KYBERNETIKA — VOLUME *44* (2008), NUMBER 3, PAGES 400 – 415

STABILITY ESTIMATING IN OPTIMAL STOPPING PROBLEM

Elena Zaitseva

We consider the optimal stopping problem for a discrete-time Markov process on a Borel state space X. It is supposed that an unknown transition probability $p(\cdot|x)$, $x \in X$, is approximated by the transition probability $\widetilde{p}(\cdot|x)$, $x \in X$, and the stopping rule $\widetilde{\tau}_*$, optimal for $\widetilde{p}$, is applied to the process governed by p. We found an upper bound for the difference between the total expected cost, resulting when applying $\widetilde{\tau}_*$, and the minimal total expected cost. The bound given is a constant times $\sup_{x\in X} \|p(\cdot|x) - \widetilde{p}(\cdot|x)\|$, where $\|\cdot\|$ is the total variation norm.

Keywords: discrete-time Markov process, optimal stopping rule, stability index, total variation metric, contractive operator, optimal asset selling

AMS Subject Classification: 60G40

1. SETTING OF THE PROBLEM AND MOTIVATION

In this paper we find an upper bound for stability index (or index of robustness in other terminology) in the problem of optimal stopping of a general discrete-time Markov process. In the setting under consideration the transition probability of the Markov process is assumed to be known only approximately. We follow the approach and the definition of stability index proposed in [8]. Upper bounds for such index were obtained for various classes of Markov control processes, for instance, in [7–11, 15, 16]. A different view on stability estimating of control processes is given, for example, in [4, 5, 18]. For the optimal stopping and for the related total cost minimization, the stability problem has been still open.

Let us consider the classical optimal stopping problem for a discrete-time Markov process on a general Borel state space $(X, \mathcal{B}_X)$. The process $\{x_t, t = 0, 1, 2, \ldots\}$ is specified by a given *initial state* $x_0 \in X$ and a *transition probability* $p(B|x)$, $B \in \mathcal{B}_X$, $x \in X$. We denote by P_{x_0} the corresponding probability on the product space $\Omega = X^\infty$, and by $\mathcal{T}_{x_0}$, the set of all stopping times (with respect to the natural filtration of this space), such that $P_{x_0}(\tau < \infty) = 1$ for each $\tau \in \mathcal{T}_{x_0}$.

Two measurable *bounded* functions $c_0 : X \to [0, \infty)$, $r : X \to [0, \infty)$ are given, and

$$\overline{c} := \sup_{x\in X} c_0(x), \qquad \overline{r} := \sup_{x\in X} r(x). \tag{1.1}$$

Any $\tau \in \mathcal{T}_{x_0}$ defines a *stopping rule* of the process $\{x_t\}$ in such a way that the process is stopped at τ with the revenue equal to $r(x_\tau)$ and with the costs $c_0(x_0), c_0(x_1), \ldots, c_0(x_{\tau-1})$ paid for the extension of observations till the moment τ (i. e. for "nonstopping" until τ).

To deal with a minimization problem, we define the *cost* of the stopping rule $\tau \in \mathcal{T}_{x_0}$ (with the initial state x_0) as follows:

$$W(x_0, \tau) := \mathrm{E}_{x_0}\left[\sum_{t=0}^{\tau-1} c_0(x_t) - r(x_\tau)\right], \tag{1.2}$$

where E_{x_0} is the expectation with respect to P_{x_0}, and, by the convention, $\sum_{t=0}^{-1}[\quad] := 0$.

The *value function* of the optimal stopping problem is

$$W_*(x_0) := \inf_{\tau \in \mathcal{T}_{x_0}} W(x_0, \tau), \quad x_0 \in X. \tag{1.3}$$

Since $\tau \equiv 0$ is one of the admissible stopping rules, we get that $W_*(x_0) \in [-\overline{r}, 0]$, $x_0 \in X$. Under assumptions made in Section 2, there exists an *optimal stopping rule* τ_*, that is

$$W(x_0, \tau_*) = W_*(x_0), \quad x_0 \in X.$$

We consider the following way to approximate the rule τ_* in the situation when we know the transition probability $p(B|x)$ *only approximately.* It is supposed that instead of an *unknown* "real" transition probability $p(B|x)$ we deal with some its approximation $\widetilde{p}(B|x)$, $B \in \mathcal{B}_X$, $x \in X$, obtained, for example, from statistical estimations or (and) theoretical simplifications. Also, assuming the existence of an optimal stopping rule $\widetilde{\tau}_*$ for the Markov process $\{\widetilde{x}_t, t = 0, 1, 2, \ldots\}$ with the transition probability $\widetilde{p}(B|x)$ and the initial state x_0, we admit that one can find $\widetilde{\tau}_*$ in order to try the latter as a reasonable approximation to the unknown "real" optimal rule τ_*.

Similarly to (1.2) and (1.3) the stopping rule $\widetilde{\tau}_*$ is defined as follows:

$$\widetilde{W}(x_0, \widetilde{\tau}_*) = \widetilde{W}_*(x_0) = \inf_{\tau \in \widetilde{\mathcal{T}}_{x_0}} \widetilde{W}(x_0, \tau), \tag{1.4}$$

where

$$\widetilde{W}(x_0, \tau) := \widetilde{\mathrm{E}}_{x_0}\left[\sum_{t=0}^{\tau-1} c_0(\widetilde{x}_t) - r(\widetilde{x}_\tau)\right], \tag{1.5}$$

and $\widetilde{\mathrm{E}}_{x_0}$ is the expectation with respect to the probability $\widetilde{P}_{x_0}$ corresponding to the process $\{\widetilde{x}_t, t = 0, 1, 2, \ldots\}$, and $\widetilde{\mathcal{T}}_{x_0}$ is the set of all stopping times satisfying: $\widetilde{P}_{x_0}(\tau < \infty) = 1$.

Having in the mind the application of the "approximating" stopping rule $\widetilde{\tau}_*$ to the real process $\{x_t, t = 0, 1, 2, \ldots\}$ (given by p), we will measure the quality of approximation by the following *stability index* [8]:

$$\Delta(x_0) := W(x_0, \widetilde{\tau}_*) - W(x_0, \tau_*) \geq 0, \tag{1.6}$$

where W is the expected cost defined in (1.2). This index evaluates the excess of the cost over the minimal value $W(x_0, \tau_*)$ when applying the stopping rule $\widetilde{\tau}_*$.

Remark 1.1. Under the assumptions of Section 2, we will get that $\widetilde{\tau}_* \in \mathcal{T}_{x_0}$, that makes consistent the definition in (1.6).

