A Note on the Optimal Replacement Problem

MICHAEL N. KATEHAKIS
Rutgers University
Department of MSIS
180 University Ave., Newark NJ 07102
USA

KARTIKEYA S. PURANAM
Rutgers University
Department of MSIS
180 University Ave., Newark NJ 07102
USA

Abstract: We study the following model for a system the state of which is continuously observed. The set of possible states is a finite set \( \{0, \ldots, L\} \), where larger values represent increased states of deterioration from the “new condition” represented by state \( 0 \), to the “totally inoperative condition” of state \( L \). Whenever the system changes state a decision has to be made as to whether it is renovated or it is left unattended. Whenever the system enters a state \( i < L \) and the decision to renovate is taken, then a cost \( c \) is incurred and its state, immediately, changes to a fixed state \( l \). If the system enters state \( L \) then it must be renovated at an increased cost \( c + A \). There is no cost whenever the decision to leave it unattended is taken in a state \( i < L \); in this case the next state will be state \( j \) with probability \( p_{ij} \) and the sojourn time in state \( i \) is a random variable with distribution function \( F_{ij}(\cdot) \). We provide necessary conditions under which optimal policies are of the “control limit” type. The results herein generalize those of Derman [1] when the state sojourn times are distributed according to a general state dependent distribution and a renovation may not result in a “new system” i.e., the renovation state \( l \) maybe \( l > 0 \).

Key–Words: Markovian Decision Processes, Reliability, Nonlinear Optimization.

1 Introduction

In this note we study the following model for a system the state of which is continuously observed. The set of possible states is a finite set \( \mathcal{S} = \{0, \ldots, L\} \), where larger values represent increased states of deterioration from the “new condition” (state 0) to the “totally inoperative condition” (state \( L \)). When the system changes state a decision has to be made as to whether it is renovated or it is left unattended. Whenever the system enters a state \( i < L \) and the decision to renovate is taken, then a cost \( c(>0) \) is incurred and its state, immediately, changes to a fixed state \( l \). If the system enters state \( L \) then it must be renovated at an increased cost \( c + A \). There is no cost if the decision to leave it unattended is taken when it enters a state \( i < L \). In this case the next state will be state \( j \) with probability \( p_{ij} \) and the sojourn time in state \( i \) is a random variable with distribution function \( F_{ij}(\cdot) \). We provide a necessary condition under which optimal policies are of the “control limit” type. The results herein generalize those of Derman [1] when the state sojourn times are distributed according to a general state dependent distribution and a renovation may not necessarily result in a “new system” i.e., the renovation state \( l \) maybe \( l > 0 \).

2 Semi Markov formulation and analysis.

Let \( a = 0 \) denote the action of no renovation and \( a = 1 \) denote the action to perform a renovation. The transition probabilities \( p_{ij} \) are such that \( p_{ij} = 0 \) if \( i > j \) and \( p_{il} = 1 \) if \( i = L \) and \( j = l \). We consider three cost criteria. The long run average cost, which for a policy \( \pi \) is represented by \( v^\pi(i) \) where \( i \) is the initial state. The infinite horizon discounted cost \( V^\pi_0(i) \) and the finite horizon discounted cost \( V^\pi_N(i) \). Note that \( V^\pi_0(i,0) = 0, i = 0, \ldots, L \).

Define \( g(i,j,\alpha) \) as follows

\[
g(i,j,\alpha) = \int_0^\infty e^{-\alpha t} dF_{ij}(t). \tag{1}
\]

Following the notation in Ross [3] the finite horizon discounted cost function can be written as

\[
V_\alpha(i,N) = \begin{cases} V^\alpha_0(i,N), & \text{if } i \neq L \\ A + V^\alpha_N(l,N), & \text{if } i = L \end{cases}
\tag{2}
\]
where
\[ V_\alpha^0(i, N) = \sum_{j=0}^{L} p_{ij} V_\alpha^\pi(j, N-1) g(i, j, \alpha), \]
and
\[ V_\alpha^1(l, N) = c + \sum_{j=0}^{L} p_{lj} V_\alpha^\pi(j, N-1) g(l, j, \alpha). \]

To prove the optimality of a control limit policy, we impose the following conditions.

