = 1$, theorems 6.17 and 6.18 in Ross(1970) can be applied. They can be rewritten as the theorem below for the model.

THEOREM 2 Let v_i be the maximum of a total expected reward over an infinite time horizon with $0 \leqq \beta < 1$ (i.e., $v_i \triangleq v_i^*$), and suppose that $v_i - v_0$ is bounded in β for all i. Then there exist u_i and g satisfying

$$g + u_i = \max_{\mathcal{H}}\{R_i(H) + \sum_j Q_{ij}(H)u_j\}$$

where $g = \lim_{\beta \to 1-}(1 - \beta)v_i$, which provides the maximum of a per-period expected reward over an infinite time horizon when $\beta = 1$. Furthermore we have $u_i = \lim_{n\to\infty}(v_i^{\beta(n)} - v_0^{\beta(n)})$ for some sequence $\beta(n)\to 1-$ where v_i^{β} represents v_i associated with a discount factor $\beta < 1$.

REMARK 2 (see Ross(1970, pp.147) The u_i inherits a structural form which v_i shares for all $\beta < 1$.

For example, if v_i is increasing in i for all $\beta < 1$, so also is u_i. The optimal stationary strategy $H^{*\infty}$, attaining the above g, can be obtained in the almost same way as in case of $\beta < 1$ (Howard(1960)). From now on in general a limit of a function f(t) as $t\to\infty$, denoted by f, will be simply referred to as "limit f".

3. TWO TYPES OF LINEAR MODELS

This section will apply the general discussions developed in the previous sections to two special cases below.

<u>case 1</u> $g(i,x,\theta) = g(i,x)\theta$ where θ is a scalar random variable with a finite expectation E and where $g(i,x)$ is either (a) strictly increasing or (b) strictly decreasing in x for all i. Here define $\Delta g(i,x) = g(i,x) - g(i,x-1)$.

<u>case 2</u> $g(i,x,\theta) = \theta_0+\theta_1+\theta_2+....+\theta_x$ for $0 \leqq x \leqq b(i)$, where $a(i) = 0$, $\theta_0 = 0$, and θ is an m-vector $\theta = (\theta_1,\theta_2,...,\theta_m)$ with $b(i) \leqq m$ such that θ_x is decreasing in x. Here let F_x and E_x denote the distribution function and expectation (or expectation operator), respectively, of a random variable θ_x.

First define

$$z(i,x,t) = \beta \sum_s q_s v_{j(i,x,s)}(t-1) \quad (= \beta \ {}_{j \in I} p_{ij} v_j(t-1)) \tag{28}$$

$$\Delta z(i,x,t) = z(i,x,t) - z(i,x-1,t)$$

$$c(i,x,t) = \begin{cases} -\Delta z(i,x,t)/\Delta g(i,x) & \text{for case 1} \\ -\Delta z(i,x,t) & \text{for case 2} \end{cases} \tag{29}$$

Then we have

$$v(i,x,t,\theta) = \begin{cases} g(i,x)\theta + z(i,x,t) & \text{for case 1} \\ \theta_0+\theta_1+...+\theta_x + z(i,x,t) & \text{for case 2} \end{cases} \tag{30}$$

DEFINITION 4 For case 1 define $T(c,h) = E((\theta-c)I(\theta-h>0))$, $T(c) = T(c,c)$ $(= E([\theta-c]^{+}))$, and $\tilde{T}(c) = c + T(c)$ for any real numbers c and h. Similarly for case 2 let $T_x(c,h) = E_x((\theta_x-c)I(\theta_x-h>0))$, $T_x(c) = T_x(c,c)$, $\tilde{T}_x(c) = c + T_x(c)$.

DEFINITION 5 Let θ^γ, $\gamma=0,\pm1,\pm2,\ldots$, be mass points of an one-dimensional discrete distribution where θ^γ is strictly increasing in γ. Then for any real number c let $<c>$ be the maximum mass points less than or equal to c, in other words, $<c> = \max\{\theta^\gamma;\theta^\gamma \leq c\}$. If θ^γ are all integers, then we have $<c> = [c]$.

LEMMA 5 $T(c) \geq T(c,h)$ for any real numbers c and h. If F is a discrete function, then we have $T(c) = T(c,<c>)$ and $\{\theta;c < \theta \leq c'\} = \{\theta;<c> < \theta \leq <c'>\}$ for any two real numbers c and c'.

PROOF The former half of the lemma is evident from $(\theta-c)I(\theta-c>0) \geq (\theta-c)I(\theta-h_x>0)$ (see lemma 4). The latter one is clear from slightly examining the definition of $<c>$.

Then we have the next theorem.

THEOREM 3 In both cases suppose that $v(i,x,t,\theta)$ is concave with respect to x for any i, any t, and all θ. Then

(a) In case 1a, (7) becomes

$$v_i(t) = g(i,a(i))E + z(i,a(i),t) + \sum_{A'(i)} \Delta g(i,x)T(c(i,x,t)) \quad (31)$$

and (8) is reduced to an interval

$$D(i,x,t) = \{\theta; c(i,x,t) < \theta \leqq c(i,x+1,t)\} \qquad x \text{ in } A(i) \tag{32}$$

where $c(i,a(i),t) = -\infty$ and $c(i,b(i)+1,t) = +\infty$.

(b) In case 1b, (7) can be arranged as,

$$v_i(t) = g(i,b(i))E + z(i,b(i),t) - \textstyle\sum_{A'(i)} \Delta g(i,x) T(c(i,x,t)) \tag{33}$$

and (8) is converted into an interval

$$D(i,x,t) = \{\theta; c(i,x+1,t) \leqq \theta < c(i,x,t)\} \qquad x \text{ in } A(i) \tag{34}$$

where $c(i,a(i),t) = +\infty$ and $c(i,b(i)+1,t) = -\infty$.

(c) In case 2, (7) is transformed into

$$v_i(t) = z(i,0,t) + \textstyle\sum_{A'(i)} T_x(c(i,x,t)) \tag{35}$$

and (8) is reduced to

$$D(i,x,t) = \{\theta; c(i,x,t) < \theta_x,\ \theta_{x+1} \leqq c(i,x+1,t)\} \quad x \text{ in } A(i) \tag{36}$$

where $c(i,a(i),t) = -\infty$ and $c(i,b(i)+1,t) = +\infty$.

(d) Suppose that F is a continuous distribution. Then in cases 1a, 1b, and 2, $c(i,x,t)$ are, respectively, increasing, decreasing, and increasing in x, hence (32) and (34) are reduced to intervals, respectively, $(c(i,x,t), c(i,x+1,t)]$ and $[c(i,x+1,t), c(i,x,t))$. When F is discrete, it is for $\langle c(i,x,t)\rangle$ that the above monotonicities are true, and (32) and (34) are converted into, respectively, $(\langle c(i,x,t)\rangle, \langle c(i,x+1,t)\rangle]$ and $[\langle c(i,x+1,t)\rangle, \langle c(i,x,t)\rangle)$.

PROOF (a,b,c) (31), (33), and (35) can be directly derived from (7). (32), (34), and (36) can be immediately obtained from (8).

(d) We shall prove this for only case 1a with a continuous distribution function F. For simplicity let the set D(i,x,t) be written as $D_x = \{c_x < \theta \leqq c_{x+1}\}$, $a \leqq x \leqq b$ where $c_a = -\infty$ and $c_{b+1} = +\infty$. If $c_{a+2} < c_{a+1}$, it must be from the mutual exclusiveness of D_x's that $c_b \leqq c_{b-1} \leqq \cdots \leqq c_{a+3} \leqq c_{a+2}$. Therefore we have $c_b < c_{a+1}$, which yields a contradiction that D_a and D_b are not exclusive. Thus we must have $c_{a+1} \leqq c_{a+2}$. Similarly it can be verified that $c_{a+2} \leqq c_{a+3}$, $c_{a+3} \leqq c_{a+4}$, and so on. Thus we have an increasingness of c_x in x, or of c(i,x,t) in x. Similarly for case of a discrete distribution F. Q.E.D.

THEOREM 4 Suppose that $v(i,x,t,\theta)$ is not always concave with respect to x. Then

(a) In case 1a(1b), if c(i,x,t) is increasing (decreasing) in x on A'(i) for any i and any t, then it follows that an optimal partition (5) is provided by (32) ((34)) and that (31) ((33)) holds.

(b) In case 2 if c(i,x,t) is increasing in x on A'(i) for any i and any t, then it follows that an optimal partition (5) is given by (36) and that (35) becomes true.

PROOF (a) For any two sequences $m_1, m_2, \ldots, m_x$ and $n_1, n_2, \ldots, n_x$ where m_x are all either positive or negative, if n_x/m_x is increasing (decreasing) in x, then we have

$$n_1/m_1 \leqq (\geqq) \ (n_1+n_2+\ldots+n_x)/(m_1+m_2+\ldots+m_x) \leqq (\geqq) \ n_x/m_x.$$

Let us prove only case 1a using the inequality. For simplicity let z(i,x,t), g(i,x), c(i,x,t), D(i,x,t), a(i), and b(i) be denoted by z_x, g_x, c_x, D_x, a, and b, respectively, and define $c_{xs} = c_{sx} =$

$(z_s - z_x)/(g_x - g_s)$ for $x \neq s$. Then clearly $c_x = c_{x,x-1} = c_{x-1,x}$. Letting $n_x = z_{x-1} - z_x$ and $m_x = g_x - g_{x-1}$, we have

$$c_{xs} = (n_{s+1} + n_{s+2} + \ldots + n_x)/(m_{s+1} + m_{s+2} + \ldots + m_x) \qquad s < x$$

$$c_{xs'} = (n_{x+1} + n_{x+2} + \ldots + n_{s'})/(m_{x+1} + m_{x+2} + \ldots + m_{s'}) \qquad x < s'$$

Now in the case the optimal partition (5) can be arranged as follows: $D_a = \{\theta; (g_a - g_{s'})\theta \geqq z_{s'} - z_a$ for $a < s' \leqq b\}$, $D_x = \{\theta; (g_x - g_s)\theta > z_s - z_x$ and $(g_x - g_{s'})\theta \geqq z_{s'} - z_x$ for $a \leqq s < x < s' \leqq b\}$ for $a < x < b$, and $D_b = \{\theta; (g_b - g_s) > z_s - z_b$ for $a \leqq s < b\}$. These can be transformed into $D_a = \{\theta; -\infty < \theta \leqq c_{as'}$ for $a < s' \leqq b\}$, $D_x = \{\theta; c_{xs} < \theta \leqq c_{xs'}$ for $a \leqq s < x < s' \leqq b\}$ for $a < x < b$, and $D_b = \{\theta; c_{bs} < \theta < +\infty$ for $a \leqq s < b\}$. Now since $c_x = n_x/m_x$ is increasing in x from the assumption, we get $c_{xs} \leqq n_x/m_x = c_{x,x-1} = c_x$ for $s < x$ and $c_{xs'} \geqq n_{x+1}/n_{x+1} = c_{x,x+1} = c_{x+1}$ for $x < s'$. Hence D_x's can be reduced to the sets $D_a = \{\theta; -\infty < \theta \leqq c_{a+1}\}$, $D_x = \{\theta; c_x < \theta \leqq c_{x+1}\}$ for $a < x < b$, and $D_b = \{\theta; c_b < \theta \leqq +\infty\}$. Thus (32) becomes true. Rearranging (15) by substituting (32) into yields (31).

(b) is clear from $\Delta v(i,x,t,\theta) = \theta_x - c(i,x,t)$, which becomes concave with respect to x.

REMARK 3 If A(i) is a set of only two elements, then $v(i,x,t,\theta)$ can be deemed concave as well as convex on A(i) and c(i,x,t) may be regarded as being increasing as well as decreasing on A'(i). Therefore in the case it follows that monotonicity reqirements of $v(i,x,t,\theta)$ in theorem 3 and of c(i,x,t) in theorems 4 can be ignored.

In the proof of theorem 4a, a strict monotonicity of g(i,x) was essential. The corollary established below provides a necessary condition under which the theorem becomes true even when $g(i,x)$ is only monotone. For a function $g(i,x)$ which is increasing (decreasing) in x for all i define $v_i(t)$ and $v_i^h(t)$ to be the total expected rewards associated with, respectively, $g(i,x)$ and $g^h(i,x) \triangleq g(i,x) + hx$ ($\triangleq g(i,x) - hx$) where h is a positive real number such that $g^h(i,x)$ becomes strictly increasing (strictly decreasing) in x for all i. An existence of such h can be easily shown. Let $\mathcal{G}$ be now a space of functions $g(i,x)$ which are strictly increasing (strictly decreasing) in x for all i.

COROLLARY 3 In case 1a (case 1b) suppose that on , $c(i,x,t)$ is independent of $g(i,x)$ for all t and is increasing (decreasing) in x and that $v_i^h(t-1) \rightarrow v_i(t-1)$ as $h \rightarrow 0+$ for a fixed t for any function $g(i,x)$ increasing (decreasing) in x for all i. Then we have $v_i^h(t) \rightarrow v_i(t)$ as $h \rightarrow 0+$, and (31) ((33)) and (32) ((34)) provides, respectively, fundamental equations and an optimal partition for the process associated with the increasing (decreasing) function $g(i,x)$, using $c(i,x,t)$ defined for any $g(i,x)$ on .

