A MARKOVIAN DECISION MODEL
FOR A JOINT REPLACEMENT AND STOCKING PROBLEM

BY

CYRUS DERMAN AND GERALD J. LIEBERMAN

TECHNICAL REPORT NO. 93
September 9, 1966

SUPPORTED BY THE ARMY, NAVY, AIR FORCE AND NASA UNDER
CONTRACT Nonr-225(53) (NR-042-002)
WITH THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
A MARKOVIAN DECISION MODEL
FOR A JOINT REPLACEMENT AND STOCKING PROBLEM

by

Cyrus Derman and Gerald J. Lieberman

TECHNICAL REPORT NO. 93
September 9, 1966

Supported by the Army, Navy, Air Force, and NASA under
Contract Nonr-225(53)(MR-042-002)
with the Office of Naval Research

Gerald J. Lieberman, Project Director

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
A MARKOVIAN DECISION MODEL
FOR A JOINT REPLACEMENT AND STOCKING PROBLEM

C. Derman and G. J. Lieberman

1. Introduction and Summary

Two important classes of problems encountered by industry and government are those dealing with replacement policies and ordering policies. These problems are usually treated separately. A particular model which treats the replacement and stocking problems as one can be described as follows:

A company has a bin initially containing $N$ transistors, each of which has a nominal rated capacity. A transistor is taken from the bin and placed into a device. At the beginning of a period (e.g., a period may be a day) the performance level of the transistor is measured, and it continues to operate at this level throughout the period, provided that it does not fail. Associated with the performance level during the period is a known expected cost (or "value"). This expected cost can take into account the probability distribution of failure during the period, some utility associated with the performance level, and the storage costs ascribed to the remaining transistors in the bin. If the transistor fails during the period, it must be replaced by another transistor from the bin prior to the beginning of the next period (say, overnight). If the transistor does not fail during the period, a decision is made at the end of the period to either leave the transistor in the device (if its performance level is considered satisfactory) or to replace it (if
unsatisfactory) with another transistor from the bin. If a replacement is made, the old transistor is scrapped. When a new transistor is placed into service, there exists a probability distribution over its possible performance levels. All transistors are assumed to possess the same performance level probability distribution. If no replacement is made, the transistor remains in service performing at the same level. That is, the only deterioration assumed, here, is that of failure. Decisions as to whether to replace or not are made at the end of every period. When no transistors remain in the bin a restocking must take place; i.e., a purchase of \( N \) transistors is made. It will be assumed that the device will be inoperative for one period during this restocking. At this point a cost (or expected cost) is incurred. This cost may include the cost of ordering \( N \) transistors, the charge for subsequently installing the \( N \) transistors, and the cost of the device being inoperative for one period. The idle period may be interpreted as the time required for routine maintenance of the entire device carried out each time the supply of transistors runs out. The company's problem is to select a policy that specifies the decisions to be made at the end of each period and also determines the number \( N \) of transistors to be stocked. It is assumed that the same number \( N \) will be ordered each time a new batch of transistors is ordered.

This problem can be formulated as a denumerable state Markovian decision model. It will be shown that the optimal decision policy has a particular structure. In obtaining the optimal policy, this structure can be exploited so that an optimal policy can be obtained in a
finite number of operations. For the general denumerable state Markovian decision model, it is not clear that a computational procedure exists for obtaining a solution in a finite number of steps.

Section 2 formulates the general Markovian decision model. Theorem 1 provides conditions under which an optimal policy exists and be of the purely deterministic type. Theorem 1 is only a slightly stronger assertion than that proved in [1]. However, the method of proof is different and suggests a certain interesting physical interpretation.

Section 3 is concerned with the specific model formulated above. For a fixed \( N \), Theorem 1 is employed. Furthermore, the structure of the optimal rule is obtained. These structural results are used to provide a method by which the optimal value of \( N \) can also be obtained.

Section 4 presents a numerical example for which the results of the paper are used to obtain, for the transistor problem, the optimal decision rule and the optimal number of transistors to be ordered. The example provided is somewhat conveniently chosen in order that the solution may be obtained by hand calculations. In general, the method of solution would require the use of a computer.

