KYBERNETIKA — VOLUME *42* (2006), NUMBER 6, PAGES 647 – 664

DISCOUNTED MARKOV CONTROL PROCESSES INDUCED BY DETERMINISTIC SYSTEMS[1]

Hugo Cruz-Suárez and Raúl Montes-de-Oca

This paper deals with Markov Control Processes (MCPs) on Euclidean spaces with an infinite horizon and a discounted total cost. Firstly, MCPs which result from the deterministic controlled systems will be analyzed. For such MCPs, conditions that permit to establish the equation known in the literature of Economy as Euler's Equation (EE) will be given. There will be also presented an example of a Markov Control Process with deterministic controlled system where, to obtain the optimal value function, EE applied to the value iteration algorithm will be used. Secondly, the MCPs which result from the perturbation of deterministic controlled systems with a random noise will be dealt with. There will be also provided the conditions which allow to obtain the optimal value function and the optimal policy of a perturbed controlled system, in terms of the optimal value function and the optimal policy of deterministic controlled system corresponding. Finally, several examples to illustrate the last case mentioned will be presented.

Keywords: discounted Markov control process, deterministic control system, Euler equation, deterministic control system perturbed by a random noise

AMS Subject Classification: 90C40, 93E20

1. INTRODUCTION

This paper deals with Markov Control Processes (MCPs) with an infinite horizon and a discounted total cost (see [2] and [9]). There will be considered the MCPs with states and controls spaces (both of the spaces are Euclidean), taking into account the (possibly) unbounded cost function and the (possibly) noncompact control sets.

In the first part of the paper, MCPs which result from deterministic controlled systems (i. e. systems where, for every state and admissible control, the probability of transition is concentrated in just one state) will be dealt with.

For such MCPs, the differentiability conditions of the second order will be presented in the dynamic of the system as well as in the cost function. This will permit to establish the Euler's Equation (EE) applied to the value iteration algorithm (see [2] and [9]).

[1] This work was partially supported by Benemérita Universidad Autónoma de Puebla (BUAP) under grant VIEP-BUAP 38/EXC/06-G and by Consejo Nacional de Ciencia y Tecnología (CONACyT) under grant 51222.

The EE is known in the applications of MCPs to economic models, where it is established and solved (in some cases empirically) (see [4] pp. 600–601, [5] pp. 196–198, [7] pp. 12–14, 36–37, [11] p. 159, [12] p. 49, and [15] pp. 97–100).

To illustrate this part of the article, two examples will be given. One of them is a deterministic problem with quadratic cost. The other example is the problem of deterministic economic growth, raised in [15] pp. 93–96. In the best of our knowledge, the latter example has not been solved by means of the EE applied to the value iteration algorithm.

In the second part of the article, the MCPs which result from the suitable perturbation of deterministic controlled systems with random noise will be considered. For deterministic controlled systems taken into account in this part, the functional form of the optimal value function is supposed to be known, as well as the existence of an optimal policy.

In this part, there will be established the conditions which permit a Markov control process with controlled perturbed system to obtain its optimal value function in terms of optimal value function of deterministic controlled system corresponding. Moreover, in this case, the optimal policies of both systems (i. e. the perturbed and the deterministic ones) coincide.

Three examples will be presented in this part. The first two correspond to the stochastic versions of deterministic controlled systems which have been previously presented in this paper. They are a stochastic problem with quadratic cost and a problem of the stochastic economic growth. (It is important to mention that this problem of the stochastic economic growth has been raised in [15] pp. 275–277, and in this reference a solution to the problem is presented without a proof). The third one is a non-convex example.

The paper is organized as follows. In Section 2, the theory of MCPs with deterministic controlled systems is presented. In Sections 3 and 4, the Euler's Equation is established and the deterministic examples are given. In Section 5, the theory of MCPs with perturbed by random noise controlled systems is presented, and in Section 6, the examples of those are given.

2. DETERMINISTIC CONTROL SYSTEMS

The basic control model $(X, A, \{A(x) : x \in X\}, Q, c)$ consists of the *state space* X, the *control* (or *action*) *space* A, $A(x)$ is a nonempty subset of A, for every $x \in X$, the *transition law* Q, and the *cost function* c.

Both X and A are (nonempty) Borel spaces of $\mathbb{R}^p$ and $\mathbb{R}^m$ ($p, m \geq 1$ are integers), respectively.

Let us denote by $\mathcal{B}(X)$ and $\mathcal{B}(A)$ the Borel σ -algebras of X and A, respectively.

For each $x \in X$, $A(x) \in \mathcal{B}(A)$ and represents the set of *admissible controls* to state x. It is supposed that the set $\mathbb{K} := \{(x, a) | x \in X, a \in A(x)\}$, is measurable in $X \times A$.

The transition law $Q(B|x, a)$, where $B \in \mathcal{B}(X)$ and $(x, a) \in \mathbb{K}$ is a stochastic kernel on X given $\mathbb{K}$ (i. e. $Q(\cdot\,|x, a)$ is a probability measure on X for every $(x, a) \in \mathbb{K}$, and $Q(B|\cdot)$ is a measurable function on $\mathbb{K}$ for every $B \in \mathcal{B}(X)$).

In Sections 2, 3 and 4 of the paper, it will be assumed that the transition law Q is given by a difference equation of the type:

$$x_{t+1} = F(x_t, a_t), \tag{2.1}$$

$t = 0, 1, \ldots$, where $F : \mathbb{K} \to X$ is a given measurable function. Here $\{x_t\}$ and $\{a_t\}$ are the state and control sequences respectively. Observe that in this case Q is given by $Q(B\,|x, a) = I_B(F(x, a))$, for all $B \in \mathcal{B}(X)$ and $(x, a) \in \mathbb{K}$, where $I_B(\cdot)$ denotes the indicator function of B. Note also that Q will be denoted by Q_F to emphasize that Q is induced by F.

The cost function c is a nonnegative measurable function on $\mathbb{K}$.

$\mathbb{F}$ denotes the set of measurable functions $f : X \to A$ such that $f(x) \in A(x)$, for all $x \in X$.

A *Markov policy* is a sequence $\pi = \{f_t\}$ such that $f_t \in \mathbb{F}$, for $t = 0, 1, \ldots$ The set of all Markov policies is denoted by $\mathbb{M}$.

A Markov policy $\pi = \{f_t\}$ is said to be *stationary* if f_t is independent of t, i.e. $f_t = f \in \mathbb{F}$, for all $t = 0, 1, \ldots$; in this case, π is denoted by f and refers to $\mathbb{F}$ as the set of stationary policies.