PROOF Let us prove only case 1a. For any function $g(i,x)$ increasing in x for all i clearly we have $v^h(i,x,t,\theta) \triangleq g^h(i,x)\theta + z^h(i,x,t) \rightarrow v(i,x,t,\theta) = g(i,x)\theta + z(i,x,t)$ as $h \rightarrow 0+$, noticing $z^h(i,x,t) \triangleq \sum_s q_s v^h_{j(i,x,s)}(t-1) \rightarrow z(i,x,t) = \sum_s q_s v_{j(i,x,s)}(t-1)$ as $h \rightarrow 0+$. Thus it is follows from (7) that $v_i^h(t) \rightarrow v_i(t)$ as $h \rightarrow 0+$. Now since $c^h(i,x,t)$ is independent of h from the assumption, define $c(i,x,t) = c^h(i,x,t)$, which is increasing in x. Therefore from theorem 4a we have $v_i^h(t) =$

$g(i,a(i))E + ha(i)E + z^h(i,a(i),t) + \sum_{A'(i)}(\Delta g(i,x) + h)T(c(i,x,t))$. Here h approaching zero yields (31). In order to show that (32) provides an optimal partition for the process with the increasing $g(i,x)$, it will suffice to show that an arrangement of the right-hand of (15) by substituting (32) into yields the right-hand of (31). This substitution produces $g(i,a(i))E + z(i,a(i),t) + \sum_{A'(i)}E((\Delta g(i,x)\theta + \Delta z(i,x,t))I(\theta>c(i,x,t)))$. The last term of the expression can be replaced by $\sum_{A'(i)}\lim_{h\to 0+}E(((\Delta g(i,x)+h)\theta + \Delta z^h(i,x,t))I(\theta>c(i,x,t)))$, which can be converted into $\sum_{A'(i)}\lim_{h\to 0+}(\Delta g(i,x)+h)\ E((\theta-c(i,x,t))I(\theta>c(i,x,t))) = \sum_{A'(i)}\Delta g(i,x)T(c(i,x,t))$ by noticing $c(i,x,t) = c^h(i,x,t) = -\Delta z^h(i,x,t)/(\Delta g(i,x)+h)$ from the assumption. Consequently we have the right-hand of (31). Q.E.D.

Policy iteration algorithm

Here we shall show how a Howard's policy iteration algorithm is applied to an infinite time horizon version of the linear models. First let us discuss the case of $\beta < 1$. Here assume that a limit $v(i,x,\theta)$ associated with the optimal stationary strategy $H^{*\infty}$ has been proved to be concave with respect to x for all i and all θ. Then define, using the v_i associated with the $H^{*\infty}$

$$c(i,x) = \begin{cases} -\Delta z(i,x)/\Delta g(i,x) & \text{for case 1} \\ -\Delta z(i,x) & \text{for case 2} \end{cases} \tag{37}$$

$$z(i,x) = \sum_s q_s v_{j(i,x,s)} \tag{38}$$

Then in the last policy improvement routine where the optimality is attained, clearly we have $h(x,\theta) = \Delta v(i,x,\theta)$ due to the concavity assumption of $v(i,x,\theta)$, where $h(x,\theta) = \Delta g(i,x)(\theta-c(i,x))$ for case 1 and $h(x,\theta) = \theta_x-c(i,x)$ for case 2. From theorem 3d, if F is continuous, then the optimal partition $D(i,x)$ are given by the intervals $(c(i,x), c(i,x+1)]$ for csae 1a, $[c(i,x+1), c(i,x))$ for case 1b, and a set $\{\theta; c(i,x) < \theta_x, \theta_{x+1} \leqq c(i,x+1)\}$ for case 2. Here $c(i,x)$ is increasing in x for cases 1a and 2 and is decreasing in x for case 1b where $c(i,a(i)) = -\infty$ and $c(i,b(i)+1) = +\infty$ for cases 1a and 2 and $c(i,a(i)) = +\infty$ and $c(i,b(i)+1) = -\infty$ for case 1b. In case of a discrete F, the above discussions are true only if $c(i,x)$ is replaced by $\langle c(i,x)\rangle$. Then that an optimal partiton is given in the above forms implies that a k(i)-partition emploied in any policy improvement routine may be confined to $(h(x),h(x+1)]$ for case 1a, to $[h(x+1),h(x))$ for case 1b, and to $\{\theta; h(x) < \theta_x, \theta_{x+1} \leqq h(x+1)\}$ for case 2, where $h(x)$ is an increasing function with $h(a(i)) = -\infty$ and $h(b(i)+1) = +\infty$ for cases 1a and 2 and is a decreasing function with $h(a(i)) = +\infty$ and $h(b(i)+1) = -\infty$ for case 2. Here note that in case of a discrete F, the above $h(x)$ takes only values on a set of mass points. When the right-hand of (23) is rearranged by substituting the above $H(x)$ into, it is immediately seen that a policy improvement routine can be reduced to

$$
\begin{aligned}
&\text{for case 1a} \quad && \max \textstyle\sum_{A'(i)} \Delta g(i,x)T(c(i,x),h(x)) \\
& && \text{subject to } h(x) \text{ increasing in } x \\
&\text{for case 1b} \quad && \max \textstyle\sum_{A'(i)} -\Delta g(i,x)T(c(i,x),h(x)) \qquad (39) \\
& && \text{subject to } h(x) \text{ decreasing in } x \\
&\text{for case 2} \quad && \max \textstyle\sum_{A'(x)} T_x(c(i,x),h(x)) \\
& && \text{subject to } h(x) \text{ increasing in } x
\end{aligned}
$$

In cases 1a and 2 (case 1b), if $c(i,x)$ determined by v_i from any value determination operation is increasing (decreasing) in x, then clearly we have from lemma 5 a solution $h(x) = c(i,x)$ to the above maximization problems. On the contrary, if the above monotonicity of $c(i,x)$ cannot be garanteed, then the solution must be sought by use of a dynamic programming. Now define $E[b] = E(\theta I(\theta \leqq b))$ for any real number b. Then for case 1a we have from (19) and (20), noticing $(-\infty, h(x)] + H(x) = (-\infty, h(x+1)]$

$$R_i(H) = \sum_{A(i)} g(i,x)(E[h(x+1)] - E[h(x)]) \tag{40}$$

$$Q_{ij}(H) = \sum_{A(i)} p_{ij}(x)(F(h(x+1)) - F(h(x))) \tag{41}$$

In case 1b they are given by multipling -1 by the right-hands of the above two expressions. In case 2, using $\{\theta; \theta_x \leqq h(x)\} + H(x) = \{\theta; \theta_{x+1} \leqq h(x+1)\}$, we get

$$R_i(H) = \sum_{A(i)} (E_x - E_x[h(x)]) \tag{42}$$

$$Q_{ij}(H) = \sum_{A(i)} p_{ij}(x)(F_{x+1}(h(x+1)) - F_x(h(x))) \tag{43}$$

where let $F_0(.) = 0$, $F_{b(i)+1}(.) = 1$, $E_0 = 0$, and $E_0[.] = 0$. For case of $\beta = 1$, the above discussions are all true only if v_i is replaced by u_i.

4. SLIGHT EXTENSION TO A CASE THAT STATE AND ACTION SPACES ARE INTERVALS

The section will attempt to make a slight extension of our model to a case that both state and action spaces are given by finite closed intervals, respectively, $I = [0,N]$ and $A(i) = [a(i),b(i)]$ where N, a(i), and b(i) are all real numbers such as $N > 0$ and $a(i) \leqq b(i)$. Furthermore let a state transition law be given in a general form of $j(i,x,s)$, in which a disturbance s is a random variable with a given continuous distribution Q(s). Here we shall assume that $v_i(0)$, $g(i,x,\theta)$, and $j(i,x,s)$ are all any times differentiable with respect to i and x. Then (3) can be converted into

$$v(i,x,t,\theta) = g(i,x,\theta) + \beta\int v_{j(i,x,s)}(t-1)dQ(s) \tag{44}$$

LEMMA 6 Consider a differentiable function v(x) defined on a closed interval $A = [0,k]$ where k is a nonnegative real number, and let $v = \max_A v(x)$. (a) If v(x) is concave, then $v = v(0) + \int_A [dv(x)/dx]^+ dx = v(k) + \int_A [-dv(x)/dx]^+ dx$ where the values of $dv(x)/dx$ at $x = 0$ and $x = k$ represent, respectively, right and left differential coefficients. If there exists such a x that $dv(x)/dx = 0$ on A, then v is attained at the x. If $dv(x)/dx$ is nonnegative (nonpositive) for all x in A, then the maximizing x is given by $x = k$ $(x = 0)$. If v(x) is strictly concave, then the maximizing x is uniquely determined.

(b) If v(x) is convex, then exactly the same things as lemma 2b hold.

PROOF Evident

Now for any $t \geqq 1$ suppose that $v_i(t-1)$ is differentiable with respect to i. Then clearly $v(i,x,t,\theta)$ becomes also differenciable with respect to x. Furthermore if the $v(i,x,t,\theta)$ is concave with respect to x for all i and all θ, we have from the above lemma,

$$\begin{aligned} v_i(t) &= E(v(i,a(i),t,\theta)) + \int_{A(i)} E([\partial v(i,x,t,\theta)/\partial x]^+)dx, \text{ or} \\ &= E(v(i,b(i),t,\theta)) + \int_{A(i)} E([-\partial v(i,x,t,\theta)/\partial x]^+)dx \end{aligned} \qquad (45)$$

Now we shall slightly modify lemma 3, theorem 1, and corollary 1 by replacing all the discrete functions defined there by the corresponding continuous functions defined above. Then it will be immediately understood that the modified lemma 3 becomes also true. By using the lemma the colloraly below is easily verified.

COROLLARY 4 Under the modification above, statements (a) to (h) in theorem 1 become true. In addition if $v_i(t)$ is differentiable with respect to i for all t, then statements (a') to (h') in corollary 1 also hold.

5. SOME APPLICATIONS

5.0 Preliminaries

This section will demonstrate by taking some examples that our approach in the previous sections has a wide applicability to a certain class of decision processes which we encounter frequently in different managerial, economic situations. Some of these decision processes and its related topics have been already studied in detail by many authers where a calculus was emploied as a method for optimization or an optimality principle in a dynamic programming was applied in the very direct way. However, in this section readers will appreciate an usefullness of our approach in a sence that it brings about an even more systematic approach to them as compared with the conventional ones. The lemma below will be needed to inspect some properties of an optimal partition which characterizes an optimal decision strategy for problems discussed in the subsequent subsections.

LEMMA 7 Let a and b be functions of θ and c be a real number. Then

(a) $[a-c]^+$ is decreasing in c. $c+[a-c]^+$ and $[a-c]^+-[b-c]^+$ with $a \leqq b$ are increasing in c.

(b) We have $[a-c]^+ \leqq [a]^+ \leqq c+[a-c]^+$ for any nonnegative c and $a-b \leqq [a-c]^+-[b-c]^+ \leqq 0$ for any c if $a \leqq b$.

(c) $T(c)$ and $\tilde{T}(c)$ are, respectively, decreasing and increasing in c.

(d) If $F(0) = 0$, then $T(c) \leqq E \leqq \tilde{T}(c)$ for any nonnegative number c.

(e) We have $E_x - E_{x-1} \leqq \Delta T_x(c) \triangleq T_x(c) - T_{x-1}(c) \leqq 0$, and $\Delta T_x(c)$ is increasing in c.

PROOF (a) and (b) are clear, from which (c) and (d) are easily verified. An increasingness of $\Delta T_x(c)$ in c is clear from $\Delta T_x(c) = E((\theta_x-c)^+-(\theta_{x-1}-c)^+)$ where $\theta_x \leqq \theta_{x-1}$ and where E is an expectation operator of random vector θ.

Below we shall state in advance some things, as remarks, which will be convenient to be assumed, to be defined, and to be noticed throughout this section.

REMARK 4 Whenever an infinite time horizon process is discussed, set $v_i(0) = 0$ for all i and assume a finiteness of the observation space Ω. Then if θ is of one dimension, there always exists a sufficiently large number M such as $T(M) = 0$, hence $\tilde{T}(M) = M$.

REMARK 5 Very often there exists a case with multiple action spaces, each labeled by k. Let the set of the labels k be denoted by K which is countable, possibly infinite. Here assume that an action space with label k appears at each point in time with probability p_k, $\sum_k p_k = 1$. From now on, to all notations which may depend on k the label k shall be attached as a subscript like p_k, E_k, and so on. If a distribution function of an observation, immediate reward, and state transition law may depend on k, it naturally follows that both $v(i,x,t,\theta)$ and $v(i,x,\theta)$ may be also dependent on k. Then $v_i(t)$ must be rewritten as $v_i(t) = \sum_k p_k v_{ik}(t)$ where $v_{ik}(t) = E_k(v_k(i,t,\theta))$. In the case since $v_i(t)$ inherits structural forms (i.e., monotonicity, concavity, or convexity) which $v_{ik}(t)$ takes all alike over K, it is easily seen that if the left-hand proposition of any statement in theorem 1 or corollary 1 is true for all k, then so also is its corresponding right-hand one.

REMARK 6 Our discussion for each of the problems in the section, except for one in subsection 4.1, will be mainly devoted to showing which statement in theorem 1 or colloary 1 holds true for the problem and how its fundamental expressions are expressed. For the problems in subsections 4.2 to 4.6 The next two things will be left to readers: How are optimal partitions expressed θ, and What meanings have the optimal partitons as an optimal decision strategy. They will not be so difficult to show. Throughout the section, except for the latter half of subsection 4.1, we shall confine both state and action indexes, i and x, to integers.

REMARK 7 Except for subsection 4.7, we shall always define a difference $y_i(t) = \beta(v_i(t) - v_{i-1}(t))$. Throughout the section define $T(+\infty) = 0$ and $\tilde{T}(-\infty) = E$, and let expectations of random variables defined will be always assumed to be finite. Let a limit of y_i as $\beta \to 1-$ be denoted newly by y_i if exists. Then it will be easily seen that the limit is given by $y_i = u_i - u_{i-1}$.

REMARK 8 When $g(i,x,\theta)$ is independent of θ in case 1, it may be treated as a special case of $g(i,x,\theta) = g(i,x)\theta$ such that θ is a random variable having an unit distribution F with parameter $\lambda = 1$, that is, $F(\theta) = 0$ if $\theta < 1$ and $F(\theta) = 1$ if $1 \leqq \theta$. Then for any real number g we have $T(g) = 1 - g$ and $\tilde{T}(g) = 1$ for $g < 1$, and $T(g) = 0$ and $\tilde{T}(g) = g$ for $1 \leqq g$.

