MANAGEMENT SCIENCE
Vol. 20, No. 11, July, 1974
Printed in U.S.A.

CONVERGENCE RESULTS AND APPROXIMATIONS FOR OPTIMAL (s, S) POLICIES*†

ARIE HORDIJK AND HENK TIJMS

Mathematisch Centrum, Amsterdam, Holland

In this paper we consider the dynamic inventory model with a discrete demand and no discounting. We verify a conjecture of Iglehart about the asymptotic behaviour of the minimal total expected cost. To do this, we give for the denumerable state dynamic programming model a number of conditions under which the minimal total expected cost for the n-stage model minus n times the minimal average cost has a finite limit as $n \to \infty$. For a positive demand distribution we establish a turnpike theorem which states that for all n sufficiently large the optimal n-stage policy (s_n, S_n) is average cost optimal. Further, we show that the computation of the (s_n, S_n) policies supplies monotonic upper and lower bounds on the minimal average cost. Also, the average cost of the (s_n, S_n) policy lies between the corresponding bounds. For a positive demand distribution these bounds converge as $n \to \infty$ to the minimal average cost.

1. Introduction

We consider the single-item dynamic inventory model with a discrete demand and no discounting. A fixed set-up cost, a linear purchase cost, convex holding and shortage costs, backlogging of excess demand, and a zero lead time are assumed. To derive asymptotic properties of this model, we discuss in §2 the asymptotic behaviour of the minimal total expected cost for the denumerable state dynamic programming model. We give in §3 a number of known results for the inventory model that will be needed in the sequel. In §4 we prove that for a positive demand distribution the minimal total expected cost for the n-period inventory model minus n times the minimal average expected cost per period has a finite limit as $n \to \infty$ which can be explicitly given up to a constant. For a continuous demand this result was first proved by Iglehart [5] for the case of no set-up cost and was conjectured by him for the case of a positive set-up cost. In §5 we establish under the assumption of a positive demand distribution a turnpike theorem which states that for all n sufficiently large the optimal n-stage policy (s_n, S_n) is also average cost optimal. Further, we show that the recursive method to compute the optimal n-stage policies (s_n, S_n) supplies monotonic upper and lower bounds on the minimal average cost. Moreover, the average cost of the (s_n, S_n) policy lies between the corresponding upper and lower bound. When the demand distribution is positive these bounds converge as $n \to \infty$ to the minimal average cost.

2. The Asymptotic Behaviour of the Minimal Total Expected Cost for Denumerable State Dynamic Programming

Consider a dynamic system which at times $t = 1, 2, \cdots$ is observed to be in one of a possible number of states. Let $\mathcal{I}$ denote the set of all possible states. We assume $\mathcal{I}$ to be *denumerable*. After observing state i, an action a must be chosen from a *finite* set $A(i)$ of possible actions. If the system is in state i at time t and action a is chosen, then, regardless of the history of the system, two things happen: (i) we incur an (ex-

* Processed by Associate Professor Eric V. Denardo, Departmental Editor for Dynamic Programming and Inventory Theory; received May 1973.

† This paper is registered as Mathematical Centre Report BW 22/73.

pected) cost $c(i, a)$, and (ii) at time $t + 1$ the system will be in state j with probability $p_{ij}(a)$. The costs $c(i, a)$ and the transition probabilities $p_{ij}(a)$ are assumed to be known. We suppose that the costs $c(i, a)$ are *nonnegative*. No further boundedness condition is imposed on the costs.

Denote by X_t and Δ_t, $t = 1, 2, \cdots$, the sequences of states and actions. A *policy* R for controlling the system is any (possibly randomized) rule which for each t specifies which action to take at time t given the current state X_t and the history $(X_1, \Delta_1, \cdots, X_{t-1}, \Delta_{t-1})$. A stationary policy f is a rule that for each i selects an action $f(i) \in A(i)$ such that always action $f(i)$ is taken whenever the system is in state i. Observe that $\{X_t\}$ is a stationary Markov chain when a stationary policy is used. For any state i and policy R, let

$$\phi(i, R) = \liminf_{n\to\infty} (1/n) \sum_{t=1}^{n} E_R \{c(X_t, \Delta_t) \mid X_1 = i\},$$

where E_R denotes the expectation under policy R. Observe that $\phi(i, R)$ exists ($+\infty$ is admitted), since $c(i, a) \geqq 0$. When the limit exists $\phi(i, R)$ is the long run average expected cost per unit time when the initial state is i and policy R is used. A policy R^* is called *average cost optimal* if $\phi(i, R^*) = \inf_R \phi(i, R)$ for all $i \in \mathcal{I}$.