The aim of the paper is to find an inequality (an upper bound of the stability index) of the form:

$$\Delta(x_0) \leq K\, d(p, \widetilde{p}), \tag{1.7}$$

where K is an explicitly calculated constant that does not depend on $x_0 \in X$,

$$d(p, \widetilde{p}) = \sup_{x \in X} \|p(\cdot|\, x) - \widetilde{p}(\cdot|\, x)\|, \tag{1.8}$$

and $\|\cdot\|$ is the *total variation norm.*

To prove (1.7) (Theorem 2.1 in Section 2), we use the geometric ergodicity of the processes $\{x_t\}$, $\{\widetilde{x}_t\}$ and the conditions that guarantee the existence of optimal stopping rules τ_* and $\widetilde{\tau}_*$ that order to stop at first passage time to certain subsets S and, respectively, $\widetilde{S}$ of X.

In Section 3 we give simple examples of "unstable" optimal stopping models, in which the stability index remains to be greater than a positive constant while the measure of disturbance $d(p, \widetilde{p})$ approaches zero. Also we apply the inequality of Theorem 2.1 (Section 2) to evaluate the stability in the asset selling problem. For this problem, in particular, we compare the cases of independent and weakly dependent offers.

2. ASSUMPTIONS AND RESULT

It is well-known (see, for instance, [20], Chapt. II, § 14) that the value functions in (1.3) and (1.4) satisfy the following optimality equations:

$$W_*(x) = \min\left\{-r(x), c_0(x) + \int_X W_*(y)\, p(\mathrm{d}y|x)\right\}, \tag{2.1}$$

$$\widetilde{W}_*(x) = \min\left\{-r(x), c_0(x) + \int_X \widetilde{W}_*(y) \widetilde{p}(\mathrm{d}y|x)\right\}, \tag{2.2}$$

$x \in X$.

We define the Borel subsets S and $\widetilde{S}$ of X (possibly empty) by setting:

$$S := \big\{x \in X :\ W_*(x) = -r(x)\big\}, \tag{2.3}$$

$$\widetilde{S} := \big\{x \in X :\ \widetilde{W}_*(x) = -r(x)\big\}. \tag{2.4}$$

Also we set

$$\tau_* := \text{a moment of the first entrance of a process in the set } S, \tag{2.5}$$

$$\widetilde{\tau}_* := \text{a moment of the first entrance of a process in the set } \widetilde{S}. \tag{2.6}$$

Relationship (2.5) defines two stopping times, respectively, for the process $\{x_t\}$ and $\{\widetilde{x}_t\}$, when it is applied to these processes. Similarly, (2.6) determines two corresponding stopping times. To this end, for example, the statement $\tau_* \in \mathcal{T}_{x_0}$ means the application of (2.5) to $\{x_t\}$ with $P_{x_0}(\tau < \infty) = 1$.

It is known, and it will be seen later in this paper, that if $\tau_* \in \mathcal{T}_{x_0}, \widetilde{\tau}_* \in \widetilde{\mathcal{T}}_{x_0}$, then these stopping times generate optimal stopping rules (for the processes $\{x_t\}$ and $\{\widetilde{x}_t\}$, respectively).

Assumption 1.

(a) The processes $\{x_t\}$ and $\{\widetilde{x}_t\}$ have stationary probabilities π and, respectively, $\widetilde{\pi}$.

(b) There exist constants δ, $0 \le \delta < 1$, and $M < \infty$ such that

$$\sup_{x \in X} \|p^t(\cdot|x) - \pi(\cdot)\| \le M\delta^t, \quad t = 1, 2, \dots, \tag{2.7}$$

$$\sup_{x \in X} \|\widetilde{p}^t(\cdot|x) - \widetilde{\pi}(\cdot)\| \le M\delta^t, \quad t = 1, 2, \dots, \tag{2.8}$$

where p^t and $\widetilde{p}^t$ are the t-step transition probabilities for the processes $\{x_t\}$ and $\{\widetilde{x}_t\}$, respectively, and

$$\begin{aligned} \|\mu - \nu\| &:= 2 \sup_{B \in \mathcal{B}_X} |\mu(B) - \nu(B)| \\ &\equiv \sup_{\varphi: \|\varphi\|_\infty \le 1} \left| \int_X \varphi(x)\, \mathrm{d}\mu - \int_X \varphi(x)\, \mathrm{d}\nu \right| \end{aligned} \tag{2.9}$$

is the total variation norm.

Assumption 2. There exists $\alpha > 0$ such that

$$\pi(S) \ge \alpha,\ \pi(\widetilde{S}) \ge \alpha,\ \widetilde{\pi}(S) \ge \alpha,\ \widetilde{\pi}(\widetilde{S}) \ge \alpha. \tag{2.10}$$

Lemma 2.1. Under Assumptions 1 and 2, $\tau_* \in \mathcal{T}_{x_0}$, $\widetilde{\tau}_* \in \widetilde{\mathcal{T}}_{x_0}$ and $\widetilde{\tau}_* \in \mathcal{T}_{x_0}$ for all $x_0 \in X$ and, consequently, the stopping rules $\tau_*, \widetilde{\tau}_*$ in (2.5), (2.6) are optimal respectively for the processes $\{x_t\}$ and $\{\widetilde{x}_t\}$.

Remark 2.1.

(a) The stability index in (1.6) is well defined due to Lemma 2.1.

(b) Example 4 of Section 3 shows that in general the stopping rules defined in (2.5), (2.6) may be not optimal.

(c) See, for instance, [14] for conditions sufficient for (2.7), (2.8) which are given in terms of stochastic Lyapunov functions.

We are ready to formulate the main result of the paper.

Theorem 2.1. Let Assumptions 1 and 2 hold. Then

$$\sup_{x_0 \in X} \Delta(x_0) \leq K \sup_{x \in X} \|p(\cdot|x) - \widetilde{p}(\cdot|x)\|, \tag{2.11}$$

where

$$K = \frac{2}{3} \frac{\max\{\overline{c}, \overline{r}\}}{\alpha} N(N+1), \tag{2.12}$$

and

$$N = \left[\frac{\log\left(\frac{\alpha}{2M}\right)}{\log(\delta)} \right] + 2. \tag{2.13}$$

Remark 2.2.

(a) The constants $\overline{c}$ and $\overline{r}$ in (2.12) come from (1.1). In (2.13) the symbol $[\cdot]$ denotes the integer part. If $\delta = 0$, then in (2.12) $N := 2$.