For each policy $\pi \in \mathbb{M}$ and initial state $x \in X$, it is defined that

$$w(\pi, x) = \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t).$$

$w(\pi, x)$ is called the *total discounted cost*, where $\alpha \in (0, 1)$ is the discount factor.

The optimal control problem for a deterministic control system consists of determining a policy $\pi^* \in \mathbb{M}$, such that

$$w(\pi^*, x) = \inf_{\pi \in \mathbb{M}} w(\pi, x),$$

$x \in X$, and π^* will be called a (*deterministic*) *optimal policy.* The function W defined by

$$W(x) = \inf_{\pi \in \mathbb{M}} w(\pi, x),$$

$x \in X$, will be called the (deterministic) *optimal value function.*

Remark 2.1. In applications of MCPs to Economics, it is common to present the following optimization problem (see [4, 6, 10, 13, 14] and [15]). Take $\widehat{X} \subset \mathbb{R}^p$. Consider a multifunction which goes from $\widehat{X}$ to $\widehat{X}$, i.e., for every $x \in \widehat{X}$, there is a nonempty subset of $\widehat{X}$, denoted by $\Gamma(x)$ (Γ is known as the technological correspondence). $U : Graph(\Gamma) \to \mathbb{R}$ is the reward function (it is assumed that $U \le 0$ or $U \le \tau$, where τ is a real fixed number, and $Graph(\Gamma) := \{(x, y) \in \widehat{X} \times \widehat{X} : x \in \widehat{X}$, $y \in \Gamma(x)\}$), and $\alpha \in (0, 1)$ is the discount factor. Then the maximization problem is posed as follows:

$$\max_{\{x_{t+1}\}} \sum_{t=0}^{\infty} \alpha^t U(x_t, x_{t+1})$$

subject to $x_{t+1} \in \Gamma(x_t)$, $t = 0, 1, \ldots$, $x_0 \in \widehat{X}$ fixed.

In fact, it is possible to establish this maximization problem in terms of a discounted MCP, taking $X = A = \widehat{X}$, $A(x) = \Gamma(x)$, $x \in X$, $x_{t+1} = F(x_t, a_t) = a_t$ and $c(x_t, a_t) = -U(x_t, x_{t+1})$, $t = 0, 1, \ldots$

Definition 2.2. The *value iteration functions* are defined as

$$W_n(x) = \min_{a \in A(x)} \left[c(x, a) + \alpha W_{n-1}(F(x, a)) \right], \tag{2.2}$$

for all $x \in X$ and $n = 1, 2, \ldots$, with $W_0(\cdot) = 0$.

Remark 2.3. Under certain Assumptions (see Remark 2.6 b) below), it is possible to demonstrate that for each $n = 1, 2, \ldots$, there exists a stationary policy $g_n \in \mathbb{F}$ such that the minimum in (2.2) is attained, i. e.

$$W_n(x) = c(x, g_n(x)) + \alpha W_{n-1}(F(x, g_n(x))), \tag{2.3}$$

$x \in X$.

Lemma 2.4. Under certain assumptions (see Remark 2.6 a) for a) and c) of this Lemma, and Remark 2.6 b) for b) of this Lemma), it results that:

a) The optimal value function W is a solution for the following equation (known as the *Dynamic Programming Equation*):

$$W(x) = \min_{a \in A(x)} \{ c(x, a) + \alpha W(F(x, a)) \}, \tag{2.4}$$

$x \in X$.

b) There exists $g \in \mathbb{F}$ such that $g(x) \in A(x)$ attains the minimum in the right-hand side of (2.4), i. e.,

$$W(x) = c(x, g(x)) + \alpha W(F(x, g(x))), \tag{2.5}$$

$x \in X$ and g is optimal.

c) For every $x \in X$, $W_n(x) \to W(x)$ as $n \to +\infty$, with W_n as in (2.2).

Remark 2.5. For a deterministic control system where the transition law is induced by a continuous function F, it is direct to prove that the transition law is weakly continuous, i. e. $\int v(y) Q_F(\mathrm{d}y \,|x, a) = v(F(x, a))$ is a continuous function of $(x, a) \in \mathbb{K}$ for every $v \in \{\zeta : X \to \mathbb{R} : \zeta \text{ is a bounded continuous function}\}$. This property of the weak-continuity in the transition law Q_F has been used as an assumption in some of the conditions given in a), and b) in Remark 2.6.

Remark 2.6.

a) For bounded costs, see conditions (a), (b) and (c) in Theorem 2.8 in [8] p. 23; for unbounded costs, see conditions in Theorems 3, 4, 5, and 6 in [13].

b) See Conditions 3.3.3 (a), 3.3.3 (b), and 3.3.3 (c1) or Conditions 3.3.4 (a) and 3.3.4 (b1), and Theorem 3.3.5 in [9] pp. 27–30.

Throughout the paper, deterministic control systems for which (2.2), (2.3) and (a), (b) and (c) of Lemma 2.4 hold are considered.

3. THE EULER'S EQUATION

Let R, Y and Z be Euclidean spaces.

For any set $B \subset R$, a point $x \in B$ is called an *interior point* of B if there exists an open set U such that $x \in U \subset B$. The *interior* of B is the set of all interior points of B and is denoted by $int(B)$.

Let $\widehat{R}$ be a subset of $R \times Y$.

Let $\theta : \widehat{R} \to Z$ be a function. Suppose that $\theta = \theta(\upsilon, \eta)$ is twice differentiable. The partial derivatives with respect to the variables υ and η are denoted by θ_ν and θ_η, respectively. The notation for the second partial derivatives of θ with respect to υ and η are $\theta_{\upsilon\upsilon}$ and $\theta_{\eta\eta}$, respectively.

Besides, for $\phi = 1, 2$, define $C^\phi(\widehat{R}; Z) := \{\theta : \widehat{R} \to Z | \text{the first } \phi \text{ derivatives of } \theta$ exist and are continuous$\}$.

Now, let $(X,\ A,\ \{A(x) : x \in X\}, Q_F, c)$ be a fixed Markov control model that satisfies the description of the previous section.

In this section, a functional equation, known in the literature of Economics as the Euler's Equation (EE) (see [4, 5, 7, 11, 12] and [15]) will be deduced. The EE characterizes the derivative of W_n and, using this information in different situations (see Section 4 below), it is possible to determine W_n explicitly (integrating its derivative) and, later, to determine W taking the limit of W_n, when n tends to ∞.