2. The Markovian Decision Model

The Markovian decision model is concerned with a dynamic system observed periodically at times \( t = 0, 1, \ldots \) and classified into one of a possible number of states. Let \( I \) denote the set of possible states. After an observation, one of a possible number of decisions is made. Let \( K_i \) denote the number of available decisions at state \( i \), \( i \in I \). When the system is in state \( i \), \( i \in I \), and decision \( k \) (\( k=1,2,\ldots,K_i \))
is made, it is assumed that there is a number $w_{ik}$ which denotes the cost or expected cost of being in state $i$ and making decision $k$.

Let $\{Y_t\}$ and $\{\Delta_t\}$, $t = 0, 1, \ldots$ denote the sequence of states and decisions. A fundamental assumption of the Markovian decision model is that there is a set of numbers $\{q_{ij}(k)\}$, $k = 1, 2, \ldots, K_i$; $i, j \in I$, such that

$$P(Y_{t+1} = j \mid Y_0, \Delta_0; Y_1, \Delta_1; \ldots; Y_t = i, \Delta_t = k) = q_{ij}(k)$$

for every "history" $h_t = (Y_0, \Delta_0; Y_1, \Delta_1; \ldots; Y_t = i, \Delta_t = k)$. A rule $R$ for controlling the system is a collection of functions $\{D_k(h_{t-1}, Y_t)\}$, where

$$0 \leq D_k(h_{t-1}, Y_t) \leq 1, \sum_k D_k(h_{t-1}, Y_t) = 1,$$

and $D_k(h_{t-1}, Y_t)$ denotes the probability of making decision $k$, given $h_{t-1}$ and $Y_t$. Thus, for any $Y_0 = i$ and a rule $R$, $\{Y_t\}$ is a stochastic process. In particular, if $D_k(\cdot)$ is of the form $D_k(h_{t-1}, Y_t = i) = D_{ik}$, $k = 1, \ldots, K_i$, $i \in I$, then $\{Y_k\}$ is a Markov process. If $D_{ik} = 0$ or 1 the rule is said to be of the purely deterministic stationary type.

Let

$$W_t = w_{ik} \text{ if } Y_t = i, \Delta_t = k.$$ 

Assume $Y_0 = i$; and set

$$\Psi_R(i) = \lim_{T \to \infty} \frac{\sum_{t=0}^T E(W_t)}{T+1},$$

where if the limit does not exist the expression on the right is replaced by the lim sup. The problem of interest is to find that rule $R$ such that $\Psi_R(i)$ is minimized.
If \( I \) is taken to be denumerable, as will be assumed here, it is possible that no optimal rule will exist. It is also possible that an optimal rule may exist but not be of the purely deterministic stationary type. In [1], Theorem 1 given below was proved. It provides conditions under which an optimal rule will exist and be of the purely deterministic stationary type. The statement given here is slightly stronger (it is not assumed that \( K_i < \infty \)) and the proof is different from that given in [1].

Theorem 1:

If \( \{w_{ik}\} \) is bounded, and if there exists a bounded set of numbers \( \{g, v_i\}, i \in I, \) such that

\[
g + v_i = \min_k \{w_{ik} + \sum_{j \in I} q_{ij}(k)v_j\}, \quad i \in I,
\]

then there exists a purely deterministic stationary rule, \( R_0 \), which is optimal. \( R_0 \) is: make decision \( k_i \) at state \( i, i \in I, \) where \( k_i \) is the decision that minimizes the right side of (1).

Proof:

First, consider the modified problem of minimizing

\[
E \left\{ \sum_{t=0}^{T} \{w_t + v_{T+1}\} \right\}.
\]

In the proof, one can think of the \( v \)'s as salvage costs incurred at time \( T+1 \), the salvage cost being a function of the state at time \( T+1 \). This is related to an approach used in [2]. Using the method of backward induction, one can see that the optimal decisions to make at time \( T \) are, in fact, \( k_i, i \in I, \) by virtue of (1). That is, the expression
within brackets on the right side of (1) is precisely the conditional expected terminal costs associated with decision $k$ and state $i$ at time $T$. Using decisions $k_i$, $i \in I$ at time $T$, the conditional expected remaining costs associated with time $T-1$ and state $i$ are, from (1),

$$w_{ik} + \sum_{j \in I} q_{ij}(k)(g + v_j) = g + (w_{ik} + \sum_{j \in I} q_{ij}v_j), \quad i \in I.$$ 