**Condition I:** For any increasing function \( h \) on \( S = \{0, 1, \ldots, L\} \), the function
\[ \xi(i) = \sum_{j=k}^{L} p_{ij} h(j) \]
is an increasing function of \( j \).

Condition I is *Condition A* in Derman [1], who established that it is equivalent to the following: **Condition II:**

For each \( k = 0, 1, \ldots, L \), the function
\[ r_k(i) = \sum_{j=k}^{L} p_{ij} \quad i = 0, 1, \ldots, L - 1. \]  
(3)
is increasing in \( i \).

**Condition III:** For each \( i = 0, 1, \ldots, L \), the function \( g(i, j, \alpha) \) is a non-decreasing function of \( j \).

We next state and prove the main result of the paper.

**Theorem I** If conditions I and III hold or equivalently if conditions II and III hold then the optimal policy for all three cost criteria considered is of the control limit type.

**Proof:** \( V_\alpha(i, 0) = 0 \) is a increasing function by definition. Using conditions II and III it follows that \( V_\alpha^0(i, 1) \) is increasing in \( i \). Using equation (2) it follows that there exists an \( i_1^\ast = i_1^\ast(\alpha) \leq L - 1 \) such that an optimal policy: \( \pi_{i_1^\ast, \alpha}^\pi \) for \( V_\alpha(i, 1) \) is specified by the actions: \( \pi_{i_1^\ast, \alpha}^\pi(i) = 0 \), for \( i < i_1^\ast \) and \( \pi_{i_1^\ast, \alpha}^\pi(i) = 1 \) for \( i > i_1^\ast \). Also, from equation (2) and conditions II and III it follows that \( V_\alpha(i, 1) \) is an increasing function in \( i \).

Using the same arguments and induction it follows that \( V_\alpha(i, N) \) is also increasing in \( i \) and that an optimal policy: \( \pi_{i_N^\ast, \alpha}^\pi \) for \( V_\alpha(i, N) \) is specified by the actions: \( \pi_{i_N^\ast, \alpha}^\pi(i) = 0 \), for \( i < i_N^\ast \) and \( \pi_{i_N^\ast, \alpha}^\pi(i) = 1 \) for \( i > i_N^\ast \), for some constant \( i_N^\ast = i_N^\ast(\alpha) \leq L - 1 \).

For the infinite horizon case note that the following hold, c.f. Ross [3].
\[ V_\alpha(i) = \lim_{N \to \infty} V_\alpha(i, N), \]  
(4)
and
\[ V_\alpha(i) = \begin{cases} V_\alpha^0(i), & \text{if } i < L, \\ V_\alpha^1(L), & \text{if } i = L. \end{cases} \]  
(5)

where
\[ V_\alpha^0(i) = \sum_{j=0}^{L} p_{ij} V_\alpha^\pi(j, i), \]
\[ V_\alpha^1(i) = \begin{cases} c + \sum_{j=0}^{L} p_{ij} V_\alpha^\pi(j, i), & \text{if } i < L, \\ c + a + \sum_{j=0}^{L} p_{ij} V_\alpha^\pi(j, i), & \text{if } i = L. \end{cases} \]

From equation (4) and the fact that \( V_\alpha(i, N) \) are increasing in \( i \), it follows that \( V_\alpha(i) \) is increasing in \( i \).

Using this, equation (5), and conditions II and III we can conclude that there is a number \( i^\ast = i_0^\ast \leq L - 1 \) such that
\[ V_\alpha(i) = \begin{cases} V_\alpha^0(i), & \text{if } i < i^\ast, \\ V_\alpha^1(i), & \text{if } i \geq i^\ast, \end{cases} \]  
(6)
i.e., the infinite horizon optimal policy \( \pi^\ast_\alpha \) is specified by the actions \( \pi^\ast_\alpha(i) = 0 \), (do not renovate) for \( i < i^\ast \) and \( \pi^\ast_\alpha(i) = 1 \) (renovate) for \( i > i^\ast \).