Let $G^n(x, a) := c(x, a) + \alpha W_{n-1}(F(x, a))$, $n = 1, 2, \ldots, (x, a) \in \mathbb{K}$.

As usual, for each $x \in X$, and $n = 1, 2, \ldots$, the *first-order condition* for the optimality of $G^n(x, \cdot)$ is defined by means of the $\widehat{a} \in int(A(x))$ such that $G^n_a(x, \widehat{a}) = 0$.

The transpose of a vector γ, denoted by γ^{tr} is considered like a vertical vector.

Assumption 1.

a) $c \in C^2(int(\mathbb{K}); \mathbb{R})$.

b) $F \in C^2(int(\mathbb{K}); X)$.

c) For each $n = 1, 2, \ldots$, and $x \in X$, $G^n_{aa}(x, \cdot)$ is positive definite (see Remark 3.1 b) below).

d) For each $n = 1, 2, \ldots$, $g_n(x) \in int(A(x))$, for every $x \in X$ (see Remark 3.1 a) below).

Remark 3.1.

a) As $A(x) \subset A$, $x \in X$ is open, then, obviously, $g_n(x) \in int(A(x))$, $x \in X$. In addition, in [4] p. 599 and [11] p. 155, for certain class of growth economic models, conditions which guarantee that the minimizers of the optimality equation belong to the interior of the admissible action sets are presented.

b) Observe that Assumption 1 c) implies that $G^n(x, \cdot)$ is a strictly convex function, for each $x \in X$ and $n = 1, 2, \ldots$ (see [4], p. 260), and also notice that g_n is unique for each $n = 1, 2, \ldots$

Lemma 3.2. Under Assumption 1, for each $n = 1, 2 \ldots$, $W_n \in C^2(int(X); \mathbb{R})$ and $g_n \in C^1(int(X); A)$.

Proof. The proof will be made by induction. Consider $n = 1$. In this case,

$$W_1(x) = \min_{a \in A(x)} c(x, a),$$

$x \in X$. Since $g_1(x) \in int(A(x))$, $x \in X$, the minimizer g_1 satisfies the first-order condition for the optimality of c, i.e. $c_a(x, g_1(x)) = 0$, $x \in X$. Then, using the Implicit Function Theorem (see [4], pp. 210–211) and Assumptions 1 a) and 1 c) it results that $g_1 \in C^1(int(X); A)$. On the other hand, $W_1(x) = c(x, g_1(x))$, $x \in X$, then

$$W_1'(x) = c_x(x, g_1(x)) + c_a(x, g_1(x)) g_1'(x), \tag{3.1}$$

$x \in X$. Using the first-order condition for the optimality of c, in (3.1) it is obtained that $W_1'(x) = c_x(x, g_1(x))$, $x \in X$. Since $c \in C^2(int(\mathbb{K}); \mathbb{R})$ and $g_1 \in C^1(int(X); A)$, $W_1 \in C^2(int(X); \mathbb{R})$. Now, suppose that $W_{n-1} \in C^2(int(X); \mathbb{R})$, for some integer $n > 1$. Hence, combining this fact with Assumptions 1 a) and 1 b), it results that $G^n \in C^2(int(\mathbb{K}); \mathbb{R})$. On the other hand, the first-order condition for the optimality of G^n is given by $G^n_a(x, g_n(x)) = 0$, $x \in X$ (see (2.2) and the definition of G^n). By Assumption 1 c) and the Implicit Function Theorem (see [4], pp. 210–211), it is obtained that $g_n \in C^1(int(X); A)$. Then $W_n'(x) = G^n_x(x, g_n(x)) + G^n_a(x, g_n(x)) g_n'(x)$, $x \in X$. Using the first-order condition in the last equality it results that $W_n'(x) = G^n_x(x, g_n(x))$, $x \in X$. Since $G^n \in C^2(int(\mathbb{K}); \mathbb{R})$ and $g_n \in C^1(int(X); A)$, it is concluded that $W_n \in C^2(int(X); \mathbb{R})$. □

The following Assumption will be used to obtain the EE (see (3.2) below).

Assumption 2.

a) $X, A \subset \mathbb{R}^p$.

b) There is the inverse of the matrix F_a. It will be denoted by F_a^{-1}.

c) The function $H(x, a) := \left(c_x - c_a F_a^{-1} F_x^{tr}\right)(x, a), (x, a) \in \mathbb{K}$ is invertible in the second variable with the inverse $H^{-1} : X \times H[\mathbb{K}] \to A$, where $H[\mathbb{K}] := \{H(x, a) : (x, a) \in \mathbb{K}\}$.

Remark 3.3. Observe that the function H in Assumption 2 c) is given in terms of the derivatives of the cost function c, of the dynamic of the system F and of its inverse F^{-1}.
In Theorem 3.4 below, for each n, the functions H and H^{-1} will be used to present the derivatives of the value iteration function W_n and the maximizer g_n, respectively. In Example 4.4, H and H^{-1} will be explicitly obtained and Assumption 2 c) will be verified for these functions (see Lemma 4.5 below).

Theorem 3.4. Under Assumptions 1 and 2, the derivatives of the value functions W_n, $n = 1, 2, \ldots$, satisfy the *Euler's Equation*:

$$\begin{array}{c}\alpha W'_{n-1}\left(F\left(x, H^{-1}(x, W'_n(x))\right)\right) F_a(x, H^{-1}(x, W'_n(x))) \\ +c_a(x, H^{-1}(x, W'_n(x))) = 0,\end{array} \tag{3.2}$$

$x \in X$.

P r o o f. Observe that for each positive integer n, the first-order condition for the optimality of G^n is

$$\alpha W'_{n-1}(F(x, g_n(x)))F_a(x, g_n(x)) + c_a(x, g_n(x)) = 0, \tag{3.3}$$

$x \in X$, $g_n \in \mathbb{F}$ and $n = 1, 2, \ldots$ On the other hand, derivating (2.3) it is obtained that

$$\begin{aligned} W'_n(x) &= c_x(x, g_n(x)) + c_a(x, g_n(x))g'_n(x) \\ &\quad +\alpha W'_{n-1}(F\left(x, g_n\left(x\right)\right))\left[F_x\left(x, g_n\left(x\right)\right) + F_a\left(x, g_n\left(x\right)\right) g'_n(x)\right], \end{aligned}$$

$x \in X$.