Let $v_0(i)$, $i \in \mathcal{I}$, be an arbitrary function such that $\sum_j p_{ij}(a) v_0(j)$ is *finite* for all i and a *and is bounded from below* in i and a. Define for $n = 1, 2, \cdots$

$$v_n(i) = \min_{a\in A(i)} \{c(i, a) + \sum_{j\in\mathcal{I}} p_{ij}(a) v_{n-1}(j)\} \quad \text{for} \quad i \in \mathcal{I}.$$

Observe that for each n the function $v_n(i)$ exists, since $c(i, a) \geqq 0$. To determine the asymptotic behaviour of $v_n(i)$, we introduce the following assumptions.

ASSUMPTION 1. There is a finite constant g and a finite function $v(i)$, $i \in \mathcal{I}$, such that

(i) $\sum_j p_{ij}(a) v(j)$ is absolutely convergent for all i and a, and

$$(1) \qquad g + v(i) = \min_{a\in A(i)} \{c(i, a) + \sum_{j\in\mathcal{I}} p_{ij}(a) v(j)\} \quad \text{for all} \quad i \in \mathcal{I}.$$

(ii) $E_R\{v(X_n) \mid X_1 = i\}$ is finite for all i, R and n, and $(1/n) E_R\{v(X_n) \mid X_1 = i\} \to 0$ as $n \to \infty$ for all i and R.

Let $F_{\text{opt}} = \{f \mid f$ is a stationary policy such that $f(i)$ minimizes the right-hand side of (1) for all $i \in \mathcal{I}\}$. By the remark following the proof of Theorem 1 in [10] we have $\inf_R \phi(i, R) = g$ for all i and each policy from F_{opt} is average cost optimal. Hence the minimal average cost is independent of the initial state and equals g.

ASSUMPTION 2. The function $v_1(i) - v(i)$, $i \in \mathcal{I}$, is bounded.

ASSUMPTION 3. For each stationary policy f the associated Markov chain $\{X_t\}$ is nondissipative, that is, the set of positive recurrent states is not empty and from each initial state the set of positive recurrent states will be reached with probability one.

ASSUMPTION 4. For each policy $f \in F_{\text{opt}}$ holds that each state which is positive recurrent under policy f is aperiodic.

ASSUMPTION 5. For each average cost optimal stationary policy the associated Markov chain $\{X_t\}$ has no two disjoint closed sets.

THEOREM 1. (a) *If part* (i) *of Assumption* 1 *and Assumption* 2 *are satisfied, then there is a finite constant* B *such that* $| v_n(i) - ng - v(i) | \leqq B$ *for all* $n \geqq 1$ *and all* $i \in \mathcal{I}$.

(b) *If the Assumptions* 1–4 *are satisfied, then* $\lim_{n\to\infty} \{v_n(i) - ng - v(i)\}$ *exists for all* i *and is bounded in* i. *This limit is independent of* i *if, in addition, Assumption* 5 *is satisfied.*

A proof of Theorem 1 can be found in [3]. The case of a finite state space has been studied in [1], [2], [8] and [12]; the proof in [3] borrows from [8] and [12].

3. The Inventory Model and Preliminaries

We consider an inventory model in which the demands $\xi_1, \xi_2, \cdots$ for a single item in periods $t = 1, 2, \cdots$ are independent random variables having a common probability distribution $\phi(j) = P\{\xi_t = j\}$ $(j = 0, 1, \cdots; t = 1, 2, \cdots)$. We assume that $\mu = E\xi_t$ is finite and positive. Any unfilled demand in a period is completely backlogged. At the beginning of each period the stock on hand is reviewed. At each review an order may be placed for any positive integral amount of stock. An order, when placed, is immediately delivered (the case of a fixed positive lead time can be reduced to the case of a zero lead time, see [11]). The demand in each period takes place after review and delivery (if any). The stock on hand may take on any integral value, where a negative value indicates the existence of a backlog. The following costs are involved. The cost of ordering j units is $K\delta(j) + c \cdot j$, where $K \geqq 0$, $c \geqq 0$, $\delta(0) = 0$, and $\delta(j) = 1$ for $j \geqq 1$. Let $L(k)$ be the expected holding and shortage costs in a period when k is the amount of stock on hand at the beginning of that period just after any additions to stock. We assume that $L(k)$ is nonnegative and *convex*, i.e. $L(k + 1) - L(k) \geqq L(k) - L(k - 1)$ for all k. For convenience it is assumed that both $L(k) \to \infty$ and $ck + L(k) \to \infty$ as $|k| \to \infty$. Finally, future costs are not discounted.