Thus, the optimal decisions at time $T-1$ are also $k_i$, $i \in I$. Continuing in this manner, the conditional expected remaining costs, given state $i$ at period $T-r$, are simply

$$rg + (w_{ik} + \sum_{j \in I} q_{ij}v_j), \quad i \in I.$$ 

Therefore, the optimal decisions at every period $t = 0, 1, \ldots, T$ are to make decision $k_i$, when in state $i$, $i \in I$. Note that this is rule $R_0$ and it is optimal for every $T = 0, 1, \ldots$. For any arbitrary rule $R$, we have that

$$\frac{1}{T+1} E_R \left\{ \sum_{t=0}^{T} W_t + v_{Y_{T+1}} \right\} \geq \frac{1}{T+1} E_{R_0} \left\{ \sum_{t=0}^{T} W_t + v_{Y_{T+1}} \right\}.$$ 

Letting $T \to \infty$, and since the $v_i$'s are bounded, it follows that

$$\lim_{T \to \infty} \inf_{T} \frac{1}{T+1} \sum_{t=0}^{T} E_R(W_t) \geq v_R(i), \quad i \in I.$$ 

This proves the theorem.

It is readily seen that $g = \Psi_{R_0}(i), \quad i \in I$. 
3. **Specific Model**

In the case considered here the state space \( I \) consists of the states \( 0 \) and \( i = (n,s), \ n = 1, 2, \ldots, N; \ s = 1, 2, \ldots \). The first subscript of \( i \) denotes the number of components on hand (i.e., in storage or in use) and the second denotes the performance level of the component in use. State \( 0 \) denotes the state where no components are on hand.

At each state \( i = (n,s) \), there are only two decisions that can be made, i.e., \( K_i = 2 \). Decision one is to leave the component in service; decision two is to remove the component in service at the end of the period. It is assumed that a component that does not leave service will fail during the period before the next observation with probability \( \alpha > 0 \). If it does not fail, it will continue to operate at the same performance level. If a component fails, it will be replaced at the end of the period. When a new component is placed in service (either by making decision two, or by making decision one and subsequently having a failure) there is a probability \( f_s \) that the new component will perform at level \( s \). It is assumed that \( f_s > 0, \ s = 1, 2, \ldots \) and \( \sum_{s=1}^{\infty} f_s = 1 \). Thus,

\[
\begin{align*}
q^{(1)}_{(1,s),(1,s)} &= 1 - \alpha \\
q^{(1)}_{(1,s),0} &= \alpha \\
q^{(2)}_{(1,s),0} &= 1
\end{align*}
\]

and

7
\[
\begin{aligned}
q^{(1)}_{(n,s),(n,s)} &= 1 - \alpha \\
q^{(1)}_{(n,s),(n-1,s')} &= \alpha f_s' \\
q^{(2)}_{(n,s),(n-1,s')} &= f_s'
\end{aligned}
\]

When in state 0, \( K_0 = 1 \), i.e., a purchase of \( N \) components is made. Thus,

\[
q^{(1)}_{0,(N,s)} = f_s, \quad s = 1, 2, \ldots
\]

It will be assumed that the costs \( w_{(n,s)k} \) are functions only of the state and not of the decision, i.e., \( w_{(n,s)k} = w_{(n,s)} \). Thus, if at the beginning of a period the component is in state \( (n,s) \), \( w_{(n,s)} \) can be interpreted as

\[
w_{(n,s)} = (1 - \alpha)\omega_{(n,s)} + \alpha \int_0^1 v_{(n,s)}(\tau) b_{(n,s)}(\tau) \, d\tau,
\]

where \( \omega_{(n,s)} \) is the cost (or value) of having the component perform for a full period at level \( s \) when there are \( n \) components on hand, \( v_{(n,s)}(\tau) \) is the cost (or value) of having the component perform for a fraction \( \tau \) of a period and then break down when it was performing at level \( s \) when there are \( n \) components on hand, and \( b_{(n,s)}(\tau) \) is the density function for the time at which a failure will occur, given that a failure occurs during the period. For the time being, it will only be assumed that \( \{w_{(n,s)}\} \) are bounded with \( \{w_{(n,s)}\} \) non-decreasing in \( s \). Later, more will be assumed. Let \( w_0^N \) denote the cost associated with state 0.
From [1] it follows that within the foregoing assumed structure an optimal rule for minimizing expected average cost per unit time exists, and is of the purely deterministic stationary type. Moreover, there exist bounded numbers \( \{g, v_i\}, \ i \in I \), where \( u_0 \) can be taken to be zero, satisfying (1) of Theorem 1. In the optimal policy let, for \( i = (n, s) \),

\[
\beta_i = \begin{cases} 
\alpha & \text{if decision 1 is made}, \\
1 & \text{if decision 2 is made}.
\end{cases}
\]

