Sample-Path and Variance Minimization of Markov Control Processes with Average Cost Criteria[1]

Onésimo Hernández-Lerma[2] Oscar Vega-Amaya[3] Guadalupe Carrasco[4]

Abstract

This paper studies several average cost criteria for Markov control processes on *Borel* spaces, with possibly *unbounded costs.* Under suitable hypotheses it is shown: (i) the existence of a *sample path average cost* (SPAC-) optimal stationary policy; (ii) a stationary policy is SPAC-optimal if and only if it is *expected average cost* (EAC-) optimal; and (iii) within the class of stationary SPAC-optimal (equivalently EAC-optimal) there exists one with minimal limiting *average variance.*

Key words.(discrete-time) Markov control processes, average cost criteria, sample-path average cost, expected average cost, canonical policies, average variance.

AMS subject classification. 93E20, 90C40.

1 Introduction

There is a huge literature dealing with discrete-time Markov control processes (MCPs) with *average cost* (AC) criteria (see, for instance, [1, 3, 10, 11, 22] and their extensive bibliographies). However, most of the related papers study the *expected average cost* (EAC) criterion, specially, the denumerable state space case or the bounded costs case. In contrast, for the *sample-path average cost* (SPAC) criterion there are a lot fewer works, among which, we can mention the pioneering works by Mandl [18, 19, 20] for the *finite* state MCPs, the works by Borkar [4] and Cavazos-Cadena and Fernández-Gaucherand [5] for the *countable* state case; whereas for MCPs on *Borel* spaces we only know of [1, Theorem 6.3(v), (vi)], [2] for bounded costs and [17, 23, 24] for unbounded costs. Finally, to the best of our knowledge, for the variance minimization the only previous work are those by Mandl [18, 19, 20] for *finite* state MCPs, and by Kurano [16] for the Borel case with *bounded costs.* The reader should be warned, however, that several authors have studied a "variance minimization problem" for a quantity which is not the "real" variance—see, for instance, [14, 22]. Here, our definition of limiting average variance (10) is the one used in the *Central Limit Theorem for Markov Chains*, as in [6, Corollaire 7.III.2] or [21, Theorem 17.0.1].

In the present paper, we consider MCPs on *Borel* spaces with possibly *unbounded costs* and show, under suitable conditions, the following results: (i) the existence of a SPAC-optimal stationary policy; (ii) a stationary policy is SPAC-optimal if and only if it is EAC-optimal; (iii) within the class of SPAC-optimal stationary policies there exists one with minimal limiting average variance.

2 Average cost criteria

We shall consider the usual discrete-time, stationary, Markov control model $(\mathbf{X}, \mathbf{A}, \{A(x) : x \in \mathbf{X}\}, Q, C)$ with state space $\mathbf{X}$, control or action space $\mathbf{A}$—both sets being Borel spaces—, transition law Q on $\mathbf{X}$ given $\mathbf{K}$ and the one-step cost function $C : \mathbf{K} \to \mathbf{R}$, where $\mathbf{K} := \{(x, a) : x \in \mathbf{X}, a \in A(x)\}$ is the admissible state-action set. The class of all admissible control policies $\pi = \{\pi_n\}$ is denoted by Π, whereas $\mathbf{F}$ stands for the class of all stationary (deterministic) polices. Following a usual convention, $\mathbf{F}$ will be identified with the class of measurable mappings $f : \mathbf{X} \to \mathbf{A}$ satisfying the constraint $f(x) \in A(x)$, $\forall x \in \mathbf{X}$. As it is well-known, for each policy $\pi \in \Pi$ and "initial state" $x \in \mathbf{X}$, there exists a probability measure P_x^π on the canonical measurable space $(\Omega, \mathcal{F})$ which governs the evolution of the state-action processes $\{(x_n, a_n)\}$. For a full description of the Markov control model see, for instance, [1, 3, 10, 11].

For notational ease, for a stationary policy $f \in \mathbf{F}$ we write

$$C_f(x) := C(x, f(x)) \text{ and } Q_f(\cdot|x) := Q(\cdot|x, f(x)). \quad (1)$$

[1]This research was partially supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT, México) grants 3115P-E9608 and 28309E.

