ROLLING HORIZON PROCEDURES FOR THE SOLUTION OF AN OPTIMAL REPLACEMENT PROBLEM OF $\mathfrak{n}$-MACHINES WITH RANDOM HORIZON

Rocio Ilhuicatzi–Roldán and Hugo Cruz-Suárez
Facultad de Ciencias Físico-Matemáticas
Benemérita Universidad Autónoma de Puebla, México

ABSTRACT
In this paper a system consisting of independently operating $\mathfrak{n}$-machines is considered, where each machine follows a stochastic process of deterioration, which is associated with a cost function. It is assumed that the system is observed at discrete time and the objective function is the total expected cost. Also, it is considered that the horizon of the problem is random with an infinite support. In this case, the optimal replacement problem with a random horizon is modelled as a non-homogeneous optimal control problem with an infinite horizon and solved by means of the rolling horizon procedure. Then, a replacement policy is provided, which approximate the optimal solution. Finally, a numerical example through a program in Maple is presented.
KEYWORDS: Optimal Stochastic Control; Dynamic Programming; Markov Decision Process; Optimal Replacement Problem.
MSC: 90C39.

RESUMEN
En este artículo se considera un sistema formado por $\mathfrak{n}$-máquinas que operan independientemente, en donde cada máquina sigue un proceso estocástico de deterioro, el cual es asociado con una función de costo. Se supone que el sistema es observado en tiempo discreto y que la función objetivo es el costo total esperado. También, se considera que el horizonte del problema es aleatorio con soporte infinito. En este caso, el problema de reemplazo óptimo con horizonte aleatorio es modelado como un problema de control no homogéneo con horizonte infinito y resuelto por medio del procedimiento de horizonte rodante. Entonces se determina una política de reemplazo, la cual aproxima a la solución óptima. Finalmente, se presenta un ejemplo numérico mediante un programa elaborado en Maple.

1. INTRODUCTION

In industrial processes the deterioration of electronic components or machines is common. Then it is important to provide replacement strategies for the optimization of these systems. The problem of optimal replacement is modeled in different ways. For example, in Sheti *et al.* (2000)[6] is studied the problem considering a single machine, which follows a deterioration stochastic process with various quality levels in continuous time. Also, the optimal replacement has been implemented based on future technological advances, in this case, a non-stationary process is considered and the optimal decision is characterized using

forecast horizon approach (see Nair and Hopp (1992)[5]). On the other hand, in Childress and Durango-Cohen (2005)[3] the problem is studied with $\mathfrak{n}$-machines considering two heuristics rules of replacement that make possible the search of optimal policies. These rules are as follow: the first suggests that a machine is replaced only if all older machines are replaced and the second one indicates that in any stage machines of the same age are either all kept or all replaced.
This work was motivated by the model proposed in Bertsekas (1987) [2], which involves a single machine that follows a Markov process of deterioration with D possible levels, where D is a positive integer. The process of deterioration is associated with an operation cost of the machine at each level. The machine is observed in discrete time and, depending on the deterioration level, the following situations are possible: leaving operate an additional period of time or, replaced it with a cost $R > 0$. Then, an optimal replacement policy, which minimizes the total cost, is provided. In this model it is considered a finite horizon of operation.
Now, in this paper a system consisting of $\mathfrak{n}$-machines with independent deterioration processes is studied, assuming that the system is operating over a random horizon. This novel consideration in the model is due to it is possible that external situations obligate to conclude the process before expected, for example, bankruptcy of the firm in an economic model (see Puterman (1994)[8], p. 125). The support of the distribution of the random horizon can be finite or infinite. In this paper the case with an infinite support is considered. The optimal replacement problem with a random horizon is modeled through Markov decision processes as a nonhomogeneous optimal control problem with an infinite horizon. Then, the optimal solution is approximated by means of the rolling horizon procedure and a theoretical result about Markov Decision Process with a random horizon is presented and applied to verify that the method is efficient.
This paper is organized as follows. Firstly, in Section 2 the basic theory of Markov decision processes and the rolling horizon procedure are presented. Afterwards, the problem of optimal replacement in a system with $\mathfrak{n}$-machines and random horizon is described and then, an algorithm for the solution is proposed in Section 3. For solving numerical cases a program in Maple is elaborated. Finally, in Section 4, some numerical results are illustrated.