Then (1) takes the form

\[
(2a) \quad g = w_0^N + \sum_{s=1}^\infty f_s v(n,s)
\]

\[
(2b) \quad g + v(l,s) = w(l,s) + (1 - \beta(l,s))v(l,s), \quad s = 1, \ldots
\]

\[
(2c) \quad g + v(n,s) = v(n,s) + (1 - \beta(n,s))v(n,s) + \beta(n,s) \sum_{s'=1}^\infty f_{s'} v(n-1,s')
\]

\[
= v(n,s) + v(n,s) - \beta(n,s)v(n,s) - \sum_{s'=1}^\infty f_{s'} v(n-1,s')
\]

\[
\quad n = 2, 3, \ldots, N; \quad s = 1, 2, \ldots
\]

From (2b), \( v(l,s) \) can be solved explicitly yielding

\[
(3) \quad v(l,s) = \frac{w(l,s) - g}{\beta(l,s)}, \quad s = 1, 2, \ldots
\]

From (2c), it is seen that

\[
(4) \quad v(n,s) = \frac{w(n,s) - g}{\beta(n,s)} + \sum_{s'=1}^\infty f_{s'} v(n-1,s')
\]

\[
\quad \text{for } n = 2, 3, \ldots, N; \quad s = 1, 2, \ldots
\]
The following theorem indicates the structure of the optimal rule.

Theorem 2:

Under the stated assumptions an optimal purely deterministic stationary rule $R_0$ is such that there are numbers $s_1, s_2, \ldots, s_n$ where

$$
\beta(n,s) = \begin{cases} 
\alpha & \text{if } s < s_n \\
1 & \text{if } s \geq s_n 
\end{cases}
$$

for $n = 1, 2, \ldots, N$.

Proof:

Consider first $n = 1$. Suppose $\beta(1,s) = 1$. Then from (2b) $v(1,s) \geq 0$, for otherwise the right side of (2b) would not be minimal and (1) would not hold. However, from (3) and since $\{w_{1,s}\}$ is non-decreasing in $s$, $v(1,s+1) \geq 0$ and hence $\beta(1,s+1) = 1$. From this the assertion of the theorem follows for $n = 1$. Now consider $n = 2, \ldots, N$, and suppose $\beta(n,s) = 1$. From (2c) and (4), a similar argument holds which completes the proof of the theorem.

Instead of using a policy iteration procedure as given in [1] for obtaining an optimal rule, Theorem 2 suggests that the procedure should also take into account the structure of the rule.

Theorem 2 was obtained under the assumption that $N$ is fixed. At state 0 only one decision, namely, to order $N$ components, can be made. A more interesting question is that of also deciding on the optimal value of $N$ if one exists. To consider this question, the following assumptions are made:
(i) \( w(n,s) = w(1,s) + h(n), \ s = 1, 2, \ldots; n = 1, 2, \ldots, \)

where \( h(n) \) is an increasing function of \( n \) such that \( \lim_{n \to \infty} h(n) = \infty. \)

(ii) For each \( m, m' = 1, 2, \ldots \) there exists finite number \( \bar{N}_{m,m'} \)
such that

\[
\frac{w_0^N + \sum_{n=m}^{N} h(n)}{N + m'} = \infty \text{ for all } N > \bar{N}_{m,m'}.
\]

Now let \( g^N, s_1^N, \ldots, s_n^N \) be as defined in Theorem 2 with the superscript
indicating the dependence on \( N \). That is, \( g^N \) is the optimal expected
average cost per unit time for \( N \) fixed and \( s_n^N \) is such that
\( \beta(n,s) = \alpha \) or \( 1 \) according to whether \( s < s_n^N \) or \( s \geq s_n^N, n = 1, \ldots, N. \)

Lemma 1.