[2]Departamento de Matemáticas, CINVESTAV-IPN, Apartado Postal 14-740, México D.F. 07000, México. **E-mail:** ohernand@math.cinvestav.mx

[3]Departamento de Matemáticas, Universidad de Sonora, Blvd. Transversal y Rosales s/n, Hermosillo, Sonora, México. **E-mail:** ovega@gauss.mat.uson.mx

[4]Departamento de Matemáticas, Facultad de Ciencias, UNAM, Ciudad Universitaria, México D.F. 04510, México. **E-mail:** gcl@hp.fciencias.unam.mx

For $n = 1, 2, \cdots$, let

$$J_n^0(\pi, x) := \sum_{t=0}^{n-1} C(x_t, a_t) \tag{2}$$

be the *n-stage sample path cost* when using the policy π, given the initial state $x \in \mathbf{X}$. Then, the long-run *sample average cost* (SPAC) is defined as

$$J_0(\pi, x) := \limsup_{n\to\infty} \frac{1}{n} J_n^0(\pi, x). \tag{3}$$

Definition 2.1. A policy $\pi^* \in \Pi$ is said to be *sample path average cost optimal* (briefly, SPAC-optimal) if there exists a constant $\widehat{\rho}$ such that

$$J^0(\pi^*, x) = \widehat{\rho} \quad P_x^{\pi^*} - \text{a.s.} \ \ \forall x \in \mathbf{X},$$

and

$$J_0(\pi, x) \geq \widehat{\rho} \quad P_x^{\pi} - \text{a.s.} \ \ \forall x \in \mathbf{X}, \pi \in \Pi.$$

The constant $\widehat{\rho}$ is called the *optimal sample path average cost.*

The "expected" analogues of (2) and (3) are, respectively, the *n-stage expected cost*

$$J_n(\pi, x) := E_x^{\pi} \sum_{t=0}^{n-1} C(x_t, a_t), \tag{4}$$

and the long-run *expected average cost* (EAC)

$$J(\pi, x) := \limsup_{n\to\infty} \frac{1}{n} J_n(\pi, x). \tag{5}$$

Among the EAC-related optimality concepts we are interested in are the following.

Definition 2.2.(a) A policy $\pi^* \in \Pi$ is said to be *expected average cost* (EAC-) *optimal* if

$$J(\pi^*, x) = J^*(x) \quad \forall x \in \mathbf{X},$$

where

$$J^*(x) := \inf_{\pi\in\Pi} J(\pi, x) \quad x \in \mathbf{X},$$

is the *optimal expected average cost.*

(b) A stationary policy $f_* \in \mathbf{F}$ is called *canonical* if there exists a constant ρ_* and measurable function $h_* : \mathbf{X} \to \mathbf{R}$ such that for all $x \in \mathbf{X}$

$$\rho_* + h_*(x) = \min_{a\in A(x)} \left[C(x, a) + \int_{\mathbf{X}} h_*(y) Q(dy|x, a) \right], \tag{6}$$

and $f_*(x)$ attains the minimum on the right-hand side of (6) for every $x \in \mathbf{X}$; that is, using the notation in (1),

$$\rho_* + h_*(x) = C_{f_*}(x) + \int_{\mathbf{X}} h_*(y) Q_{f_*}(dy|x) \ \ \forall x \in \mathbf{X}. \tag{7}$$

If (6)-(7) are satisfied, it is said that (ρ_*, h_*, f_*) is a *canonical triplet* [25]. When (ρ_*, h_*, f_*) is a canonical triple and, additionally, the function h_* satisfy the condition

$$\lim_{n\to\infty} \frac{1}{n} E_x^{\pi} h_*(x_n) = 0 \ \ \forall x \in \mathbf{X}, \pi \in \Pi,$$

it is easy to see that f_* is EAC-optimal and also that ρ_* is the optimal expected average cost for every initial state $x \in \mathbf{X}$, that is,

$$J(f_*, x) = J^*(x) = \rho_* \ \ \forall x \in \mathbf{X}. \tag{8}$$

Thus, in this case, we have $\mathbf{F}_{cp} \subset \mathbf{F}_{eac}$, where $\mathbf{F}_{cp}$ is the class of canonical polices and $\mathbf{F}_{eac}$ is the class of stationary EAC-optimal polices.