We now give a number of known results for this inventory model.

(a) *The finite period model.* Let Z be the set of all integers. Define $v_0(i) = 0$ for all $i \in Z$, and for $n = 1, 2, \cdots$, let

$$(2) \quad v_n(i) = \inf_{k \geqq i}\{c \cdot (k - i) + K\delta(k - i) + L(k) + \textstyle\sum_{j=0}^{\infty} v_{n-1}(k - j)\phi(j)\}, \quad i \in Z.$$

The choice $v_0(i) \equiv 0$ can be interpreted as follows. In the finite period model it is assumed that stock left over at the end of the final period has no value and backlogged demand remaining at the end of the final period is satisfied at a cost zero. Scarf [11] proved that, for each $n = 1, 2, \cdots$,

$$(3) \quad \begin{aligned} v_n(i) &= -ci + K + G_n(S_n) \quad &\text{for} \quad i < s_n, \\ &= -ci + G_n(i) \quad &\text{for} \quad i \geqq s_n, \end{aligned}$$

where $G_n(k) = ck + L(k) + \sum_{j=0}^{\infty} v_{n-1}(k - j)\phi(j)$, S_n is the smallest integer which minimizes the K-convex function $G_n(k)$, and s_n is the smallest integer satisfying $G_n(s_n) \leqq K + G_n(S_n)$. Hence the right-hand side of (2) is minimal for $k = S_n$ when $i < s_n$ and for $k = i$ when $i \geqq s_n$. The quantity $v_n(i)$ is the minimal total expected cost for the n-period model when the initial state is i, and $v_n(i)$ is achieved by the following policy of the (s, S) type: If at the beginning of period t the stock on hand $j < s_t$, order $S_t - j$ units; otherwise, do not order in period t $(t = 1, \cdots, n)$. Finally, the integers s_n and S_n are bounded ([4], [5], [7], [14]).

(b) *The infinite period model.* We first introduce some notation. Let the renewal quantity $m(j)$ be defined by $m(j) = \phi(j) + \sum_{k=0}^{j} \phi(j - k)m(k)$, $j = 0, 1, \cdots$, and let $M(j) = \sum_{k=0}^{j} m(k)$. Let s and S be integers with $s \leqq S$. The (s, S) policy is a stationary policy of the following form: if, at review, the stock on hand $i < s$, then $S - i$ units are ordered; otherwise, no order is placed. When an (s, S) policy is used the sequence of stock levels at the beginning of subsequent periods just before review is a Markov chain that has a unique stationary probability distribution [5], [13], [14], say $\{q_j(s, S)\}$. Clearly, $q_j(s, S) = 0$ for $j > S$, and

$$(4) \quad q_j(s, S) = \textstyle\sum_{i=-\infty}^{s-1} q_i(s, S)\phi(S - j) + \sum_{i=s}^{S} q_i(s, S)\phi(i - j) \quad \text{for all } j,$$

where $\phi(k) = 0$ for $k < 0$. We note that $\sum jq_j(s, S)$ is finite. Denote by $a(s, S)$ the long run average expected cost per period when an (s, S) policy is used. The quantity $a(s, S)$ is independent of the initial stock and is given by [5], [13], [14]

$$(5)\quad \begin{aligned} a(s, S) &= \textstyle\sum_{j=-\infty}^{s-1} \{c\cdot(S - j) + K + L(S)\}q_j(s, S) + \sum_{j=s}^{S} L(j)q_j(s, S) \\ &= \{L(S) + \textstyle\sum_{k=0}^{S-s} L(S - k)m(k) + K\}/\{1 + M(S - s)\} + c\mu. \end{aligned}$$

