KYBERNETIKA — VOLUME *45* (2009), NUMBER 5, PAGES 737 – 754

APPROXIMATION, ESTIMATION AND CONTROL OF STOCHASTIC SYSTEMS UNDER A RANDOMIZED DISCOUNTED COST CRITERION

Juan González-Hernández, Raquiel R. López-Martínez and J. Adolfo Minjárez-Sosa

The paper deals with a class of discrete-time stochastic control processes under a discounted optimality criterion with random discount rate, and possibly unbounded costs. The state process $\{x_t\}$ and the discount process $\{\alpha_t\}$ evolve according to the coupled difference equations $x_{t+1} = F(x_t, \alpha_t, a_t, \xi_t)$, $\alpha_{t+1} = G(\alpha_t, \eta_t)$ where the state and discount disturbance processes $\{\xi_t\}$ and $\{\eta_t\}$ are sequences of i.i.d. random variables with densities ρ^ξ and ρ^η respectively. The main objective is to introduce approximation algorithms of the optimal cost function that lead up to construction of optimal or nearly optimal policies in the cases when the densities ρ^ξ and ρ^η are either known or unknown. In the latter case, we combine suitable estimation methods with control procedures to construct an asymptotically discounted optimal policy.

Keywords: discounted cost, random rate, stochastic systems, approximation algorithms, density estimation

AMS Subject Classification: 93E10, 93E20, 90C40

1. INTRODUCTION

We are concerned with a class of discrete-time stochastic control processes under a discounted optimality criterion with random discount rate, and possibly unbounded costs. The state and discount processes evolve according to the coupled difference equations:

$$x_{t+1} = F(x_t, \alpha_t, a_t, \xi_t), \tag{1}$$

$$\alpha_{t+1} = G(\alpha_t, \eta_t), \tag{2}$$

for $t = 0, 1, \dots,$ where F and G are known continuous functions, x_t, α_t, and a_t represent the state, the discount rate, and the control at time t, respectively. Moreover, the state and discount disturbance processes $\{\xi_t\}$ and $\{\eta_t\}$ are independent sequences, each one consisting of independent and identically distributed (i.i.d.) random variables with density functions ρ^ξ and ρ^η, respectively.

Our work is a sequel to [6] in which, applying a standard dynamic programming approach, it has been shown the existence of discounted optimal policies for the system (1) – (2). However, it is worth noting that [6] does not present computational procedures neither of such policies nor of the value function.

The main objective in this paper is to introduce three approximation algorithms for the control problem's value function V^* that lead up to the construction of optimal or nearly optimal policies. In the first one, V^* is approximated by means of bounded-costs discounted programs. That is, we consider a sequence of bounded costs $\{c_n\}$ converging to the original cost c. Then we show that the sequence of corresponding value functions U_n converges to V^*. Next, we combine this procedure with a value iteration scheme yielding a second approximation algorithm of V^*. The third one, is the well-known policy iteration method, which is an approximation procedure on the control space.

Finally, we consider the case when the state and discount disturbance processes $\{\xi_t\}$ and $\{\eta_t\}$ are observable with unknown densities ρ^ξ and ρ^η. In this context, our approach consists in combining suitable density estimation methods of the unknown join density $\rho := \rho^\xi \rho^\eta$ with a particular algorithm to approximate V^* by means of bounded costs. Next, this procedure is used to construct an asymptotically discounted optimal policy $\bar{\pi} = \{\bar{f}_t\}$ (where $\bar{f}_t$ depends of the estimators ρ_t of ρ), and, moreover, to show the convergence, in the sense of Schäl [15], of the sequence of minimizers $\{\bar{f}_t\}$ to an optimal stationary policy f_∞.

The motivation to study systems of the form (1) – (2) comes from its applications to economic and financial models. Indeed, usually, when a discounted optimality criterion is applied, the discount factor is assumed to be fixed or constant in the course of the process. However (see e. g., [1, 4, 7, 12, 13, 14, 17]), since the discount factor is typically a function of interest rates, which in turn are random variables, this assumption might be too strong or unrealistic. In these cases, we have a time-varying random discount factor that can be represented as in (2).

The remainder of the paper is organized as follows. In Section 2 we introduce the control model we will deal with. Next, Section 3 contains the assumptions on the control model and some preliminary results on the randomized discounted criterion, which are used in Sections 4 and 5 to obtain the approximation algorithms to the value functions. Finally, in Section 6, we present an estimation and control procedure to construct adaptive polices.

Notation. Given a Borel space X (that is, a Borel subset of a complete and separable metric space) its Borel sigma-algebra is denoted by $\mathcal{B}(X)$, and "measurable", for either sets or functions, means "Borel measurable". We denote by $\mathbb{B}(X)$ the space of real-valued bounded measurable functions on X with the supremum norm $\|u\| := \sup_{x\in X} |u(x)|$, and $L(X)$ stands for the family of lower semi-continuous (l.s.c.) functions on X bounded below. Let X and Y be Borel spaces. Then a stochastic kernel $Q(dx \mid y)$ on X given Y is a function such that $Q(\cdot \mid y)$ is a probability measure on X for each fixed $y \in Y$, and $Q(B \mid \cdot)$ is a measurable function on Y for each fixed $B \in \mathcal{B}(X)$.

2. MARKOV CONTROL MODEL

Control model. The Markov control processes associated to the system (1)–(2) is specified by the elements

$$\mathcal{M} := \left(X, \Gamma, A, \mathbb{R}^{k_1}, \mathbb{R}^{k_2}, F, G, \rho^{\xi}, \rho^{\eta}, c\right) \tag{3}$$

satisfying the following conditions. The state space X and the action space A are Borel spaces. The set $\Gamma := [\alpha^*, \infty)$, $\alpha^* > 0$, is the discount rate space. For each pair $(x, \alpha) \in X \times \Gamma$, $A(x, \alpha)$ is a nonempty Borel subset of A denoting the set of admissible controls when the system is in state x and a discount rate α is imposed. The set

$$\mathbb{K} = \{(x, \alpha, a) : x \in X, \alpha \in \Gamma, a \in A(x, \alpha)\} \tag{4}$$

of admissible state-discount-action triplets is assumed to be a Borel subset of the Cartesian product of X, Γ and A. The dynamics of the system is defined by the coupled difference equations (1)–(2), where $F : X \times \Gamma \times A \times \mathbb{R}^{k_1} \to X$ and $G : \Gamma \times \mathbb{R}^{k_2} \to \Gamma$ are continuous functions, and $\{\xi_t\}$ and $\{\eta_t\}$ are independent sequences of i.i.d. random variables in $\mathbb{R}^{k_1}$ and $\mathbb{R}^{k_2}$, respectively. Finally, the cost-per-stage $c(x, \alpha, a)$ is a measurable real-valued function on $\mathbb{K}$, possibly unbounded.