In the following section we show, among other things, the existence of optimal stationary polices in the sense of Definitions 2.1 and 2.2, and moreover, the existence of a stationary policy that minimizes the limiting average variance in the class $\mathbf{F}_{eac}$. In other words, for each $f \in \mathbf{F}$ and $x \in \mathbf{X}$, consider the *n-stage variance*

$$V_n(f, x) := E_x^f \left[J_n(f, x) - J_n^0(f, x) \right]^2 \tag{9}$$

and the limiting *average variance*

$$V(f, x) := \limsup_{n\to\infty} \frac{1}{n} V_n(\pi, x). \tag{10}$$

Then, we shall prove that there exists a stationary policy $\widehat{f}$ which is EAC-optimal and

$$V(\widehat{f}, x) = \inf_{f\in\mathbf{F}_{eac}} V(f, x) \ \ \forall x \in \mathbf{X}. \tag{11}$$

3 Main results

We first require two sets of hypotheses. The first one, Assumption 3.1, is a combination of the usual continuity/compactness requirements (to ensure, for instance, the existence of "measurable minimizers") together with a growth condition on the one-step cost C.

Assumption 3.1. For each state $x \in \mathbf{X}$:

(a) $A(x)$ is a compact subset of $\mathbf{A}$;

(b) $C(x, \cdot)$ is lower semicontinuous on $A(x)$;

(c) $\int_{\mathbf{X}} u(y)Q(dy|x, \cdot)$ is continuous on $A(x)$ for every bounded measurable function u on $\mathbf{X}$;

(d) there exists a measurable function $W \geq 1$ on $\mathbf{X}$ and a constant r_1 such that

(d1) $|C(x,a)| \leq r_1 W(x) \quad \forall (x,a) \in \mathbf{K}$, and

(d2) $\int_{\mathbf{X}} W(y)Q(dy|x, \cdot)$ is continuous on $A(x)$.

The second set of hypotheses we need is to guarantee that the MCP has a nice "stable" behavior, uniformly in $\mathbf{F}$. Here, to fix ideas we shall impose Assumptions 3.2 and 3.4 below, which are an adaptation to MCPs of a Lyapunov-like condition used in Markov chain theory—see [7] or [21, p. 367]. However, the reader should keep in mind that, as noted in Remark 4.1, there are other hypotheses that yield the same stable behavior.

Assumption 3.2. For each stationary policy $f \in \mathbf{F}$:

(a) There exists positive constants $B_f < 1$ and $b_f < \infty$, and a petite subset K_f of $\mathbf{X}$ such that [using the notation (1)]

$$\int_{\mathbf{X}} W(y)Q_f(dy|x) \leq B_f W(x) + b_f \mathbf{I}_{K_f}(x) \quad \forall x \in \mathbf{X},$$

where $W \geq 1$ is the function in Assumption 3.1(d), and $\mathbf{I}_K(\cdot)$ denotes the indicator function of K;

(b) The state processes $\{x_n\}$—which under $f \in \mathbf{F}$ is a Markov chain with transition kernel Q_f [11, Proposition 2.3.5]—is φ-irreducible and aperiodic, for some σ-finite measure φ on $\mathbf{X}$.

To state some consequences of Assumption 3.2 let us first introduce the following notation: $B_W(\mathbf{X})$ denotes the normed linear space of measurable functions u on $\mathbf{X}$ with a finite W-*norm* $||u||_W$, which is defined as

$$||u||_W := \sup_{x \in \mathbf{X}} |u(x)|/W(x). \tag{12}$$

We shall write $\int_{\mathbf{X}} u(y)\mu(dy)$ as $\mu(u)$, i.e.,

$$\mu(u) := \int_{\mathbf{X}} u(y)\mu(dy).$$

Remark 3.3. [21, Theorem 16.0.1] Under Assumption 3.2, for each stationary policy $f \in \mathbf{F}$ we have:

(a) The Markov chain $\{x_n\}$ induced by f is positive Harris-recurrent and, moreover, its unique invariant probability measure μ_f is such that $\mu_f(W) < \infty$;

(b) $\{x_n\}$ is W*–geometrically ergodic*; that is, there exist positive constants $\gamma < 1$ and $M_f < \infty$ such that

$$\left| \int_{\mathbf{X}} u(y)Q_f^n(dy|x) - \mu_f(u) \right| \leq ||u||_W M_f \gamma_f^n W(x) \tag{13}$$

for each $u \in B_W(\mathbf{X}), x \in \mathbf{X}$, and $n = 0, 1, \cdots$.

The next assumption concerns the constants M_f and γ_f in (13).