Let g be defined as

$$g = \min\{a(s, S) \mid s \leqq S,\ s, S \in Z\}.$$

The constant g exists and is finite. Now fix finite integers s^* and S^* with $s^* \leqq S^*$ such that $g = a(s^*, S^*)$ and $L(s^* - 1) \geqq g - c\mu \geqq L(s^*)$. Such integers exist ([5], [6], [13]). From definition the (s^*, S^*) policy is average cost optimal among the class of the (s, S) policies. However, the (s^*, S^*) policy is also average cost optimal among the class of all possible policies ([5], [6], [13]). Hence the minimal average expected cost is independent of the initial stock and equals g. Define the finite function $v(i)$, $i \in Z$, by

$$(6)\quad \begin{aligned} v(i) &= -c\cdot(i - s^* + 1), && i < s^*, \\ &= L(i) + \textstyle\sum_{k=0}^{i-s^*} L(i - k)m(k) - \{g - c\mu\}\{1 + M(i - s^*)\}, && i \geqq s^*. \end{aligned}$$

Then ([5], [13]) (in [5] the continuous demand version is given),

$$(7)\quad \begin{aligned} g + v(i) = \min\nolimits_{k \geqq i}\{&c\cdot(k - i) + K\delta(k - i) \\ &+ L(k) + \textstyle\sum_{j=0}^{\infty} v(k - j)\phi(j)\}, \qquad i \in Z, \end{aligned}$$

where the right-hand side of (7) is minimized by $k = S^*$ for $i < s^*$ and by $k = i$ for $i \geqq s^*$.

4. The Asymptotic Behaviour of the Minimal Total Expected Cost for the Inventory Model

In this section we shall prove that if $\phi(i) > 0$ for all i sufficiently large, then $v_n(i) - ng$ has a finite limit as $n \to \infty$ for all i. To do this, we shall define a Markovian decision model which has both the same probabilistic structure and the same cost structure as the inventory model under consideration. Choose finite integers L and U such that $L < s_n \leqq S_n \leqq U$ for all n and $L < s^* \leqq S^* \leqq U$. Consider now the Markovian decision model defined by (cf. §2),

$$\begin{aligned} \mathcal{I} &= \{i \mid i \text{ integer},\ i \leqq U\}, \\ A(i) &= \{a \mid a \text{ integer},\ \max(i, L) \leqq a \leqq U\}, \qquad (i \in \mathcal{I}), \\ c(i, a) &= c\cdot(a - i) + K\delta(a - i) + L(a), \qquad \text{and} \\ p_{ij}(a) &= \phi(a - j), \qquad (a \in A(i);\ i, j \in \mathcal{I}). \end{aligned}$$

By (2), (3) and (7) we have $v_n(i) = \min_a\{c(i, a) + \sum_j p_{ij}(a)v_{n-1}(j)\}$ for all $i \in \mathcal{I}$ and all $n \geqq 1$, and $g + v(i) = \min_a\{c(i, a) + \sum_j p_{ij}(a)v(j)\}$ for all $i \in \mathcal{I}$. Since $v(i)$ is linear for $i < s^*$ and $\mu < \infty$, the series $\sum_j v(k - j)\phi(j)$ converges absolutely for all k. By (3) and (6), $v_1(i) - v(i)$ is bounded in $i \in \mathcal{I}$. Hence part (i) of Assumption 1 and Assumption 2 are satisfied. For this Markovian decision model the state X_t at time t denotes the stock on hand just before ordering in period t and the action Δ_t at time t denotes the stock on hand just after ordering in period t. Since excess demand is backlogged, we have $X_{t+1} = \Delta_t - \xi_t$ for $t \geqq 1$. Further, $X_t \leqq U$ and $L \leqq \Delta_t \leqq U$

for all $t \geqq 1$. Since $v(i)$ is linear for $i < s^*$ and $\mu = E\xi_t$ is finite, it now follows that $E_R\{v(X_n) \mid X_1 = i\}$ is bounded in n for each policy R and each i. so part (ii) of Assumption 1 is also satisfied. Suppose now that $\phi(i) > 0$ for all i sufficiently large. Then, for each stationary policy, the associated Markov chain $\{X_t\}$ has a nonempty set of aperiodic positive recurrent states, a finite number of transient states and no two disjoint closed sets. Hence the Assumptions 3–5 are also satisfied; so, by part (b) of Theorem 1, there is a finite constant γ such that $v_n(i) - ng - v(i)$ converges as $n \to \infty$ to γ for all $i \leqq U$. Since U can be chosen arbitrarily large, we have proved the next theorem.