(b) In some specific cases the stability bound (2.11) might be inaccurate. For instance, in Example 1 of Section 3 $\Delta(x_0) = o(\varepsilon)$ as $\varepsilon \to 0$ (see Remark 3.2), where $\varepsilon := \sup_{x \in X} \|p(\cdot|x) - \widetilde{p}(\cdot|\, x)\|$. It even could be that $\Delta(x_0) = 0$ in spite of $\varepsilon > 0$. On the other hand, it is easy to give simple examples of chains on two-point state space X (satisfying the hypothesis of Theorem 2.1), such that the stability index $\Delta(x_0)$ on the left-side of (2.11) is of order ε times a constant.

3. EXAMPLES

3.1. Example 1. Stability estimating in the asset selling problem

In the recent years new important applications related to this old problem have appeared. These are, for example, the models of optimization of an option exercising times (see, for instance, [1, 21]) and of optimization of a risk process stopping to recalculate premiums ([13, 17]).

In the classical version of the problem the state space is a bounded interval $[0, L]$, an initial state is $\widetilde{x}_0 = 0$, while the Markov process $\{\widetilde{x}_t,\ t = 1, 2, \ldots\}$ is a sequence of i.i.d. random variables taking values in $[0, L]$ with a given distribution function $\widetilde{F}$. These random variables represent offers received from period to period for the asset on sale. If an owner of the asset accepts at time t the offer $\widetilde{x}_t$, then he/she gets a revenue equal to $r(\widetilde{x}_t) = \widetilde{x}_t$; otherwise the constant holding cost $c_0 > 0$ is paid for expecting the next offer. It is assumed that the past offers can be accepted at any future period and that

$$\mathrm{E}\widetilde{x}_1 > c_0. \tag{3.1}$$

As is well-known (see, for instance, [2], Sec. 6.3), the optimal stopping rule $\widetilde{\tau}_*$ is determined by (2.6) with $\widetilde{S} = [\widetilde{x}_*, L]$ in (2.4). Here the number $\widetilde{x}_* < L$ can be found (for the case of a continuous $\widetilde{F}$) as a solution to the equation

$$\widetilde{x}_* = \widetilde{x}_* \widetilde{F}(\widetilde{x}_*) + \int_{\widetilde{x}_*}^{L} y \,\mathrm{d}\widetilde{F}(y) - c_0. \tag{3.2}$$

Now we suppose that the hypothesis that the offers are independent is only a theoretical simplification of a “real” situation, in which successive offers $x_1, x_2, \ldots$ form a Markov process on the state space $X = [0, L]$. We let that this process satisfies Assumptions 1 and 2. The two last inequalities in (2.10) turn into the following ones:

$$\widetilde{F}(S) := P(\widetilde{x}_1 \in S) \geq \alpha, \quad 1 - \widetilde{F}(\widetilde{x}_*) \geq \alpha. \tag{3.3}$$

Note that for $c_0 > 0$ equation (3.2) provides that $1 - \widetilde{F}(\widetilde{x}_*) > 0$.

Conditions (2.8) trivially hold if we choose $\widetilde{\pi} = \widetilde{F}$ and $\delta = 0$. Consequently, we can apply Theorem 2.1, and, taking into account that by (3.1) $\max(\overline{c}, \overline{r}) \leq L$, we get that

$$\Delta(0) \leq \frac{4L}{3\alpha} N(N+1) \sup_{x \in [0,L]} \sup_{B \in \mathcal{B}_{[0,L]}} |p(B|x) - \widetilde{F}(B)|, \tag{3.4}$$

where $p(B|x)$ is the transition probability of the “real offer process” $\{x_t,\ t \geq 1\}$.

Remark 3.1. The right-hand side of (3.4) is “small” when $\{x_t, t \geq 1\}$ consists of some sort “weakly dependent” random variables with the marginal distributions close to $\widetilde{F}$.

Let us consider two particular cases. First is the simplest one, when the random variables $x_1, x_2, \ldots$ are independent and identically distributed with a distribution function F, for which $\widetilde{F}$ serves as a known approximation. Assuming that $\mathrm{E}x_1 > c_0$ and that F is continuous, we see again that the optimal stopping rule τ_* is determined by (2.3), (2.5) with $S = [x_*, L],\ x_* < L$, and

$$x_* = x_* F(x_*) + \int_{x_*}^{L} y \, \mathrm{d}F(y) - c_0. \tag{3.5}$$

Assumption 1 holds trivially with $\delta = 0$. Further, assuming, for instance, that

$$[\gamma, L] \subset \mathrm{supp}(F) \ \text{ and } \ [\gamma, L] \subset \mathrm{supp}(\widetilde{F}) \quad \text{for some} \quad \gamma < L,$$

we can choose a suitable $\alpha > 0$ to satisfy Assumption 2.

Hence the inequality (3.4) turns into the following one:

$$\Delta(0) \leq \frac{8L}{\alpha} \sup_{B \in \mathcal{B}_{[0,L]}} |F(B) - \widetilde{F}(B)|. \tag{3.6}$$

Remark 3.2. The direct calculations show that if $F = U[0,1], \widetilde{F} = U[0, 1-\varepsilon]$, then the right-hand side of (3.6) is $\frac{8}{\alpha}\,\varepsilon$ while $\Delta(0) = o(\varepsilon)$ as $\varepsilon \to 0$. The arguments given in Remark 2.2 (b) of Section 2 refer to this particular case.

To consider the second particular case, let $\xi_1, \xi_2, \ldots$ be i.i.d. random variables with the density f satisfying on $[0, L]$ the Lipschitz condition with the constant ρ. Let $\widetilde{x}_t := \xi_t,\ t \geq 1$, while the “real” process is defined by the equations: $x_t = \varepsilon x_{t-1} + (1-\varepsilon)\xi_t,\ t \geq 1$, where $\varepsilon \in (0, 1/2)$. It is easy to check that the ergodicity condition (2.7) is fulfilled with $\delta=\varepsilon$, $M=4L^2\rho$ and π being the distribution of the random variable $(1-\varepsilon)\sum_{k=1}^{\infty}\varepsilon^{k-1}\xi_k$. For $c_0>0$ Assumption 2 is satisfied under certain mild restrictions on f.

Denoting by $\widetilde{F}$ the distribution function of $\widetilde{x}_1=\xi_1$, we have:

$$\sup_{x\in[0,L]}\|p(\cdot|\,x)-\widetilde{F}(\cdot)\|=\sup_{x\in[0,L]}\int_0^L\left|\frac{1}{1-\varepsilon}f\left(\frac{y-\varepsilon x}{1-\varepsilon}\right)-f(y)\right|\mathrm{d}y\le 2\varepsilon(1+3\rho L^2).$$

Thus inequality (3.4) provides the following stability bound: $\Delta(0)\le K\varepsilon$, where K is an explicitly calculated constant.