If (i) holds, then

\[
s_n^N = \min\{s | h(n) + w(1,s) - g_n^N \geq 0\}, \ n = 2, 3, \ldots, N.
\]

Consequently, \( s_1^N \geq s_2^N \geq \cdots \geq s_n^N. \)

Proof:

The fact that \( s_n^N \) is as asserted for \( n = 2, 3, \ldots, N \) follows from the proof of Theorem 2.

Lemma 2:

If \( g_1^N \leq g_2^N \), then \( s_1^N \leq s_2^N, \ n = 1, 2, \ldots, \min(N_1,N_2) \)

Proof:

Follows from Lemma 1.
Lemma 3:

If (i) and (ii) hold, then there exist an \( N^* < \infty \) such that \( g^N_{\min} < g^N \).

Proof:

Assume that the minimum is not attained. Then there exists an \( N \) arbitrarily large such that \( g^N < g^1 \). Let \( \bar{m}^N \) denote the smallest value of \( n \) such that \( h(n) + w(1,1)^N g^N \geq 0 \). From Lemma 1, and (i), \( \bar{m}^N \leq \bar{m}^1 \). \( N \) can be taken to be larger than \( \bar{m}^1 \). However, from the fact that \( g^N \) can be expressed as the ratio of the expected total cost between returns to state 0 to the expected length of a return to state 0, it is easy to see that

\[
g^N \geq \frac{w^N_0 + \sum_{n=\bar{m}^1}^N h(n)}{\bar{m}^N + \bar{m}^1 / \alpha}.
\]

However, from (ii), for \( N \) sufficiently large, \( g^N > g^1 \). This is a contradiction proving the lemma.

Lemma 4:

If (i) and (ii) hold, then \( N^* \leq M \), where \( M \) is the smallest integer greater than \( \max(N \bar{m}^N, \bar{m}^1 / \alpha) \) such that

\[
w^M_0 + \sum_{n=\bar{m}^1}^M h(n) > g^1.
\]

Proof:

Since \( g^N_{\min} \leq g^1 \), then \( \bar{m}^{N*} \leq \bar{m}^1 \). If \( N^* > M \), it follows from the argument of Lemma 3 that \( g^N_{\min} > g^1 \). Since this is a contradiction \( N^* \leq M \).

Lemmas 1 to 4 when summarized yield the following theorem.

Theorem 3:

If (i) and (ii) hold, then there exists a value of \( N = N^* \) such that

\[
g^N_{\min} = \min_N g^N, \quad s^N_{\min} \leq s^1_{\min} \quad \text{and} \quad N^* \leq M \quad \text{where} \quad M \quad \text{is defined in Lemma 4}.
\]
4. Outline of Computations

Theorem 1 asserts that a rule of the purely deterministic type will be optimal. Theorem 2 indicates the simple structure of an optimal rule.

For a given \( N \) the optimal rule can be obtained by the policy improvement procedure described in [1]. For the special problem treated here the policy improvement procedure has a simple form. Let \( R \) be any initial rule of the purely deterministic type described by the numbers \( s_1, \ldots, s_N \). Given \( R, g^N_R \), the expected average cost, can be computed. Let \( s_n \) be defined as the smallest \( s \) such that \( w_{n,s} > g^N_R \) for \( n = 1, \ldots, N \). Let rule \( R' \) be defined by the numbers \( s_1', \ldots, s_N' \). As a result of this iteration \( g^N_R \) \( \leq g^N_{R'} \). The iteration proceeds until \( s_n = s_n' \) \( (n = 1, \ldots, N) \). At this point \( g^N_R \) is obtained. For the problem under consideration only a finite number of iterations is required to obtain the optimal rule \( R \).