Interpretation. The control model $\mathcal{M}$ has the following interpretation. At stage t, the system is in the state $x_t = x \in X$ and the discount factor $\alpha_t = \alpha \in \Gamma$ is imposed. The controller chooses a control $a_t = a \in A(x, \alpha)$. As a consequence of this the following happens: 1) a cost $c(x, \alpha, a)$ is incurred, and 2) the system moves to a new state $x_{t+1} = x'$ and a new discount factor $\alpha_{t+1} = \alpha'$ is imposed according to the transition laws:

$$\begin{aligned} P_1(B|x, \alpha, a) &:= \int_{\mathbb{R}^{k_1}} 1_B\left(F(x, \alpha, a, s)\right) \rho^{\xi}(s)\, \mathrm{d}s, \qquad B \in \mathcal{B}(X), \\ P_2(D|\alpha) &:= \int_{\mathbb{R}^{k_2}} 1_D\left(G(\alpha, s)\right) \rho^{\eta}(s)\, \mathrm{d}s, \qquad D \in \mathcal{B}(\Gamma). \end{aligned}$$

Once the transition to state x' occurs, the process is repeated.

Control policies. We define the space of admissible histories up to time t by $\mathbb{H}_0 := X \times \Gamma$ and $\mathbb{H}_t := (\mathbb{K} \times \mathbb{R}^{k_1} \times \mathbb{R}^{k_2})^t \times X \times \Gamma$, $t \geq 1$. A generic element of $\mathbb{H}_t$ is written as $h_t = (x_0, \alpha_0, a_0, \xi_0, \eta_0, \ldots, x_{t-1}, \alpha_{t-1}, a_{t-1}, \xi_{t-1}, \eta_{t-1}, x_t, \alpha_t)$. A control policy is a sequence $\pi = \{\pi_t\}$ of measurable functions $\pi_t : \mathbb{H}_t \to A$ such that $\pi_t(h_t) \in A(x_t, \alpha_t)$ for all $h_t \in \mathbb{H}_t$, $t \geq 0$. We denote by Π the set of all control policies.

Let $\mathbb{F}$ be the family of measurable functions $f : X \times \Gamma \to A$ such that $f(x, \alpha) \in A(x, \alpha)$ for every $(x, \alpha) \in X \times \Gamma$. A sequence $\{f_t\}$ of functions $f_t \in \mathbb{F}$ is called a Markov policy. A Markov policy $\{f_t\}$ is said to be stationary if $f_t = f$ for all $t \geq 0$ and some $f \in \mathbb{F}$. In this case we use the notation

$$c(x, \alpha, f) := c(x, \alpha, f(x, \alpha)) \quad \text{and} \quad F(x, \alpha, f, s) := F(x, \alpha, f(x, \alpha), s)$$

for all $x \in X$, $\alpha \in \Gamma$ and $s \in \mathbb{R}^{k_1}$.

Performance index. We assume that the costs are exponentially discounted with accumulative random discounted rates. That is, a cost C incurred at stage t is equivalent to a cost $C\exp(-S_t)$ at time 0, where $S_t = \sum_{i=0}^{t-1} \alpha_i$ if $t \geq 1$, $S_0 = 0$. In this sense, when using a policy $\pi \in \Pi$, given the initial state $x_0 = x$ and the initial discount factor $\alpha_0 = \alpha$, we define the total expected discounted cost (with random discount rates) as

$$V(\pi, x, \alpha) := E^{\pi}_{(x,\alpha)} \left[\sum_{t=0}^{\infty} \exp(-S_t) c(x_t, \alpha_t, a_t) \right], \tag{5}$$

where $E^{\pi}_{(x,\alpha)}$ denotes the expectation operator with respect to the probability measure $P^{\pi}_{(x,\alpha)}$ induced by the policy π, given $x_0 = x$ and $\alpha_0 = \alpha$. (see, e. g., [3] for the construction of $P^{\pi}_{(x,\alpha)}$)

The optimal control problem associated to the control model $\mathcal{M}$ is then to find an optimal policy $\pi^* \in \Pi$ such that $V(\pi^*, x, \alpha) = V^*(x, \alpha)$ for all $(x, \alpha) \in X \times \Gamma$, where

$$V^*(x, \alpha) := \inf_{\pi \in \Pi} V(\pi, x, \alpha) \tag{6}$$

is the optimal value function.

Remark 2.1. From (2), observe that $\{\exp(-S_t)\}$ is a sequence of random variables (not independent) representing the rate of discount at each stage t. Moreover, if $\alpha_t = \alpha$ for all $t \geq 0$ and some $\alpha \in (0, \infty)$, the performance index (5) reduces to the usual $\beta-$discounted cost criterion with $\beta = \exp(-\alpha)$.

3. ASSUMPTIONS AND PRELIMINARY RESULTS

Observe that we can write the system (1)−(2) as

$$y_{t+1} = H(y_t, a_t, \chi_t), \quad t = 0, 1, \dots, \tag{7}$$

where, letting $Y := X \times \Gamma$, $\mathbb{R}^k := \mathbb{R}^{k_1} \times \mathbb{R}^{k_2}$, $y_t^T := (x_t, \alpha_t)$, and $\chi_t := (\xi_t, \eta_t)$, $H : Y \times A \times \mathbb{R}^k \to Y$ is a continuous function defined as

$$H(y_t, a_t, \chi_t) := (F(x_t, \alpha_t, a_t, \xi_t), G(\alpha_t, \eta_t))^T,$$

and $\{\chi_t\}$ is a sequence of i.i.d. $\mathbb{R}^k-$ valued random variables on a probability space $(\Omega, \mathcal{F}, P)$, with common density function $\rho(\cdot) = \rho^{\xi}(\cdot)\rho^{\eta}(\cdot)$.

In what follows, the probability space $(\Omega, \mathcal{F}, P)$ is fixed.

Now, for notational convenience, we put the control model $\mathcal{M}$ in the form

$$\left(Y, A, \{A(y) \subset A | y \in Y\}, \mathbb{R}^k, H, \rho, c\right).$$

Thus, the transition law determined by the system equation (7) is the stochastic kernel on Y given $\mathbb{K} = \{(y,a) : y \in Y, a \in A(y)\}$ (see (4)) defined as

$$Q(B|y,a) := \int_{\mathbb{R}^k} 1_B\left(H(y,a,s)\right)\rho(s)\,\mathrm{d}s, \quad B \in \mathcal{B}(Y). \tag{8}$$

To guarantee the existence of "measurable minimizers", we will suppose the following continuity and compactness conditions.