5.1 Purchasing problem

Now let N units of a material have to be purchased within a given number of days so as to minimize the total expected purchase price paid. It is assumed that per-unit prices of the material at subsequent days $\theta, \theta', \theta'', \ldots$ are independent identically distributed positive random variables having a known distribution F with an expectation E. Suppose that the total immediate price paied by purchasing x units when a per-unit price is θ is given by $p(x,\theta)$, assumed as a matter of course that $p(0,\theta) = p(x,0) = 0$, $p(1,\theta) = \theta$, and $p(x,\theta)$ is increasing in x. In the section only the next two cases will discussed: one is a quantity discount case that $p(x,\theta)$ is concave with respect to x for all θ and the other is a quantity premium case where $p(x,\theta)$ is convex with respect to x for all θ. It has been already pointed out by Morris (1959) that if $p(x,\theta) = x\theta$, optimal is a single procurement policy, that is, a policy of buying either nothing or the entire quantity required if the purchase decision is made. The section will prove by using corollary 1 that, in quantity discount case, optimal is a single procurement policy, while, in quantity premium case, optimal is a multiple procurement policy which means that the amount purchased may depend on a price quoted. (The above procurement problem has different versions and many related topics of interest, which have been already investigated in detail by Simon(1957), Morris(1959), Sakaguchi(1961), Arrow(1962), Hayes(1969), Hockman(1973), Lippman and MaCall(1976), and so on.)

Now a state of the above purchasing process may be described by the number of units having been already purchased so far, i, hence a state space becomes $I = \{0, 1, \ldots, N\}$. Let x denote the number of units to purchase, which represents an action for the decision problem. Then clearly we have $0 \leqq x \leqq N-i$, hence $a(i) = 0$ and $b(i) = N-i$. Associated with taking an action x, the immediate reward is provided by $g(i,x,\theta) = -p(x,\theta)$ and state transition law by $j(i,x) = i+x$, hence $j(u) = u$. Now let $v_i(t)$ represent -1 time the mimimum total expected purchase price paid starting with state i at day t. Then we have $v(i,x,t,\theta) = -p(x,\theta) + \beta v_{i+x}(t-1)$. Here it is evident from the objective of this problem that the final conditions are given by for all i

$$v_i(0) = -E(p(N-i,\theta)) \tag{46}$$

which is convex (concave) in quantity discount case (quantity premium case).

quantity discount case

In the case statement (f') in corollary 1 is true. Then since $v(i,x,t,\theta)$ is convex with respect to x, the optimal partition is given by (11) with $V(i,t,\theta) = -p(N-i,\theta) + \beta v_N(t-1) - \beta v_i(t-1)$. Hence a single procurement policy becomes optimal. From (9) we have $v_i(t) = -E(p(N-i,\theta)) + \beta v_N(t-1) + E([\beta v_i(t-1)+p(N-i,\theta)-\beta v_N(t-1)]^+)$ for all i and all t, by use of which it can be easily verified by induction that for all t we have $v_N(t) = 0$ and $v_i(t) \leqq 0$ for all i. Then the above expression can be reduced to

$$v_i(t) = -E(p(N-i,\theta)) + E([p(N-i,\theta) + \beta v_i(t-1)]^+) \tag{47}$$

, and we have $V(i,t,\theta) = -p(N-i,\theta)-\beta v_i(t-1)$. Since $V(i,t,\theta)$ is decreasing in θ, from corollary 2b we have $D(i,0,t) = [d(i,t),+\infty)$ and $D(i,N-i,t) = (-\infty,d(i,t))$. From (47) we have $v_i(t) \geqq -E(p(N-i,\theta)) + E(p(N-i,\theta)+\beta v_i(t-1)) = \beta v_i(t-1) \geqq v_i(t-1)$, notcing $[y]^+ \geqq y$ for any real number y. Hence $v_i(t)$ is increasing in t for all i. Thus since it follows that $V(i,t,\theta)$ is decreasing in t, we have an decreasingness of $d(i,t)$ in t. Now from (46) clearly $v_i(0)$ is increasing in i. Suppose $v_i(t-1)$ is increasing in i for any $t \geqq 1$. Then letting $\Delta p(x,\theta) = p(x,\theta) - p(x-1,\theta)$, we have $v_i(t) - v_{i-1}(t) \geqq E(\Delta p(N-i+1,\theta)) + E([p(N-i,\theta)+\beta v_{i-1}(t-1)]^+ - [p(N-i+1,\theta)+ \beta v_{i-1}(t-1)]^+) \geqq E(\Delta p(N-i+1,\theta)) - E(\Delta p(N-i+1,\theta)) = 0$, noticing lemma 7b. Hence by induction $v_i(t)$ is increasing in i for all t. However it is unfortunate that the monotonicity of $V(i,t,\theta)$ in i cannot be always derived from only the increasingness of $v_i(t)$ in i.

If $p(x,\theta) = p(x)\theta$ where $p(0) = 0$, $p(1) = 1$, and $p(x)$ is concave as well as increasing, then (47) can be reduced to $c(i,t+1) = \beta(E - T(c(i,t)))$ for $t \geqq 1$ where $c(i,t) = -\beta v_i(t-1)/p(N-i) \geqq 0$. Since $c(i,1) = E$, it follows by induction that $c(i,t)$ is independent of i for all t. Hence letting $c(t) = c(i,t)$, we have for all $t \geqq 1$

$$c(t+1) = \beta(E - T(c(t))) \tag{48}$$

where $c(1) = E$ and $c(t)$ is decreasing in t due to an increasingness of $v_i(t)$ in t. Furthermore from (11) we get $D(i,0,t) = [c(t),+\infty)$ and $D(i,N-i,t) = (-\infty,c(t))$, which are independent of i.

<u>Quantity premium case</u>

Statement (c') in corollary 1 holds for the case. Therefore an optimal partition is provided by (8), which means that a multiple procurement policy may be optimal in this case. Then from (7) we get for all i and all t

$$v_i(t) = -E(p(N-i,\theta)) + \sum_{x=1}^{N-i} E([\Delta p(x,\theta) - y_{i+x}(t-1)]^+) \qquad (49)$$

where $v_N(t) = 0$ for all t can be proved inductively in the same manner as in a quantity discount case above and where $y_i(t)$ is decreasing in i for all t due to a concavity of $v_i(t)$ with respect to i for all t. If $p(x,\theta)$ is strictly convex with respect to x for all θ, since $v(i,x,t,\theta)$ becomes also strictly concave with respect to x for all i, all t, and all θ, from lemma 2a the optimal x is uniquely determined for all i, all t, and all θ. Now from (49) we have $\beta^{-1}y_i(t) = E(\Delta p(N-i+1,\theta) - [\Delta p(N-i+1,\theta)-y_N(t-1)]^+) + \sum_{x=1}^{N-i} E([\Delta p(x,\theta)-y_{i+x}(t-1)]^+ - [\Delta p(x,\theta) - y_{i+x-1}(t-1)]^+)$. Since $y_i(0) = E(\Delta p(N-i+1,\theta)) \geqq 0$ and $y_i(t)$ is decreasing in i for all t, it follows from applying induction to the above expression that $y_i(t) \geqq 0$ for all i and all t. The above expression can be also rearranged as $\beta^{-1}y_i(t) = E(\Delta p(N-i+1,\theta)) - E([\Delta p(1,\theta)-y_i(t-1)]^+) + \sum_{x=1}^{N-i} E([\Delta p(x,\theta)-y_{i+x}(t-1)]^+ - [\Delta p(x+1,\theta) -y_{i+x}(t-1)]^+)$. For convenience let the second plus third terms of the expression be denoted by U_t, that is, $\beta^{-1}y_i(t) = E(\Delta p(N-i+1,\theta)) + U_t$ where U_t is nonpositive and is an increasing function of $y_i(t-1)$ and $y_{i+x}(t-1)$ from lemma 7b. Since we have $\beta^{-1}y_i(1) \leqq E(\Delta p(N-i+1,\theta)) = y_i(0)$ due to $U_t \leqq 0$, it follows that $y_i(1) \leqq y_i(0)$. Next by using the relationship of $\beta^{-1}(y_i(t)-y_i(t-1)) = U_t - U_{t-1}$, it can be easily shown that $y_i(t) \leqq y_i(t-1)$ for all i if $y_i(t-1) \leqq y_i(t-2)$ for all i. Hence by induction $y_i(t)$ becomes decreasing in t for all i. Consequently

$\Delta v(i,x,t,\theta) = -\Delta p(x,\theta) + y_{i+x}(t-1)$ is decreasing in i and in t. If $\Delta p(x,\theta)$ is increasing in θ for all x, then since $\Delta v(i,x,t,\theta)$ becomes decreasing in θ, it follows from corollary 2a that d(i,x,t) is decreasing in i, x, and t and that the optimal partition is given by intervals $D(i,x,t) = [d(i,x+1,t), d(i,x,t))$ where $d(i,0,t) = +\infty$ and $d(i,N-i+1,t) = -\infty$. A case that $\Delta p(x,\theta)$ is decreasing in θ is not worth discussions, because if so, we have an unreality of $\Delta p(x,\theta) = 0$ for all x and all θ due to $\Delta p(x,0) = 0$ and $\Delta p(x,\theta) \geqq 0$ for all x and all θ, in other words, $p(x,\theta)$ is independent of the number of units purchased, x.

If $p(x,\theta) = p(x)\theta$ where $p(0) = 0$, $p(1) = 1$, and p(x) is strictly concave, then (49) can be reduced to

$$v_i(t) = -p(N-i)E + \sum_{x=1}^{N-i} \Delta p(x) T(c(i,x,t)) \qquad (50)$$

$$c(i,x,t) = y_{i+x}(t-1)/\Delta p(x) \qquad (51)$$

where c(i,x,t) is nonnegative and is decreasing in i, x, and t.

<u>A version that both state and action spaces are intervals</u>

Consider the following continuous version of the above purchasing problem. Assume that the commodity can be bought in any amount, that a commodity purchased deteriorates with day where commodityies deteriorated are immediately thrown away, that a capacity of the storage space is of one unit (of the commodity), that at the terminating date of the process, $t = 0$, there must be stocked just one unit which has not deteriorated, and that a per-unit price θ has a known distribution F with $F(0) = 0$ and $0 < F(c) < 1$ for any positive c, hence we have $E > T(c) >$ for any positive c. Furthermore let F be assumed for $T(c)$ defined by use of the F to be any times differenciable. Now let i denote the amount of the commodity in stock at each days. Then the set of all possible i can be defined by the interval $I = [0,1]$, which provides a state space for the problem. Now we shall assume that the deterioration phenomena above can be incorporated into a state transition law $j(i,x)$ by setting it to be increasing in i and x, to be concave with respect to both i and x, and to be less than or equal to $i+x$. In the problem clearly we have $0 \leqq x \leqq 1-i$, in other words, an action space $A(i) = [0,1-i]$ when in state i, hence $a(i) = 0$ and $b(i) = 1-i$, where the upper bound of x is from the srorage space restriction. Now let an immediate cost for purchasing x units when a per-unit price is θ be in proportion to x. Then the immediate reward for taking an action x is given by $g(i,x,\theta) = -x\theta$. Hence we have $v(i,x,t,\theta) = -x\theta + \beta v_{j(i,x)}(t-1)$. It is clear just from the meaning of the problem that the final value is provided by

$$v_i(0) = -(1-i)E \tag{52}$$

which is concave, increasing, and any times differenciable with respect to i. Now suppose that $v_i(t-1)$ is concave, increasing, and any times differnciable with respect to i for any t. Then a mere examination of the above things could lead immediately to that statement (a) in corollary 4 is true. Therefore it follows that $v_i(t)$ is concave with respect to i. Then from (45) we have

$$v_i(t) = -(1-i)E + \beta v_{j(i,1-i)}(t-1) + \int_0^{1-i} T(c(i,x,t))dx \tag{53}$$

$$c(i,x,t) = \beta v'_{j(i,x)}(t-1)\partial j(i,x)/\partial x \tag{54}$$

in which $v_i'(t-1)$ represents a derivative of $v_i(t-1)$ with respect to i. Thus $v_i(t)$ becomes also differentiable from (53). Consequently if an increasingness of $v_i(t)$ in i can be verified, then it follows by induction that $v_i(t)$ is concave, increasing, and any times differentiable in i for all t, hence that (53) and (54) holds for all i, x, and t. The optimal purchasing policy at a day of per-unit price θ becomes as follows, given a present state i: If there exists x such that $d \triangleq \partial v(i,x,t,\theta)/\partial x = -\theta + c(i,x,t) = 0$ on $A(i) = [0,1-i]$, buy x units, if $d < 0$ on A(i), buy nothing, and if $d > 0$ on A(i), buy 1-i units. Here notice that c(i,x,t) becomes nonnegative and decreasing in x from (54).

Now we shall apply the above results to the case that j(i,x) depends only on $u = i+x$, hence let $j(u) = j(i,x)$. Here assume that j(u) is strictly increasing and is strictly concave. Then (53) can be reduced to

$$v_i(t) = -(1-i)E + \beta v_{j(1)}(t-1) + \int_i^1 T(c(u,t))du \tag{55}$$

and (54) become

$$c(u,t) \triangleq c(i,x,t) = \beta v_{j(u)}'(t-1)j'(u) \tag{56}$$

where $j'(u) = dj(u)/du > 0$. Now assume that $v_i(t-1)$ is strictly increasing in i for any t. Then c(u,t) becomes positive from (56). Now differentiating (55) with respect to i yields

$$v_i'(t) = E - T(c(i,t)) \tag{57}$$

from which it follows that $v_i'(t) > 0$ due to $c(i,t) > 0$. Thus $v_i(t)$ is strictly increasing in i, hence for all t by induction. Consequently it follows that (55), hence (57) holds for all t. Hence from (56) c(i,t) becomes also strictly decreasing in i for all t.

If $j(u) = (1-\alpha)u$, $0 \leqq \alpha \leqq 1$ in which α denotes a per-period, per-unit deterioration rate of the commodity, then we have $c(i,1) = \beta(1-\alpha)E$, which is independent of i. From this, (56), and (57), it can be proved inductively that both $v_i'(t)$ and c(i,t) are also independent of i for all t. Hence in the case a single procurement policy becomes potimal. Then letting $v(i) = v_i(t)$, (57) can be reduced to $v'(t) = E - T((1-\alpha)\beta v'(t-1))$ which implies that this case is equivalent to one with discount factor $(1-\alpha)\beta$ and without deterioration.