3.2. Counterexamples

Example 2. In Example 1 with $L=1$, let $\varepsilon\in(0,1)$ be a small enough number, $\widetilde{x}_1,\widetilde{x}_2,\dots$ be i.i.d. random variables with the distribution $\widetilde{F}=U[0,1]$ (the uniform distribution on the segment $[0,1]$) and, finally, let $x_1,x_2,\dots$ be i.i.d. random variables with the distribution $F=U[0,1-\varepsilon+\varepsilon^2]$. We choose also $c_0=\varepsilon^2/2$.

The solution to equation (3.2) is $\widetilde{x}_*=1-\varepsilon$. Thus the optimal for $\{\widetilde{x}_t\}$ stopping rule $\widetilde{\tau}_*$ is to stop at first instant t when $\widetilde{x}_t\ge 1-\varepsilon$. Applying this rule to the "real" process $\{x_t\}$, we find that (see (1.2))

$$W(0,\widetilde{\tau}_*)=\mathrm{E}_0[c_0\widetilde{\tau}_*-x_{\widetilde{\tau}_*}]\to-\frac{1}{2}\tag{3.7}$$

as $\varepsilon\to 0$.

This is due to the fact that for $x_0=0$, under the probability P_0, the random variable $\widetilde{\tau}_*$ is geometric with the parameter equal to $\varepsilon^2=2c_0$.

On the other hand, resolving equation (3.5) we get $x_*=1-\delta-\sqrt{2c_0}\sqrt{1-\delta}$, $\delta=\varepsilon-\varepsilon^2$, or $1-\delta-x_*=\varepsilon\sqrt{1-\varepsilon+\varepsilon^2}\sim\varepsilon=\sqrt{2c_0}$ as $\varepsilon\to 0$.

Hence the stopping time τ_* has the geometric distribution with parameter of order $\sqrt{2c_0}$ (under the probability P_0). Taking into account the expression for W given in (3.7), we find that $W(0,\tau_*)\to -1$ as $\varepsilon\to 0$, and the stability index $\Delta(0)$ in (1.6) is *greater than* $\frac{1}{3}$ *for all small enough* $\varepsilon>0$. On the other hand,

$$\sup_{x\in[0,1]}\|p(\cdot|x)-\widetilde{p}(\cdot|x)\|=2\sup_{B\in\mathcal{B}_{[0,1]}}|F(B)-\widetilde{F}(B)|\to 0\qquad\text{as}\qquad\varepsilon\to 0.$$

Note that Assumption 2 does not hold (while Assumption 1 does). In this counterexample we have chosen the cost function c_0 dependent on "the proximity parameter" ε. Actually, this changes the original problem setting. The next example is free from such shortcoming.

Example 3. Let $X=\{0,1,2,3\}$, $x_0=0$ be the initial state, $\varepsilon\in(0,1)$ and the transition probability of the Markov chain $\{x_t\}$ be defined as follows: $p(1|0)=p(2|1)=1$, $p(2|2)=1-\varepsilon$, $p(3|2)=\varepsilon$, $p(3|3)=1$. The "approximating chain" $\{\widetilde{x}_t\}$ is defined by the transition probabilities:

$$\widetilde{p}(1|0)=\widetilde{p}(2|1)=\widetilde{p}(2|2)=\widetilde{p}(3|3)=1.$$

The functions c_0 and r are such that $c_0(1) = 1$, $c_0(x) = 0$ for $x \neq 1$ and $r(3) = 3$, $r(x) = 0$ for $x \neq 3$.

Under the transition probabilities p the state "3" is accessible from the state "0" for a finite, with probability 1, time. Moreover, on the "route" from "0" to "3" the "penalized" state "1" is visited only once.

Since $c_0(1) = 1 < r(3) = 3$, the optimal rule τ_* is to stop at "3" with the optimal (minimal) cost $W_*(0) = W(0, \tau_*) = 1 - 3 = -2$. On the other hand, as far as "3" is not attainable from "0" under the transition probability $\widetilde{p}$, and $r(x) = 0$ for $x \neq 3$, the rule $\widetilde{\tau}_*$, optimal for the approximating model, is to stop at $t = 0$.

Consequently, $W(0, \widetilde{\tau}_*) = 0$, and $\Delta(0) = W(0, \widetilde{\tau}_*) - W_*(0) = 2$ for all $\varepsilon > 0$. Along with this we see that

$$\max_{x \in X} \|p(\cdot|x) - \widetilde{p}(\cdot|x)\| \to 0 \qquad \text{as} \qquad \varepsilon \to 0.$$

Note that in this example the "approximating chain" contains two recurrent classes and condition (2.8) is violated.

Example 4. Modifying Example 2, let $c_0 = 0$, $x_1, x_2 \dots$ be i.i.d. random variables with the uniform distribution $F = U[0, 1]$ and, for every $\varepsilon \in (0, 1)$, $\widetilde{x}_1, \widetilde{x}_2, \dots$ be i.i.d. random variables distributed as follows (with the distribution function denoted by $\widetilde{F}$):

$$\widetilde{x}_1 = \begin{cases} x_1 \text{ with probability } 1 - \varepsilon, \\ 1 \text{ with probability } \varepsilon. \end{cases}$$

We fix the initial states $x_0 = \widetilde{x}_0 = 0$. It is easy to see that $W_*(x) = \widetilde{W}_*(x) = -1$, $x \in X = [0, 1]$, and in (2.3), (2.4) $S = \widetilde{S} = \{1\}$ (since $r(x) = x$). Thus (2.5), (2.6) define the stopping rules τ_* and $\widetilde{\tau}_*$ that order to stop on the first passage at the point 1. Surely, $\widetilde{\tau}_* \in \widetilde{\mathcal{T}}_0$, and it is optimal for $\{\widetilde{x}_t\}$ ($\widetilde{W}(0, \widetilde{\tau}_*) = -1$). On the other hand, $\tau_*, \widetilde{\tau}_* \notin \mathcal{T}_0$ (and they take the value $+\infty$ with P_0-probability 1), $W(0, \tau_*) = 0 \neq W_*(0)$ (supposing that in (1.2) $r(x_\tau) := 0$ if $\tau = \infty$), and there does not exist an optimal stopping rule for the process $\{x_t\}$. However, for every $\delta > 0$ there is a rule τ_δ such that $W(0, \tau_\delta) = -1 + \delta$. In spite of the fact that

$$\sup_{B \in \mathcal{B}_{[0,1]}} |F(B) - \widetilde{F}(B)| \to 0 \qquad \text{as} \qquad \varepsilon \to 0,$$

the usage of $\widetilde{\tau}_*$ as "an approximation" to τ_δ is senseless, since $W(0, \widetilde{\tau}_*) = 0$.