Assumption 3.1. a) For each $y \in Y$, the set $A(y)$ is compact. Moreover, the multifunction $y \to A(y)$ is upper semi-continuous.
b) The one-stage cost c is nonnegative and belongs to $L(\mathbb{K})$.
c) The function $(y,a) \to H(y,a,s)$ is continuous on $\mathbb{K}$ for every $s \in \mathbb{R}^k$.

Upper semi-continuity of $y \to A(y)$ means: for each open set $A' \subset A$, the set $\{y \in Y : A(y) \subset A'\}$ is open in Y.

Remark 3.2. a) Assumption 3.1(c) implies that the function

$$\bar{v}(y,a) := \int_{\mathbb{R}^k} v\left(H(y,a,s)\right)\mu(\mathrm{d}s) \tag{9}$$

is continuous on $\mathbb{K}$ for any continuous and bounded function v on Y and any probability distribution μ on $\mathbb{R}^k$. Hence, the transition law (8) is weakly continuous.

b) For each $u \in L(Y)$, we define the operator T by

$$Tu(y) := \min_{a\in A(y)} \left\{c(y,a) + \exp(-\alpha)\int_{\mathbb{R}^k} u(H(y,a,s))\rho(s)\,\mathrm{d}s\right\}.$$

From Assumption 3.1, T maps $L(Y)$ into itself. We also define the sequence of function $\{v_n\}$ in $L(Y)$ as $v_0(\cdot) = 0$ and $v_n = Tv_{n-1}$ for $n = 1, 2, \ldots$. That is, for $n \geq 1$ and $y \in Y$,

$$v_n(y) := \min_{a\in A(y)} \left\{c(y,a) + \exp(-\alpha)\int_{\mathbb{R}^k} v_{n-1}(H(y,a,s))\rho(s)\,\mathrm{d}s\right\}.$$

A first consequence of Assumption 3.1, which is stated in [6], with adequate changes, is the following.

Theorem 3.3. Suppose that Assumption 3.1 holds. If $V^*(y) < \infty$ for every $y \in Y$, then

a) $v_n \nearrow V^*$;

b) V^* is the minimal solution in $L(Y)$ that satisfies the optimality equation:

$$V^*(y) = \min_{a\in A(y)} \left\{ c(y,a) + \exp(-\alpha) \int_{\mathbb{R}^k} V^*(H(y,a,s))\rho(s)\,\mathrm{d}s \right\}, \quad \forall\, y \in Y. \tag{10}$$

Equivalently $TV^* = V^*$.

c) There exists a stationary policy $\{f^*\}$, $f^* \in \mathbb{F}$, such that $f^*(y) \in A(y)$ attains the minimum in (10), i. e.,

$$V^*(y) = c(y, f^*) + \exp(-\alpha) \int_{\mathbb{R}^k} V^*(H(y,f^*,s))\rho(s)\,\mathrm{d}s, \quad \forall\, y \in Y,$$

and moreover, the stationary policy $\{f^*\}$ is optimal.

4. APPROXIMATIONS

Part a) of Theorem 3.3 gives an approximation procedure of the value function V^* which is known as *Value Iteration*. Now, we present other types of approximation algorithms.

Approximation via bounded costs. Let $\{c_n(y,a)\}$ be a sequence of nonnegative, bounded and l.s.c. functions on $\mathbb{K}$ such that $c_n \nearrow c$, and we define the sequence of performance indices with their corresponding value functions as:

$$U_n(\pi, y) := E_y^\pi \sum_{t=0}^{\infty} \exp(-S_t) c_n(y_t, a_t), \tag{11}$$

$$U_n(y) := \inf_{\pi\in\Pi} U_n(\pi, y), \quad y \in Y. \tag{12}$$

We also define, for each $n > 0$ and $u \in L(Y)$, the operator $T_n : L(Y) \to L(Y)$ as

$$T_n u(y) = \min_{a\in A(y)} \left\{ c_n(y,a) + \exp(-\alpha) \int_{\mathbb{R}^k} u(H(y,a,s))\rho(s)\,\mathrm{d}s \right\}, \quad \forall\, y \in Y. \tag{13}$$

Observe that from Theorem 3.3, if the one-stage cost c is replaced by the bounded cost c_n, then U_n is the *unique bounded* function in $L(Y)$ satisfying

$$U_n = T_n U_n, \quad n > 0. \tag{14}$$

Following the ideas in Remark 3.2(b), we define the corresponding value iteration functions as: $u_0 := 0$, and

$$u_n(y) = \min_{a\in A(y)} \left\{ c_n(y,a) + \exp(-\alpha) \int_{\mathbb{R}^k} u_{n-1}(H(y,a,s))\rho(s)\,\mathrm{d}s \right\}, \; n > 0, \; y \in Y. \tag{15}$$

That is, $u_n = T_n u_{n-1}$ for $n > 0$.

Theorem 4.1. Under Assumptions 3.1, each of the sequences U_n and u_n is monotone increasing and converges to the value function V^*.

The proof of Theorem 4.1 is based in the following result which is a consequence of Lemma 4.2.4 in [9].

Lemma 4.2. Let u and u_n, $n \geq 1$, be functions in $L(\mathbb{K})$. If $u_n \nearrow u$, then

$$\lim_{n\to\infty} \min_{A(y)} u_n(y,a) = \min_{A(y)} u(y,a), \quad \forall\, y \in Y.$$

P r o o f. [Theorem 4.1] Since $c_n \nearrow c$, from (11) and (12), $\{U_n\}$ is an increasing sequence in $L(Y)$. Thus there exists a function $u \in L(Y)$ such that

$$U_n \nearrow u. \tag{16}$$

Therefore, to prove $U_n \nearrow V^*$, it is sufficient to show

$$u = V^*. \tag{17}$$

To this end, first observe that from (11) and (12) $U_n \leq V^*$ for all n, and from (16) we obtain $u \leq V^*$. On the other hand, letting $n \to \infty$ in (14), Lemma 4.2 yields $u = Tu$, i.e., u satisfies the optimality equation. Therefore, from Theorem 3.3(b), $u \geq V^*$ which proves (17). Finally, applying similar arguments we can also obtain $u_n \nearrow V^*$. □

Policy iteration. First observe that for any stationary policy $f \in \mathbb{F}$, applying the Markov property, the corresponding cost $V(f,y)$, $y \in Y$, satisfies

$$V(f,y) = c(y,f) + \exp(-\alpha) \int_{\mathbb{R}^k} V(f, H(y,a,s))\rho(s)\,\mathrm{d}s, \qquad y \in Y. \tag{18}$$