THEOREM 2. *If $\phi(i) > 0$ for all i sufficiently large, then there is a finite constant γ such that* $\lim_{n\to\infty}\{v_n(i) - ng\} = v(i) + \gamma$ *for all $i \in Z$.*

This result was first proved in [5] for the case of $K = 0$. The next example shows that $v_n(i) - ng$ may diverge when the condition of Theorem 2 is not satisfied. Suppose that $\phi(1) = 1$, $c = 0$, $K = 1$, $L(1) = 0$, and $L(k) = 2\mid k \mid$ for $k \neq 1$. Then, $v_{2n-1}(-1) = v_{2n}(-1) = n$ for all $n \geqq 1$, and $g = \frac{1}{2}$. Moreover, $(s_{2n-1}, S_{2n-1}) = (0, 0)$ and $(s_{2n}, S_{2n}) = (0, 1)$, where $a(0, 0) = 1$ and $a(0, 1) = \frac{1}{2}$.

REMARK. In this remark we consider the choice $v_0(i) = -ci$ for all i. This choice corresponds to the case where in the finite period inventory model each unit of stock left over at the end of the final period can be salvaged with a return of c and each unit of backlogged demand remaining at the end of the final period is satisfied at a cost of c. For this case, let $v'_n(i)$ be the minimal total expected cost for the n-period model. The inventory model with a salvage cost c and a salvage value c can be reduced to an equivalent model with no salvage cost and no salvage value (see [14, pp. 528–529]). Using this reduction it is easily verified that Theorem 2 also holds for the choice $v_0(i) = -ci$ provided that we replace $v_n(i)$ by $v'_n(i)$. Moreover, the assumption $L(k)$ is convex can be weakened to $-L(k)$ is unimodal (cf. [6], [13] and [15]).

5. A Turnpike Planning Horizon Theorem and Approximations

We first prove the following turnpike planning horizon theorem.

THEOREM 3. *If $\phi(i) > 0$ for all i sufficiently large, then there is a finite integer n_0 such that for all $n \geqq n_0$ the (s_n, S_n) policy is average cost optimal.*

PROOF. Since $\{s_n\}$ and $\{S_n\}$ are bounded sequences of integers, there is an integer n_0 such that (s_r, S_r) is a limit point of the sequence $\{s_n, S_n\}$ for each $r \geqq n_0$. Fix now an (s, S) policy such that $(s_n, S_n) = (s, S)$ for some $n \geqq n_0$. Choose a sequence $\{n_k\}$ with $n_k \to \infty$ as $k \to \infty$ such that $(s_{n_k}, S_{n_k}) = (s, S)$ for all k. By Theorem 2 there is a finite constant γ such that $v_n(i) - ng$ converges as $n \to \infty$ to $v(i) + \gamma$ for all i. Subtracting $n_k g$ from both sides of (3) with n replaced by n_k and letting $k \to \infty$, we find

$$(8)\quad \begin{aligned} v(i) + \gamma &= c\cdot(S - i) + K + L(S) - g + \textstyle\sum_{j=0}^{\infty} v(S - j)\phi(j) + \gamma, && \text{for} \quad i < s, \\ &= L(i) - g + \textstyle\sum_{j=0}^{\infty} v(i - j)\phi(j) + \gamma, && \text{for} \quad s \leqq i \leqq S. \end{aligned}$$

The derivation of this equality involves an interchange of limit and summation which is justified by the fact that $\sum_j v(k - j)\phi(j)$ is absolutely convergent for all k and, for some finite constant B, $\mid v_n(i) - ng \mid \leqq v(i) + B$ for all $n \geqq 1$ and all $i \leqq S$ (see part (a) of Theorem 1 and § 4). It is now standard to prove that $a(s, S) = g$. To do this, multiply both sides of (8) with the stationary probability $q_i(s, S)$ and sum over i. Using (4) and (5), we then find $g = a(s, S)$, so the (s, S) policy is average cost optimal. This derivation of $g = a(s, S)$ involves an interchange of the order of summation which is justified by the fact that $\sum_j v(j)q_j(s, S)$ is absolutely convergent.

We note that for the discounted cost criterion an analogous turnpike planning horizon theorem holds without the assumption that $\phi(i) > 0$ for all i sufficiently large ([4] and [14, pp. 530–531]; see also [7]). Further, we note that Theorem 3 implies that the cycling found in the examples given on p. 695 in [16] must stop after a finite number of iterations.