To calculate \( g^N_R \) for a given \( N \) and \( R \), one can use the well-known relationship

\[
(5) \quad \frac{g^N_R}{E(\text{Cost of a Cycle})} = \frac{E(\text{Length of a Cycle})}{W_0 + \sum_{n=1}^{N} \left( \frac{1}{\alpha} \sum_{s=1}^{s_n-1} r_s W(n,s) + \sum_{s=s_n}^{s_n} r_s W(n,s) \right)}
\]

where a cycle is defined as the behavior of the process between returns to the state 0. Using this formula, one gets

\[
g^N_R = \frac{W_0 + \sum_{n=1}^{N} \left( \frac{1}{\alpha} \sum_{s=1}^{s_n-1} r_s W(n,s) + \sum_{s=s_n}^{s_n} r_s W(n,s) \right)}{1 + \sum_{n=1}^{N} \left( \frac{1}{\alpha} \sum_{s=1}^{s_n-1} r_s + \sum_{s=s_n}^{s_n} r_s \right)}
\]
The policy improvement procedure is first carried out for $N=1$. At this point $M(\text{of Lemma } h)$ can be obtained. Successively, $g^2, \ldots, g^M$ can be obtained using the policy improvement procedure for each $N=2, \ldots, M$. The smallest of the numbers $g^1, \ldots, g^M$ indicate the optimal value of $N$.

Example

An aerospace company uses a silicon n-p-n U.H.F. transistor in one of its major assemblies. A critical measurement for this transistor is its emitter to base capacitance. This characteristic ranges from 1.2 (good) to 1.4 (bad). Suppose

$$w(n,s) = w(1,s) + h(n)$$

where $h(n) = 4(n-1)$, i.e., the holding cost is 4 units per transistor, and

$$w(1,s) = 100(1.4 - \frac{s}{2s-1}).$$

Also suppose

$$f_s = \left(\frac{1}{2}\right)^s \quad s = 1, 2, 3, \ldots$$

$$\alpha = .1$$

and $w_0^N = 20N + 140$ where the 140 represents the cost due to idle time (note that this is greater than $w(1,s)$ for all $s$ so idle time is not desireable). For $N = 1$ a rule $R^1_{s_1}$, which uses $s_1$ as its critical value has by (5) an associated $g^1_{R^1_{s_1}}$ of
\[
G_{S_1}^1 = \left[ 20 + 140 + 1000\left( 1.4 - 1.4\left( \frac{1}{2} \right)^{S_1-1} \right) \right] - 200 \frac{1}{2} \frac{4}{3} \left[ 1 - \left( \frac{1}{4} \right)^{S_1+1} \right] \\
+ 140 \frac{1}{2} \left( \frac{1}{2} \right)^{S_1-1} - 20 \frac{1}{2} \frac{4}{3} \left( \frac{1}{4} \right)^{S_1-1} \\
\left[ 1 + 10 - 10 \left( \frac{1}{2} \right)^{S_1-1} + \left( \frac{1}{2} \right)^{S_1-1} \right] \\
= \frac{1426.7 - 1260\frac{1}{2}^{S_1-1} + 120\left( \frac{1}{4} \right)^{S_1-1}}{1 - 9\left( \frac{1}{2} \right)^{S_1-1}}.
\]

If we let \( S_1 = 1 \), we have
\[
G_{S_1}^1 = 143.2.
\]

Now, there is no \( S \) such that \( w_{(1,S)} > 143.2 \), and so the policy improvement technique tells us to next let \( S_1 = \infty \). Now, when \( S_1 = \infty \),
\[
G_{S_1}^1 = 129.8.
\]

The smallest value of \( S \) such that \( w_{(1,S)} = 140 - \frac{20}{2^{S_1-1}} \) is greater than 129.8 is \( S = 2 \). When
\[
S_1 = 2, \quad G_{S_1}^1 = 127.2
\]

and the smallest \( S \) such that \( w_{(1,S)} \) is greater than 127.2 is again \( S = 2 \). Therefore, by the policy improvement technique, we have that for \( N = 1 \) the optimal choice for \( S_1 \) is \( S_1 = 2 \) and \( G_{S_1}^1 = 127.2 \).

We could determine \( M (M = 10) \) and use the policy improvement method to determine \( G^2, \ldots, G^M \) and then choose the smallest. Instead we shall make use of Lemmas 1 and 2. We first note that \( h(n) + w_{(1,1)} > G^1 \) when \( n \geq 3 \) (since \( w_{(1,1)} + h(n) = 120 + 4(n-1) \) and \( G^1 = 127.2 \)).
For any $N^*$ with $g_{N^*}^{N^*} \leq g_1^1$ we have by Lemma 1 that $s_{2}^{N^*} = s_{4}^{N^*} = \ldots = s_{N^*}^{N^*} = 1$; and by Lemma 2 that $s_{1}^{N^*} \leq s_{1}^{1} - 2$. It also follows from Lemma 1 that (for any $N$) $s_{1}^{N} \geq s_{2}^{N} \geq \ldots \geq s_{N}^{N}$ (since $h(n)$ is increasing). Therefore either $N = 1$ is optimal or there exists an $N^*$ with