Let $f_0 \in \mathbb{F}$ be a stationary policy with a finite valued cost $w_0(\cdot) := V(f_0,\cdot) \in L(Y)$. Then from (18),

$$w_0(y) = c(y,f_0) + \exp(-\alpha) \int_{\mathbb{R}^k} w_0(H(y,f_0,s))\rho(s)\,\mathrm{d}s, \qquad y \in Y. \tag{19}$$

Now, let $f_1 \in \mathbb{F}$ be such that

$$\begin{aligned} & c(y,f_1) + \exp(-\alpha) \int_{\mathbb{R}^k} w_0(H(y,f_1,s))\rho(s)\,\mathrm{d}s \\ = \quad & \min_{a\in A(y)} \left\{ c(y,a) + \exp(-\alpha) \int_{\mathbb{R}^k} w_0(H(y,a,s))\rho(s)\,\mathrm{d}s \right\}. \end{aligned} \tag{20}$$

That is,

$$c(y,f_1)+\exp(-\alpha)\int_{\mathbb{R}^k} w_0(H(y,f_1,s))\rho(s)\,\mathrm{d}s = Tw_0(y), \tag{21}$$

where T is the operator defined in Remark 3.2(b). We define $w_1(\cdot)=V(f_1,\cdot)$.

In general, we define a sequence $\{w_n\}$ in $L(Y)$ as follows. Given $f_n\in\mathbb{F}$, suppose that $w_n(\cdot):=V(f_n,\cdot)\in L(Y)$, and let $f_{n+1}\in\mathbb{F}$ be such that

$$\begin{aligned} &c(y,f_{n+1})+\exp(-\alpha)\int_{\mathbb{R}^k} w_n(H(y,f_{n+1},s))\rho(s)\,\mathrm{d}s = Tw_n(y)\\ =\;& \min_{a\in A(y)}\left\{c(y,a)+\exp(-\alpha)\int_{\mathbb{R}^k} w_n(H(y,a,s))\rho(s)\,\mathrm{d}s\right\}. \end{aligned}$$

Theorem 4.3. Under Assumption 3.1, there exists a measurable nonnegative function $w\geq V^*$ such that $w_n\searrow w$, and $Tw=w$. Moreover, if w satisfies

$$\lim_{n\to\infty} E_y^\pi\left[\exp(-S_n)w(y_n)\right]=0 \quad \forall\,\pi\in\Pi,\ y\in Y, \tag{22}$$

then $w=V^*$.

To prove Theorem 4.3 we need the following two results. The first one, Lemma 4.4(a) is stated in [6], whereas the Lemma 4.4(b) is the Lemma 3.3 in [11].

Lemma 4.4. a) Under Assumption 3.1, if $u:Y\to\mathbb{R}$ is a measurable function such that Tu is well defined, $u\leq Tu$, and

$$\lim_{n\to\infty} E_y^\pi\left[\exp(-S_n)u(y_n)\right]=0 \quad \forall\,\pi\in\Pi,\ y\in Y,$$

then $u\leq V^*$.

b) If u_n is a sequence of function on $\mathbb{K}$ such that $u_n\searrow u$, then

$$\lim_{n\to\infty}\inf_{a\in A(y)} u_n(y,a)=\inf_{a\in A(y)} u(y,a),\quad y\in Y.$$

Proof. [Theorem 4.3] According to Lemma 4.4(a), it is sufficient to show the existence of a function $w\geq V^*$ such that $w_n\searrow w$ and $Tw=w$.

From (19) – (21),

$$\begin{aligned} w_0(y) &\geq \min_{a\in A(y)}\left\{c(y,a)+\exp(-\alpha)\int_{\mathbb{R}^k} w_0(H(y,a,s))\rho(s)\,\mathrm{d}s\right\}\\ &= c(y,f_1)+\exp(-\alpha)\int_{\mathbb{R}^k} w_0(H(y,f_1,s))\rho(s)\,\mathrm{d}s. \end{aligned} \tag{23}$$

Iterating this inequality, a straightforward calculation shows that

$$w_0(y) \geq V(f_1, y) = w_1(y), \quad y \in Y.$$

In general, similar arguments yield

$$w_n \geq Tw_n \geq w_{n+1}, \quad n = 1, 2, \ldots \tag{24}$$

Hence, there exists a nonnegative measurable function w such that $w_n \searrow w$. In addition, since $w_n \geq V^* \quad \forall\, n$, we have $w \geq V^*$. Now, applying Lemma 4.4(b) to (24), we have $w \geq Tw \geq w$, which yields $w = Tw$. $\square$

An obvious sufficient condition for (22) is that

(C1) for some constant m, $0 \leq c(y,a) \leq m \quad \forall\, (y,a) \in \mathbb{K}$.

Indeed, under bounded costs, $\{w_n\}$ is a bounded sequence which in turn implies (since $w_n \searrow w$) the boundeness of the function w. This fact clearly yields (22).

Other less obvious sufficient conditions are the following.

(C2) There exist a measurable function $W : Y \to [1, \infty)$ and constants $M > 0$, $\beta \in (0,1)$, and $b < \infty$, such that for all $(y,a) \in \mathbb{K}$ (see Assumption 6.2 below),

$$\sup_{A(y)} c(y,a) \leq MW(y)$$

and

$$\int_{\mathbb{R}^k} W[H(y,a,s)]\rho(s)\,\mathrm{d}s \leq \beta W(y) + b.$$

(C3) $\lim_{n\to\infty} E_y^{\pi}\left[\exp(-S_n)V(\pi', y)\right] = 0 \quad \forall\, \pi, \pi' \in \Pi,\; y \in Y.$

In fact, using similar arguments as is Proposition 4.3.1 in [9] , we have the following relations

$$\text{(C1)} \Longrightarrow \text{(C2)} \Longrightarrow \text{(C3)} \Longrightarrow \text{(22)}.$$

5. APPROXIMATION AND ESTIMATION

In this section we study the approximation problem of V^* for the control system (1)–(2) assuming that the state and discount disturbance processes $\{\xi_t\}$ and $\{\eta_t\}$ are observable with *unknown* densities ρ^ξ and ρ^η, respectively. In this context, in contrast with the interpretation of the control model $\mathcal{M}$ introduced in Section 2, before choosing the action a_t at time t, the controller gets an estimate ρ_t of the unknown density $\rho := \rho^\xi \rho^\eta$ and combines this with the history of the system to select a control $a = a_t(\rho_t)$.

Our approach consists in combining suitable density estimation methods of ρ with a particular case of approximation of V^* via bounded costs treated in the previous section.