Now, substituting (3.3) in the last equation and using the definition of H, it results that $W'_n(x) = H(x, g_n(x))$, $x \in X$. Then, by the invertibility of H (see Assumption 2 c)) it yields that

$$g_n(x) = H^{-1}(x, W'_n(x)), \tag{3.4}$$

$x \in X$. Substituting (3.4) in (3.3), the result follows. □

4. DETERMINISTIC EXAMPLES

In this section two examples will be presented. These examples are going to illustrate the theory of the previous section.

Example 4.1. Let $X = A = A(x) = \mathbb{R}$, $x \in X$. The cost function and the dynamic of the system are $c(x, a) = x^2 + \frac{1}{2}a^2$, and $x_{t+1} = \left(a_t^2 + 1\right)^{-1/4}$, $t = 0, 1, \ldots$, $(x, a) \in \mathbb{K}$, respectively.

Lemma 4.2. For Example 4.1, for each $n = 1, 2, \ldots$

a) W_n is convex;

b) $W_n \in C^2(int(X); \mathbb{R})$ and $g_n \in C^1(int(X); A)$;

c) The value functions W_n satisfy

$$W_n(x) = x^2 + D_n, \tag{4.1}$$

$x \in X$, and

$$D_n - \alpha D_{n-1} = \alpha, \tag{4.2}$$

with $D_0 = 0$.

Proof.

a) The proof is made by induction and is similar to the proof of Lemma 6.2 in [3].

b) The proof will be made by induction. For $n = 1$, $W_1(x) = x^2$, and $g_1(x) = 0$, $x \in X$. Then, $W_1 \in C^2(X;\mathbb{R})$ and $g_1 \in C^1(X;A)$. Suppose that $W_n \in C^2(X;\mathbb{R})$ and $g_n \in C^1(X;A)$, for $n > 1$. Since W_n is convex for $n = 1, 2, \ldots$, then for each $x \in X$,

$$G^{n+1}_{aa}(x, a) > 0,$$

$a \in A(x)$. Then the result follows by Lemma 3.2.

c) Let $n > 1$. In this case the equation (3.3) is

$$g_n(x) = \frac{\alpha}{2} g_n(x) W'_{n-1}\left(\frac{1}{\sqrt[4]{(g_n(x))^2 + 1}}\right)\left((g_n(x))^2 + 1\right)^{-5/4}, \tag{4.3}$$

$x \in X$. On the other hand,

$$W_n(x) = x^2 + (1/2)\,(g_n(x))^2 + \alpha W_{n-1}\left(1/\sqrt[4]{(g_n(x))^2 + 1}\right),$$

$x \in X$. Derivating W_n and using (4.3) it results that $W'_n(x) = 2x$, $x \in X$. Hence, $W_n(x) = x^2 + D_n$, $x \in X$ and $D_n \in \mathbb{R}$. Substituting W_n and W_{n-1} in (3.2) yields (4.2). □

Lemma 4.3. For Example 4.1, the optimal value function and the optimal policy are given by

$$W(x) = x^2 + \frac{\alpha}{1-\alpha}, \tag{4.4}$$

and

$$g(x) = 0, \tag{4.5}$$

respectively, $x \in X$.

Proof. By c) of the previous lemma, $W_n(x) = x^2 + D_n$, $n = 1, 2, \ldots$, $x \in X$, and D_n is a constant real to determine. The sequence D_n is convergent to $D = \alpha/(1-\alpha)$, due to $W_n(x) \to W(x)$, $x \in X$ and (4.2). Then, when $n \to \infty$ in (4.1), (4.4) results. Finally, substituting (4.4) in (2.4), results that the first-order condition is

$$g(x) - \frac{\alpha g(x)}{\left((g(x))^2 + 1\right)^{3/2}} = 0,$$

$x \in X$, where the only solution is $g(x) = 0$, $x \in X$. Hence (4.5) results. □

Example 4.4. An Economic Growth Model (see [13] and [15]). Let $\delta \in (0,1)$ be a fixed number. Take $X = A = [0,1)$, and $A(0) = \{0\}$, $A(x) = (0, x^\delta]$, $x \in X$, $x \neq 0$. The dynamic of the system is given by the difference equation $x_{t+1} = x_t^\delta - a_t$, $t = 0, 1, \ldots$, and the reward function is defined by $r(x,a) = \ln a$, if $x \in (0,1)$, $a \in (0, x^\delta]$, and $r(0,0) = -\infty$. Suppose that $0 < \alpha\delta < 1$.

Notice that in the example $Q(\{x^\delta - a\}\,|x,a) = 1$ if $x \in (0,1), a \in (0, x^\delta]$, and $Q(\{0\}\,|0,0) = 1$. Moreover, for each $n = 1, 2, \ldots$, $W_n(0) = -\infty$, $g_n(0) = 0$, $W(0) = -\infty$, and $g(0) = 0$. (Observe that this example can be considered as a case of a non-negative cost, taking $c(x,a) = -r(x,a) = -\ln a \geq 0$, $(x,a) \in \mathbb{K}$.)

The rest of the solution of the example is given in the following lemmas.

Lemma 4.5. For Example 4.4,

a) W_n is concave for each $n = 1, 2, \ldots$;

b) $W_n \in C^2(int(X); \mathbb{R})$ and $g_n \in C^1(int(X); A)$ for each $n = 1, 2, \ldots$;

c) The value functions W_n, $n = 2, 3, \ldots$ satisfy the functional equation

$$\frac{W_n'(x)}{\delta x^{\delta-1}} = \alpha W_{n-1}\left(x^{\delta-1} \frac{W_n'(x)x - \delta}{W_n'(x)} \right), \tag{4.6}$$

$x \in (0,1)$. Moreover,

$$W_n'(x) = (\delta/x) \sum_{k=0}^{n-1} (\alpha\delta)^k, \tag{4.7}$$

$n = 1, 2, \ldots, 0 < x < 1$.