2. BASIC THEORY

Markov Decision Processes

Let $(X, A, \{A(x) : x \in X\}, Q, c)$ be a Markov decision or control model, which consists of the state space X, the action set A (X and A are Borel spaces), a family $\{A(x) : x \in X\}$ of nonempty measurable subsets $A(x)$ of A, whose elements are the feasible actions when the system is in state $x \in X$. The set $\mathbb{K} := \{(x, a) : x \in X, \ \ a \in A(x)\}$ of the feasible state-action pairs is assumed to be a measurable subset of $X \times A$. The following component is the transition law Q, which is a stochastic kernel on X given $\mathbb{K}$. Finally, $c : \mathbb{K} \to \mathbb{R}$ is a measurable function called the cost per stage function.
A policy is a sequence $\pi = \{\pi_t : t = 0, 1, \ldots\}$ of stochastic kernels π_t on the control set A given the history $\mathbb{H}_t$ of the process up to time t ($\mathbb{H}_t = \mathbb{K} \times \mathbb{H}_{t-1}$, $t = 1, 2, \ldots$, $\mathbb{H}_0 = X$). The set of all policies is denoted by Π.
$\mathbb{F}$ denotes the set of measurable functions $f : X \to A$ such that $f(x) \in A(x)$, for all $x \in X$. A deterministic Markov policy is a sequence $\pi = \{f_t\}$ such that $f_t \in \mathbb{F}$, for $t = 0, 1, 2, \ldots$.
Let $(\Omega, \mathcal{F})$ be the measurable space consisting of the canonical sample space $\Omega = \mathbb{H}_\infty := (X \times A)^\infty$ and $\mathcal{F}$ as the corresponding product σ-algebra. The elements of Ω are sequences of the form $\omega =$

$(x_0, a_0, x_1, a_1, \ldots)$ with $x_t \in X$ and $a_t \in A$ for all $t = 0, 1, 2, \ldots$. The projections x_t and a_t from Ω to the sets X and A are called state and action variables, respectively.
Let $\pi = \{\pi_t\}$ be an arbitrary policy and μ be an arbitrary probability measure on X called the initial distribution. Then, by the theorem of C. Ionescu-Tulcea (see Hernández-Lerma and Lasserre (1996) [4]), there is a unique probability measure P_μ^π on $(\Omega, \mathcal{F})$ which is supported on $\mathbb{H}_\infty$, i.e., $P_\mu^\pi(\mathbb{H}_\infty) = 1$. The stochastic process $(\Omega, \mathcal{F}, P_\mu^\pi, \{x_t\})$ is called a discrete-time Markov control process or a Markov decision process.
The expectation operator with respect to P_μ^π is denoted by E_μ^π. If μ is concentrated at the initial state $x \in X$, then P_μ^π and E_μ^π are written as P_x^π and E_x^π, respectively.

Markov Decision Processes with Random Horizon

Let $(\Omega', \mathcal{F}', P)$ be a probability space and let $(X, A, \{A(x) \mid x \in X\}, Q, c)$ be a Markov decision model with a planning horizon τ, where τ is considered as a random variable on $(\Omega', \mathcal{F}')$ with the probability distribution $P(\tau = t) = \rho_t$, $t = 0, 1, 2, ..., T$, where T is a positive integer or $T = \infty$. Define the performance criterion as

$$\jmath^\tau(\pi, x) := E\left[\sum_{t=0}^{\tau} c(x_t, a_t)\right],$$

$\pi \in \Pi$, $x \in X$, where E denotes the expected value with respect to the joint distribution of the process $\{(x_t, a_t)\}$ and τ. Then, the optimal value function is defined as

$$J^\tau(x) := \inf_{\pi \in \Pi} \jmath^\tau(\pi, x), \tag{2.1}$$

$x \in X$. The optimal control problem with a random horizon is to find a policy $\pi^* \in \Pi$ such that $\jmath^\tau(\pi^*, x) = J^\tau(x)$, $x \in X$, in which case, π^* is said to be optimal.