Assumption 3.4. $M := \sup_f M_f$ and $\gamma := \sup_f \gamma_f$ are such that $M < \infty$ and $\gamma < 1$.

Assumptions 3.1, 3.2 and 3.4 have been used in [12] and [24] to study several undiscounted cost criteria, such as overtaking optimality, bias optimality, and others. In particular, the following result was established.

Remark 3.5. [12, Theorem 3.5], [24, Theorem 4.5.3]. Under Assumptions 3.1, 3.2 and 3.4, there exists a canonical triplet (ρ_*, h_*, f_*), where h_* is a function in $B_W(\mathbf{X})$ that satisfies (6)-(7); hence (8) holds.

In others words, we already have a canonical policy $f_* \in \mathbf{F}_{cp} \subset \mathbf{F}_{eac}$. It turns out that by suitably strengthening Assumption 3.1 we can also obtain SPAC-optimal policies with minimal average variance in $\mathbf{F}_{eac}$. Thus, consider:

Assumption 3.6. There exists a constant r_2 such that

$$C^2(x,a) \leq r_2 W(x) \qquad \forall (x,a) \in \mathbf{K}.$$

We can now state our first main result.

Theorem 3.7.([13]) Suppose that Assumptions 3.1, 3.2, 3.4 and 3.6 are satisfied, and let ρ_* be as in Remark 3.5. Then:

(a) For each $\pi \in \Pi$ and $x \in \mathbf{X}$

$$J^0(\pi, x) \geq \liminf_{n \to \infty} \frac{1}{n} J_n^0(\pi, x) \geq \rho_* \quad P_x^\pi - \text{a.s.}; \tag{14}$$

(b) A stationary policy is EAC-optimal if and only if it is SPAC-optimal; hence [by (14) and Remark 3.5] there exists a SPAC-optimal policy $f_* \in \mathbf{F}$ and ρ_* is the optimal sample path average cost , that is, $\rho_* = \widehat{\rho}$ where $\widehat{\rho}$ is the constant in Definition 2.1.

It is worth noting that Theorem 3.7(b) and the second inequality in (14) state that $f_* \in \mathbf{F}$ is in fact *strong SPAC-optimal*, where "strong " means that $f_* = \pi^*$ satisfies Definition 2.1 (with $\widehat{\rho} = \rho_*$) when the "lim-sup" SPAC in (3) is replaced by "lim-inf " SPAC

$$\underline{J}^0(\pi, x) := \liminf_{n\to\infty} \frac{1}{n} J_n^0(\pi, x).$$

On the other hand, denoting by $\mathbf{F}_{spac} \subset \mathbf{F}$ the class of SPAC-optimal stationary policies, we may rewrite the first statement in Theorem 3.7(b) as

$$\mathbf{F}_{eac} = \mathbf{F}_{spac}.$$

The reader should note that *a priori* neither one of the relations $\mathbf{F}_{eac} \subset \mathbf{F}_{spac}$ and $\mathbf{F}_{eac} \supset \mathbf{F}_{spac}$ is obvious!

To state our second main result we need some notation: For each $x \in \mathbf{X}$, let $A^*(x)$ be the subset of control actions $a \in A(x)$ that attain the minimum in (6); that is, $a \in A^*(x)$ if

$$\rho_* + h_*(x) = C(x, a) + \int_{\mathbf{X}} h_*(y) Q(dy|x, a)$$

Observe that, by (7), a policy $f \in \mathbf{F}$ is canonical ($f \in \mathbf{F}_{cp}$) if and only if $f(x) \in A^*(x)$ for all $x \in \mathbf{X}$. Moreover, consider the function Φ on $\mathbf{K}$ defined as

$$\Phi(x, a) := \int_{\mathbf{X}} h_*^2(y) Q(dy|x, a) - \left[\int_{\mathbf{X}} h_*(y) Q(dy|x, a)\right]^2.$$

As in (1), for $f \in \mathbf{F}$ and $x \in \mathbf{X}$ we write

$$\Phi_f(x) := \Phi(x, f(x)).$$

With this notation we can state our variance-minimization result as follows.