P r o o f.

a) Similar to the proof of Lemma 6.2 in [3].

b) The proof will be made by induction. For $n = 1$, it results that

$$W_1(x) = \max_{a \in (0, x^\delta]} \ln a = \delta \ln x,$$

$x \in (0,1)$. Then, $W_1 \in C^2(int(X); \mathbb{R})$ and $g_1 \in C^1(int(X); A)$. Suppose that $W_n \in C^2(int(X); \mathbb{R})$ and $g_n \in C^1(int(X); A)$, for $n > 1$. Then,

$$\begin{aligned} W_{n+1}(x) &= \max_{a \in (0, x^\delta]} \left[\ln a + \alpha W_n(x^\delta - a) \right], \\ &= \max_{a \in (0, x^\delta)} \left[\ln a + \alpha W_n(x^\delta - a) \right], \end{aligned}$$

$x \in (0,1)$, because $W_{n+1}(x^\delta - x^\delta) = W_{n+1}(0) = -\infty$. Hence, the optimal policy $g_{n+1}(x) \in (0, x^\delta), x \in X$. On the other hand, $G_{aa}^{n+1}(x,a) = -a^{-2} + \alpha W_n''(x^\delta - a) < 0, (x,a) \in \mathbb{K}$, since W_n is concave for $n = 1, 2, \ldots$ Then the result follows by Lemma 3.2.

c) Since $H(x,a) = \left(\delta x^{\delta-1}\right)/a, (x,a) \in \widehat{\mathbb{K}} := \{(x,a) : x \in (0,1), a \in (0,x^\delta)\}$ and $H^{-1}(x,u) = \left(\delta x^{\delta-1}\right)/u, x \in (0,1)$ and $u \in (\delta/x, \infty)$, Assumption 2 holds and, using Theorem 3.4, (4.6) follows.

Now, for $n = 1$, it is obtained that

$$W_1(x) = \max_{a \in (0,x^\delta]} \ln a = \delta \ln x,$$

$x \in (0,1)$, and $W_1'(x) = \delta/x, x \in (0,1)$. For $n > 1$, suppose that $W_{n-1}'(x) = (\delta/x)\sum_{k=0}^{n-2}(\alpha\delta)^k$, $x \in (0,1)$. Then, substituting W_{n-1}' in (4.6), (4.7) results. □

Lemma 4.6. For Example 4.4, the optimal value function and the optimal policy are given by

$$W(x) = \frac{\delta}{1-\alpha\delta} \ln x + M, \tag{4.8}$$

and

$$g(x) = (1-\alpha\delta)x^\delta, \tag{4.9}$$

respectively, where $x \in (0,1)$ and

$$M = \frac{1}{1-\alpha}\left[\ln(1-\alpha\delta) + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta)\right]. \tag{4.10}$$

Proof. By Lemma 4.5, it results that $W_n(x) = \delta\sum_{k=0}^{n-1}(\alpha\delta)^k \ln x + M_n$, $n = 1, 2, \ldots$, $x \in (0,1)$, and M_n is a real constant to determine. The sequence $\{M_n\}$ is convergent to a real number M, because $W_n(x) \to W(x)$, $x \in (0,1)$, and $\sum_{k=0}^{\infty}(\alpha\delta)^k = 1/(1-\alpha\delta)$. Then, (4.8) follows. Substituting (4.8) in (2.4), it results that

$$\frac{\delta}{1-\alpha\delta}\ln x + M = \max_{a \in (0,x^\delta)}\left[\ln a + \frac{\alpha\delta}{1-\alpha\delta}\ln(x^\delta - a) + \alpha M\right], \tag{4.11}$$

$x \in (0,1)$. The first-order condition for the right-hand side of (4.11) is $1/g(x) = \alpha\delta/\left[(1-\alpha\delta)\left(x^\delta - g(x)\right)\right]$, $x \in (0,1)$. Solving this equation for g, it results that

$$g(x) = (1-\alpha\delta)x^\delta,$$

$x \in (0,1)$. Substituting g in (4.11), it is obtained that

$$\frac{\delta}{1-\alpha\delta}\ln x + M = \ln(1-\alpha\delta) + \delta\ln x + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta) + \frac{\alpha\delta^2}{1-\alpha\delta}\ln x + \alpha M,$$

$x \in (0,1)$. Then, $M = \ln(1-\alpha\delta) + \ln(\alpha\delta)\alpha\delta/(1-\alpha\delta) + \alpha M$. Solving this for M, (4.10) is obtained. □

5. DETERMINISTIC CONTROL SYSTEMS PERTURBED BY A RANDOM NOISE

As in Section 2, consider a deterministic problem with space state X, space control A ($X \subseteq \mathbb{R}^p$ and $A \subseteq \mathbb{R}^m$, $p, m \geq 1$ are integers), with admissible sets $A(x) \subset A$, $x \in X$. Suppose that the dynamic of the system is given by the difference equation $x_{t+1} = F(x_t, a_t)$, $t = 0, 1, \ldots$, where $F : \mathbb{K} \to X$ is a given measurable function, and the cost function $c : \mathbb{K} \to \mathbb{R}$ is a nonnegative and measurable function.

Now, consider a stochastic control system with the same: state space X, the control space A, the admissible sets $A(x)$, $x \in X$, and the cost function c, but with the following dynamic of the system:

$$x_{t+1} = L(F(x_t, a_t), \xi_t), \tag{5.1}$$

$t = 0, 1, \ldots$, where $\{\xi_t\}$ is a sequence of i.i.d. random elements taking values in a Borel space $S \subset \mathbb{R}^k$ ($k \geq 1$ is an integer) with density function Δ. $L : X \times S \to X$ is assumed to be a measurable function.

Observe that in this case the transition law Q_L is given by

$$Q_L(B \,|x, a) = \int I_B(L(F(x, a), s))\Delta(s)\, \mathrm{d}s, \tag{5.2}$$

$B \in \mathcal{B}(X)$ and $(x, a) \in \mathbb{K}$. Also notice that the sets $\mathbb{F}$ and $\mathbb{M}$ are the same for the deterministic system $(X, A, \{A(x) : x \in X\}, Q_F, c)$ and for the stochastic system $(X, A, \{A(x) : x \in X\}, Q_L, c)$ (see Section 2).

A *control policy* is a sequence $\pi = \{\pi_t\}$ such that for each $t = 0, 1, \ldots$, $\pi_t(\cdot\,|h_t)$ is a stochastic kernel on $\mathcal{B}(A)$, given the history $h_t = (x_0, a_0, \ldots, x_{t-1}, a_{t-1}, x_t)$, and which satisfies the constraint $\pi_t(A(x_t)\,|h_t) = 1$. The class of all policies is denoted by Π.

Notice that $\mathbb{F} \subset \mathbb{M} \subset \Pi$.

For each policy π and initial state $x_0 = x \in X$, a probability measure P_x^π is defined on the space $\Omega = (X \times A)^\infty$ in a canonical way. E_x^π denotes the corresponding expectation operator.

For each policy π and initial state $x \in X$, it is defined that

$$v(\pi, x) := E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t) \right]. \tag{5.3}$$

$v(\pi, x)$ is called the *expected total discounted cost*, where $\alpha \in (0, 1)$ is the discount factor.