Let $\chi_0, \chi_1, \ldots, \chi_{n-1}$ be independent r.v.'s (observed up to time $n-1$) with the unknown density ρ. We consider the control model

$$\mathcal{M}_n = \left(Y, A, \{A(y) \subset A | y \in Y\}, \mathbb{R}^k, H, \rho_n, \bar{c}_n\right)$$

satisfying the following conditions. The space Y, the control space A and the function H are define as before (see (3) and (7)); $\rho_n(s) := \rho_n(s; \chi_0, \chi_1, \ldots, \chi_{n-1})$, $s \in \mathbb{R}^k$, is an estimator of ρ such that, for some $\gamma > 0$,

$$\mathrm{E} \int_{\mathbb{R}^k} |\rho_n(s) - \rho(s)| \, \mathrm{d}s = O(n^{-\gamma}) \text{ as } n \to \infty; \tag{25}$$

and, finally, $\bar{c}_n : \mathbb{K} \to \mathbb{R}$ is the truncated cost defined as

$$\bar{c}_n(y, a) := \min\{c(y, a), n\}, \ (y, a) \in \mathbb{K}. \tag{26}$$

Estimators satisfying (25) are given, for instance, in [2, 5], see also Example 5.3.

Remark 5.1. a) Observe that, for each n, the cost $\bar{c}_n$ satisfies Assumption 3.1(b). Moreover

$$\bar{u}_n \nearrow V^* \quad \text{as} \quad n \to \infty, \tag{27}$$

where $\{\bar{u}_n\}$ is the sequence of l.s.c. functions on $\mathbb{B}(Y)$ defined in (15) corresponding to the cost $\bar{c}_n$. That is, $\bar{u}_0 = 0$, and for $n \geq 1$

$$\bar{u}_n(y) = \min_{a \in A(y)} \left\{ \bar{c}_n(y, a) + \exp(-\alpha) \int_{\mathbb{R}^k} \bar{u}_{n-1}(H(y, a, s)) \rho(s) \, \mathrm{d}s \right\}, \quad y \in Y. \tag{28}$$

b) Since $\bar{c}_n(y, a) \leq n$ for each $n \geq 0$, it is easy to see that

$$\bar{u}_n(y) \leq \sum_{k=1}^{n} k(\exp(-\alpha))^{n-k} \leq n \sum_{k=0}^{\infty} \exp(-\alpha^* k) = \frac{n}{1 - \exp(-\alpha^*)} \tag{29}$$

for every $y \in Y$.

For each fixed $t \geq 0$ and $n > 0$, we define the sequence $\{V_n^{\rho_t}\}$ of l.s.c. function on $\mathbb{B}(Y)$ as:

$$\begin{aligned} V_0^{\rho_t} &= 0; \\ V_n^{\rho_t}(y) &= \min_{a \in A(y)} \left\{ \bar{c}_n(y, a) + \exp(-\alpha) \int_{\mathbb{R}^k} V_{n-1}^{\rho_t}(H(y, a, s)) \rho_t(s) \, \mathrm{d}s \right\}. \end{aligned} \tag{30}$$

In addition, for an arbitrary real number $\nu \in (0, \gamma/2)$, we define a sequence $\{n_t\}$ of integer numbers as $n_t := [t^\nu]$, where γ is as (25) and $[x]$ represents the integer part of x. Recall that a function u on Y $\|u\| = \sup_{y \in Y} |u(y)|$.

Theorem 5.2. Suppose that Assumption 3.1 holds. Then:

a) $\mathrm{E}\left\|V^{\rho_t}_{n_t} - \bar{u}_{n_t}\right\| \to 0$ as $t \to \infty$;

b) For each $y \in Y$,

$$\mathrm{E}\left|V^{\rho_t}_{n_t}(y) - V^*(y)\right| \to 0 \quad \text{as} \quad t \to \infty.$$

P r o o f. Observe that part (b) is a consequence of (a) and the fact that $\bar{u}_n \nearrow V^*$ (see Theorem 4.1). Indeed, for each $y \in Y$ and $t \geq 0$,

$$\left|V^{\rho_t}_{n_t}(y) - V^*(y)\right| \leq \left|V^{\rho_t}_{n_t}(y) - \bar{u}_{n_t}(y)\right| + \left|\bar{u}_{n_t}(y) - V^*(y)\right|. \tag{31}$$

Hence, part (b) follows by taking expectation on both sides of (31) and letting $t \to \infty$.

Therefore, to complete the proof of Theorem 5.2, it only remains to prove part (a). To do this, by (28) and (30), by adding and subtracting the term

$$\exp(-\alpha) \int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\rho_t(s)\,\mathrm{d}s,$$

we have, for each $t \geq 0$,

$$\begin{aligned}
\left|V^{\rho_t}_{n_t}(y) - \bar{u}_{n_t}(y)\right| \leq \sup_{a\in A(y)} &\left\{\int_{\mathbb{R}^k} \left|V^{\rho_t}_{n_t-1}(H(y,a,s)) - \bar{u}_{n_t-1}(H(y,a,s))\right| \rho_t(s)\,\mathrm{d}s \right.\\
&\left. + \int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\left|\rho_t(s) - \rho(s)\right|\,\mathrm{d}s\right\}\\
&\leq \left\|V^{\rho_t}_{n_t-1} - \bar{u}_{n_t-1}\right\| + \left\|\bar{u}_{n_t-1}\right\| \int_{\mathbb{R}^k} \left|\rho_t(s) - \rho(s)\right|\,\mathrm{d}s,
\end{aligned}$$

which implies

$$\left\|V^{\rho_t}_{n_t} - \bar{u}_{n_t}\right\| \leq \left\|V^{\rho_t}_{n_t-1} - \bar{u}_{n_t-1}\right\| + \left\|\bar{u}_{n_t-1}\right\| \int_{\mathbb{R}^k} \left|\rho_t(s) - \rho(s)\right|\,\mathrm{d}s.$$

Iterating the latter inequality we have (recall $V^{\rho_t}_0 = \bar{u} = 0$)

$$\left\|V^{\rho_t}_{n_t} - \bar{u}_{n_t}\right\| \leq \left(\left\|\bar{u}_0\right\| + \cdots + \left\|\bar{u}_{n_t-1}\right\|\right) \int_{\mathbb{R}^k} \left|\rho_t(s) - \rho(s)\right|\,\mathrm{d}s.$$

Now, using the fact that $\{\bar{u}_n\}$ is an increasing sequence, (29) yields

$$\begin{aligned}
\left\|V^{\rho_t}_{n_t} - \bar{u}_{n_t}\right\| &\leq n_t \left\|\bar{u}_{n_t-1}\right\| \int_{\mathbb{R}^k} \left|\rho_t(s) - \rho(s)\right|\,\mathrm{d}s\\
&\leq \frac{n_t^2}{1-\exp(-\alpha^*)} \int_{\mathbb{R}^k} \left|\rho_t(s) - \rho(s)\right|\,\mathrm{d}s, \quad t \geq 0.
\end{aligned} \tag{32}$$