5.2 Investment problem

Here we shall consider the following investment decision process, which is an application of a target attacking problem by Mastran and Thomas(1973) to a management decision problem. Suppose that there exist N million dollars which can be invested in subsequent investment opportunities over a given finite time horizon. Postulate that every point in time just one opportunity arises which has a net profit of θ million dollars, gained if it results successfully. Here net profits from subsequent opportunities are assumed to be independent positive random variabls, each having a common distribution F with an expectation E. If x million dollars are invested in an opportunity with a net profit θ, then an immediate net profit of $p(x,\theta)$ can be obtained from it where $p(0,\theta) = p(x,0) = 0$ and $p(x,\theta)$ is increasing and concave with respect to x for all θ.

Our objective to the process is then to maximize the total expected net profit gained over the horizon. Now a state of the investment process may be described by the amount of dollars having been invested so far, hence the state space becomes $I = \{0,1,\ldots,N\}$. Clearly we have $0 \leqq x \leqq N-i$ if in state i, that is, $a(i) = 0$ and $b(i) = N-i$. Then associated with investing x million dollars in an opportunity of net profit θ, the immediate reward becomes $g(i,x,\theta) = p(x,\theta)$ and the state transition law is given by $j(i,x) = i+x$, hence $j(u) = u$. Now let $v_i(t)$ be the amount of the total expected net profit starting with state i at time t. Then we have $v(i,x,t,\theta) = p(x,\theta) + \beta v_{i+x}(t-1)$ where

$$v_i(0) = E(p(N-i,\theta)) \qquad (58)$$

which is decreasing and concave in i. Since statement (b') in corollary 1 holds for the problem, it follows that $y_i(t)$ for all t are decreasing in i as well as nonpositive for all i. Then (7) becomes for all i and all t,

$$v_i(t) = \beta v_i(t-1) + \sum_{x=1}^{N-i} E([\Delta p(x,\theta) + y_{i+x}(t-1)]^+) \tag{59}$$

from which we have $v_{i-1}(t) \geqq \beta v_{i-1}(t-1) + \sum_{x=1}^{N-i} E([\Delta p(x,\theta) + y_{i+x}(t-1)]^+)$, noticing $y_{i+x}(t-1) \leqq y_{i+x-1}(t-1)$ and $[c]^+ \geqq 0$ for any c. Thus from the inequality and (59) we obtain $\beta^{-1} y_i(t) \geqq y_i(t-1)$, hence $y_i(t) \geqq y_i(t-1)$. Consequently $y_i(t)$ is increasing in t for all i. Then if $\Delta p(x,\theta)$ is increasing in θ, since $\Delta v(i,x,t,\theta) = \Delta p(x,\theta) + y_{i+x}(t-1)$ is also increasing in θ, it follows from corollary 2a that d(i,x,t) is decreasing in i and t and is increasing in x. If $p(x,\theta) = (1-(1-p)^x)\theta$ where p represents the probability that an opportunity will result in success if one million dollars are invested in it (The p corresponds to a single shot hit probability in Mastran and Thomas (1973)), then (59) can be converted into

$$v_i(t) = \beta v_i(t-1) + \sum_{x=1}^{N-i} p(1-p)^{x-1} T(c(i,x,t)) \tag{60}$$

$$c(i,x,t) = -y_{i+x}(t-1)/p(1-p)^{x-1} \tag{61}$$

where it is easy to check that c(i,x,t) is nonnegative and is increasing both in i and in x and is decreasing in t.

A version

In the above model suppose that m (a fixed positive integer) investment opportunities are presented every point in time where net profits of them, denoted by $\theta_1, \theta_2, \ldots, \theta_m$, are mutually independent random variables with a known common distribution F. In the model let x denote the amount of million dollars invested in the whole of m opportunities arising. Now we shall here adopt a policy of investing either nothing or just one million dollar in each opportunity. Then clearly we have $0 \leqq x \leqq \min\{N-i,m\}$ when in state i, hence $a(i) = 0$ and $b(i) = \min\{N-i,m\}$. Under the above investment policy, if it has been decided that x ($\leqq b(i)$) million dollars was invested when in state i, they are to be invested in the first x ones of m opportunities in order of net profit sizes. Now we shall rearrange $\theta_1, \theta_2, \ldots, \theta_m$ as $\theta_1 \geqq \theta_2 \geqq \cdots \geqq \theta_m$. Let p(z), an increasing nonnegative function, be an expected net profit from investing one million dollar in an opportinity with net profit z. Then investing x million dollars in x ones of m opportunities with net profit vector $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$ yields an immediate reward $g(i,x,t) = \sum_{r=1}^{x} p(\theta_r)$, $0 \leqq x \leqq b(i)$. In the version clearly we have

$$v_i(0) = \sum_{r=1}^{b(i)} E(p(\theta_r)) \tag{62}$$

which is decreasing as well as concave with respect to i, and a state transition law becomes the same as in the original process. Consequently since statement (b') in corollary 1 holds also for the model, it follows that for all t $y_i(t)$ are nonpositive and decreasing in i. Then for all t we have from (7)

$$v_i(t) = \beta v_i(t-1) + \sum_{x=1}^{b(i)} E_x([p(\theta_x) + y_{i+x}(t-1)]^+) \tag{63}$$

An increasingness of $y_i(t)$ in t can be verified in exactly the same way as in the original process. If $p(\theta_x) = p\theta_x$ where p (> 0) is a probability of an opportuinity being succsessfull when one million dollars is invested, then (63) becomes

$$v_i(t) = \beta v_i(t-1) + p\sum_{x=1}^{b(i)} T_x(c(i,x,t)) \tag{64}$$

$$c(i,x,t) = -y_{i+x}(t-1)/p \tag{65}$$

where c(i,x,t) is a nonnegative function which is increasing in both i and x and is decreasing in t.

5.3 Warehouse problem

Consider a business of repeating a pair of transactions, one of buying some amount of a commodity at a certain day and the other of selling some of it at a future day. In the transaction process the difference of the total purchasing number of units (of the commodity) and total selling number of units are stocked in a warehouse, assumed to be of a storage capacity of N units where N is a given positive integer. Here assume that every day there comes a pair of offers, one of selling b units and the other of buying s units, where both b and s are non-negative fixed integers. At each day let only either one of the two transactions (i.e., buying and selling) be permitted. Postulate that selling and buying prices per unit are equal every day and that the per-unit price, denoted by $\theta(> 0)$, be a random variable independent from day to day according to a known distribution F with an expectation E. Moreover assume that no commission is charged on any transaction. The objective of the transaction process is to maximize a total expected marginal profit (, or an expectation of the total selling price minus the total buying price) over a given number of days. For the problem a decision variable is given by the number of units purchased, $b'(\leq b)$, or the number of units sold, $s'(\leq s)$. The decision variables can be compiled into only one variable x such as $-s \leq x \leq b$ where $0 < x$ implyies to purchase x units and $x < 0$ to sell $-x$ units. Below for convenience, even when $x < 0$, we shall use an expression of "purchase x units".

Now a state of the above buying-selling process may be characterized by the number of units being stocked in the warehouse, i, hence the state space becomes $I = \{0,1,\ldots,N\}$. Now when in state i, the x must satisfy the inequality $-\min\{i,s\} \leqq x \leqq \min\{N-i,b\}$. Thus we have $a(i) = -\min\{i,s\}$ and $b(i) = \min\{N-i,b\}$. Then associated with buying x units, the immediate reward is given by $g(i,x,\theta) = -x\theta$, assumed to be in proportion to the amount of units purchased, and the state transition law becomes $j(i,x) = i+x$, hence $j(u) = u$. Therefore we have $v(i,x,t,\theta) = -x\theta + \beta v_{i+x}(t-1)$. Now suppose that $v_i(0)$ is <u>concave</u> with respect to i. Then since statement (c') in corollary 1 holds for the process, from (29) and (33) we have for all t

$$v_i(t) = -b(i)E + \beta v_{i+b(i)}(t-1) + \sum_{x=a(i)+1}^{b(i)} T(c(i,x,t)) \qquad (66)$$

where $c(i,x,t) = y_{i+x}(t-1)$, which is decreasing both in i and in x for all t because of a decreasingness of $y_i(t)$ in i for all t due to a concavity of $v_i(t)$ for all t. Now notice that the vector $(b(i), b(i-1), a(i), a(i-1))$ equals

(a) $(\ b\ ,\ \ b\ ,\ -s\ ,\ -s\ )$ if $b \leqq N-i$ and $s < i$

(b) $(\ N-i\ ,\ N-i+1\ ,\ -s\ ,\ -s\ )$ if $b > N-i$ and $s < i$

(c) $(\ b\ ,\ \ b\ ,\ -i\ ,\ -i+1\ )$ if $b \leqq N-i$ and $s \geqq i$

(d) $(\ N-i\ ,\ N-i+1\ ,\ -i\ ,\ -i+1\ )$ if $b > N-i$ and $s \geqq i$

Then the next expressions can be easily derived from (66)

$$\beta^{-1} y_i(t) = \begin{cases} \tilde{T}(y_{i+b}(t-1)) - T(y_{i-s}(t-1)) & \text{for (a)} \\ E - T(y_{i-s}(t-1)) & \text{for (b)} \\ \tilde{T}(y_{i+b}(t-1)) & \text{for (c)} \\ E & \text{for (d)} \end{cases} \qquad (67)$$

Now for convenience define $y_i(t) = +\infty$ for $i \leqq 0$ and $y_i(t) = -\infty$

for i > N. Then (67) can be compiled into the only one expression below (Note remark 7),

$$\beta^{-1}y_i(t) = \hat{T}(y_{i+b}(t-1)) - T(y_{i-s}(t-1)) \qquad 1 \leqq i \leqq N \tag{68}$$

Furthermore we shall assume that $v_i(0)$ is <u>increasing</u> in i, in other words, $y_i(0) \geqq 0$ for all i. Then by use of induction it can be easily shown from (67) and lemma 7c that $y_i(t) \geqq 0$ for all i and all t, in other words, $v_i(t)$ is increasing in i for all t.

Next we shall consider the case of an infinite horizon. In the case let M be a sufficiently large positive number greater than E such that $\hat{T}(M) = M$ (see remark 4). If $y_i(t-1) \leqq \beta M < M$ for any t, then from (67) we have $\beta^{-1}y_i(t) \leqq M$, or $y_i(t) \leqq \beta M < M$, hence by induction for all t. Consequently limits $v_i - v_0$ becomes also bounded in β, and it follows from theorem 2 that there exists g, the maximum expected marginal profit gained per day over an infinite time horizon if β = 1. Then since t approaching infinity in (66) with i = N leads to $v_N = \beta v_N + \sum_{x=a(N)+1}^{0} T(y_{N+1})$, we get

$$g = \sum_{x=a(N)+1}^{0} T(y_{N+x}) \tag{69}$$

<u>Three examples</u>

Now suppose that the vector (s,b) is a random variable with a joint probability p(s,b) where $\sum_{s,b} p(s,b) = 1$. Then from remark 5 we have $v_i(t) = \sum_{s,b} p(s,b) \times$(right-hand of (66), dependent on s and b), which yields $\beta^{-1}y_i(t) = \sum_{s,b} p(s,b) \times$(right-hand of (67) or (68), dependent on s and b). Noticing this we shall apply the above results to the following three special cases, where let a(i) and b(i) associated with (s,b) be denoted by $a_{sb}(i)$ and $b_{sb}(i)$, respectively.

Example 1 Let $p(s,0) \geqq 0$ and $p(s,b) = 0$ for $b > 0$

Example 2 Let $p(0,b) \geqq 0$ and $p(s,b) = 0$ for $s > 0$

Here for simplicity let $p_s = p(s,0)$ for ex.1 and $p_b = p(0,b)$ for ex.2, and suppose that both examples are of finite time horizon.

In ex.1, no selling offer is made, in other words, no units can be bought at all. This means that the example is equivalent to the problem of selling out inventory in hand so as to maximize the total expected selling price over a given number of days.

On the other hand, ex.2 is a case that no buying offer is made, or no units in stock can be sold. Then we shall set up the objective of the example to mimimize the total expected purchase price paied under the following terminating conditions, provided that N units must be purchased in all over a given number of days: Suppose that i units have been already purchased up to day 1 and an offer of selling b units are made at day 0. Then if $b \geqq N-i$, $N-i$ units are bought at a current price quoted at day 0. Conversaly if $b < N-i$, first b units are purchased at a current price θ at day 0 and then the additional $N-i-b$ units are bought at a higher price $h\theta$ than the current price θ, where h is a fixed number greater than 1.

For ex.1 we have $a_{s0}(i) = -\min\{i,s\}$, $b_{s0}(i) = 0$ and for ex.2 $a_{0s}(i) = 0$, $b_{0b}(i) = \min\{N-i,b\}$. Final values $v_i(0)$ becomes from the above terminating conditions

ex.1 $$v_i(0) = \sum_s p_s \min\{s,i\} E$$

ex.2 $$v_i(0) = -\sum_{b=0}^{N-i-1} p_b (b+h(N-i-b))E - (N-i)\sum_{b=N-i}^{\infty} p_b E$$

Thus we have

ex.1
$$y_i(0) = \sum_{s=i}^{\infty} p_s E \geqq 0$$
$$y_i(0) - y_{i-1}(0) = -p_{i-1}E \leqq 0$$

ex.2
$$y_i(0) = h\sum_{b=0}^{N-i} p_b E + \sum_{b=N-i+1}^{\infty} p_b E \geqq 0$$
$$y_i(0) - y_{i-1}(0) = -(h-1)p_{N-i+1}E \leqq 0$$

Hence for both examples it follows that $y_i(0)$ is nonnegative and decreasing in i, or $v_i(0)$ is increasing and concave in i. Consequently so also becomes $y_i(t)$ for all t. Then from (68) we have for all i and all t,

ex.1
$$\beta^{-1}y_i(t) = y_i(t-1) + \sum_{s=i}^{\infty} p_s T(y_i(t-1)) + \sum_{s=0}^{i-1} p_s (T(y_i(t-1)) - T(y_{i-s}(t-1)))$$

ex.2
$$\beta^{-1}y_i(t) = y_i(t-1) + \sum_{b=N-i+1}^{\infty} p_b (E - \tilde{T}(y_i(t-1))) + \sum_{b=0}^{N-i} p_b (\tilde{T}(y_{i+b}(t-1)) - \tilde{T}(y_i(t-1)))$$

It can be seen from the just above two expressions that $y_i(t) \geqq y_i(t-1)$ for ex.1 and $y_i(t) \leqq y_i(t-1)$ for ex.2, respectively. Thus it follows that $y_i(t)$ is increasing in t for ex.1 and is decreasing in t for ex.2.