Remark 3.3. To prove the stability inequality in the next section, we reduce the optimal stopping problem to the problem of minimization of the total expected cost (nonnegative one) for a "derived" Markov control process. Example 4 is given to show that, in general, such reduction does not provide an equivalent problem. Indeed, as it is easily seen from the definitions given in Section 4, the stationary policy

$$f(x) = \begin{cases} \text{to stop if } x = 1, \\ \text{to continue the process if } x \neq 1 \end{cases}$$

is *optimal* for the "derived" Markov control process which corresponds to $\{x_t\}$. But for the original stopping rule optimization problem this policy is the worst one.

4. THE PROOFS

In order to prove Theorem 2.1, we first pass to an equivalent (under Assumptions 1 and 2) problem of minimization of the total expected (nonnegative) cost for the Markov control process $\{z_t\}$ described below. Such a transformation is standard. On the other side, the key step in the proof of inequality (2.11) is showing certain contractive properties of the operators defined below in (4.14), (4.15).

The control process $\{z_t\}$ is specified by $(Z, A, A(z), q, c)$, where:

- $Z := X \cup \{\infty\}$ is the state space, and "∞" denotes an absorbing state where the process "lives" once it has been stopped;
- $A = \{1, 2\}$ is the action space, the action "1" prescribes to stop the process, while the action "2" orders to continue observations of the trajectory of $\{x_t\}$;
- $A(z) \equiv A$ is the set of admissible actions at a state $z \in Z$;
- $q(D|z, a), D \in \mathcal{B}_Z,\ z \in Z,\ a \in A(z)$, is the transition probability of the control process $\{z_t\}$ defined as follows:

$$q(D|z,2) := \begin{cases} p(D'|z) & \text{if} \quad z \in X, \\ 1 & \text{if} \quad z = \infty \ \text{ and } \ \infty \in D, \\ 0 & \text{if} \quad z = \infty \ \text{ and } \ \infty \notin D, \end{cases} \tag{4.1}$$

  where $D' = D \backslash \{\infty\}$;

$$q(D|z,1) := \begin{cases} 1 & \text{if} \quad \infty \in D,\ z \in Z, \\ 0 & \text{if} \quad \infty \notin D,\ z \in Z; \end{cases} \tag{4.2}$$

- $c : Z \times A \to [0, \infty)$ is the one-stage cost function defined as:

$$c(z,2) := \begin{cases} c_0(z) & \text{if} \quad z \in X, \\ 0 & \text{if} \quad z = \infty; \end{cases} \tag{4.3}$$

$$c(z,1) := \begin{cases} \overline{r} - r(z) & \text{if} \quad z \in X, \\ 0 & \text{if} \quad z = \infty \end{cases} \tag{4.4}$$

(the number $\overline{r}$ was defined in (1.1)).

This definition means that z-component of the trajectory $\{(z_t, a_t),\ t = 0, 1, 2, \ldots\}$ coincides with $\{x_t,\ t = 0, 1, 2, \ldots\}$ (with the payments $c_0(x_t)$) until the first application of the "stopping action" $a = 1$. At such instant τ the nonnegative "payoff" $\overline{r} - r(x_\tau)$ is made, and the process moves to the absorbing state "∞".

Similarly, replacing p by $\widetilde{p}$, we define the "approximating" Markov control process $(Z, A, A(z), \widetilde{q}, c)$. The trajectories of this process we denote by $\{(\widetilde{z}_t, \widetilde{a}_t), t = 0, 1, 2, \ldots\}$.

For a given initial state $z \in Z$ and a control policy π (see, e. g. [3, 6, 12] for the definition), for the both processes, the total expected costs are defined as follows:

$$V(z, \pi) := \mathrm{E}_z^{\pi} \left[\sum_{t=0}^{\infty} c(z_t, a_t) \right], \tag{4.5}$$

$$\widetilde{V}(z, \pi) := \widetilde{\mathrm{E}}_z^{\pi} \left[\sum_{t=0}^{\infty} c(\widetilde{z}_t, \widetilde{a}_t) \right]. \tag{4.6}$$

The corresponding value functions

$$V_*(z) := \inf_{\pi \in \sqcap} V(z, \pi), \quad \widetilde{V}_*(z) := \inf_{\pi \in \sqcap} \widetilde{V}(z, \pi), \quad z \in Z,$$

are finite and nonnegative. Here $\sqcap$ denotes the class of all control policies.

Since $c \geq 0$, the following optimality equation takes place (see, for instance, [2, 19]):

$$V_*(z) = \inf_{a \in A} \left\{ c(z, a) + \int_Z V_*(y)\, q(\mathrm{d}y | z, a) \right\}, \quad z \in Z, \tag{4.7}$$

or, in view of (4.1), (4.4), and, because of $V_*(\infty) = 0$,

$$V_*(x) = \min \left\{ \overline{r} - r(x), c_0(x) + \int_X V_*(y)\, p(\mathrm{d}y | x) \right\} \tag{4.8}$$

for $z = x \in X$.

Comparing (2.1) with (4.8), we see that $V_*(x) = W_*(x) + \overline{r}$, $x \in X$, and that the stationary policy

$$f_*(z) := \begin{cases} 1 & \text{if} \quad z \in S \\ 2 & \text{if} \quad z \in X \backslash S \\ 1 & \text{if} \quad z = \infty \end{cases} \tag{4.9}$$

minimizes the right-hand side of (4.7). Therefore (see [19]), the policy f_* is V-optimal, i. e. $V(z, f_*) = V_*(z)$, $z \in Z$. Meanwhile, if we choose $z = x_0 \in X$, then the policy f_* in (4.9) generates the stopping rule τ_* defined in (2.5). Moreover, (2.7) and (2.10) ensure that the stopping time τ_* is P_{x_0}-almost surely finite, i. e. $\tau_* \in \mathcal{T}_{x_0}$ (see, e. g. [14], § 16.2). Collating (4.5) with (1.2), we get that

$$V(x_0, f_*) = \mathrm{E}_{x_0}^{f_*} \left[\sum_{t=0}^{\tau_* - 1} c_0(x_t) + \overline{r} - r(x_{\tau_*}) \right] = \overline{r} + W(x_0, \tau_*). \tag{4.10}$$

Therefore the stopping rule given in (2.5) is W-optimal, that proves the corresponding part of Lemma 2.1.