Taking expectation on both sides of (32), we get

$$\mathrm{E}\left\|V_{n_t}^{\rho_t}-\bar{u}_{n_t}\right\| = O(t^{2\nu})O(t^{-\gamma}) = O(t^{2\nu-\gamma}) \to 0 \quad \text{as} \quad t\to\infty,$$

since, by definition of n_t,

$$\frac{n_t^2}{1-\exp\left(-\alpha^*\right)} = O(t^{2\nu}) \quad \text{as} \quad t\to\infty,$$

and $\nu<\gamma/2$. This proves the desired results. □

Example 5.3. The following example shows a sequence of densities f_n that satisfies equation (25). Let g be a strictly positive function such that $\int_\theta^\infty g(x)\,\mathrm{d}x<\infty$ for some θ in $\mathbb{R}$. Let us define the density function

$$f(x) := \frac{g(x)}{\int_\theta^\infty g(y)\,\mathrm{d}y},\ x\in(\theta,\infty).$$

Let $\{x_1, x_2, \dots, \}$ be a sequence of random variables i.i.d. with this distribution. Let $\theta_n := \min\{x_1, x_2, \dots, x_n\}$ for $n=1,2,\dots$ and

$$f_n(x) := \frac{g(x)}{\int_{\theta_n}^\infty g(y)\,\mathrm{d}y},\ x\in(\theta_n,\infty).$$

Then

$$|f(x)-f_n(x)| = \begin{cases} f(x), & \text{if } \theta<x\le\theta_n \\ f_n(x)-f(x) & \text{if } x>\theta_n \end{cases}$$

Hence

$$\int_\theta^\infty |f(x)-f_n(x)|\,\mathrm{d}x = 2\int_\theta^{\theta_n} f(x)\,\mathrm{d}x = 2F(\theta_n)$$

where F is the distribution function corresponding to f. Therefore,

$$\begin{aligned}\mathrm{E}\left(2F(\theta_n)\right) = 2\int_\theta^\infty F(y)f_{\theta_n}(y)\,\mathrm{d}y \quad &= \quad 2\int_\theta^\infty nF(y)(1-F(y))^{n-1}f(y)\,\mathrm{d}y \\ = \frac{2}{n+1} = O(n^{-\gamma}) \qquad &\text{for} \quad 0<\gamma\le 1.\end{aligned}$$

6. ESTIMATION AND CONTROL

We consider again the setting of the previous section. That is, we assume that the density $\rho=\rho^\xi\rho^\eta$ is unknown and the estimator ρ_n is applied. Then we must combine this estimation scheme with control procedures in order to construct optimal policies. It worth noting that, as the discounted index (5) depends strongly on the controls selected at the first stages (precisely when the information about the density ρ is deficient), we cannot ensure, in general, the existence of such policies (see, e. g., [8, 16]). Therefore, the optimality of the resulting policy in the combination of estimation and control will be studied in the following sense.

Let $\Phi : \mathbb{K} \to \mathbb{R}$ be the discrepancy function defined as

$$\Phi(y,a) := c(y,a) + \exp(-\alpha) \int_S V^* \left(H(y,a,s)\right) \rho(s) \, \mathrm{d}s - V^*(y), \ (y,a) \in \mathbb{K}$$

Observe that (10) is equivalent to

$$\min_{a \in A(y)} \Phi(y,a) = 0, \quad y \in Y,$$

and in addition a stationary policy $\{f^*\}$ is optimal if

$$\Phi(y, f^*) = 0, \quad y \in Y.$$

This fact motivates the following definition.

Definition 6.1. A policy $\pi = \{f_t\}$, $f_t \in \mathbb{F}$, is called pointwise asymptotically discounted optimal for the control model $\mathcal{M}$ if, for each $y \in Y$,

$$\lim_{t \to \infty} \mathrm{E} \, \Phi(y, f_t) = 0.$$

In this section we use the algorithm in Theorem 5.2 to show the existence of a policy $\bar{\pi} = \left\{\bar{f}_t\right\}$ ($\bar{f}_t$ depending of the estimators ρ_t) which is (expected) asymptotically discounted optimal. Furthermore, we show that the sequence of minimizers $\left\{\bar{f}_t\right\}$ converges, in the sense of Schäl [15], to an optimal stationary policy f_∞ for the control model $\mathcal{M}$.

The existence of such policies is guaranteed by Assumption 3.1 (see Theorem 3.3 and Remark 3.2(a)). That is, for each $t > 0$, there exists $\bar{f}_t \equiv \bar{f}_t^{\rho_t} \in \mathbb{F}$ such that $\bar{f}_t(y) \in A(y)$ attains the minimum in (30) with n_t instead of n. Thus, for each $t > 0$,

$$V_{n_t}^{\rho_t}(y) = \bar{c}_{n_t}(y, \bar{f}_t) + \exp(-\alpha) \int_{\mathbb{R}^k} V_{n_t - 1}^{\rho_t}(H(y, \bar{f}_t, s)) \rho_t(s) \, \mathrm{d}s, \quad y \in Y, \tag{33}$$

where the minimization is done for every $\omega \in \Omega$. Now, using the compactness of the sets $A(y)$, by a result of Schäl [15], which is reproduced in [9, Proposition D7], there is a stationary policy $\{f_\infty\}$ for the control model $\mathcal{M}$ such that, for each $y \in Y$, $f_\infty(y) \in A(y)$ is an accumulation point of $\left\{\bar{f}_t\right\}$. Hence, for every $y \in Y$, there exists a subsequence $\{t_i\}$ of t ($t_i = t_i(y)$) such that $f_{t_i}(y) \to f_\infty(y)$ as $i \to \infty$. Moreover, observe that from (33), letting $t_i = i$ for notational convenience, we have for each $i > 0$,

$$V_{n_i}^{\rho_i}(y) = \bar{c}_{n_i}(y, \bar{f}_i) + \exp(-\alpha) \int_{\mathbb{R}^k} V_{n_i - 1}^{\rho_i}(H(y, \bar{f}_i, s)) \rho_i(s) \, \mathrm{d}s, \quad y \in Y. \tag{34}$$

In order to show the optimality of the policies $\bar{\pi}$ and $\{f_\infty\}$, we need the following technical requirements on the control model $\mathcal{M}$.