Example 3 Let $N \geqq 2$ and $p_{1,1} = 1$ and assume that a time horizon is infinite where let an observation space be finite. Then we have $a(i) = -\min\{i,1\}$ and $b(i) = \min\{N-i,1\}$, and for limit y_i we have from (67)

$$\beta^{-1}y_1 = \tilde{T}(y_2)$$

$$\beta^{-1}y_i = \tilde{T}(y_{i+1}) - T(y_{i-1}) \qquad 1 < i < N$$

$$\beta^{-1}y_N = \qquad E - T(y_{N-1})$$

If $\beta = 1$, then from (69) we get the maximum of the expected marginal profit per day, $g = T(y_N)$, using a limit y_N (see remark 7). A straightforward way for determining the g is to directly solve a system of the above nonlinear equations with $\beta = 1$. A more elegant one is to apply a Howard's argorithm, by use of which the solution can be obtained in the finite number of steps. We shall explain this by using a case of $N = 2$ where an observation space is given by $\Omega = \{1,2,3\}$ with probabilities $f_i = P\{\theta=i\} > 0$, $i=1,2,3$, $f_1+f_2+f_3 = 1$. In the case a state space is given by $I = \{0,1,2\}$ and action spaces become $A(0) = \{0,1\}$, $A(1) = \{-1,0,1\}$, and $A(2) = \{-1,0\}$. Now for convenience define a vector $p_i(x) = (p_{i0}(x), p_{i1}(x), p_{i2}(x))$ for $i = 0,1,2$ and for x in $A(i)$. Then clearly we have $p_0(0) = (1,0,0)$, $p_0(1) = (0,1,0)$, $p_1(-1) = (1,0,0)$, $p_1(0) = (0,1,0)$, $p_1(1) = (0,0,1)$, $p_2(-1) = (0,1,0)$, $p_2(0) = (0,0,1)$. Furthermore notice that the example can be regarded as being of case 1b with $g(i,x) = -x$. Then from (40) and (41) $R_i(H)$ and $Q_{ij}(H)$ can be expressed as follows, using a given i-dependent decreasing function $h(x)$ on $A'(i)$ which takes only integers on Ω where $h(0) = +$ and $h(2) = -\infty$ for state 0, $h(-1) = +\infty$ and $h(2) = -\infty$ for state 1, and $h(-1) = +\infty$ and $h(1) = -\infty$ for state 2:

$$R_0(H) = -E[h(1)]$$

$$R_1(H) = E - E[h(1)] - E[h(0)]$$

$$R_2(H) = E - E[h(0)]$$

$Q_{00}(H) = 1 - F(h(1))$, $Q_{01}(H) = F(h(1))$, $Q_{02}(H) = 0$

$Q_{10} = 1 - F(h(0))$, $Q_{11} = F(h(0)) - F(h(1))$, $Q_{12}(H) = F(h(1))$

$Q_{20}(H) = 0$, $Q_{21}(H) = 1 - F(h(0))$, $Q_{22}(H) = F(h(0))$

A value determination operation in a Howard's algorithm can be done by solving (23) using the above $R_i(H)$ and $Q_{ij}(H)$ for a given H. The subsequent policy improvement routine is carried out by solving the next maximization problems (see (39)): for i = 0 maximize $T(c(0,1),h(1))$, for i = 1 maximize$\{T(c(1,0),h(0)) + T(c(1,1),h(1))\}$ subject to $h(0) \geqq h(1)$, and for i = 2 maximize $T(c(2,0),h(0))$. When i = 0 and i = 2, from lemma 5 the above maximization can be attained at $h(1) = [c(0,1)]$ and $h(0) = [c(2,0)]$, respectively. When i =1, if $c(1,0) \geqq c(1,1)$, then clear we have the solution $h(0) = [c(1,0)]$ and $h(1) = [c(1,1)]$, otherwise it can be obtained by using a dynamic programming technique as follows: Set $L_0(h) = \max_{g \geqq h} T(c(1,0),g)$ where g, h = 1,2,3 and $L_1 = \max_{h=1,2,3}\{T(c(1,1),h) + L_0(h)\}$. For all h determine g = g(h) attaining $L_0(h)$, and then h=h* attaining L_1. Thus the maximing h(1) and h(0) are given by h* and g(h*), respectively.

5.4 Inventory problem

The section here attempts to apply our approach to an inventory problem for a material with a fluctuating price, where the material is assumed to be consumed at a constant rate of s units per day. The objective is to minimize the total expected purchaseing price paid over an infinite time horizon, given a warehouse in which the maximum N units of the material can be stocked. Let both N and s be fixed positive integers such as $N \geqq s$. Assume that per-unit prices at successive days θ, θ', ... are independent positive random variables from a known distribution F with an expectation E. Let an immediate purchase price for buying x units when a price is θ be given by $p(x,\theta)$ which is an increasing and concave function in x for all θ with $p(0,\theta) = p(x,0) = 0$ and $p(1,\theta) = \theta$. Furthermore assume that no shortage can be permitted and that a delivery of the material placed an order for is instantaneous. A state of the purchasing process may be described by the number of units in stock, i, hence the state space is given by $I = \{0,1,\ldots,N\}$. Then when in state i, the number of units purchased, x, must satisfy the inequality $[s-i]^+ \leqq x \leqq N-i$, hence we have $a(i) = [s-i]^+$ and $b(i) = N-i$. Associated with taking an action x, the immediate reward becomes $g(i,x,\theta) = -p(x,\theta)$ and the state transition law is provided by $j(i,x) = i+x-s$, hence $j(u) = u - s$. Then (3) becomes $v(i,x,t,\theta) = -p(x,\theta) + \beta v_{i+x-s}(t-1)$. For this problem it is easily seen by noticing remark 4 that statement (c') of corollary 1 is true. Therefore we obtain from (7)

$$v_i(t) = -E(p(N-i,\theta)) + \beta v_{N-s}(t-1) + \sum_{A'(i)} E([\Delta p(x,\theta) - y_{i+x-s}(t-1)]^+) \qquad (70)$$

where $y_i(t)$ is decreasing in i for all t due to a concavity of $v_i(t)$ for all t. Now suppose that $\Delta p(x,\theta)$ is increasing in θ. Then since $\Delta v(i,x,t,\theta) = -\Delta p(x,\theta) + y_{i+x-s}(t-1)$ is decreasing in θ, from corollary 2a d(i,x,t) becomes decreasing in both i and x and is increasing in s, hence so also dose a limit d(i,x). Below we shall show a boundedness of limit $v_i - v_0$ in β for all t. First from (70) we get

$$
\begin{aligned}
\beta^{-1} y_i(t) &= E(\Delta p(N-i+1,\theta)) \\
&\quad + \textstyle\sum_{A'(i)} E([\Delta p(x,\theta) - y_{i+x-s}(t-1)]^+ \\
&\qquad\qquad - [\Delta p(x+1,\theta) - y_{i+x-s}(t-1)]^+) \\
&\quad + \begin{cases} 0 & \text{for } 1 \leqq i \leqq s \\ -E([\Delta p(1,\theta) - y_{i-s}(t-1)]^+) & \text{for } s < i \leqq N \end{cases}
\end{aligned}
$$

Since the second and third terms of the above expression is nonpositve, we have $\beta^{-1} y_i(t) \leqq E(\Delta p(N-i+1,\theta))$ for all i and all t. Thus we have an upper boundedness of $\beta^{-1} y_i(t)$ in i for all t. The expression above can be also rearranged as follows, letting $A''(i) = \{x; [s-i]^+ + 1 < x \leqq N-i\}$,

$$
\begin{aligned}
\beta^{-1} y_i(t) &= E([\Delta p([s-i]^+ + 1,\theta) - y_{i+[s-i]^+ - s+1}(t-1)]^+) \\
&\quad + E(\Delta p(N-i+1,\theta) - [\Delta p(N-i+1,\theta) - y_{N-s}(t-1)]^+) \\
&\quad + \textstyle\sum_{A''(i)} E([\Delta p(x,\theta) - y_{i+x-s}(t-1)]^+ \\
&\qquad\qquad - [\Delta p(x,\theta) - y_{i+x-s-1}(t-1)]^+) \\
&\quad + \begin{cases} 0 & \text{for } 1 \leqq i \leqq s \\ -E([\Delta p(1,\theta) - y_{i-s}(t-1)]^+) & \text{for } s < i \leqq N \end{cases}
\end{aligned}
$$

Now suppose that $y_i(t-1) \geqq 0$ for any t and all i. Then since the second and third terms of the above expression becomes nonnegative, it follows that

$$\beta^{-1}y_i(t) \geqq \begin{cases} E([\Delta p(s-i+1,\theta)-y_1(t-1)]^+) & 1\leqq i\leqq s \\ E([\Delta p(1,\theta)-y_{i-s+1}(t-1)]^+ - [\Delta p(1,\theta)-y_{i-s}(t-1)]^+) & s<i\leqq N \end{cases}$$

the right-hands of which becomes nonnegative for all i. Consequently $y_i(t)$ becomes nonnegative, hence for all t by induction. In other words, $y_i(t)$ is lower bounded in β for all t. Thus for limit $\beta^{-1}y_i$ we have $0 \leqq \beta^{-1}y_i = v_i - v_{i-1} \leqq E(\Delta p(N-i+1,\theta))$, that is, a boundedness of $v_i - v_0$ in β. Therefore from theorem 2 it follows that there exists g where -g is the minimum of the expected purchase price paid per day over an infinite number of days if $\beta = 1$. Since from (70) we have $v_N = \beta v_{N-i}$ for limit v_i, it follows that

$$-g = y_N + y_{N-i} + \cdots + y_{N-s+1} \tag{71}$$

A case that the consumption rate s is a random variable has been already studied by Fabian et al.(1959) by using an approach based on a calculus, where a shortage cost was assumed to be finite. A possibility of applying our approach to their model was discussed by Yagishita (1981) under some additional conditions.

5.5 Sequential stochastic assignment problem

Here we shall demonstrate that applying theorem 4a and corollary 3 brings about another approach to a so-called sequential stochastic assignment problem by Derman et al.(1972). This problem is defined as follows: Assume that t jobs arrive at a groupe of t men(workers) one by one. They are called, respectively, g(1)-man, g(2)-man,....., g(t)-man where g(x) is assumed to be an increasing function in x. If g(x)-man is assigned to the first arriving job of value θ, then the reward of $g(x)\theta$ is gained. A value of each arriving job is assumed to be an independent positive random variable having a known distribution F with an expectation E. A man assigned to a job is unavairable for future assignment. The objective to the model is to maximize the total expected reward obtained. We shall refer to the decision process as h(t)-process where $h(t) = (g(1),g(2),...,g(t))$, t-vector. Let $v_*(t)$, $v_i(t)$, and $v_{ix}(t-1)$ denote the maximum total expected rewards for, respectively, h(t)-, $h_i(t)$-, and $h_{ix}(t-1)$-processes, where $h_i(t)$ represents a t-vector resulting from removing the i-th element of a vector h(t+1) and $h_{ix}(t-1)$ is a (t-1)-vector from eliminating the x-th element of the vector $h_i(t)$, denoted by $g_i(x)$. A state of a h(t)-process at a starting point in time can be completely characterized by the vector h(t). Then for simplicity we shall denote the starting states of h(t)-, $h_i(t)$-, and $h_{ix}(t-1)$-processes be denoted by, respectively, *, i, and ix. Now notice that the h(t)-process can be regarded as case 1a with an action space $A(*) = \{x\text{=integer}; 1 \leqq x \leqq t\}$ (hence $a(*) = 1$ and $b(*) = t$), immediate reward $g(*,x)\theta = g(x)\theta$, and state transition law $j(*,x) = x$. Then clearly we have $v(*,x,t,\theta) = g(x)\theta + z(*,x,t)$ where $z(*,x,t) = v_x(t-1)$.