Similarly we define the stationary policy $\widetilde{f}_*$ for the process $\{\widetilde{z}_t\}$ (see (2.4) for the definition of the set $\widetilde{S}$):

$$\widetilde{f}_*(z) := \begin{cases} 1 & \text{if} \quad z \in \widetilde{S} \\ 2 & \text{if} \quad z \in X \setminus \widetilde{S} \\ 1 & \text{if} \quad z = \infty. \end{cases} \tag{4.11}$$

Using the same arguments we find that the policy $\widetilde{f}_*$ is $\widetilde{V}$-optimal, and that the stopping rule, defined in (2.6), is generated by $\widetilde{f}_*$. Moreover, in view of Assumptions 1 and 2, $\widetilde{\tau}_* \in \widetilde{\mathcal{T}}_{x_0}$, and, by the equality analogous to (4.10) (see (1.5)), the stopping rule $\widetilde{\tau}_*$ is $\widetilde{W}$-optimal.

On the other hand, by (2.7) and (2.10) we obtain that $\widetilde{\tau}_* \in \mathcal{T}_{x_0}$ (that completes the proof of Lemma 2.1), and, in view of (4.3), (4.4) and (4.5), we get that

$$V(x_0, \widetilde{f}_*) = \mathrm{E}_{x_0}^{\widetilde{f}_*} \left[\sum_{t=0}^{\widetilde{\tau}_*-1} c_0(x_t) + \overline{r} - r(x_{\widetilde{\tau}_*}) \right]. \tag{4.12}$$

Comparing (4.10), (4.12) with (1.2) and (1.6) we conclude that the stability index in (1.6) can be rewritten in the following way:

$$\Delta(x_0) = V(x_0, \widetilde{f}_*) - V(x_0, f_*) \geq 0, \tag{4.13}$$

where $f_*, \widetilde{f}_*$ are the stationary policies that are optimal, respectively, for the control processes $\{z_t\}$ and $\{\widetilde{z}_t\}$.

Remark 4.1. Returning to Example 4 in the previous section, we can see that, for the corresponding to this example "derived" control process, $V_* \equiv 0$, and, since in (4.9) $S = \{1\}$, the application of the policy f_* means never stop (with probability 1). However, by (4.5) and (4.8), $V(z, f_*) = 0 = V_*(z)$, i. e. the policy f_* is V-optimal for the control process $\{z_t\}$. As we had noted early, in this example the stopping rule, defined by f_*, is the worst one among all possible stopping rules.

Let B be the Banach space of all bounded measurable functions $u : Z \to \mathbb{R}$ such that $u(\infty) = 0$. The space B is equipped with the uniform norm $\|u\| := \sup_{z \in Z} |u(z)| \equiv \sup_{x \in X} |u(x)|$. Let $\mathcal{F} = \{f_*, \widetilde{f}_*\}$, where the stationary policies f_* and $\widetilde{f}_*$ were defined in (4.9) and (4.11). For every $f \in \mathcal{F}$ we define the operators $T_f : B \to B$, $\widetilde{T}_f : B \to B$ by the formulas:

$$T_f u(z) := c(z, f(z)) + \int_Z u(y)\, q(\mathrm{d}y | f(z)), \tag{4.14}$$

$$\widetilde{T}_f u(z) := c(z, f(z)) + \int_Z u(y) \widetilde{q}(\mathrm{d}y | f(z)), \tag{4.15}$$

where $u \in B,\ z \in Z$. Since $c(\infty, a) = 0$ and $q(\{\infty\} | \infty, a) = 1$, $a \in A$, the operators T_f and $\widetilde{T}_f$ map B into B.

Lemma 4.1. Let Assumptions 1 and 2 hold, and N be the integer from (2.13). Then for every $f \in \mathcal{F}$,

$$\|T_f^N u - T_f^N v\| \le \left(1 - \frac{3}{4}\,\alpha\right)\ \|u - v\| \tag{4.16}$$

and

$$\|\widetilde{T}_f^N u - \widetilde{T}_f^N v\| \le \left(1 - \frac{3}{4}\,\alpha\right)\ \|u - v\| \tag{4.17}$$

for all $u, v \in B$. The constant $\alpha > 0$ was specified in Assumption 2.

P r o o f. Let us prove, for example, (4.16) for $f = \widetilde{f}_*$. We denote by $\{z_t\}$ the Markov process on Z with the transition probability $q(\cdot|z) \equiv q(\cdot|z, \widetilde{f}_*(z))$, $z \in Z$, where $q(\cdot|z, a)$ was defined in (4.1) and (4.2), and $\widetilde{f}_*$ is the stationary policy from (4.11). By the Markov property and by (4.14), we get that for every integer $n \ge 1$

$$T_f^n u(z) = R_{f,n}c(z, f(z)) + \int_Z u(y)\, q^n\big(\mathrm{d}y|z, f(z)\big), \tag{4.18}$$

$z \in Z$, $u \in B$, where the function $R_{f,n}c$ does not involve u, and q^n is the n-step transition probability of $\{z_t\}$.

Since $u(\infty) = v(\infty) = 0$, we can write:

$$\begin{aligned} \|T_f^n u - T_f^n v\| &= \sup_{x\in X}\left|\int_Z u(y)\, q^n\big(\mathrm{d}y|x, f(x)\big) - \int_Z v(y)\, q^n\big(\mathrm{d}y|x, f(x)\big)\right| \\ &= \sup_{x\in X}\left|\int_X \big[u(y) - v(y)\big] q^n\big(\mathrm{d}y|x, f(x)\big)\right| \\ &\le \|u - v\| \sup_{x\in X} q^n\big(X|x, f(x)\big). \end{aligned} \tag{4.19}$$

Choosing N as in (2.13), we have for $m = N - 1$ that $M\delta^m \le \frac{\alpha}{2}$, and, in view of (2.7) and (2.9),

$$\sup_{x\in X}\sup_{B\in\mathcal{B}_X} \big|p^m(B|x) - \pi(B)\big| \le \frac{\alpha}{4}. \tag{4.20}$$

By (4.20) and (2.10), for all $x \in X$,

$$p^m(\widetilde{S}|x) \ge \frac{3\alpha}{4}, \tag{4.21}$$

where the set $\widetilde{S}$ was defined in (2.4) and was used in (4.11) to define the policy $f = \widetilde{f}_*$.

For every $x \in \widetilde{S}$ $f(x) = 1$ and, by (4.2), $q\big(\{\infty\}|x, 1\big) = 1$. Therefore, we have the following inequality (recall that "∞" is an absorbing state):

$$q^{m+1}\big(\{\infty\}|x, f(x)\big) \ge p^m(\widetilde{S}|x),\ x \in X.$$

Thus from (4.21) it follows that

$$q^N\big(\{\infty\}|x, f(x)\big) \geq \frac{3\alpha}{4}, \quad x \in X.$$

Finally, $q^N\big(X|x, f(x)\big) = 1 - q^N\big(\{\infty\}|x, f(x)\big)$, and it is enough to take $n = N$ in (4.19). □

P r o o f o f T h e o r e m 2.1. It is easy to show (see [8]) that for the stability index given in (4.13) the following inequality holds true:

$$\Delta(x_0) \leq 2 \max_{f \in \mathcal{F}} \left| V(x_0, f) - \widetilde{V}(x_0, f) \right|, \tag{4.22}$$

where V and $\widetilde{V}$ are defined, respectively, in (4.5) and (4.6).