Assumption 6.2. a) For every $y \in Y$, the one-stage cost $c(y,a)$ is nonnegative and continuous in $a \in A(y)$. Moreover, there exist a measurable function $W : Y \to [1,\infty)$ and constants $M > 0$, $\beta \in (0,1)$, and $b < \infty$, such that

$$\sup_{A(y)} c(y,a) \le MW(y)$$

and

$$\int_{\mathbb{R}^k} W[H(y,a,s)]\rho(s)\,\mathrm{d}s \le \beta W(y) + b.$$

b) For each $y \in Y$ and $v \in \mathbb{B}(Y)$, the function $a \to \int_{\mathbb{R}^k} v[H(y,a,s)]\rho(s)\,\mathrm{d}s$ is continuous and bounded on $A(y)$.

c) For each $y \in Y$, the function $a \to \int_{\mathbb{R}^k} W[H(y,a,s)]\rho(s)\,\mathrm{d}s$ is continuous on $A(y)$.

Remark 6.3. Under Assumption 6.2, $\sup_{A(y)} \bar{c}_n(y,a) \le MW(y)$ for all $y \in Y$, $n \ge 0$. Furthermore, straightforward calculation shows that, for each $n \ge 0$, the function $\bar{u}_n$ defined in (28) satisfies

$$\bar{u}_n(y) \le M'W(y), \qquad y \in Y, \tag{35}$$

for some constant $M' < \infty$, which in turns yields

$$V^*(y) \le M'W(y), \qquad y \in Y. \tag{36}$$

We can now state our result as follows.

Theorem 6.4. Under Assumptions 3.1 and 6.2, the policy $\bar{\pi} = \{\bar{f}_t\}$ is pointwise asymptotically discounted optimal. That is, for each $y \in Y$,

$$\mathrm{E}\,\Phi(y,\bar{f}_t) \to 0, \quad \text{as} \quad t \to \infty.$$

In addition, the stationary policy $\{f_\infty\}$ is optimal for the control model $\mathcal{M}$.

The proof of Theorem 6.4 is consequence of the following lemma.

Lemma 6.5. Suppose that Assumption 3.1 and 6.2 hold. Then, for each $y \in Y$,

$$\mathrm{E} \sup_{a \in A(y)} |\Phi(y,a) - \Phi_t(y,a)| \to 0 \text{ as } t \to \infty, \tag{37}$$

where Φ_t is the approximate discrepancy function defined as

$$\Phi_t(y,a) := \bar{c}_{n_t}(y,a) + \exp(-\alpha)\int_{\mathbb{R}^k} V^{\rho_t}_{n_t-1}(H(y,a,s))\rho_t(s)\,\mathrm{d}s - V^{\rho_t}_{n_t}(y). \tag{38}$$

(Note that, from (30), Φ_t is nonnegative.) Moreover, for each $y \in Y$,

$$\mathrm{E}\left|\int_{\mathbb{R}^k} V^*(H(y,\bar{f}_t,s))\rho(s)\,\mathrm{d}s - \int_{\mathbb{R}^k} V^{\rho_i}_{n_i-1}(H(y,\bar{f}_t,s))\rho_t(s)\,\mathrm{d}s\right| \to 0 \tag{39}$$

as $t \to \infty$.

Proof. By definition of Φ and Φ_t, adding and subtracting the terms $\int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\rho(s)\,\mathrm{d}s$ and $\int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\rho_t(s)\,\mathrm{d}s$, we have, for each $y \in Y$ and $t \geq 0$,

$$\sup_{a\in A(y)} |\Phi(y,a) - \Phi_t(y,a)| \leq c_t + d_t$$

$$+\sup_{a\in A(y)} \left|\int_{\mathbb{R}^k} V^*(H(y,a,s))\rho(s)\,\mathrm{d}s - \int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\rho(s)\,\mathrm{d}s\right|$$

$$+\sup_{a\in A(y)} \int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\,|\rho_t(s) - \rho(s)|\,\mathrm{d}s$$

$$+\sup_{a\in A(y)} \int_{\mathbb{R}^k} \left|\bar{u}_{n_t-1}(H(y,a,s)) - V^{\rho_t}_{n_t-1}(H(y,a,s))\right| \rho_t(s)\,\mathrm{d}s, \tag{40}$$

where $c_t := \sup_{a\in A(y)} |c(y,a) - c_{n_t}(y,a)|$ and $d_t := \left|V^*(y) - V^{\rho_t}_{n_t-1}(y)\right|$.

On the other hand, from Assumption 6.2(a) and Theorem 5.2 (b) we obtain, as $t \to \infty$,

$$c_t \to 0 \quad \text{and} \quad \mathrm{E}\,\eta_t \to 0, \tag{41}$$

for each $y \in Y$. Moreover, from Remark 6.3 and an extension of Fatou's lemma (see Lemma 8.3.7 in [10]), it is easy to see that for each $y \in Y$,

$$\sup_{a\in A(y)} \left|\int_{\mathbb{R}^k} V^*(H(y,a,s))\rho(s)\,\mathrm{d}s - \int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\rho(s)\,\mathrm{d}s\right| \to 0 \text{ as } t \to \infty. \tag{42}$$

Now, from (29) and (32) we get a.s., for every $y \in Y$.

$$l_t^{(1)} := \sup_{a\in A(y)} \int_{\mathbb{R}^k} \bar{u}_{n_t-1}(H(y,a,s))\,|\rho_t(s) - \rho(s)|\,\mathrm{d}s$$

$$\leq \frac{n_t}{1-\exp(-\alpha^*)} \int_{\mathbb{R}^k} |\rho_t(s) - \rho(s)|\,\mathrm{d}s \tag{43}$$

and

$$\begin{aligned} l_t^{(2)} &:= \sup_{a \in A(y)} \int_{\mathbb{R}^k} \left| \bar{u}_{n_t - 1}(H(y,a,s)) - V_{n_t-1}^{\rho_t}(H(y,a,s)) \right| \rho_t(s)\, \mathrm{d}s \\ &\le \frac{(n_t - 1)^2}{1 - \exp(-\alpha^*)} \int_{\mathbb{R}^k} |\rho_t(s) - \rho(s)|\, \mathrm{d}s \\ &\le \frac{n_t^2}{1 - \exp(-\alpha^*)} \int_{\mathbb{R}^k} |\rho_t(s) - \rho(s)|\, \mathrm{d}s. \end{aligned} \tag{44}$$

Thus, taking expectation on both sides of (43) and (44), the definition of n_t together (25) implies, for each $y \in Y$,

$$\mathrm{E}\, l_t^{(1)} = O(t^{\nu}) O(t^{-\gamma}) \to 0 \ \text{ as } \ t \to \infty; \tag{45}$$

and

$$\mathrm{E}\, l_t^{(2)} = O(t^{2\nu}) O(t^{-\gamma}) \to 0 \ \text{ as } \ t \to \infty. \tag{46}$$

Hence, (37) follows from (40) – (46).