Assumption 1. *For each $x \in X$ and $\pi \in \Pi$ the induced process $\{(x_t, a_t) \mid t = 0, 1, 2, \ldots\}$ is independent of τ.*

Remark 1. *Observe that, under Assumption 1,*

$$\begin{aligned} E\left[\sum_{t=0}^{\tau} c(x_t, a_t)\right] &= E\left[E\Big[\sum_{t=0}^{\tau} c(x_t, a_t) \mid \tau\Big]\right] \\ &= \sum_{n=0}^{T} E_x^\pi\left[\sum_{t=0}^{n} c(x_t, a_t)\right]\rho_n \\ &= \sum_{t=0}^{T}\sum_{n=t}^{T} E_x^\pi\left[c(x_t, a_t)\right]\rho_n \\ &= E_x^\pi\left[\sum_{t=0}^{T} P_t\, c(x_t, a_t)\right], \end{aligned}$$

$\pi \in \Pi$, $x \in X$, where $P_k = \sum_{n=k}^{T} \rho_n = P(\tau \geq k)$, $k = 0, 1, 2, \ldots, T$. Thus, the optimal control problem with a random horizon τ is equivalent to a nonhomogeneous optimal control problem with a horizon $T + 1$ and a null terminal cost.

In the case $T < +\infty$ the problem is solved using the dynamic programming (see Theorem 1).

Assumption 2. *(a) The one-stage cost c is lower semicontinuous, nonnegative and inf-compact on $\mathbb{K}$ (for definitions, see Bertsekas (1987) [2], p. 146).*

(b) Q is either strongly continuous or weakly continuous (see Hernández-Lerma and Lasserre (1996) [4], p. 28).

Theorem 1. *Let $J_0, J_1, ..., J_{T+1}$ be the functions on X defined by*

$$J_{T+1}(x) := 0$$

and for $t = T, T-1, ..., 0,$

$$J_t(x) := \min_{a \in A(x)} \left[P_t c(x,a) + \int_X J_{t+1}(y) Q(dy \mid x,a) \right], \qquad x \in X. \tag{2.2}$$

Under Assumption 2, these functions are measurable and for each $t = 0, 1, ..., T$, there is $f_t \in \mathbb{F}$ such that $f_t(x) \in A(x)$ attains the minimum in (2.2) for all $x \in X$; i.e.

$$J_t(x) = P_t c(x, f_t(x)) + \int_X J_{t+1}(y) Q(dy \mid x, f_t(x)),$$

$x \in X$ and $t = 0, 1, ..., T$. Then, the deterministic Markov policy $\pi^* = \{f_0, ..., f_T\}$ is optimal and the optimal value function is given by

$$J^\tau(x) = \jmath^\tau(\pi^*, x) = J_0(x),$$

$x \in X$.

The proof of previous theorem is similar to the proof of Theorem 3.2.1 in Hernández-Lerma and Lasserre (1996) [4].
Let

$$U_t = \frac{J_t}{P_t},$$

$t \in \{0, 1, 2, \ldots, T\}$. Then, (2.2) is equivalent to

$$U_t(x) := \min_{a \in A(x)} \left[c(x,a) + \alpha_t \int_X U_{t+1}(y) Q(dy \mid x,a) \right], \tag{2.3}$$

where

$$\alpha_t = \frac{P_{t+1}}{P_t}, \qquad t \in \{0, 1, 2, \ldots, T\}. \tag{2.4}$$

Now consider that $T = +\infty$. In this case,

$$\jmath^\tau(\pi, x) = E_x^\pi \left[\sum_{t=0}^{\infty} P_t \, c(x_t, a_t) \right], \tag{2.5}$$

$\pi \in \Pi$ and $x \in X$.

For each $n = 0, 1, 2, \ldots$, defines

$$v_n^\tau(\pi, x) := E_x^\pi \left[\sum_{t=n}^{\infty} \prod_{k=n}^{t} \alpha_{k-1} c(x_t, a_t) \right], \tag{2.6}$$

$\pi \in \Pi$, $x \in X$ and

$$V_n^\tau(x) := \inf_{\pi \in \Pi} v_n^\tau(\pi, x), \tag{2.7}$$

$x \in X$. $v_n^\tau(\pi, x)$ is the expected total cost from time n onwards, applied to (2.5), where the initial condition $x_n = x$ is given and x is a generic element of X.