Note that for every $z \in Z$

$$0 \leq V(z, f_*) = V_*(z) \leq \overline{r}, \tag{4.23}$$
$$0 \leq \widetilde{V}(z, \widetilde{f}_*) = \widetilde{V}_*(z) \leq \overline{r},$$

because of $V(z, \tau_0) = \widetilde{V}(z, \tau_0) = \overline{r} - r(x_0)$ for the stopping rule $\tau_0 \equiv 0$. We will also show that $V(\cdot, \widetilde{f}_*), \widetilde{V}(\cdot, f_*) \in B$. By (4.3) and (4.4), $0 \leq c(z, a) \leq \max\{\overline{c}, \overline{r}\}$. In view of (4.6) the boundedness, for example, of the function $\widetilde{V}(\cdot, f_*)$ would follow if

$$\sup_{x \in X} \widetilde{\mathrm{E}}_x \tau_*(x) < \infty, \tag{4.24}$$

where $\tau_*(x) := \inf\{t \geq 0 : \widetilde{x}_t \in S\}$ given that $\widetilde{x}_0 = x$. Under ergodicity conditions (2.8) and (2.10), inequality (4.24) follows from Theorem 16.2.2 in [14], § 16.2.

For any stationary policy f (particularly, for $f \in \mathcal{F}$) the expected total costs satisfy the equations (see, e. g. [3], § 9.4):

$$V_f = T_f V_f, \qquad \widetilde{V}_f = \widetilde{T}_f \widetilde{V}_f, \tag{4.25}$$

where $V_f(\cdot) := V(\cdot, f), \widetilde{V}_f(\cdot) := \widetilde{V}(\cdot, f)$, and the operators $T_f, \widetilde{T}_f$ were defined in (4.14), (4.15). Let us fix, for example, $f = f_*$ in the term under the sign of maximum in (4.22). By (4.25),

$$\begin{aligned} &\big|V(x_0, f_*) - \widetilde{V}(x_0, f_*)\big| = \big|\widetilde{T}^N_{f_*} \widetilde{V}_{f_*}(x_0) - T^N_{f_*} V_{f_*}(x_0)\big| \\ \leq \ &\big|\widetilde{T}^N_{f_*} \widetilde{V}_{f_*}(x_0) - \widetilde{T}^N_{f_*} V_{f_*}(x_0)\big| + \big|T^N_{f_*} V_{f_*}(x_0) - \widetilde{T}^N_{f_*} V_{f_*}(x_0)\big|. \end{aligned} \tag{4.26}$$

Simplifying the notation, we write:

$$f := f_*, \ V := V_f, \ \widetilde{V} := \widetilde{V}_f, \ T := T_f, \ \widetilde{T} := \widetilde{T}_f.$$

Then from (4.26),

$$\|V - \widetilde{V}\| \leq \|\widetilde{T}^N \widetilde{V} - \widetilde{T}^N V\| + \|T^N V - \widetilde{T}^N V\|,$$

and, by (4.17),

$$\|V - \widetilde{V}\| \leq \frac{4}{3\alpha} \|T^N V - \widetilde{T}^N V\|. \tag{4.27}$$

Applying the induction and the Fubini Theorem, we obtain from (4.14) and (4.15) the following specification of (4.18):

$$T^N V(z) = c\big(z, f(z)\big) + \int_Z c\big(y, f(y)\big) q\big(\mathrm{d}y|z, f(z)\big) + \int_Z c\big(y, f(y)\big) q^2\big(\mathrm{d}y|z, f(z)\big)$$
$$+ \cdots + \int_Z c\big(y, f(y)\big) q^{N-1}\big(\mathrm{d}y|z, f(z)\big) + \int_Z V(y)\, q^N\big(\mathrm{d}y|z, f(z)\big), \quad z \in Z,$$

and a similar expression for $\widetilde{T}^N V$.

Thus, from (4.23) and the fact that $0 \leq c\big(z, f(z)\big) \leq \widetilde{c} := \max\{\bar{c}, \bar{r}\}$, it follows that

$$\|T^N V - \widetilde{T}^N V\| \leq \widetilde{c} \sum_{n=1}^{N} \sup_{z \in Z} \sup_{\varphi \in B_1} \left| \int_Z \varphi(y) \left[q^n \left(\mathrm{d}y|z, f(z)\right) - \widetilde{q}^n \left(\mathrm{d}y|z, f(z)\right)\right] \right|, \tag{4.28}$$

where $B_1 := \{\varphi \in B : \ 0 \leq \varphi \leq 1\}$.

Denoting by $Q_n(f)$ the nth summand on the right-hand side of (4.28), we will verify by induction that for $n = 1, 2, \ldots$

$$\begin{aligned} Q_n(f) \quad &\leq \quad n \sup_{z \in Z} \sup_{\varphi \in B_1} \left| \int_Z \varphi(y)\, q(\mathrm{d}y|z, f(z)) - \int_Z \varphi(y) \widetilde{q}(\mathrm{d}y|z, f(z)) \right| \\ &\leq \quad \frac{n}{2} \sup_{x \in X} \|p(\cdot|x) - \widetilde{p}(\cdot|x)\|. \end{aligned} \tag{4.29}$$

The last inequality in (4.29) is true because of the following facts:

1) $\varphi(\infty) = 0$, so we can replace $\int_Z$ by $\int_X$.

2) For any $a \in A$ $q(\{\infty\}|\infty, a) = \widetilde{q}(\{\infty\}|\infty, a) = 1$, hence instead of $\sup_{z \in Z}$ we can use $\sup_{x \in X}$.

3) For every $x \in X$ if $f(x) = 1$, then $q(\{\infty\}|x, f(x)) = \widetilde{q}(\{\infty\}|x, f(x)) = 1$, and for such x the difference between the integrals in the intermediate term in (4.29) is equal to zero.

4) By (4.1) (and, analogously, for $\widetilde{q}$)

$$q(D|x, 2) = p(D|x), \qquad \widetilde{q}(D|x, 2) = \widetilde{p}(D|x)$$

for every $x \in X$, $D \in \mathcal{B}_X$.