A Markov control process for which the transition law is given by (2.1) will be referred to as a *deterministic control system.* In the case of the transition law given by (5.1), it will be referred to as a *stochastic control system.*

The stochastic optimal control problem consists of determining a policy $\pi^* \in \Pi$, such that

$$v(\pi^*, x) = \inf_{\pi \in \Pi} v(\pi, x),$$

$x \in X$. π^* will be called an *optimal policy*. The function V defined by

$$V(x) := \inf_{\pi \in \Pi} v(\pi, x), \tag{5.4}$$

$x \in X$, will be called the *optimal value function*.

Let $(X, A, \{A(x) : x \in X\}, Q_F, c)$ be a fixed deterministic control system with corresponding stochastic control system $(X, A, \{A(x) : x \in X\}, Q_L, c)$. Besides, let W be the optimal value function and g be the optimal policy for the deterministic control system.

Assumption 3.

a) The optimal value function W and the optimal policy g are known for the deterministic control system.

b) There exist a non-negative measurable function $h : S \to \mathbb{R}$, such that

$$\int W(y) Q_L(\mathrm{d}y \,|x, a) = W(F(x, a)) + \int h(s)\Delta(s)\,\mathrm{d}s,$$

$(x, a) \in \mathbb{K}$.

c) There exists $\vartheta \in [0, 1)$ and $N \geq 0$ such that

$$\int W(y) Q_L(\mathrm{d}y \,|x, a) \leq \vartheta W(x) + N,$$

$(x, a) \in \mathbb{K}$.

Remark 5.1.

a) Assumption 3 a) has been analyzed in the first part of this paper (see Section 3). W and g can be obtained by means of the Euler's equation or, directly, by means of the value iteration algorithm.

b) Assumption 3 b) can be checked directly once that W is known.

c) Notice that if c is bounded with upper bound $\widehat{N}$, then Assumption 3 c) holds with $\vartheta = 0$ and $N = \widehat{N}/(1 - \alpha)$.

Definition 5.2. $M(X)^+$ denotes the cone of non-negative measurable functions on X, and, for every $U \in M(X)^+$, TU is the function on X defined as

$$TU(x) := \min_{a \in A(x)} \left[c(x, a) + \alpha \int U(y) Q_L(\mathrm{d}y \,|x, a) \right],$$

$x \in X$. T is known as the *dynamic programming operator* associated to $(X, A, \{A(x) : x \in X\}, Q_L, c)$.

Definition 5.3. Let $\mathcal{R} \subset M(X)^+$, $\mathcal{R}$ is invariant with respect to the the dynamic programming operator (see Definition 5.2), if for each $U \in \mathcal{R}$, $TU \in \mathcal{R}$.

Lemma 5.4. Let $\mathcal{L} := \{W + \lambda : \lambda \in \mathbb{R}\}$. Then under Assumption 3, $\mathcal{L}$ is invariant with respect to the dynamic programming operator T.

Proof. Take $U \in \mathcal{L}$. Hence, $U(x) = W(x) + \lambda$, $x \in X$ for some $\lambda \in \mathbb{R}$. Then

$$\begin{aligned} TU(x) &= \min_{a\in A(x)} \left[c(x,a) + \alpha \int U(y) Q_L(\mathrm{d}y \,|x,a) \right] \\ &= \min_{a\in A(x)} \left[c(x,a) + \alpha E[U(L(F(x,a),\xi))] \right], \\ &= \min_{a\in A(x)} \left[c(x,a) + \alpha E[W(L(F(x,a),\xi)) + \lambda] \right], \\ &= \min_{a\in A(x)} \left[c(x,a) + \alpha W(F(x,a)) \right] + \overline{\lambda}, \\ &= W(x) + \overline{\lambda}, \end{aligned}$$

$x \in X$, where $\overline{\lambda} = \alpha\left(E\left[h(\xi)\right] + \lambda\right)$, and

$$E[h(\xi)] = \int h(s)\Delta(s)\,\mathrm{d}s.$$

Therefore, $TU \in \mathcal{L}$. □

Lemma 5.5. Under Assumption 3, $T\overline{W} = \overline{W}$, where

$$\overline{W}(x) := W(x) + \frac{\alpha}{1-\alpha} E\left[h(\xi)\right],$$

$x \in X$.

Proof. Note that in the proof of the previous lemma, $TU = U$ if and only if $\lambda = [\alpha/(1-\alpha)]\, E\left[h(\xi)\right]$. Then $T\overline{W} = \overline{W}$. This concludes the proof of Lemma 5.5. □

Theorem 5.6. Under Assumption 3, the value function of the stochastic control system is $\overline{W}$, i.e. $V = \overline{W}$. Moreover, $V(x) = v(g,x)$, $x \in X$ (hence g is an optimal policy for the stochastic control system).

Proof. Since $T\overline{W} = \overline{W}$ and g is an optimal policy for the deterministic problem, it results that

$$\overline{W}(x) = c(x,g(x)) + \alpha \int \overline{W}(y) Q_L(\mathrm{d}y \,|x,g(x)),$$

$x \in X$. Iterating the last equality, it results that

$$\overline{W}(x) = E_x^g \left[\sum_{t=0}^{n-1} \alpha^t c(x_t,a_t) \right] + \alpha^n E_x^g \overline{W}(x_n),$$

for all $n \geq 1, x \in X$. Since, $\overline{W}$ is nonnegative, $\overline{W}(x) \geq E_x^g \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t)\right]$, $n \geq 1, x \in X$. Then, when $n \to \infty$ in the last inequality, it is obtained that

$$\overline{W}(x) \geq v(g, x), \tag{5.5}$$

$x \in X$. Therefore, $\overline{W} \geq V$.