We shall now demonstrate that the recursive method to compute the optimal n-stage policies (s_n, S_n) yields approximations both for the minimal average cost and for an average cost optimal (s, S) policy. To prove this, we shall first specify the bounds on s_n and S_n. Let $\mathbf{s}$ be the smallest integer for which $c\mathbf{s} + L(\mathbf{s}) \leqq K + \min_k \{ck + L(k)\}$. Define $\mathbf{S}$ as the smallest integer for which $L(k)$ is minimal, and let $\bar{S}$ be the smallest integer not less than $\mathbf{S}$ for which $L(\bar{S} + 1) \geq K + L(\mathbf{S})$. Observe that $\mathbf{s}$, $\mathbf{S}$ and $\bar{S}$ exist, since both $ck + L(k) \to \infty$ and $L(k) \to \infty$ as $|k| \to \infty$. Then [7], [14], $\mathbf{s} \leqq s_n \leqq S_n \leqq \bar{S}$ for all n and, moreover, there is an average cost optimal (s, S) policy such that $\mathbf{s} \leqq s \leqq S \leqq \bar{S}$.

THEOREM 4. *For any $n \geq 2$, let $r_n = \min(s_{n-1}, s_n) - 1$, and let*

$$\begin{aligned} L_n &= \min \{v_n(i) - v_{n-1}(i) \mid r_n \leqq i \leqq \bar{S}\}, \\ U_n &= \max \{v_n(i) - v_{n-1}(i) \mid r_n \leqq i \leqq \bar{S}\}, \\ U'_n &= \max \{v_n(i) - v_{n-1}(i) \mid r_n \leqq i \leqq S_n\}. \end{aligned}$$

Then,

(a) *$L_n \leqq g \leqq a(s_n, S_n) \leqq U'_n \leqq U_n$ for all $n \geqq 2$.*

(b) *L_n is nondecreasing and U_n is nonincreasing in n.*

(c) *If $\phi(i) > 0$ for all i sufficiently large, then both L_n, U_n and U'_n converge as $n \to \infty$ to g.*

PROOF. (a) Let $F = \{(s, S) \mid \mathbf{s} \leqq s \leq S \leqq \bar{S}\}$. Then $(s_n, S_n) \in F$ for all $n \geqq 1$, and $g = a(s, S)$ for some $(s, S) \in F$. Fix now $n \geqq 2$. By (2) and (3) we have for any (s, S) policy

$$(9) \qquad \begin{aligned} v_n(i) &\leqq c\cdot(S - i) + K + L(S) + \textstyle\sum_{j=0}^{\infty} v_{n-1}(S - j)\phi(j) \quad &&\text{for} \quad i < s, \\ &\leqq L(i) + \textstyle\sum_{j=0}^{\infty} v_{n-1}(i - j)\phi(j) &&\text{for} \quad i \geqq s, \end{aligned}$$

with equality for all i when $(s, S) = (s_n, S_n)$. Choose now an (s, S) policy from F, By (3), $v_n(i) - v_{n-1}(i)$ is constant for $i \leqq r_n$. Hence $v_n(i) - v_{n-1}(i) \geqq L_n$ for $i \leqq \bar{S}$. and so, by (9),

$$\begin{aligned} v_{n-1}(i) + \mathrm{L}_n &\leqq c\cdot(S - i) + K + L(S) + \textstyle\sum_{j=0}^{\infty} v_{n-1}(S - j)\phi(j) \quad &&\text{for} \quad i < s \\ &\leqq L(i) + \textstyle\sum_{j=0}^{\infty} v_{n-1}(i - j)\phi(j) &&\text{for} \quad s \leqq i \leqq S. \end{aligned}$$

Multiplying both sides of this inequality by $q_i(s, S)$, summing over i, and using the relations (4) and (5), we find $L_n \leqq a(s, S)$. Hence $L_n \leqq g$, since the (s, S) policy was arbitrarily chosen from F and $g = a(s, S)$ for some $(s, S) \in F$. Consider now the (s_n, S_n) policy. Since $v_n(i) - v_{n-1}(i) \leqq U'_n$ for all $i \leqq S_n$ and the equality sign holds in (9) for all i when $s = s_n$ and $S = S_n$, it follows that

$$\begin{aligned} v_{n-1}(i) + U'_n &\geqq c\cdot(S_n - i) + K + L(S_n) + \textstyle\sum_{j=0}^{\infty} v_{n-1}(S_n - j)\phi(j) \quad &&\text{for} \quad i < s_n, \\ &\geqq L(i) + \textstyle\sum_{j=0}^{\infty} v_{n-1}(i - j)\phi(j) &&\text{for} \quad s_n \leqq i \leqq S_n. \end{aligned}$$

Multiplying both sides of this inequality by $q_i(s_n, S_n)$, summing over i, and using the relations (4) and (5), we find $U'_n \geqq a(s_n, S_n)$. This completes the proof of (a).