Theorem 3.8.([13]) Suppose that Assumptions 3.1, 3.2, 3.4 and 3.6 are satisfied. Then there exists a constant $\sigma_*^2 \geq 0$, a canonical policy $f_* \in \mathbf{F}_{cp}$, and a function $V^*(\cdot)$ in $B_W(\mathbf{X})$ such that, for each $x \in \mathbf{X}$,

$$\begin{aligned} \sigma_*^2 + V^*(x) &= \min_{a \in A^*(x)} \left[\Phi(x, a) + \int_{\mathbf{X}} V^*(y) Q(dy|x, a)\right] \\ &= \Phi_{f_*}(x) + \int_{\mathbf{X}} V^*(y) Q_{f_*}(dy|x). \end{aligned}$$

Furthermore, f_* satisfies (11) and $V(f_*, \cdot) = \sigma_*^2$; in fact,

$$V(f_*, x) = \mu_{f_*}(\Phi_{f_*}) = \sigma_*^2 \qquad \forall x \in \mathbf{X},$$

and

$$\sigma_*^2 \leq V(f, x) \qquad \forall f \in \mathbf{F}_{eac}, x \in \mathbf{X}.$$

4 Additional comments

In this section we briefly discuss alternative forms of the stability Assumptions 3.2 and 3.4, as well as examples related to our main results.

In addition to the space of functions $B_W(\mathbf{X})$ with the W-norm (12), where $W \geq 1$ is the "weight" function in Assumption 3.1(d), we shall consider the normed linear space $M_W(\mathbf{X})$ of the finite signed measures with a finite W-norm

$$||\mu||_W := \int_{\mathbf{X}} W(y) |\mu|(dy) \tag{15}$$

where $|\mu| := \mu^+ + \mu^-$ denotes the *total variation of* μ.

Remark 4.1. Suppose that Assumption 3.1 holds. Then from the proofs in [12, Theorem 3.5] and [24, Theorem 4.5.3] one can see that the result mentioned in Remark 3.5 holds provided that:

I. The W-geometric ergodicity (13) holds, with constants M_f and γ_f that satisfy Assumption 3.4; and

II. The transition kernel Q_f is φ-irreducible for each $f \in \mathbf{F}$, where φ is a σ-finite measure on $\mathbf{X}$ independent of $f \in \mathbf{F}$.

These two conditions were obtained in §3 from Assumptions 3.2 and 3.4. However, there are other ways of getting I. For example, Assumptions 3.2 and 3.4 may be replaced by the following hypotheses used in [8, 9].

(a) For each stationary policy $f \in \mathbf{F}$ the kernel Q_f admits a unique invariant probability measure μ_f;

(b) There exists a probability measure ν in $M_W(\mathbf{X})$ and a positive constants $\alpha < \infty$ and $\beta < 1$ for which the following holds: for each $f \in \mathbf{F}$ there exists a measurable function $0 \leq l_f \leq 1$ such that for all $x \in \mathbf{X}$ and $B \in \mathcal{B}(\mathbf{X})$,

(b1) $Q_f(B|x) \geq l_f(x)\nu(B)$

(b2) $\nu(l_f) \geq \alpha$, and $\nu(W) = ||\nu||_W < \infty$,

(b3) $\int_{\mathbf{X}} W(y) Q_f(dy|x) \leq \beta W(x) + l_f(x)\nu(W)$.

These conditions, (a) and (b), are an adaptation to MCPs of ideas used by Kartashov [15] to obtain the W-geometric ergodicity in (13). In fact, under (a) and (b), our present Assumption 3.4 is satisfied and one can also obtain estimates of the constants M and γ—see [15, Theorem 3.6] and [8, Lemmas 3.3 and 3.4]—in Assumption 3.4.

Similarly, instead of (a) and (b), one could adapt to MCPs the "contraction" property in [15, Corollary 2.1], which in our notation would be of the form $||\theta Q_f||_W \leq$

$\rho||\theta||_W$ for every signed measure θ in $M_W(\mathbf{X})$ with $\theta(\mathbf{X}) = 0$, for some positive constant $\rho < 1$. On the other hand, as it was already noted in [12, Remark 2.10], if the cost function $C(x,a)$ is *bounded*, then the "weight" function $W \geq 1$ may be bounded and (13) can be obtained from *Doeblin's condition*—see [21, Theorem 16.0.2].