Finally, applying similar arguments as in (40) and (42) – (46) we get (39). □

Proof. [Theorem 6.4] First observe that from (33), (38) and definition of the policy $\bar{\pi}$ we conclude

$$\Phi_t \left(y, \bar{f}_t(y) \right) = 0 \ \ y \in Y,$$

which in turn yields

$$\begin{aligned} \Phi \left(y, \bar{f}_t(y) \right) &= \left| \Phi \left(y, \bar{f}_t(y) \right) - \Phi_t \left(y, \bar{f}_t(y) \right) \right| \\ &\le \sup_{a \in A(y)} \left| \Phi(y,a) - \Phi_t\left(y,a\right) \right|, \quad y \in Y. \end{aligned}$$

Then, the optimality of $\bar{\pi}$ follows by taking expectation on both sides of this inequality and using Lemma 6.5.

On the other hand, to prove the optimality of the policy $\{f_\infty\}$ (see (34)), we fix an arbitrary $y \in Y$ and observe that for each $i > 0$,

$$\begin{aligned} \int_{\mathbb{R}^k} V_{n_i-1}^{\rho_i}(H(y,\bar{f}_i,s)) \rho_i(s)\, \mathrm{d}s &= \left[\int_{\mathbb{R}^k} V_{n_i-1}^{\rho_i}(H(y,\bar{f}_i,s)) \rho_i(s)\, \mathrm{d}s \right. \\ &\quad \left. - \int_{\mathbb{R}^k} V^*(H(y,\bar{f}_i,s)) \rho(s)\, \mathrm{d}s \right] + \int_{\mathbb{R}^k} V^*(H(y,\bar{f}_i,s)) \rho(s)\, \mathrm{d}s. \end{aligned}$$

Taking expectation and lim inf as $i \to \infty$ on both sides of this equality, from (39) we get

$$\liminf_{i\to\infty} \mathrm{E} \int_{\mathbb{R}^k} V^{\rho_i}_{n_i-1}(H(y,\bar{f}_i,s))\rho_i(s)\,\mathrm{d}s \geq \liminf_{i\to\infty} \mathrm{E} \int_{\mathbb{R}^k} V^*(H(y,\bar{f}_i,s))\rho(s)\,\mathrm{d}s.$$

Thus, since $\bar{f}_i \to f_\infty$, from the lower semicontinuity of V^* (Theorem 3.3), the continuity of H (Assumption 3.1), and Fatou's Lemma,

$$\liminf_{i\to\infty} \mathrm{E} \int_{\mathbb{R}^k} V^{\rho_i}_{n_i-1}(H(y,\bar{f}_i,s))\rho_i(s)\,\mathrm{d}s \geq \int_{\mathbb{R}^k} V^*(H(y,f_\infty,s))\rho(s)\,\mathrm{d}s.$$

On the other hand, taking expectation and lim inf as $i \to \infty$ in (34) (see (41)), we obtain

$$c(y,f_\infty) + \exp(-\alpha) \int_{\mathbb{R}^k} V^*(H(y,f_\infty,s))\rho(s)\,\mathrm{d}s \leq V^*(y). \tag{47}$$

Finally, as y was arbitrary, by (10), the equality holds in (47) for every $y \in Y$. Hence (see Theorem 3.3(c)) $\{f_\infty\}$ is optimal for the control model $\mathcal{M}$. □

ACKNOWLEDGEMENT

This work was partially supported by Consejo Nacional de Ciencia y Tecnología (CONACyT) under Grant 46633.

(Received March 7, 2008.)

REFERENCES

[1] H. Berument, Z. Kilinc, and U. Ozlale: The effects of different inflation risk premiums on interest rate spreads. Physica A *333* (2004), 317–324.

[2] L. Devroye and L. Györfi: Nonparametric Density Estimation the L_1 View. Wiley, New York 1985.

[3] E. B. Dynkin and A. A. Yushkevich: Controlled Markov Processes. Springer–Verlag, New York 1979.

[4] A. Gil and A. Luis: Modelling the U. S. interest rate in terms of I(d) statistical model. Quart. Rev. Economics and Finance *44* (2004), 475–486.

[5] R. Hasminskii and I. Ibragimov: On density estimation in the view of Kolmogorov's ideas in approximation theory. Ann. Statist. *18* (1990), 999–1010.

[6] J. González-Hernández, R. R. López-Martínez, and R. Pérez-Hernández: Markov control processes with randomized discounted cost in Borel space. Math. Meth. Oper. Res. *65* (2007), 27–44.

[7] S. Haberman and J. Sung: Optimal pension funding dynamics over infinite control horizon when stochastic rates of return are stationary. Insurance Math. Econom. *36* (2005), 103–116.

[8] O. Hernández-Lerma: Adaptive Markov Control Processes. Springer–Verlag, New York 1989.

[9] O. Hernández-Lerma and J. B. Lasserre: Discrete-Time Markov Control Processes: Basic Optimality Criteria. Springer–Verlag, New York 1996.

[10] O. Hernández-Lerma and J. B. Lasserre: Further Topics on Discrete-Time Markov Control Processes. Springer–Verlag, New York 1999.

[11] O. Hernández-Lerma and W. Runggaldier: Monotone approximations for convex stochastic control problems. J. Math. Syst., Estimation, and Control *4* (1994), 99–140.

[12] P. Lee and D. B. Rosenfield: When to refinance a mortgage: a dynamic programming approach. European J. Oper. Res. *166* (2005), 266–277.

[13] R. G. Newell and W. A. Pizer: Discounting the distant future: how much do uncertain rates increase valuation? J. Environmental Economic and Management *46* (2003), 52–71.

[14] B. Sack and V. Wieland: Interest-rate smooothing and optimal monetary policy: A review of recent empirical evidence. J. Econom. Business *52* (2000), 205–228.

[15] M. Schäl: Conditions for optimality and for the limit of n-stage optimal policies to be optimal. Z. Wahrsch. Verw. Gerb. *32* (1975), 179–196.

[16] M. Schäl: Estimation and control in discounted stochastic dynamic programming. Stochastics *20* (1987), 51–71.

[17] N. L. Stockey and R. E. Lucas, Jr.: Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA 1989.

Juan González-Hernández, Departamento de Probabilidad y Estadística, IIMAS-UNAM, A.P. 20-726, 01000 México D.F. Mexico.
e-mail: juan@sigma.iimas.unam.mx

Raquiel R. López-Martínez, Facultad de Matemáticas UV, A.P. 270, 91090 Xalapa Ver. Mexico.
e-mail: ralopez@uv.mx

J. Adolfo Minjárez-Sosa, Departamento de Matemáticas, Universidad de Sonora, Rosales s/n, Centro, 83000 Hermosillo, Sonora. Mexico.
e-mail: aminjare@gauss.mat.uson.mx