First we shall discuss a case that $g(x)$ is strictly increasing in x. Then from (29) we have $c(*,x,t) = (v_{x-1}(t-1) - v_x(t-1))/(g(x) - g(x-1))$ for $2 \leqq x \leqq t$. Throughout the section let $c(*,1,t) = -\infty$ and $c(*,t+1,t) = +\infty$ for all t. If $c(*,x,t)$ is increasing in x on $A'(*)$ for any t, then we have from theorem 4a

$$v_*(t) = g(1)E + v_1(t-1) + \sum_{x=2}^{t}(g(x)-g(x-1))T(c(*,x,t)) \qquad (72)$$

For $h(1)$-process with $h(1) = (g(1))$, clearly we have $v(1) = g(1)E$. For $h(2)$-process, noticing remark 2, we have from (72) $v_*(2) = g(1)E + v_1(1) + (g(2)-g(1))T(c(*,2,2))$ in which $c(*,2,2) = (v_1(1)-v_2(1))/(g(2)-g(1)) = E$ because $v_1(1) = g(2)E$ and $v_2(1) = g(1)E$. Here it should be noted that $c(*,2,2)$ is independent of the present state $*$, in other words, independent of $h(2)$, or of "$g(*,1) \triangleq g(1)$, $g(*,2) \triangleq g(2)$". Hence let $c(2,2) = c(*,2,2) = E$. When noticing remark 7, $c(2,2)$ may be rewritten as $c(2,2) = \tilde{T}(c(1,1)) - T(c(2,1))$ where $c(1,1) = -\infty$ and $c(2,1) = +\infty$. Now define $c(1,2) = -\infty$ and $c(3,2) = +\infty$. Then clearly $c(x,2)$ can be deemed to be increasing in x $(=1,2,3)$. For any $t \geqq 2$ suppose that $c(*,x,t)$ associated with a $h(t)$-process is increasing in x on $A(*)$ as well as independent of $h(t)$. Hence for all t let $c(x,t) = c(*,x,t)$ on $A(*)$ where $c(1,t) = -\infty$ and $c(t+1,t) = +\infty$. Then (72) associated with $h_i(t)$-process can be expressed as

$$v_i(t) = g_i(x)E + v_{i1}(t-1) + \sum_{x=2}^{t}(g_i(x)-g_i(x-1))T(c(x,t)) \quad 1 \leqq i \leqq t+1 \qquad (73)$$

Now $i > x$ leads to $g_i(x) = g(x)$ and $v_{ix}(t-1) = v_{x,i-1}(t-1)$, and $i \leqq x$ to $g_i(x) = g(x+1)$ and $v_{ix}(t-1) = v_{x+1,i}(t-1)$. Then noticing that

$g(i) - g(i-1) = g_1(i-1) - g_1(i-2)$ for $3 \leqq i$, we have

$$(v_1(t)-v_2(t))/(g(2)-g(1)) = \tilde{T}(c(1,t))-T(c(2,t))$$

$$(v_{i-1}(t)-v_i(t))/(g(i)-g(i-1)) = T(c(i-1,t))-T(c(i,t)) + (v_{1,i-2}(t-1)-v_{1,i-1}(t-1))/(g_1(i-1)-g_1(i-2)) \quad 3 \leqq i \leqq t$$

$$(v_t(t)-v_{t+1}(t))/(g(t+1)-g(t)) = T(c(t,t)) + (v_{1,t-1}(t-1)-v_{1,t}(t-1))/(g_1(t)-g_1(x-1))$$

Now arranging the above expressions by noticing the induction hypothesis that $c(*,x,t)$ is independent of $h_i(t)$, we get $c(*,x,t+1) = \tilde{T}(c(x-1,t)) - T(c(x,t))$ for $2 \leqq x \leqq t+1$. Clearly the right-hand of the expression is independent of $h(t+1)$, and it is easily verified if noticing lemma 7c that $c(*,x,t+1)$ is increasing in x. Hence these becomes also true for all t by induction. Consequently letting $c(x,t+1) = c(*,x,t+1)$, we have for all t

$$c(x,t+1) = \tilde{T}(c(x-1,t)) - T(c(x,t)) \qquad 2 \leqq x \leqq t+1 \qquad (74)$$

Since for $2 \leqq x \leqq t$ we obtain $c(x,t+1) = c(x,t) - (\tilde{T}(c(x,t)) - \tilde{T}(c(x-1,t))) \leqq c(x,t)$ owing to $c(x,t) \geqq c(x-1,t)$, it follows that $c(x,t)$ is decreasing in t for all x. Now since clearly $v_*^h(1) \to v_*(1)$ as $h \to 0+$, from corollary 3 all the above results become also true in a case that $g(x)$ is an increasing function.

5.6 customer Selection Problem I

Consider the following discrete time queuing system, which is a discrete time version of a customer selection system by Miller (1969). Assume that there exist K (a positive integer) customer classes and that only one customer, belonging to one of them, arrives every period with a given probability p_k by which we mean a probability of an arriving customer being in class k. Then the probability of no customer arriving can be given by $p_0 = 1 - \sum_{k=1}^{K} p_k$. Let a customer of class k have a fixed amount of value r_k where $r_1 \geqq r_2 \geqq \cdots \geqq r_K > 0$. Assume that every customer accepted is served in a same station of a single stage according to a given probability law q_s, which means the probability of completeing services for s customers within one period, provided that a sufficient large number of customers are now in the system. Here we shall let $q_0 > 0$. Notice that the probability of services for all customers in the system being completed in one period is given by $\sum_{s=i}^{\infty} q_s$, given that i customers are in the system. Finally we shall suppose that more than $N(\geqq 1)$ customers cannot be held in the system at any instant.

Our objective in the selection process above is to maximize the total expected value from customers accepted over an infinte time horizon. A state of the system can be characterized by the number of customers in the system, i, hence its state space is given by $I = \{0,1,\ldots,N\}$. Now let x = 0 if a customer arriving is rejected and x = 1 if accepted. More generally the x denotes the number of customers accepted. Then it is clear that we have $0 \leqq x \leqq 1$ if $i < N$ and $x = 0$ if $i = N$. This means that $i < N$ leads to $a(i) = 0$ and $b(i) = 1$, and i

$= N$ to $a(i) = b(i) = 0$. Here note that both $a(i)$ and $b(i)$ are independent of the class of a customer arriving. Then associated with taking an action x if a customer of class k arrives in state i, the immediate reward is given by $g_k(i,x,\theta) = r_k x\theta$ with a random variable θ having a unit distribution of parameter $\lambda = 1$ (see remarks 5 and 8), and the state transition law can be expressed by $j(i,x,s) = [i+x-s]^+$, hence $j(u,s) = [u-s]^+$. Then we have $v_k(i,x,\theta) = g_k(x)\theta + z(i,x)$ where $g_k(x) = r_k x$ and

$$z(i,x) = \beta\sum_s q_s v_{[i+x-s]^+} \qquad x = 0,1 \tag{75}$$

which is independent of k. For the problem statement (b') of corollary 1 becomes true. Therefore it follows that a limit y_i is nonpositive as well as decreasing and that from (31) we have $v_i = p_0 z(i,0) + \sum_k q_k(z(i,0)+r_k T(c_k(i)))$ where the summation is over $1 \leqq k \leqq K$. The expression can be arranged as

$$v_i = \beta\sum_s q_s v_{[i-s]^+} + \sum_k q_k r_k T(c_k(i)) \tag{76}$$

$$c_k(i) \triangleq c_k(i,1) = -\sum_{s=0}^{i} q_s y_{i+1-s}/r_k \qquad i < N \tag{77}$$

where let $c_k(N) = +\infty$. Here $c_k(i)$ becomes nonnegative and increasing in i because of $c_k(i) - c_k(i-1) = -(q_i y_1 + \sum_{s=0}^{i-1} q_s(y_{i+1-s} - y_{i-s}))/r_k \geqq 0$. From (76) we can obtain $\beta^{-1} y_i \geqq \sum_{s=0}^{i-1} q_s y_{i-s} - \sum_k q_k r_k T(c_k(i-1))$ for $1 \leqq i \leqq N$. Now since $T(c_k(i-1)) \leqq T(0) = 1$, we have $y_i \geqq \beta(\sum_{s=0}^{i-1} q_s y_{i-s} - h)$ where $h = \sum_k p_k r_k > 0$. If $i = 1$, it follows that $y_1 \geqq \beta(q_0 y_1 - h)$, from which we have $\beta^{-1} y_1 \geqq M_1(\beta)$ where $M_1(\beta) = -h/(1-\beta q_0)$. Since $M_1(\beta)$ is nonpositive and decreasing in β, we obtain $\beta^{-1} y_1 \geqq M_1(1)$, hence y_1

$\geqq \beta M_1(1) \geqq M_1(1)$. Therefore both y_1 and $\beta^{-1}y_1$ are bounded in β. For any $i \geqq 2$ we have $y_i \geqq \beta(q_0 y_i + d_i - h)$ where $d_i = \sum_{s=1}^{i-1} q_s y_{i-s} \leqq 0$. Then we get $\beta^{-1}y_i \geqq N_i(\beta) \triangleq -(h-d_i)/(1-\beta q_0)$. If y_n and $\beta^{-1}y_n$, $n=1,2,\ldots,i-1$, are all bounded in β, then so also is d_i, hence let a lower bound of d_i be now denoted by $h_i(\leqq 0)$. Then clearly we have $N_i(\beta) \geqq M_i(\beta) \triangleq -(h-h_i)/(1-\beta q_0)$, which is nonpositive and decreasing in β. Thus we have $\beta^{-1}y_i \geqq M_i(1)$, which leads to $y_i \geqq M_i(1)$. Thus both $\beta^{-1}y_i$ and y_i are bounded in β, hence for all i by induction. Consequently it follows from theorem 2 that g exists which is the maximum of a per-period expected value over an infinite time horizon with $\beta = 1$. Now from (76) since $v_0 = \beta v_0 + \sum_k q_k r_k T(c_k(0))$ with $c_k(0) = -q_0 y_1/r_k$, it follows that

$$g = \sum_k q_k r_k T(-q_0 y_1/r_k) \tag{78}$$

<u>A version</u>

In the above model suppose that there exists only one customer class, i.e., $K = 1$, and instead that k customers of the only class arrive every period where k is a given positive integer. Furthermore values of the m customers arriving, denoted by $\theta_1, \theta_2, \ldots, \theta_k$, are assumed to be independent positive random variables with a given common distribution F. We shall here rearrange them as $\theta_1 \geqq \theta_2 \geqq \cdots \geqq \theta_k \geqq 0$. Let x ($\leq$k) represent the number of customers accepted when in state i. Then when in state i, clearly we have $0 \leqq x \leqq \min\{k, N-i\}$, hence $a(i) = 0$ and $b(i) = \min\{k, N-i\}$. In the version it is clear that the state transition law of the process is given in just the same form as in the original model, that is, $j(i,x,s) = (i+x-s)^+$ with $a(i) \leqq x \leqq$

b(i). Then for taking action x in state i we have $v(i,x,\theta) = \theta_1+\theta_2+\ldots+\theta_x + z(i,x)$ where z(i,x) is given by (75) with $a(i) \leqq x \leqq b(i)$. Since statement (b') in corollary 1 becomes also true for the version, expression (35) associated with limits v_i can be written as

$$v_i = \beta\sum_s q_s v_{(i-s)^+} + \sum_{x=1}^{b(i)} T_x(c(i,x)) \tag{79}$$

$$c(i,x) = -\sum_{s=0}^{i+x-1} q_s y_{i+x-s} \tag{80}$$

where y_i is nonpositive and decreasing in i. Hence c(i,x) is nonnegative, which is dependent only on u. Then letting c(u) = c(i,x), we have $c(u) - c(u-1) = -\sum_{s=0}^{u-2} q_s(y_{u-s}-y_{u-s-1}) - q_{u-1}y_1 \geqq 0$. Thus it follows that c(i,x) is increasing both in i and in x. From (79) we get $\beta^{-1}y_i \geqq \sum_{s=0}^{i-1} q_s y_{i-s} - \sum_{x=1}^{b(i-1)} T_x(c(i-1,x))$, from which we have $y_i \geqq \beta(\sum_{s=0}^{i-1} q_s y_{i-s} - h_i)$ where $h_i = \sum_{x=1}^{b(i-1)} E_x \geqq 0$. By use of the inequality, a boundedness of y_i and $\beta^{-1}y_i$ in β can be proved in exactly the same way as in the original model. Hence it follows from theorem 2 that g exists. From (79) since $v_0 = \beta v_0 + \sum_{x=1}^{b(0)} T_x(c(0,x))$, we have

$$g = \sum_{x=1}^{b(0)} T_x(c(0,x)) \tag{81}$$

5.7 Customer selection problem II

All discussions in the previous sections have been based on a monotonicity of $y_i(t)$ in i, in other words, on a concavity of $v_i(t)$ with respect to i, just through which a concavity or convexity of $v(i,x,t,\theta)$ has been verified. This section will give and develop one example such that while $y_i(t)$ is not always monotone, $v(i,x,t,\theta)$ becomes concave.

Consider the following discrete time bulk queuing system where let points in time be taken at an equally spaced interval on time axis. Assume that k customers arrive with certainty at the beginning of every period where values which they have are defined just in the same way as in section 4.6. Postulate that every arriving customer has the identical fixed service time of n periods, where a period means a time interval between successive two points of time. Now define a term backlog by the total periods required for completing the service for all customers in the system. Then assume that more than M periods of backlog can not be held at any instant. Here let the above n, M, and k be all fixed postive integer such that $n > 1$, $M > 1$, and $k \geq 1$. In the problem the decision maker must decide every period how many customers of k customers arriving to accept from the upper rank of their value sizes with an aim of maximizing the total expected value gained over an infinite time horizon.

Now a state of the decision system can be described by the amount of backlog at the beginning of each period but before a decision of a customer selection is made. Then notice that the system being in state i at any point in time is equvalent to that there have existed i+1 periods of backlog immediately after the decision has been made at the previous period, because one period elapcing yields one period reduction in backlog. Therefore since it must be from the assumption that $i+1 \leqq M$, we have $i \leqq N \triangleq M-1$, hence the state space becomes $I = \{0,1,\dots,N\}$. Then acceptance of x customers in state i leads to the next state $j(i,x) = [i+xn-1]^+$. Thus (28) becomes $z(i,x,t) = \beta v_{[i+xn-1]^+}(t-1)$. Now since it must be that $j(i,x) \leqq N$, it follows that the maximum x satisfying this inequality is given by $[(N-i+1)/n]$ where [ ] represents Gauss's symbol, hence the maximum permissible number of customers to accept when in state i, is given by $b(i) = \min\{k,[(N-i+1)/n]\}$. Thus we have $0 \leqq x \leqq b(i)$. Hence since it follows that a concavity or convexty of A(i) is not always assured from lemma 1, we cannot appliy theorem 1 and corollary 1 to this problem. Then we shall attempt to develop a rather direct approach to it. First we shall here define $y_i(t)$ in somewhat different form from in the previous sections as follows:

$$y_i(t) = \beta(v_{[i-n]^+}(t) - v_i(t)) \qquad 1 \leqq i \leqq N$$

$$\Delta y_i(t) = y_i(t) - y_{i-1}(t) \qquad 2 \leqq i \leqq N$$

Then (29) can be expressed as

$$c(i,x,t) = y_{i+xn-1}(t-1) \quad 1 \leqq x \leqq b(i) \qquad 1 \leqq i \leqq N \tag{82}$$

where for convenience let $c(i,x,t) = +\infty$ for $b(i) < x$. Here define $C(t) = \{c(i,x,t); 1 \leqq x \leqq b(i),\ 0 \leqq i \leqq N\}$, a set of only finite $c(i,x,t)$. Then the next relationships hold on $C(t)$: $c(i,x,t) = c(i-n,x+1,t) = c(i+n,x-1,t)$, $c(i,x,t) - c(i-1,x,t) = \Delta y_{i+xn-1}(t)$, and $c(i,x,t) - c(i,x-1,t) = \Delta y_{i+xn-1}(t) + \Delta y_{i+xn-2}(t) + \dots + \Delta y_{i+xn-n}(t)$. Now suppose that both $y_i(t-1)$ and $\Delta y_i(t-1)$ are nonnegative for all i and for any t. Then since $c(i,x,t)$ becomes increasing in both i and x from the above expressions, it follows from theorem 4b that we have

$$v_i(t) = \beta v_{(i-1)^+}(t-1) + \sum_{x=1}^{b(i)} T_x(c(i,x,t)) \qquad 0 \leqq i \leqq N \qquad (83)$$

from which we get, noticing remark 7,

$$\beta^{-1}y_1(t) = \sum_{x=1}^{\infty}(T_x(c(0,x,t)) - T_x(c(1,x,t))) \qquad (84)$$

$$\beta^{-1}y_i(t) = y_{i-1}(t-1) + \sum_{x=1}^{\infty}(T_x(c(0,x,t)) - T_x(c(i,x,t))) \qquad 2\leqq i\leqq n \quad (85)$$

$$\beta^{-1}y_i(t) = y_{i-1}(t-1) + \sum_{x=1}^{\infty}(T_x(c(i-n,x,t)) - T_x(c(i,x,t))) \qquad n<i\leqq N \quad (86)$$