For the average cost case let \( v(i) = \sup_{\alpha} \nu^\pi(i) \); note that the following holds, c.f. Ross [3].
\[ v(i) = \lim_{\alpha \to \infty} \alpha \ast V_\alpha(i). \]  
(7)

Consider a sequence \( \{\alpha_\nu\} \) with \( \lim_{\nu \to \infty} \alpha_\nu = \infty \), with \( i_\nu^\ast = i^\ast \) for all \( \nu \). Since the total number of possible states is finite, such a sequence and \( i^\ast \) exist. Let \( \pi^\ast \) be the control limit policy defined by \( i^\ast \) and let \( \pi \) is some policy that is not a control limit. Then
\[ V_{\alpha_\nu}^\pi(i) \geq V_{\alpha_\nu}^\ast(i) = V_\alpha(i) \quad \nu = 1, 2, \cdots \]  
(8)
and hence
\[ V^\pi = \lim_{\nu \to \infty} \alpha_\nu V_{\alpha_\nu}^\pi(i) \geq \lim_{\nu \to \infty} \alpha_\nu V_{\alpha_\nu}(i) = V^\pi. \]  
(9)

Hence the theorem.

**Observations**

1. If \( g(.) \) does not depend on \( j \), then for **Condition I** and **Condition II** to hold and for control limit policies to be optimal, it is required that \( g(.) \geq 0 \). This is always true since \( g(.) \) is the Laplace transform of a probability distribution function.

2. If the distribution function \( f(.) \), is not a function of \( t \), then the function \( g(i, j, \alpha) \) can be simplified to \( f(i, j)/\alpha \). Then **Condition II** will only require \( f(i, j) \) to be an increasing function in \( j \).
3 Examples

3.1 Example 1

Let the transition times between states \( i \) and \( j \) be exponentially distributed with parameter \( \mu_{ij} \), i.e., \( f(i, j) = \mu_{ij} e^{-\mu_{ij} t} \) and \( g(i, j, \alpha) = \frac{\mu_{ij}}{\mu_{ij} + \alpha} \) with, \( \mu_{ij} \leq \mu_{i+1,j} \) and \( \mu_{ij} \leq \mu_{i,j+1} \). Assume the transition probabilities to be \( p_{ij} = \frac{\mu_{ij}}{\mu_i} \) where, after uniformization (cf. [2]), we may take \( \mu_i = \sum_j \mu_{ij} = 1 \). Then conditions II and III are satisfied and by theorem 1 the system has an optimal policy that is of the control limit type.

![Figure 1: Total cost versus values of \( i^* \)](image)

3.2 Example 2

Let the transition times between states \( i \) and \( j \) have a poisson distribution with parameter \( \lambda_{ij} \), i.e., \( f(i, j) = (\lambda_{ij} t)^n e^{-\lambda_{ij} t}/n! \) and \( g(i, j, \alpha) = e^{\lambda_{ij}(e^{-\alpha} - 1)} \) with \( \lambda_{ij} \leq \lambda_{ij+1} \) and \( p_{ij} \) satisfying Condition II. Then conditions II and III are satisfied and by theorem 1 the system has an optimal policy of the control limit type.

3.3 Example 3

Let the transition times between states \( i \) and \( j \) be distributed uniformly over \([b_i, a_j]\) i.e., \( f(i, j) = 1/(a_j - b_i) \) and using remark (2) \( g(i, j, \alpha) = 1/\alpha(a_j - b_i) \) where the parameter \( a_j \) depends only on the final state \( j \) and the parameter \( b_i \) depends only on the initial state \( i \). Consider the transition probabilities \( p_{ij} \) that satisfy Condition II. Then by theorem 1 we can say that the system has an optimal policy that is of the control limit type.

References:

[1] C. Derman, “On Optimal Replacement Rules when Changes of State are Markovian”. In