5) From (2.9) it easily follows that

$$\sup_{0 \leq \varphi \leq 1} \left| \int_X \varphi(x)\, \mathrm{d}\mu - \int_X \varphi(x)\, \mathrm{d}\nu \right| = \frac{1}{2} \|\mu - \nu\|.$$

For $n=1$ the first inequality in (4.29) holds trivially. Suppose that it is true for some $n \geq 1$. Then for any $z \in Z$, $\varphi \in B_1$ we have:

$$\begin{aligned}\delta_n &= \left|\int_Z \varphi(y)\, q^{n+1}\,(\mathrm{d}y|z, f(z)) - \int_Z \varphi(y)\widetilde{q}^{n+1}\,(\mathrm{d}y|z, f(z))\right| \\ &= \left|\int_Z \varphi_1(u)\, q^n(\mathrm{d}u|z, f(z)) - \int_Z \varphi_2(u)\widetilde{q}^n(\mathrm{d}u|z, f(z))\right|,\end{aligned}$$

where

$$\begin{aligned}\varphi_1(u) &= \int_Z \varphi(y)\, q(\mathrm{d}y|u, f(u)), \\ \varphi_2(u) &= \int_Z \varphi(y)\widetilde{q}(\mathrm{d}y|u, f(u)),\ u \in Z,\end{aligned}$$

are the functions which belong to B_1 (due to the above point 2). Thus,

$$\begin{aligned}\delta_n &\leq \int_Z |\varphi_1(u) - \varphi_2(u)| q^n(\mathrm{d}u|z, f(z)) \\ &+ \left|\int_Z \varphi_2(u)\,[q^n\,(\mathrm{d}u|z, f(z)) - \widetilde{q}^n\,(\mathrm{d}u|z, f(z))]\right| \\ &\leq \sup_{u \in Z}\left|\int_Z \varphi(y)\,[q\,(\mathrm{d}y|u, f(u)) - \widetilde{q}\,(\mathrm{d}y|u, f(u))]\right| \\ &+ n \sup_{z \in Z}\sup_{\varphi \in B_1}\left|\int_Z \varphi(y)\,[q\,(\mathrm{d}y|z, f(z)) - \widetilde{q}\,(\mathrm{d}y|z, f(z))]\right|.\end{aligned}$$

From this inequality we obtain the inequality in (4.29) for $n+1$.

Combining the inequalities in (4.26), (4.27), (4.28) and (4.29), we obtain the following bound:

$$|V(x_0, f_*) - \widetilde{V}(x_0, f_*)| \leq \frac{\widetilde{c}}{3\alpha}\, N(N+1) \sup_{x \in X} \|p(\cdot|x) - \widetilde{p}(\cdot|x)\|.$$

By similar arguments we get the same upper bound for $|V(x_0, \widetilde{f}_*) - \widetilde{V}(x_0, \widetilde{f}_*)|$. Taking into account (4.22) we finally obtain the desired stability inequality (2.11).□

ACKNOWLEDGEMENT

The author is grateful to the referees for their valuable suggestions which improved the text of the paper.

This research was partially supported by SEP–CONACYT under grant CB-2005-C01-49854-F.

(Received September 20, 2007.)

REFERENCES

[1] P. Allart: Optimal stopping rules for correlated random walks with a discount. J. Appl. Prob. *41* (2004), 483–496.

[2] D. P. Bertsekas: Dynamic Programming: Deterministic and Stochastic Models. Prentice Hall, Englewood Cliffs, N. J. 1987.

[3] D. P. Bertsekas and S. E. Shreve: Stochastic Optimal Control: The Discrete Time Case. Academic Press, New York 1979.

[4] N. M. Van Dijk: Perturbation theory for unbounded Markov reward process with applications to queueing systems. Adv. in Appl. Probab. *20* (1988), 99–111.

[5] N. M. Van Dijk and K. Sladký: Error bounds for nonnegative dynamic models. J. Optim. Theory Appl. *101* (1999), 449–474.

[6] E. B. Dynkin and A. A. Yushkevich: Controlled Markov Process. Springer-Verlag, New York 1979.

[7] G. Favero and W. J. Runggaldier: A robustness results for stochastic control. Systems Control Lett. *46* (2002), 91–97.

[8] E. I. Gordienko: An estimate of the stability of optimal control of certain stochastic and deterministic systems. J. Soviet Math. *59* (1992), 891–899. (Translated from the Russian publication of 1989).

[9] E. I. Gordienko and F. S. Salem: Robustness inequality for Markov control process with unbounded costs. Systems Control Lett. *33* (1998), 125–130.

[10] E. I. Gordienko and A. A. Yushkevich: Stability estimates in the problem of average optimal switching of a Markov chain. Math. Methods Oper. Res. *57* (2003), 345–365.

[11] E. I. Gordienko, E. Lemus-Rodríguez and R. Montes-de-Oca: Discounted cost optimality problem: stability with respect to weak metrics. In press in: Math. Methhods Oper. Res. (2008).

[12] O. Hernández-Lerma and J. B. Lassere: Discrete-Time Markov Control Processes: Basic Optimality Criteria. Springer-Verlag, N.Y. 1996.

[13] U. Jensen: An optimal stopping problem in risk theory. Scand. Actuarial *J.2* (1997), 149–159.

[14] S. P. Meyn and R. L. Tweedie: Markov Chains and Stochastic Stability. Springer-Verlag, London 1993.

[15] R. Montes-de-Oca and F. Salem-Silva: Estimates for perturbations of an average Markov decision process with a minimal state and upper bounded by stochastically ordered Markov chains. Kybernetika *41* (2005), 757–772.

[16] R. Montes-de-Oca, A. Sakhanenko, and F. Salem-Silva: Estimate for perturbations of general discounted Markov control chains. Appl. Math. *30* (2003), 287–304.

[17] B. K. Muciek: Optimal stopping of a risk process: model with interest rates. J. Appl. Prob. *39* (2002), 261–270.

[18] A. Müller: How does the value function of a Markov decision process depend on the transition probabilities? Math. Oper. Res. *22* (1997), 872–885.

[19] M. Schäl: Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. Z. Wahrsch. verw. Gebiete *32* (1975), 179–196.

[20] A. N. Shiryaev: Optimal Stopping Rules. Springer-Verlag, New York 1978.

[21] A. N. Shiryaev: Essential of Stochastic Finance. Facts, Models, Theory. World Scientific Publishing Co., Inc., River Edge, N.J. 1999.

Elena Zaitseva, Universidad Autónoma Metropolitana — Unidad Iztapalapa, Av. San Rafael Atlixco # 186, Colonia Vicentina, México, D.F. 09340. México.
e-mail: lenagordi@hotmail.com