To conclude, we should mention that the examples in [9] and [12] (see also [24, Chapter 4]) hold in our present case. For instance, the example in [12, §6], consists of an inventory system with state space $\mathbf{X} := [0,\infty)$ and a compact control set $\mathbf{A} \subset \mathbb{R}$, in which the one-step cost $C(x,a)$ is *piecewise-linear* in $x \in \mathbf{X}$ and $a \in \mathbf{A}$. Therefore, as the weight function is *exponential*, say

$$W(x) := k\exp(rx) \quad \forall x \in \mathbf{X},$$

with $k, r > 0$, our Assumption 3.6 will trivially hold for some $r_2 > 0$ sufficiently large. A similar comment holds for the queueing system in [9, §5], except that this reference uses the conditions (a) and (b) in Remark 4.1 in lieu of Assumptions 3.2 and 3.4.

References

[1] A. Arapostathis, V.S. Borkar, E. Fernández-Gaucherand, M.K. Ghosh, and S.I. Marcus, *Discrete-time controlled Markov processes with average cost criterion: a survey*, SIAM J. Control Optim., 31 (1993), pp. 282–344.

[2] A.V. Asriev and V.I. Rotar, *On the asymptotic optimality in probability and almost surely in dynamic control,* Sotchastics and Stochastic Reports, 33 (1990), pp. 1-16.

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

[4] V.S. Borkar, *Topics in Controlled Markov Processes*, Pitman Res. Notes in Math. No. 240, Longman Harlow, U.K., 1991.

[5] R. Cavazos-Cadena and E. Fernández-Gaucherand, *Denumerable controlled Markov chains with average reward criterion: sample path optimality*, Math. Methods Oper. Res., 41 (1995), pp. 89–108.

[6] M. Duflo, *Méthodes Récursives Aléatoires*, Masson, Paris, 1990.

[7] P.W. Glynn and S.P. Meyn, *A Lyapunov bound for solutions of the Poisson equation*, Ann. Probab., 24 (1996), pp. 916-931.

[8] E. Gordienko and O. Hernández-Lerma, *Average cost Markov control processes with weighted norms: existence of canonical policies*, Appl. Math. (Warsaw) 23 (1995), pp. 199-218.

[9] E. Gordienko and O. Hernández-Lerma, *Average cost Markov control processes with weighted norms: value iteration*, Appl. Math. (Warsaw) 23 (1995), pp. 219-237.

[10] O. Hernández-Lerma, *Adaptive Markov Control Processes*, Springer-Verlag, New York, 1989.

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

[12] O. Hernández-Lerma and O. Vega-Amaya, *Infinite-horizon Markov control processes with undiscounted cost criteria: from average to overtaking optimality*, Appl. Math. (Warsaw) 25 (1998), pp. 153-178.

[13] O. Hernández-Lerma, O. Vega-Amaya and G. Carrasco, *Sample-path optimality and variance-minimization of average Cost Markov control processes*, to appear in SIAM J. Control Optim.

[14] Y. Huang and L.C.M. Kallenberg, *On finding optimal policies for Markov decision chains: a unifying framework for mean-variance tradeoffs*, Math. Oper. Res., 19 (1994), pp. 434–448.

[15] N.V. Kartashov, *Strong Stable Markov Chains*, VSP, Utrecht, The Netherlands, 1996.

[16] M. Kurano, *Markov decision processes with a minimum-variance criterion*, J. Math. Anal. Appl., 123 (1987), pp. 572-583.

[17] J.B. Lasserre, *Sample-path average optimality for Markov control processes*, IEEE Trans. Autom. Control (to appear).

[18] P. Mandl, *On the variance in controlled Markov chains*, Kybernetika (Prague) 7 (1971), pp. 1-12.

[19] P. Mandl, *A connection between controlled Markov chains and martingales*, Kybernetika (Prague), 9 (1973), pp. 237–241.

[20] P. Mandl, *Estimation and control in Markov chains*, Adv. Appl. Prob., 6 (1974), pp. 40–60.

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

[22] M.L. Puterman, *Markov Decision Processes*, Wiley, New York, 1994.

[23] O. Vega-Amaya, *Sample-path average optimality of Markov control processes with strictly unbounded cost*, Appl. Math. (Warsaw) 1999, to appear.

[24] O. Vega-Amaya, *Markov Control Processes in Borel Spaces: Undiscounted Cost Criteria*, Doctoral Thesis, UAM-Iztapalapa, México, 1998. (In Spanish.)

[25] A.A. Yushkevich, *On a class of strategies in general Markov decision models*, Theory Probab. Appl., 18 (1973), pp. 777-779.