Remark 2.

i) Observe that $P_t = \prod_{k=0}^{t} \alpha_{k-1}$, $t = 0, 1, 2, \ldots$, where $\alpha_{-1} = P_0 = 1$ is considered and α_k, $k = 0, 1, 2, \ldots$, is defined by (2.4).

ii) Observe that $V_0^\tau(x) = J^\tau(x)$, $x \in X$ (see (2.1)).

For $N > n \geq 0$, it is defined that

$$v_{n,N}^\tau(\pi, x) := E_x^\pi \left[\sum_{t=n}^{N} \prod_{k=n}^{t} \alpha_{k-1} c(x_t, a_t) \right], \tag{2.8}$$

with $\pi \in \Pi$, $x \in X$, and

$$V_{n,N}^\tau(x) := \inf_{\pi \in \Pi} v_{n,N}^\tau(\pi, x), \tag{2.9}$$

$x \in X$.

Assumption 3. *(a) Same as Assumption 2.*

(b) There exists a policy $\pi \in \Pi$ such that $J^\tau(\pi, x) < \infty$ for each $x \in X$.

Definition 1. $M(X)^+$ denotes the cone of nonnegative measurable functions on X, and, for every $u \in M(X)^+$, $T_n u$, $n = 0, 1, 2, \ldots$, is the operator on X defined as

$$T_n u(x) = \min_{a \in A(x)} \left[c(x, a) + \alpha_n \int_X u(y) Q(dy \mid x, a) \right],$$

$x \in X$.

The proofs of Lemmas 1 and 2 below are similar to the proofs of Lemmas 4.2.4 and 4.2.6 in Hernández-Lerma and Lasserre (1996) [4], respectively. This is why these proofs are omitted.

Lemma 1. *For every $N > n \geq 0$, let w_n and $w_{n,N}$ be functions on $\mathbb{K}$, which are nonnegative, lower semicontinuous and inf-compact on $\mathbb{K}$. If $w_{n,N} \uparrow w_n$ as $N \to \infty$, then*

$$\lim_{N \to \infty} \min_{a \in A(x)} w_{n,N}(x, a) = \min_{a \in A(x)} w_n(x, a),$$

$x \in X$.

Lemma 2. *Suppose that Assumption 2(a) holds. For every $u \in M(X)^+$, $Tu \in M(X)^+$ and, moreover, there exists $f_n \in \mathbb{F}$ such that*

$$T_n u(x) = c(x, f_n(x)) + \alpha_n \int_X u(y) Q(dy \mid x, f_n(x)),$$

$x \in X$.

Lemma 3. *Suppose that Assumption 3(a) holds and let $\{u_n\}$ be a sequence in $M(X)^+$. If $u_n \geq T_n u_{n+1}$, $n = 0, 1, 2, \ldots$, then $u_n \geq V_n^\tau$, $n = 0, 1, 2, \ldots$.*

Proof. Let $\{u_n\}$ be a sequence in $M(X)^+$ such that $u_n \geq T_n u_{n+1}$, and then by Lemma 2, there exists $f_n \in \mathbb{F}$, where

$$u_n(x) \geq c(x, f_n(x)) + \alpha_n \int_X u_{n+1}(y) Q(dy \mid x, f_n(x)),$$

$x \in X$. Iterating this inequality, it is obtained that

$$\begin{aligned} u_n(x) \quad &\geq \quad E_x^\pi \left[c(x_n, f_n(x_n)) + \sum_{t=n+1}^{N-1} \prod_{j=n+1}^{t} \alpha_{j-1} c(x_t, f_t(x_t)) \right] \\ &\quad + \prod_{j=n+1}^{N} \alpha_{j-1} E_x^\pi \big[u(x_N) \big], \end{aligned} \tag{2.10}$$