From (84) and (85) we have

$$\beta^{-1}\Delta y_2(t) = y_1(t-1) + \sum_{x=1}^{\infty}(T_x(c(1,x,t))-T_x(c(2,x,t))) \qquad (87)$$

$$\beta^{-1}\Delta y_i(t) = \Delta y_{i-1}(t-1) + \sum_{x=1}^{\infty}(T_x(c(i-1,x,t))-T_x(c(i,x,t))) \qquad 3\leqq i\leqq n \quad (88)$$

Now since $n < i$ leads always to $1 \leqq b(i-n)$ and $1 \leqq b(i-n-1)$, it follows that both $c(i-n,1,t)$ and $c(i-n-1,1,t)$ are in $C(t)$, that is, are finite. Then we have from (85) and (86)

$$\begin{aligned}\beta^{-1}\Delta y_i = {} & (\tilde{T}_1(c(i-n,1,t)) - \tilde{T}_1(c(i-n-1,1,t)))^* \\ & + \sum_{x=2}^{\infty}((T_{x-1}(\underbrace{c(i-1,x-1,t)}_{A}) - T_x(\underbrace{c(i-n-1,x,t)}_{B})) \\ & - (T_{x-1}(\underbrace{c(i,x-1,t)}_{C}) - T_x(\underbrace{c(i-n,x,t)}_{D})))^{**} \qquad n < i \leqq N \quad (89)\end{aligned}$$

It is clear from (84) to (88) that $y_i(t)$ and $\Delta y_i(t)$ are nonnegative for, respectively, $1 \leqq i \leqq N$ and $2 \leqq i \leqq n$. Next we shall show the nonnegativity of $\Delta y_i(t)$ for $n < i \leqq N$. Evidently the first term ( )* of (89) is nonnegative. The nonnegativity of the second term ( )** can be proved as follows. First notice that

(a1) b(i) is decreasing in i.

(a2) b(i-n) is greater by at most 1 than b(i).

(a3) The event of $[X-1]+1 > [X]$ dose not occur for any real number X.

The arguments A, B, C, and D in the right hand of (89) are not always in C(t), in other words, are not always finite. For this reason a somewhat cumbersome treatment as stated below will be needed for inspecting a sign of the second term. From (a1) and (a2) above we can easily obtain the next relations: (1) if $x-1 > b(i-1)$, then $x > b(i-n-1)$, $x-1 > b(i)$, and $x > b(i-n)$, (2) if $x > b(i-n-1)$, then $x > b(i-n)$, and (3) if $x-1 > b(i)$, then $x > b(i-n)$. These immediately yield, respectively, (1') if A is infinite, then so also are B, C, and D, (2') if B is infinite, then so also is D, and (3') if C is infinite, then so also is D. In addition we can prove that the joint event of (4) "$x \leqq b(i-n-1)$, $x-1 \leqq b(i)$, $b(i-n) < x$" is impossible. If the event may occur, then we have $b(i-n) < b(i-n-1)$, which leads to $\min\{k,[X]\} < \min\{k,[X+1/n]\}$ where $X = (N-i+1+n)/n$. From the inequality k must become greater than [X], that is, $k > [X]$, because $k \leqq [X]$ leads to the contradiction of $k < k$. Then we have $b(i-n) = [X]$ and $b(i) = [X-1]$. Now since we have $b(i-n) < b(i)+1$ from the assumption above, it follows that $[X] < [X-1]+1$, which is a contradiction from above (a3). Hence it follows that the joint event (4) is impossible. This means that the joint event of (4') B and C being finite as well as D being infinite cannot occur at all. It is easy from above (1') to (4') to check that the possible combinations as to finiteness or infiniteness of these four arguments are only the following five:

	A	B	C	D
(a)	Infinite	Infinite	Infinite	Infinite
(b)	Finite	Infinite	Infinite	Infinite
(c)	Finite	Infinite	Finite	Infinite
(d)	Finite	Finite	Infinite	Infinite
(e)	Finite	Finite	Finite	Finite

For each of (a) to (e) above, the second term ( )** of (89) becomes, respectively, noticing remark 6c,e

(a) 0

(b) $T_{x-1}(c(i-1,x-1,t)) \geqq 0$

(c) $T_{x-1}(c(i-1,x-1,t)) - T_{x-1}(c(i,x-1,t)) \geqq 0$

(d) $-\Delta T_x(c(i-1,x-1,t)) \geqq 0$

(e) $\Delta T_x(c(i,x-1,t)) - \Delta T_x(c(i-1,x-1,t)) \geqq 0$

Thus $\Delta y_i(t)$ becomes nonnegative for $n < i \leqq N$, hence for all i. Consequently it follows by induction that both $y_i(t)$ and $\Delta y_i(t)$ become nonnegative for all i and all t, hence so also become both limits y_i and Δy_i for all i. Now let $h_i = \sum_{x=1}^{b(i)} E_x$ $(\geqq 0)$. Then at a mere glance of (84) to (88) with attention to the inequality of $0 \leqq T_x(c) \leqq E_x$ for any nonegative number c(see lemma 7d), for limits y_i we get $0 \leqq y_1 \leqq \beta h_0$, $0 \leqq y_i \leqq \beta(y_{i-1} + h_0)$ for $2 \leqq i \leqq n$, and $0 \leqq y_i \leqq \beta(y_{i-1} + h_{i-n})$ for $n < i \leqq N$. From these inequalities the following can be recurrently derived: $0 \leqq y_i \leqq \beta d(i) \leqq d(i) \triangleq h_0 \min\{i,n\} + \sum_{x=1}^{(i-n)^+} h_x$. Thus $\beta^{-1} y_i$, or $v_{(i-n)^+} - v_i$ becomes bounded in β for all i. From this a boundedness of $v_i - v_0$ in β for all i can be easily brought about. Therefore it follows from theorem 2 that when $\beta = 1$, the maximum of the total expected value per period, g, exists. Now since $v_0 = \beta v_0 +$

$\sum_{x=1}^{b(0)} T_x(c(0,x))$ from (83) for limits v_i where $c(0,x) = y_{nx-1}$, we have

$$g = \sum_{x=1}^{b(0)} T_x(y_{xn-1}) \tag{90}$$

ACKNOWLEDGMENT

The auther wishes to extend appreciation to a lecturer, Y. Yamamoto, who contributed to the clarification and modification of mathematical aspects of the presentation.

REFERENCES

Arrow, K.J., Karlin, S., and Scarf, H., 1962. Studies in Applied Probability and Manegement Science, Stanford University Press, 148-158

Bellman, R., 1957. Dynamic Programming, Princeton University Press,

Derman, C., Lieberman, G.J. and Ross, S.M., 1972. A Sequential Stochastic Assignment Problem, Mgmt. Sci. 18, 349-355

Fabian, T., Fisher. J.L., Sasieni, M.W. and Yardeni, A. 1959. Purchasing Raw Material on a Fluctuating Market, Opns. Res. 7, 107-122

Gilbert, J.P. and Mosteller, F., 1966. Recognizing the Maximum of a Sequence, J. Amer. Statist. Assoc. 16, 35-73

Hayes, R.H., 1969. Optimal Strategies for Divestiture, Opns. Res. 17, 292-310

Hockman. E, 1973. An Optimal Stopping Problem of a Growing Inventory, Mgmt Sci. 19, 1289-1291

Howard, R.A., Dynamic Programming and Markov Processes, MIT press, (1960)

Ikuta, S., 1975. A method of Optimal Order Selection, Dissertation, Keio University

Lippman, S.A. and MaCall, J.J. 1976. Job Search in a Dynamic Economy, Journal of Economic theory, 12, 365-390

Mastran, D.V. and Thomas, C.J., 1973. Decison Rules for Attacking Targets of Opportunity, Nav. Res. Log. Quart., 20, 661-672

Miller, B.L., 1969. A Queueing Reward System with Several Customer Classes, Mgmt. Sci. 16, 234-245

Morris, W.T., 1959. Some Analysis of Purchasing Policy, Mgmt. Sci. 5, 443-452

Ross, S.M, 1970. Applied Probability Models with Optimization Applications, Holden-Day, San Francisco (1970), 144-147

Sakaguchi, M., 1961. Dynamic Programming of Some Sequential Sampling Design, J. Math. Anal. and Appl. 2, 446-466

Simon, H.A., 1957. Models of Man, Jhon Wily Sons, Inc.

Taylor, H.M., 1967. Evaluating a Call Option and Optimal Timing Strategy in the Stock Market, Mgmt. Sci. 14, 111-120

Yagishita, M., 1980, Inventory Model of a Material with a Fluctuating Price, Graduation Thesis, Aoyama Gakuin University.

Wagner, H.M., 1975. Principles of Operations Research with Applications to Managerial Decisions, Prentice-Hall

INSTITUTE OF SOCIO-ECONOMIC PLANNING

Discussion Paper Series (1974-1979.3)

No. 1 Shuntaro Shishido and Shinyasu Hoshino,
"Economic Planning Techniques in Japan," (December, 1974).

No. 2 Shuntaro Shishido and Akira Oshizaka,
"An Econometric Analysis of the Impacts of Pollution Control in Japan," (January, 1975).

No. 3 Shuntaro Shishido,
"Administrative Arrangements for Increasing Effective Planning Systems," (September, 1975).

No. 4 Koichi Mera,
"Changing Pattern of Population Distribution in Japan and its Implications to Developing Countries," (November, 1975).

No. 5 Shuntaro Shishido,
"Japan's Role in Future World Economy," (December, 1975).

No. 6 Haruo Onishi,
"An Operational Approach to a Worldwide Temporal Food Allocation and Price Determination Problem," (November, 1975).

No. 7 Shuntaro Shishido, Naoki Kitayama and Hajime Wago
"Changes in Regional Distribution of Population in Japan and Its Implications for Social Policy," (September, 1976).

No. 8 Koichi Mera,
"Population Concentration and Regional Income Disparity: A Comparative Analysis of Japan and Korea," (December, 1976).

No. 9 Hajime Eto,
"Statistical Methods to Measure the Consensus of Experts Opinions in Delphi Forecasts and Assessments," (January, 1977).

No. 10 Hajime Eto,
"Fuzzy Operational Approach to Analysis of Delphi Forecasting," (April, 1977).

No. 11 Hajime Eto,
"A Formal Approach to the Evaluation of Forecasts," (April, 1977).

No. 12 Hiroshi Atsumi,
"On Proportional Malinvaud Prices," (June, 1977).

No. 13 Atsuyuki Okabe,
"An Expected Rank-Size Rule : A Theoretical Relationship between the Rank-Size Rule and City Size Distributions," (April, 1977).

No. 14 Mamoru Kaneko,
"The Assignment Markets," (July, 1977).

No. 15 Hiroshi Atsumi,
"A Geometric Note on Global Monotonicity Theorem," (July, 1977).

No. 16 Atsuyuki Okabe,
"Some Reconsiderations of Simon's City Size Distribution Model," (July, 1977).

No. 17 Atsuyuki Okabe,
"Spatial Aggregation Bias in Trip Distribution Probabilities: The Case of the Gravity Model," (September, 1977).

No. 18 Mamoru Kaneko,
"Consideration of the Nash Social Welfare Function," (September, 1977).

No. 19 Koichi Mera and Hiroshi Ueno,
"Population Factors in Planning of Sub-national Areas: Their Roles and Implications in the Long-Run," (September, 1977).

No. 20 Haruo Onishi,
"On the Existence and Uniqueness of a Solution to an Operational Spatial Net Social Quasi-Welfare Maximization Problem," (October, 1977).

No. 21 Hajime Eto,
"Evaluation Model of Distribution Sector in Decentralized Economy," (December, 1977).

No. 22 Atsuyuki Okabe,
"Population Dynamics of Cities in a Region: Conditions for the Simultaneously Growing State," (January, 1978).

No. 23 Mamoru Kaneko,
"A Bilateral Monopoly and the Nash Solution," (January, 1978).

No. 24 Mamoru Kaneko,
"The Nash Social Welfare Function for a Measure Space of Individuals," (February, 1978).

No. 25 Hajime Eto,
"Generalized Domination and Fuzzy Domination in Preference Structure," (March, 1978).

No. 26 Atsuyuki Okabe,
"The Stable State Conditions of the Population-Dependent Migration Functions under No Population Growth," (April, 1978).

No. 27 Mamoru Kaneko,
"An Extension of the Nash Bargaining Problem and the Nash Social Welfare Function," (April, 1978).

No. 28 Hiroshi Atsumi,
"On Efficiency Prices of Competitive Programs in Closed Linear Models," (May, 1978).

No. 29 Mamoru Kaneko,
"A Measure of Inequality in Income Distribution," (June, 1978).

No. 30 Atsuyuki Okabe,
"Transportation and the Equilibrium Size of Cities in a Region," (September, 1978).

No. 31 Kozo Sasaki,
"Food Demand Matrix Derived from Additive Quadratic Model," (September, 1978).

No. 32 Yozo Ito and Mamoru Kaneko,
"Note on Linearizability of Cost Functions in Public Goods Economies," (November, 1978).

No. 33 Mamoru Kaneko,
"The Stable Sets of a Simple Game," (November, 1978).