On the other hand, let $\pi \in \Pi$ and $x \in X$. Then using the Markov-like property (see [9], p. 20) and $T\overline{W} = \overline{W}$, it results that for each $t = 0, 1, \ldots,$

$$\begin{aligned} E_x^\pi[\alpha^{t+1}\overline{W}(x_{t+1})\,|h_t, a_t] &= \alpha^t \left[\alpha \int_X \overline{W}(y) Q_L(\mathrm{d}y\,|x_t, a_t)\right], \\ &= \alpha^t \left[c(x_t, a_t) + \alpha \int_X \overline{W}(y) Q_L(\mathrm{d}y\,|x_t, a_t) - c(x_t, a_t)\right], \\ &\geq \alpha^t \left[\overline{W}(x_t) - c(x_t, a_t)\right]. \end{aligned}$$

Hence, $\alpha^t c(x_t, a_t) \geq -E_x^\pi \left[\alpha^{t+1}\overline{W}(x_{t+1}) - \alpha^t \overline{W}(x_t) \middle|\, h_t, a_t\right]$. Taking expectations and adding $t = 0, \ldots, n-1$, it is obtained that

$$E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t)\right] \geq \overline{W}(x) - \alpha^n E_x^\pi \left[\overline{W}(x_n)\right],$$

$x \in X$. Since $\overline{W}(x) = W(x) + [\alpha/(1-\alpha)]\, E\,[h(\xi)]$, $x \in X$,

$$E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t)\right] \geq \overline{W}(x) - \alpha^n E_x^\pi \left[W(x_n)\right] - \frac{\alpha^{n+1}}{1-\alpha} E\left[h(\xi)\right]. \tag{5.6}$$

$x \in X$. On the other hand, using Assumption 3 c), it results that

$$\begin{aligned} E_x^\pi[W(x_n)\,|h_{n-1}, a_{n-1}] &= \int W(y) Q_L(\mathrm{d}y\,|x_{n-1}, a_{n-1}), \\ &\leq \vartheta W(x_{n-1}) + N. \end{aligned}$$

Therefore $E_x^\pi\left[W(x_n)\right] \leq \vartheta E_x^\pi\left[W(x_{n-1})\right] + N$. Iterating this inequality yields $E_x^\pi\left[W(x_n)\right] \leq \vartheta^n W(x) + (1 + \vartheta + \cdots + \vartheta^{n-1})N$. Then it is obtained by (5.6) that

$$E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t)\right] \geq \overline{W}(x) - (\alpha\vartheta)^n W(x) - \alpha^n N \sum_{k=0}^{n-1} \vartheta^k - \frac{\alpha^{n+1}}{1-\alpha} E\left[h(\xi)\right], \tag{5.7}$$

$x \in X$. Finally, letting $n \to \infty$ in (5.7) gives

$$E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t)\right] \geq \overline{W}(x), \tag{5.8}$$

$x \in X$ (recall that $\vartheta \in [0, 1)$ and observe that $0 < \alpha\vartheta < 1$).

Now, from (5.8) and since π and x are arbitrary, $V \geq \overline{W}$.

Therefore, $\overline{W} = V$. Now, combining (5.4) and (5.5) it follows that $V(x) = v(g, x)$, $x \in X$. Therefore g is optimal for the stochastic control system. This is the end of the proof of Theorem 5.6. □

6. STOCHASTIC EXAMPLES

Example 6.1. Consider the deterministic model worked on in Example 4.1 with the following variant in the dynamic of the system $x_{t+1} = 1/\sqrt[4]{(a_t)^2+1} + \xi_t$, $t = 0, 1, \ldots$, and $\{\xi_t\}$ as a sequence of i.i.d. random elements taking values in $S = \mathbb{R}$. It is supposed that these elements have the mean equal to zero and a finite variance σ^2.

Lemma 6.3. For Example 6.1, the optimal value function V and the optimal policy g are given by

$$V(x) = x^2 + \frac{\alpha\left(\sigma^2+1\right)}{1-\alpha}, \quad \text{and} \quad g(x) = 0,$$

$x \in X$, respectively (see Example 4.1).

Proof. To prove the Lemma, it is just necessary to verify Assumption 3 and recall that the optimal value function and the optimal policy to the deterministic problem are $W(x) = x^2 + D$, and $g(x) = 0$, $x \in X$, respectively. Then

$$\begin{aligned} E\left[W(1/\sqrt[4]{a^2+1}+\xi)\right] &= E\left[(1/\sqrt[4]{a^2+1}+\xi)^2 + D\right], \\ &= \left(1/\sqrt[4]{a^2+1}\right)^2 + \sigma^2 + D, \\ &= W(1/\sqrt[4]{a^2+1}) + E\left[h(\xi)\right], \end{aligned}$$

$x \in X$, where $h(u) = s^2$, $s \in \mathbb{R}$. Assumption 3 c) is verified with $\vartheta = 0$ and $N = 1 + \sigma^2 + D$. Then using Theorem 5.6, the result follows. □

Example 6.3. A Stochastic Economic Growth Model [15]. Consider the deterministic growth model worked on in Example 4.4 with the stochastic dynamic of the system given by $x_{t+1} = \left(x_t^\delta - a_t\right)\xi_t$, $t = 0, 1, \ldots$, where $\{\xi_t\}$ is a sequence of i.i.d. random elements taking values in $S = (0, 1)$. Let $\kappa := E(\ln \xi)$, where ξ is a generic element of the sequence $\{\xi_t\}$. Suppose that κ is finite.

Lemma 6.4. For Example 6.3, the value function V and the optimal policy g are given by

$$V(x) = \frac{\delta}{1-\alpha\delta}\ln x + \frac{1}{1-\alpha}\left[\ln(1-\alpha\delta) + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta)\right] + \frac{\alpha\delta}{(1-\alpha\delta)(1-\alpha)}\kappa,$$

and $g(x) = (1-\alpha\delta)\, x^\delta$, $x \in (0, 1)$, respectively.

Proof. In this case, using the deterministic solution (see Example 4.4), it results that

$$\begin{aligned} E\left[W(\left(x^\delta - a\right)\xi)\right] &= E\left[\frac{\delta}{1-\alpha\delta}\ln\left[\left(x^\delta - a\right)\xi\right] + M\right], \\ &= \frac{\delta}{1-\alpha\delta}\ln\left(x^\delta - a\right) + M + \frac{\delta}{1-\alpha\delta}\kappa, \\ &= W(x^\delta - a) + E\left[h(\xi)\right], \end{aligned}$$

$(x,a) \in \mathbb{K}$ and $h(s) = [\delta/(1-\alpha\delta)] \ln s$, $s \in S$. Assumption 3 c) can be proved in the following way:

$$\begin{aligned}\int W((x^\delta - a)s)\Delta(s)\,\mathrm{d}s &= \int \left(\frac{\delta}{1-\alpha\delta}\ln\left[\left(x^\delta - a\right)s\right] + M\right)\Delta(s)\,\mathrm{d}s,\\ &= \frac{\delta}{1-\alpha\delta}\ln\left(x^\delta - a\right) + M + \frac{\delta}{1-\alpha\delta}\kappa,\\ &\leq \frac{\delta^2}{1-\alpha\delta}\ln x + M + \frac{\delta}{1-\alpha\delta}\kappa,\\ &= \delta\left[\frac{\delta}{1-\alpha\delta}\ln x + M\right] + \frac{\delta}{1-\alpha\delta}\kappa + M(1-\delta),\end{aligned}$$

$(x,a) \in \widehat{\mathbb{K}}$. The last inequality is due to $0 < x^\delta - a \leq x^\delta$ and $0 < \delta < 1$. Therefore, Assumption 3 holds with $\vartheta = \delta$ and $N \geq \delta\kappa/(1-\alpha\delta) + M(1-\delta)$, where N is positive. □

Now, the following Example illustrates that the theory developed in Section 5 holds for some non-convex models.