$x \in X$. Here

$$E_x^\pi \big[u(x_N) \big] = \int_X u(y) Q^N(dy \mid x_n, f_n(x_n)),$$

where $Q^N(\cdot \mid x_n, f_n(x_n))$ denotes the N-step transition kernel of the Markov process $\{x_t\}$ when the policy $\pi = \{f_k\}$ is used, beginning at stage n. Since u is nonnegative and $\alpha_k \leq 1$, recalling that $x_n = x$, it is obtained from (2.10) that

$$u_n(x) \geq E_x^\pi \left[\alpha_{n-1} c(x_n, f_n(x_n)) + \sum_{t=n+1}^{N-1} \prod_{j=n}^{t} \alpha_{j-1} c(x_t, f_t(x_t)) \right].$$

Hence, letting $N \to \infty$, it yields that

$$u_n(x) \geq v_n^\tau(\pi, x) \geq V_n^\tau(x),$$

$x \in X$. □

Theorem 2. *Suppose that Assumption 3 holds. Then, for every $n \geq 0$ and $x \in X$,*

$$V_{n,N}^\tau(x) \uparrow V_n^\tau(x) \quad \text{as} \quad N \to \infty.$$

Proof. Using the dynamic programming equation given in (2.3), that is

$$U_t(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha_t \int_X U_{t+1}(y) Q(dy \mid x,a) \right], \tag{2.11}$$

for $t = N-1, N-2, \ldots, n$, with $U_N(x) = 0$, $x \in X$, it is obtained that $V^{\tau}_{n,N}(x) = U_n(x)$ and $V^{\tau}_{s,N}(x) = U_s(x)$, $n \leq s < N$. Furthermore, it is proved by backwards induction that U_s, $n \leq s < N$, is lower semicontinuous. For $t = n$, (2.11) is written as

$$V^{\tau}_{n,N}(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha_n \int_X V^{\tau}_{n+1,N}(y) Q(dy \mid x,a) \right], \tag{2.12}$$

and $V^{\tau}_{n,N}(\cdot)$ is lower semicontinuous. Then, by the nonnegativity of c, for each $n = 0, 1, 2, \ldots$, the sequence $\{V_{n,N} : N = n, n+1, \ldots\}$ is nondecreasing, which implies that there exists a function $u_n \in M(X)^+$ such that for each $x \in X$,

$$V^{\tau}_{n,N}(x) \uparrow u_n(x),$$

as $N \to \infty$. Moreover,

$$V^{\tau}_{n,N}(x) \leq v^{\tau}_{n,N}(\pi, x) \leq v^{\tau}_{n}(\pi, x),$$

$x \in X$ and $\pi \in \Pi$, hence

$$V^{\tau}_{n,N}(x) \leq V^{\tau}_{n}(x),$$

$N > n$. Then $u_n \leq V^{\tau}_n$. Using Lemma 1 and letting $N \to \infty$ in (2.12), it is obtained that

$$u_n(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha_n \int_X u_{n+1}(y) Q(dy \mid x,a) \right], \tag{2.13}$$

$n = 0, 1, 2, \ldots$ and $x \in X$. Finally, by Lemma 3, $u_n \geq V^{\tau}_n$, obtaining that $u_n = V^{\tau}_n$ and concluding this way the proof of Theorem 2. □

Remark 3. *Theorem 2 can be applied to obtain $J^{\tau} = V^{\tau}_0$, as the limit of the sequence $\{V_{0,N}\}^{\tau}$. Nevertheless in the practice is very difficult to obtain this limit. But, Theorem 2 can be used to approximate J^{τ}.*

The Rolling Horizon Procedure

The rolling horizon procedure is the most common method employed in practice for generating solutions to nonhomogeneous optimal control problems when the horizon is infinite (see Wang and Liu (2011)[9]). The procedure fixes a horizon N, solves the corresponding N–period problem, implements the first optimal decision found, rolls forward one period and repeats from the new current state. Below, the rolling horizon algorithm is presented.

Algorithm

1. Set $m = 0$ and $n = N$.

2. Find the policy $\pi^* = (\pi^*_m, \pi^*_{m+1}, \ldots, \pi^*_{n-1})$, which is optimal for periods from m to n, and set $\widehat{\pi}_m = \pi^*_m$.