$$s_{1}^{N^*} = s_{2}^{N^*} = \ldots = s_{N^*}^{N^*} = 1,$$

or one with

$$s_{1}^{N^*} = 2, s_{2}^{N^*} = s_{3}^{N^*} = \ldots = s_{N^*}^{N^*} = 1,$$

or one with

$$s_{1}^{N^*} = s_{2}^{N^*} = 2, s_{3}^{N^*} = \ldots = s_{N^*}^{N^*} = 1$$

which is optimal.

For a rule $R$ with $s_{i}^{N} = 1, i = 1, \ldots, N$, we have that

$$g_{R}^{N} = \frac{20N + 140 + N(140 - 20 \frac{1}{2} \frac{h}{3} + 4 \frac{N(N-1)}{2}}{1 + N}$$

$$= \frac{2N^2 + 144.7N + 140}{N + 1}$$

which can be seen to be an increasing function of $N$, so its minimum is at $N = 1$ and equals 143.2.

Now for a rule $R$ with $s_{1}^{N} = 2, s_{2}^{N} = s_{3}^{N} = \ldots = s_{N}^{N} = 1$, we have that

$$g_{R}^{N} = [20N + 140 + 1000(0.7) - 200 \frac{1}{2} \frac{h}{3} + 70 - 20 \frac{1}{2} \frac{h}{3} \frac{1}{4}$$

$$+ (140 - 20 \frac{1}{2} \frac{h}{3})(N-1) + 4 \frac{N(N-1)}{2(N-1)}/[1 + 5 + \frac{1}{2} + N-1]$$

$$= \frac{2N^2 + 144.7N + 680}{N + 1\frac{1}{2}}$$

which can also be seen to be an increasing function of $N$. Its minimum is at $N = 1$ and equals 127.2.
Finally for a rule $R$ with $s_1^N = s_2^N = 2$, $s_3^N = \ldots = s_N^N = 1$ ($N \geq 2$) we have that

$$s_R^N = (20N + 140 + 2[700 - 200 \frac{1}{2} \frac{1}{3} \frac{3}{4} + 70 - 20 \frac{1}{2} \frac{1}{3} \frac{1}{4}]) + 4$$

$$+ (140 - 20 \frac{1}{2} \frac{4}{3})(N - 2) + 4(\frac{N}{2(N - 1)} - 1))/[1 + (5 + \frac{1}{2})2 + N - 2]$$

$$= \frac{2N^2 + 144.7N + 1220}{n + 10}.$$ 

This is also an increasing function of $N$ and its minimum is at $N = 2$ (since $N$ was $\geq 2$) and equals 126.5.

Since 126.5 is smaller than both of the other two minimums and is also less than $g^1$, it follows that it is optimal; i.e., $N = 2$ is optimal and $s_1 = s_2 = 2$ is the optimal rule with $\min \varphi_R(0) = g^2 = 126.5$. 

17
ACKNOWLEDGMENT

The authors would like to thank Sheldon Ross for carrying out the computations of the example, and for some helpful comments.

REFERENCES


A Markovian Decision Model for a Joint Replacement and Stocking Problem

Technical Report

Cyrus Derman
G. J. Lieberman

September 15, 1966

18
2

Nonr-225(53)
NR-042-002

TR No. 14 NSF GP-3739

Distribution of this document is unlimited

Logistics and Mathematical Statistics Br.
Office of Naval Research
Washington, D.C. 20360

Two important classes of problems encountered by industry and government are those dealing with replacement policies and ordering policies. These problems are usually treated separately. A particular model which treats the replacement and stocking problems as one is considered in this report. It is formulated as a denumerable state Markovian decision model. It is shown that the optimal decision policy has a particular structure. In attaining the optimal policy, this structure is exploited so that an optimal policy can be obtained in a finite number of steps.
stocking policy
replacement policy
ordering policy
Markovian decision model
denumerable state
policy improvement procedure

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through"

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through"

(5) "All distribution of this report is controlled. Qualified DDC users shall request through"

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (T3), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.