Example 6.5. Let ζ be a fixed positive real number. Let $X = (0,\zeta)$, $A = A(x) = [-\pi, \pi/2]$, $x \in X$. The cost and the dynamic of the system are the following: $c(x,a) = x + \sin a + 1$, and $x_{t+1} = \theta x_t$, $t = 0,1,2,\ldots$ with $0 < \theta < 1$.

Lemma 6.6. For Example 6.5, the iteration value functions $W_n, n = 1,2,\ldots$, are given by

$$W_n(x) = (1 + \alpha\theta + \cdots + (\alpha\theta)^{n-1})x, \tag{6.1}$$

$x \in X$.

P r o o f. The proof will be made by induction. For $n = 1$,

$$W_1(x) = \min_{a\in A(x)} [x + \sin a + 1] = x,$$

$x \in X$. Suppose that for $n > 1$, $W_n(x) = (1 + \alpha\theta + \cdots + (\alpha\theta)^{n-1})x$. Since the minimum of the function $\eta(a) = \sin a$, $a \in [-\pi, \pi/2]$ is uniquely attained in $a^* = -\pi/2$, then $g_{n+1}(x) = -\pi/2$ and

$$\begin{aligned}W_{n+1}(x) &= \min_{a\in A(x)} \left[x + \sin a + 1 + \alpha W_n(\theta x)\right],\\ &= \min_{a\in A(x)} \left[x + \sin a + 1 + \alpha\theta(1 + \alpha\theta + \cdots + (\alpha\theta)^{n-1})x\right],\\ &= (1 + \alpha\theta + \cdots + (\alpha\theta)^{n})x,\end{aligned}$$

$x \in X$. This completes the proof of Lemma 6.6. □

Lemma 6.7. The value function and the optimal policy are given by $W(x) = x/(1-\alpha\theta)$, $g(x) = -\pi/2$, $x \in X$, respectively.

Proof. Using Lemma 6.6 (and the fact that $0 < \alpha\theta < 1$), when n tends to ∞ in (6.1), it is obtained that $W(x) = [1/(1-\alpha\theta)]\, x$, $x \in X$. Finally, substituting W in (2.2), it results that $g(x) = -\pi/2$, $x \in X$. □

Now, suppose that the dynamic of the system is perturbed in the following form: $x_{t+1} = \theta x_t + \xi_t$, $t = 0, 1, \ldots$, where $\{\xi_t\}$ is a sequence of i.i.d. random elements taking values in $S = (0, (1-\theta)\zeta)$. It is supposed that these elements have a finite mean μ.

Lemma 6.8. For the stochastic version of Example 6.5, the value function V and the optimal policy g are given by

$$V(x) = \frac{1}{1-\alpha\theta}x + \frac{\alpha}{1-\alpha}\mu$$

and $g(x) = -\pi/2$, $x \in X$, respectively.

Proof. Using the deterministic solution, it follows that

$$\begin{aligned} E\left[W(\theta x + \xi)\right] &= E\left[\frac{\theta x}{1-\alpha\theta} + \xi\right], \\ &= W(\theta x) + E\left[h(\xi)\right], \end{aligned}$$

$x \in X$ and $h(s) = s$, $s \in S$. Note that, in this case, the cost is bounded. Then Assumption 3 c) holds. Therefore, by Theorem 5.6, the result follows. □

ACKNOWLEDGEMENT

The authors wish to thank an anonymous referee for providing us reference [7].

(Received December 15, 2005.)

REFERENCES

[1] L. M. Benveniste and J. A. Scheinkman: On the differentiability of the value function in dynamic models of economics. Econometrica *47* (1979), 727–732.
[2] D. P. Bertsekas: Dynamic Programming: Deterministic and Stochastic Models. Prentice-Hall, Englewood Cliffs, New Jersey 1987.
[3] D. Cruz-Suárez, R. Montes-de-Oca, and F. Salem-Silva: Conditions for the uniqueness of optimal policies of discounted Markov decision processes. Math. Methods Oper. Res. *60* (2004), 415–436.
[4] A. De la Fuente: Mathematical Methods and Models for Economists. Cambridge University Press, New York 2000.
[5] D. Duffie: Security Markets. Academic Press, Boston 1988.
[6] J. Durán: On dynamic programming with unbounded returns. J. Econom. Theory *15* (2000), 339–352.

[7] B. Heer and A. Maußner: Dynamic General Equilibrium Modelling: Computational Method and Application. Springer-Verlag, Berlin 2005.

[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] C. Le Van and L. Morhaim: Optimal growth models with bounded or unbounded returns: a unifying approach. J. Econom. Theory *105* (2002), 158–187.

[11] D. Levhari and T. N. Srinivasan: Optimal savings under uncertainty. Rev. Econom. Stud. *36* (1969), 153–164.

[12] L. J. Mirman: Dynamic models of fishing: a heuristic approach. In: Control Theory in Mathematical Economics (Pan-Tai Liu and J. G. Sutinen, eds.), Marcel Dekker, New York 1979, pp. 39–73.

[13] J. L. Rincón-Zapatero and C. Rodríguez-Palmero: Existence and uniqueness of solutions to the Bellman equation in the unbounded case. Econometrica *71* (2003), 1519–1555.

[14] M. S. Santos: Numerical solution of dynamic economic models. In: Handbook of Macroeconomic, Volume I (J. B. Taylor and M. Woodford, eds.), North Holland, Amsterdam 1999, pp. 311–386.

[15] N. L. Stokey and R. E. Lucas: Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, Mass. 1989.

Hugo Cruz-Suárez, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Ave. San Claudio y Río Verde, Col. San Manuel, Ciudad Universitaria, Puebla, Pue. 72570. México.
e-mail: hcs@fcfm.buap.mx

Raúl Montes-de-Oca, Departamento de Matemáticas, UAM-Iztapalapa, Ave. San Rafael Atlixco #186. Col. Vicentina. 09340. México.
e-mail: momr@xanum.uam.mx