3. Let $m = m + 1$ and $n = n + 1$.

4. Go to step 2.

The policy $\widehat{\pi} = (\widehat{\pi}_0, \widehat{\pi}_1, \widehat{\pi}_2, \ldots)$ is called a **rolling horizon policy**.

Remark 4. *The rolling horizon policy in some cases may not be optimal, an example is presented in Alden and Smith (1992)[1]. Also, in Alden and Smith (1992)[1] an error bounded theoretical is provided for the rolling horizon procedure applied to nonhomogeneous Markov decision processes with an infinite horizon.*

3. THE REPLACEMENT PROBLEM

Consider a system consisting of $\mathfrak{n}-$machines, each other with an independent stochastic process of deterioration, whose possible levels of deterioration are denoted by $1, 2, 3, \ldots, D$, where D is a positive integer. Level one denotes that the machine is in perfect condition. Suppose that deterioration level is increasing, i.e., that a machine operating at level i is better than at level $i + 1$, $i = 1, 2, 3, \ldots, D - 1$.
Let $P = (p_{i,j})_{D\times D}$ be the matrix of transition probabilities for going from level i to level j (identical for the $\mathfrak{n}$-machines). Because any machine cannot move to better level of deterioration, $p_{i,j} = 0$ if $j < i$. Let $g : \{1, 2, 3, \ldots, D\} \to \mathbb{R}$ be a known function, which will measure the cost of operation of a machine. Suppose that g is nondecreasing, i.e.

$$g(1) \leq g(2) \leq \ldots \leq g(D),$$

and that at the beginning of each period of time can be considered the following options.

a) Operate the machine k, $k = 1, 2, \ldots, \mathfrak{n}$ in a level of deterioration for this time period, or

b) replace by a new one with a fixed cost $R > 0$.

Also, consider that the system can operate over τ time periods, where τ is a random variable, which is independent of the process followed by the system with probability distribution $P(\tau = t) = \rho_t$, $t = 0, 1, 2, \ldots$, such that $E[\tau] < \infty$.

The problem consists on determining optimal replacement policies that minimize the total cost of operation of the system.

The problem is solved through theory of Markov decision processes. This requires to identify the corresponding Markov control model.

At the beginning of an arbitrary time period, the state of the system can be registered as $(d_1, d_2, \ldots, d_{\mathfrak{n}})$ where d_k, $k = 1, 2, \ldots, \mathfrak{n}$, is the level of deterioration in which the machine is operating, therefore the state space is defined by:

$$X = \{(d_1, d_2, \ldots, d_{\mathfrak{n}}) : d_k \in \{1, 2, \ldots, D\}, k = 1, 2, \ldots, \mathfrak{n}\},$$

$x \in X$, where $card(X) = D^{\mathfrak{n}}$ states.
A replacement action can be represented by $(a_1, a_2, \ldots, a_{\mathfrak{n}})$ with $a_k = 0$ or $a_k = 1$, where $a_k = 0$ means letting that the machine k operate on the level d_k and $a_k = 1$ means replacing it. At this way

$$A = A(x) = \{(a_1, a_2, \ldots, a_{\mathfrak{n}}) : a_k \in \{0, 1\}, k = 1, 2, \ldots, \mathfrak{n}\},$$

where its cardinality is $2^{\mathfrak{n}}$ actions.
For an arbitrary machine k, let $P^0 = (p^0_{i,j})_{DXD} = P$ be the transition matrix of the process of deterioration when the machine is not replaced. Let $P^1 = (p^1_{i,j})_{D\times D}$ be the transition matrix when the machine is replaced, where $p^1_{i,j} = 1$, if $j = 1$ and $p^1_{i,j} = 0$ in otherwise (safely when machine k was replaced, the machine goes to level one). Let $Q^a = (q^a_{i,j})_{D^{\mathfrak{n}}\times D^{\mathfrak{n}}}$ be the transition matrix of the state i at state j of the system, when the action $a \in A$ is taken, $i, j \in X$. For the independence of the deterioration processes of the machines, it is obtained that

$$q^{(a_1,a_2,\ldots,a_{\mathfrak{n}})}_{(i_1,i_2,\ldots,i_{\mathfrak{n}}),(j_1,j_2,\ldots,j_{\mathfrak{n}})} = p^{a_1}_{i_1,j_1} \cdot p^{a_2}_{i_2,j_2} \cdot \ldots \cdot p^{a_{\mathfrak{n}}}_{i_{\mathfrak{n}},j_{\mathfrak{n}}}.$$