No. 34 Atsuyuki Okabe,
"Spatially Constrained Clustering: Parametric and Nonparametric Methods for Testing the Spatially Homogeneous Clusters," (November, 1978).

No. 35 Ayse Gedik,
"Spatial Distribution of Population in Postwar Japan (1945-75) and Implications for Developing Countries," (November, 1978).

No. 36. Ayse Gedik,
"Sizes of Different Migration Flows in Turkey, 1965-70: Possible Future Directions and Towards Comparative Analysis," (December, 1978).

No. 37 Atsuyuki Okabe,
"An Application of the Spatially Constrained Cluster Method," (March, 1979).

No. 38 Yasoi Yasuda and Ryohei Nakamura,
"A Model of Social Dissatisfaction Function and Its Application to Regional Indicators," (March, 1979).

INSTITUTE OF SOCIO-ECONOMIC PLANNING

Discussion Paper Series (1979.4-1980.3)

No. 39 ----------,
....

No. 40 Hiroyuki Odagiri,
"Income Distribution and Growth in a Hierarchical Firm," (August, 1979).

No. 41 Koichi Mera,
"Basic Human Needs versus Economic Growth Approach for Coping with Urban-Rural Imbalances: An Evaluation Based on Relative Welfare," (April, 1979).

No. 42 Iwaro Takahashi,
"Switching Functions Constructed by Galois Extension Fields," (June, 1979).

No. 43 Takao Fukuchi,
"Growth and Stability of Multi-Regional Economy," (July, 1979).

No. 44 Atsuyuki Okabe,
"The Number of Quadrats and The Goodness-of-Fit Test of the Quadrat Method for Testing Randomness in the Distribution of Points on a Plane," (July, 1979).

No. 45 Nozomu Matsubara,
"Informational Evaluation of Decision Criteria in Situational Decision Making Model," (August, 1979).

No. 46 Mamoru Kaneko,
"The Optimal Progressive Income Tax -- The Existence and the Limit Tax Rates," (July, 1979).

No. 47 Yozo Ito and Mamoru Kaneko,
"Ratio Equilibrium in an Economy with an Externality," (August, 1979).

No. 48 Hajime Eto,
Effectiveness of Decentralization with Power Separation in Central Authority," (September, 1979).

No. 49 Yukio Oguri,
"Relocation Demand and Housing Preference of the Households of the Tokyo Metropolitan Region: A Metropolitan Residential Relocation Survey," (October, 1979).

No. 50 Hiroyuki Odagiri,
"Advertising and Welfare: A Pedagogical Note," (September, 1979).

No. 51 Satoru Fujishige,
"Lexicographically Optimal Base of a Polymatroid with respect to a Weight Vector," (September, 1979).

No. 52 Satoru Fujishige,
"A New Efficient Algorithm for Finding Shortest Paths in Networks with Arcs of Negative Length," (October, 1979).

No. 53 Hajime Eto,
"Decentralization Model with Coordination in Terms of Policy Selection," (December, 1979).

No. 54 Yoshiko Nogami,
"A Non-Regular Squared-Error loss Set-Compound Estimation Problem," (September, 1979).

No. 55 Mikoto Usui,
"Technological Capacitation and International Division of Labor," (September, 1979).

No. 56 Takao Fukuchi, Fumio Isaka and Mamoru Obayashi,
"Economic Growth and Exchange Rate Systems," (October, 1979).

No. 57 Takatoshi Tabuchi,
"Optimal Distribution of City Sizes in a Region," (November, 1979).

No. 58 Ayse Gedik,
"Descriptive Analyses of Vilage-to-Province-Center Migration in Turkey: 1965-70," (November, 1979).

No. 59 Shoichiro Kusumoto,
"Price Strategic Economic Behaviour in an Exchange Economy -- A General (Non-) Walrasian Prototype, PART 1," (November, 1979).

No. 60 Atsuyuki Okabe,
"Statistical Test of the Pattern Similarity between Two Sets of Regional Clusters," (November, 1979).

No. 61 Yukio Oguri,
"A Residential Search Routine for A Metropolitan Residential Relocation Model," (December, 1979).

No. 62 Mikoto Usui,
"Advanced Developing Countries and Japan in Changing International Economic Relationships," (December, 1979).

No. 63 Shigeru Matsukawa,
"Fringe Benefits in a Dynamic Theory of the Firm," (January, 1980).

No. 64 Takao Fukuchi,
"A Dynamic Analysis of Urban Growth," (December, 1979).

No. 65 Ryosuke Hotaka,
"A Design of the Integrated Data Dictionary Directory System," (January, 1980).

No. 66 Shoichi Nishimura,
"Monotone Optimal Control of Arrivals Distinguished by Reward and Service Time," (January, 1980).

No. 67 Yozo Ito and Mamoru Kaneko,
"A Game Theoretical Interpretation of the Stackelberg Disequilibrium," (January, 1980).

No. 68 Sho-Ichiro Kusumoto,
"Global Aspects of the Economic Integrability Theory -- Equivalence Theorems on the Hypothesis of Economic Man," (February, 1980).

No. 69 Satoru Fujishige,
"An Efficient PQ-Graph Algorithm for Solving the Graph-Realization Problem," (February, 1980).

No. 70 Koichi Mera,
"The Pattern and Pace of Urbanization and Socio-Economic Development : A Cross-Sectional Analysis of Development Since 1960," (March, 1980).

No. 71 Atsuyuki Okabe,
"A Note : Spatial Distributions Maximizing or Minimizing Geary's Spatial Contiguity Ratio," (March, 1980).

No. 72 Isao Ohashi,
"Wage Profiles and Layoffs in the Theory of Specific Training," (March, 1980).

INSTITUTE OF SOCIO-ECONOMIC PLANNING

Discussion Paper Series (1980.4-1981.3)

No. 73 --------,
....

No. 74 Seizo Ikuta,
"A Sequential Selection Process and Its Applications," (April, 1980).

No. 75 Mamoru Kaneko,
"On the Existence of an Optimal Income Tax Schedule," (April, 1980).

No. 76 Kazumi Asako,
"Heterogeneity of Labor, the Phillips Curve, and Stagflation," (April, 1980).

No. 77 Hiroyuki Odagiri,
"Worker Participation and Growth Preference: A Theory of the Firm with Two-Layer Hierarchical Structure and Profit Sharing," (June, 1980).

No. 78 Yoshimi Kuroda,
"Production Behavior of the Farm Household and Marginal Principles on Postwar Japan," (April, 1980).

No. 79 Kazumi Asako,
"Rational Expectations and the Effectiveness of Monetary Policy with a Special Reference to the Barro-Fischer Model," (May, 1980).

No. 80 Takao Fukuchi and Makoto Yamaguchi,
"An Econometric Analysis of Tokyo Metropolis," (July, 1980).

No. 81 Satoru Fujishige,
"Canonical Decompositions of Symmetric Submodular Systems," (June, 1980).

No. 82 Kazumi Asako,
"On the Simultaneous Estimation of Means and Variances of the Random Coefficient Model," (July, 1980).

N0. 83 Yoshitsugu Kanemoto,
"Price-Quantity Dynamics in a Monopolistically Competitive Economy with Small Inventory Costs," (June, 1980).

No. 84 Nozomu Matsubara,
"The N-part Partition of Risks," (July, 1980).

No. 85 Atsuyuki Okabe,
"A Static Method of Qualitative Trend Curve Analysis," (September, 1980).

No. 86 Shigeru Matsukawa,
"Dualistic Development in the Manufacturing Sector : Japan's Experience," (July, 1980).

No. 87 Hiroyuki Odagiri,
"Antineoclassical Management Motivation in a Neoclassical Economy: An Interpretation of Japan's Economic Growth," (August, 1980).

No. 88 Koichi Mera,
"City Size Distribution and Income Distribution in Space," (August, 1980).

No. 89 Yoshitsugu Kanemoto, Mukesh Eswaran and David Ryan,
"A Dual Approach to the Locational Decision of the Firm," (October, 1980).

No. 90 Hajime Eto,
"Evaluation of the Reformed Division System with Enforcement of Short-Range Corporate Strategy," (August, 1980).

No. 91 Shuntaro Shishido,
"Long-Term Forecast and Policy Implications : Simulations with a World Econometric Model (T - FAIS IV)," (September, 1980).

No. 92 Isao Ohashi,
"A Model of Labor Quality, Wage Differentials, and Unemployment," (September, 1980).

No. 93 Sho-Ichiro Kusumoto,
"The Economic Location Theory -- Revisited a Confirmation," (September, 1980).

NO. 94 Seizo Ikuta,
"A Generalization of a Sequential Selection Process by Introducing an Extended Shortage Function," (October, 1980).

No. 95 Kazumi Asako and Ryuhei Wakasugi,
"Some Findings on an Empirical Aggregate Production Function with Government Capital," (October, 1980).

No. 96 Yoshimi Kuroda and Pan A. Yotopoulos,
"A Subjective Equilibrium Model of the Agricultural Household with Demographic Behavior -- A Methodological Note --," (November, 1980).

No. 97 Atsuyuki Okabe,
"Relative Efficiency of Simple Random, Stratified Random and Systematic Sampling for Estimating an Area of a Certain Land Use," (November, 1980).

No. 98 Hideto Sato,
"Handling Summary Information in Databases: Derivability," (November, 1980).

No. 99 Yoshitsugu Yamamoto,
"Subdivisions and Triangulations induced by a Pair of Subdivided Manifolds," (December, 1980).

No. 100 Sho-Ichiro Kusumoto,
"Foundations of the Economic Theory of Location -- Transport Distance v.s. Substitution," (January,1981).

No. 101 Hideto Sato,
"Handling Summary Information in a Database: Categorization and Summarization," (January, 1981).

No. 102 Kazumi Asako,
"Utility Function and Superneutrality of Money on the Transition Path in a Monetary Optimizing Model," (February 1981).

No. 103 Yoshitsugu Yamamoto,
"A Note on Van Der Heyden's Variable Dimension Algorithm for the Linear Complementarity Problem," (February, 1981).

No. 104 Kanemi Ban,
"Estimation of Consumption Function with a Stochastic Income Stream," (February, 1981).

No. 105 Ryosuke Hotaka and Masaaki Tsubaki,
"Sentential Database Design Method," (February, 1981).

No. 106 Yoshitsugu Kanemoto,
"Housing as an Asset and Property Taxes," (February, 1981).

No. 107 Nozomu Matsubara, Jack Carpenter and Motoharu Kimura,
"Possible Application of the James-Stein Estimator to Several Regression Lines," (March, 1981).

No. 108 Shuntaro Shishido and Hideto Sato,
"An Econometric Analysis of Multi-Country Multipliers under fixed and Floating Exchange Rate Regimes," (March, 1981).

No. 109 Yasoi Yasuda and Ken Watanabe,
"An Equitable Cost Allocation of Cooperation Sewerage System as Regional Public Goods," (April 1981).

No. 110 Kazumi Asako,
"On the Optimal Short-Run Money-Supply Management under the Monetarist Long-Run Money-Supply Rule," (March, 1981).

No. 111 Yoshitsugu Yamamoto,
"A New Variable Dimension Algorithm for the Fixed Point Problem," (March, 1981).

No. 112 ----------,
....

INSTITUTE OF SOCIO-ECONOMIC PLANNING

Discussion Paper Series (1981.4-)

No. 113 Sho-Ichiro Kusumoto,
"On the Equilibrium Concepts in a General Equilibrium Theory with Public Goods and Taxes-Pareto Optimallity and Existence," (April, 1981).

No. 114 Ryosuke Hotaka,
"A Meta-Database for a Database Design Method," (May, 1981).

No. 115 Hidehiko Tanimura,
"A Minimum-Distance Location Model Central Facilities with Entropy-Maximizing Spatial Interaction," (May, 1981).

No. 116 Sho-Ichiro Kusumoto, Kanemi Ban, Hajime Wago and Kazumi Asako,
"Rational Savings, Price Expectation and Money Supply in a Growing Economy," (July, 1981).

No. 117 Sho-Ichiro Kusumoto,
"On the Equilibrium Concepts in a General Equilibrium Theory with Public Goods and Taxes II -- "Surplus" Maximum," (June, 1981).

No. 118 Hajime Eto,
"Decision-Theoretical Foundations of the Validities of Technology Forecasting Methods," (June, 1981).

No. 119 Hiroyuki Odagiri,
"Internal Promotion, Intrafirm Wage Structure and Corporate Growth," (July, 1981).

No. 120 Hajime Eto and Kyoko Makino,
"The Validity of the Simon's Firm-Size Model and its Revision," (August, 1981).

No. 121 Satoru Fujishige,
"Structures of Polytopes Determined by Submodular Functions on Crossing Families," (August, 1981).

No. 122 Hajime Eto,
"Epistemologico-Logical Approach to the Validity of Model in View of the Fuzzy System Model," (August, 1981).

No. 123 Hiroyuki Odagiri,
"R & D Expenditures, Royalty Payments, and Sales Growth in Japanese Manufacturing Corporations," (August, 1981).

No. 124 Noboru Sakashita,
"Evalution of Regional Development Policy-An Alternative Approach," (September, 1981).

No. 125 Takao Fukuchi and Noriyoshi Oguchi,
"A Generalization of Keynesian Macromodel; the STIC Model," (September, 1981).

No. 126 Kazumi Asako,
"The Penrose Effect and the Long-Run Equilibrium of a Monetary Optimizing Model: Superneutrality and Nonexistence," (September, 1981).

No. 127 Isao Ohashi,
"Optimal Properties of Wage and Layoff Policies and the Impact of Trade Unionism," (August, 1981).

No. 128 Kanemi Ban,
"A Macroeconometric Model of Japan with Rational Expectations," (September, 1981).

No. 129 Yoshitsugu Yamamoto and Kiyoshi Murata,
"A New Variable Dimension Algorithm : Extension for Separable Mappings, Geometric Interpretation and Some Applications," (October, 1981).

No. 130 Hirotaka Sakasegawa,
"Numerical Analysis on Tandem Queueing System with Blocking," (November, 1981).

No. 131 Ikuo Kabashima,
"Supportive Participatory Model of Development -Political Participation and Income Distribution in Growing Economies- ," (December, 1981).

No. 132 Seizo Ikuta,
"Multistage Decision Process with Random Observations and its Applications," (December, 1981).