(b) For any $m \geqq 1$, let $k_m(i) = S_m$ for $i < s_m$, and let $k_m(i) = i$ for $i \geqq s_m$. Then, by (2) and (3), for all $i \leqq \bar{S}$,

$$v_{n+1}(i) - v_n(i) \geqq \sum_{j=0}^{\infty} v_n(k_n(i) - j)\phi(j) - \sum_{j=0}^{\infty} v_{n-1}(k_n(i) - j)\phi(j) \geqq L_n,$$

so $L_{n+1} \geqq L_n$. The proof of $U_{n+1} \leqq U_n$ is very similar and is omitted.

(c) This assertion is an immediate consequence of Theorem 2.

We note that an analogous theorem can be established for the discounted cost criterion by using results from [9] (see also [7]).

References

1. Brown, B. W., "On the Iterative Method of Dynamic Programming on a Finite Space Discrete Time Markov Process," *Ann. Math. Statist.*, Vol. 36 (1965), pp. 1279–1285.
2. Denardo, E. V., "A Markov Decision Problem," pp. 33–68 in T. C. Hu and S. M. Robinson (eds.), *Mathematical Programming*, Academic Press, New York, 1972.
3. Hordijk, A. and Tijms, H. C., "The Asymptotic Behaviour of the Minimal Total Expected Cost in Denumerable State Dynamic Programming and an Application in Inventory Theory," Report BW 17/73, Mathematisch Centrum, Amsterdam (1973).
4. Iglehart, D. L., "Optimality of (s, S) Inventory Policies in the Infinite Horizon Dynamic Inventory Problem," *Management Science*, Vol. 9, No. 2 (1963), pp. 259–267.
5. ——, "Dynamic Programming and Stationary Analysis of Inventory Problems," Chap. 1 in H. Scarf, D. Gilford, and M. Shelly (eds.), *Multistage Inventory Models and Techniques*, Stanford University Press, Stanford, 1963.
6. Johnson, E. L., "On (s, S) Policies," *Management Science*, Vol. 15, No. 1 (1968), pp. 80–101.
7. Kalymon, B. A., "Stochastic Prices in a Single-Item Inventory Purchasing Model," *Operations Research*, Vol. 19 (1971), pp. 1434–1458.
8. Lanery, E., "Étude Asymptotique des Systèmes Markoviens à Commande," *Revue Française d'Informatique et de Recherche Operationnelle*, Vol. 1, No. 5 (1967), pp. 3–56.
9. MacQueen, J. B., "A Modified Dynamic Programming Method for Markovian Decision Problems," *J. Math. Anal. and Appl.*, Vol. 14 (1966), pp. 38–43.
10. Ross, S. M., "Arbitrary State Markovian Decision Processes," *Ann. Math. Statist.* Vol. 39 (1968), pp. 2118–2122.
11. Scarf, H., "The Optimality of (S, s) Policies in the Dynamic Inventory Problem," Chap. 13 in K. J. Arrow, S. Karlin, and P. Suppes (eds.), *Mathematical Methods in the Social Sciences*, Stanford University Press, Stanford, 1960.
12. Schweitzer, P. J., "Perturbation Theory and Markovian Decision Processes," M.I.T. Operations Research Center Technical Report No. 15, (1965).
13. Tijms, H. C., *Analysis of (s, S) Inventory Models*, Mathematical Centre Tract No. 40, Mathematisch Centrum (1972).
14. Veinott, A. F., Jr. and Wagner, H. M., "Computing Optimal (s, S) Policies," *Management Science*, Vol. 11, No. 5 (1965), pp. 525–552.
15. ——, "On the Optimality of (s, S) Inventory Policies: New Conditions and a New Proof," *J. Siam Appl. Math.*, Vol. 14 (1966), pp. 1067–1083.
16. Wagner, H. M. with O'Hagan, M. and Lundh, B., "An Empirical Study of Exactly and Approximately Optimal Inventory Policies," *Management Science*, Vol. 11, No. 7 (1965), pp. 690–723.