Also at any time period t

$$c(x_t, a_t) = \sum_{k=1}^{\mathfrak{n}} \gamma(x_{k,t}, a_{k,t}),$$

where $\gamma(x_{k,t}, a_{k,t}) = g(x_{k,t})$, if $a_{k,t} = 0$ and $\gamma(x_{k,t}, a_{k,t}) = g(1) + R$, if $a_{k,t} = 1$. In this case $x_{k,t}$ and $a_{k,t}$ represent the state and the action at time t of the machine k, respectively.

Lemma 4. *The replacement model of* $\mathfrak{n}$*–machines with random horizon satisfies Assumption 3.*

Proof. Since X and A are finite, then Assumption 3(a) holds. Moreover, there exist M such that $c(x, a) \leq M$ for each $x \in X$ and $a \in A(x)$. Hence, $\jmath^{\tau}(\pi, x) \leq M \sum_{t=0}^{\infty} P_t = M\,(1 + E[\tau])$. But, it is supposed that $E[\tau] < \infty$, then $\jmath^{\tau}(\pi, x) < \infty$ for each $\pi \in \Pi$ and $x \in X$, i.e. Assumption 3(b) holds. $\square$

Algorithm to obtain the rolling horizon policies

1. Do $n = 0$.
2. Do $t = N + 1$ and $U_t(x) = 0$ for each $x \in X$.
3. If $t = 0$, do $\widehat{f}_n = f_0$, $n = n + 1$ and return to step 2. Else, go to step 4.
4. Replace t by $t - 1$ and calculate $U_t(x)$ for each $x \in X$ by means of the equation

$$U_t(x) = \min_{a\in A(x)} \left[c(x, a) + \alpha_{t+n} \sum_{y\in X} U_{t+1}(y) q^a_{x,y} \right], \qquad ,$$

   (see 2.3), putting

$$f_t(x) = a,$$

   for some

$$a \in \arg\min_{a\in A(x)} \left[c(x, a) + \alpha_{t+n} \sum_{y\in X} U_{t+1}(y) q^a_{x,y} \right].$$

   Return to step 3.

$\widehat{f} = (\widehat{f}_0, \widehat{f}_1, \widehat{f}_2, \ldots)$ is the rolling horizon policy obtained with N fixed, where N is an integer.

4. NUMERICAL EXAMPLE

Consider the optimal replacement problem with random horizon and the following numerical values: $\mathfrak{n} = 3$, $D = 3$,

$$P = \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0 & 0.3 & 0.7 \\ 0 & 0 & 1 \end{pmatrix},$$

$g(1) = 5$, $g(2) = 7$ and $g(3) = 29$, $R = 4$, $\rho_k = P(\tau = k) = -\frac{(1-p)^{k+1}}{(k+1)\ln p}$, $k = 0, 1, 2, \ldots$, where $0 < p < 1$ with $p = .8$ (Logarithmic Distribution) and $N = 20$.

In Table 1, the rolling horizon action depending on the stage and the state of the system is presented.

Stage State	1	2	3	4	5	6	7	...
1 1 1	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
1 1 2	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
1 1 3	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	
1 2 1	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
1 2 2	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
1 2 3	0 0 1	0 0 1	0 0 1	0 0 1	0 1 0	0 0 1	0 0 1	
1 3 1	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	
1 3 2	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	
1 3 3	0 1 1	0 1 1	0 1 1	0 1 1	0 1 1	0 1 1	0 1 1	
2 1 1	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
2 1 2	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
2 1 3	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	
2 2 1	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
2 2 2	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	
2 2 3	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	
2 3 1	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	
2 3 2	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	0 1 0	
2 3 3	0 1 1	0 1 1	0 1 1	0 1 1	0 1 1	0 1 1	0 1 1	
3 1 1	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	
3 1 2	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	
3 1 3	1 0 1	1 0 1	1 0 1	1 0 1	1 0 1	1 0 1	1 0 1	
3 2 1	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	
⋮	⋮	⋮	⋮	⋮	⋮	⋮		

Table 1: Rolling horizon policies

In this numerical case, it is observed that the policy is stationary (independent of time) and it is possible determine a replacement optimal level for each machine. In this case, each machine is replaced if the machine is in level 3. The value function is obtained evaluating the rolling horizon policy in the performance criterion.

Remark 5. *By Theorem 2*

$$V^{\tau}_{0,N}(x) \uparrow J^{\tau}(x) \quad as \quad N \to \infty.$$

Then, for N^* long enough

$$V^{\tau}_{0,N^*}(x) \approx J^{\tau}(x).$$

Observe that if $|\jmath^{\tau}(\hat{\pi}, x) - V^{\tau}_{0,N^*}(x)| < \epsilon$, $x \in X$, then optimal solution is approximated by means of the rolling horizon policy efficiently.

In Table 2, the column number one shows the evaluation only of the first five actions of the rolling horizon due to the curse of the dimensionality (see Powell (2007)[7]), and in the second column it is presented an approximation of the optimal value function using Theorem 2 with $N^* = 30$ obtained from a $\epsilon = 10^{-19}$.

Initial State x	$\jmath^{\tau}(\hat{\pi}, x)$	$V^{\tau}_{0,N^*}(x)$
1 1 1	17.44943	17.45081
1 1 2	19.60531	19.60669
1 1 3	21.26398	21.26537
1 2 1	19.60531	19.60669
1 2 2	21.76119	21.76257
1 2 3	23.41986	23.42125
1 3 1	21.26398	21.26537
1 3 2	23.41986	23.42125
1 3 3	25.07854	25.07992
2 1 1	19.60531	19.60669
$\vdots$	$\vdots$	$\vdots$

Table 2: Approximation of the optimal value function

Remark 6. *The program was developed in Maple, for the goodness that comes in the manipulation of highly accurate values. For the calculation of the α_t, $t = 0, 1, 2, \ldots$, was necessary to work with up to 400 digits. Nevertheless, the program gets the rolling horizon policy over acceptable time (about 30 seconds for the example). The main problem is in the evaluation of the rolling horizon policy, where, arrays of size $D^{\mathrm{n}(k-1)}$ are required for saving the information in the stage k, making impossible the evaluation taking k large.*

RECEIVED MAY, 2012
REVISED DECEMBER, 2012

REFERENCES

[1] ALDEN, J. M. and SMITH, R. L. (1992): Rolling horizon procedures in nonhomogeneous markov decision process **Operations Research**, 40:183–194.

[2] BERTSEKAS, D. (1987): **Programming and Optimal Control** Athena Scientific, Massacchusetts.

[3] CHILDRESS, S. and DURANGO-COHEN, P. (2005): On parallel machine replacement problems with general replacement functions and stochastic deterioration **Naval Research Logistics**, 52:409–419.

[4] HERNÁNDEZ-LERMA, O. and LASSERRE, J. B. (1996): **Discrete-Time Markov Control Processes: Basic Optimality Criteria** Springer-Verlag, New York.

[5] NAIR, S. K. and HOPP, W. J. A model for equipment replacement due to technological obsolescence **European Journal of Operational Research**, 63:207–221.

[6] P., S. S., SORGER, G., and ZHOU, X. Y. (2000): Stability of real-time lot-scheduling and machine replacement policies with quality levels **IEEE Transactions on Automatic Control**, 45:1–17.

[7] POWELL, W. B. (2007): **Approximate Dynamic Programming: Solving the Curses of Dimensionality** John Wiley & Sons, New Jersey.

[8] PUTERMAN, M. (1994): **Markov Decision Process: Discrete Stochastic Dynamic Programming** John Wiley & Sons, New York.

[9] WANG, B. and LIU, D. (2011): Rolling horizon procedure for large-scale equipment replacement problems with multiple alternatives Control and Decision Conference, Chinese.