KYBERNETIKA — VOLUME *48* (2012), NUMBER 1, PAGES 83–104

AN OPTIMALITY SYSTEM FOR FINITE AVERAGE MARKOV DECISION CHAINS UNDER RISK-AVERSION

Alfredo Alanís-Durán and Rolando Cavazos-Cadena

This work concerns controlled Markov chains with finite state space and compact action sets. The decision maker is risk-averse with constant risk-sensitivity, and the performance of a control policy is measured by the long-run average cost criterion. Under standard continuity–compactness conditions, it is shown that the (possibly non-constant) optimal value function is characterized by a system of optimality equations which allows to obtain an optimal stationary policy. Also, it is shown that the optimal superior and inferior limit average cost functions coincide.

Keywords: partition of the state space, nonconstant optimal average cost, discounted approximations to the risk-sensitive average cost criterion, equality of superior and inferior limit risk-averse average criteria

Classification: 93E20, 60J05, 93C55

1. INTRODUCTION

This note is concerned with discrete-time Markov decision processes (MDPs) evolving on a finite state space. The system is driven by a risk-averse decision maker with constant risk sensitivity coefficient $\lambda > 0$, and the performance of a control policy is measured by the (superior limit) risk-sensitive average cost criterion. It is supposed that the action set is a compact metric space, and that the cost function and the transition law depend continuously on the action applied, but otherwise they are arbitrary; in particular, no communication conditions are imposed on the transition law, so that the optimal value function may not be constant. Within that framework, the following problem is addressed:

- To characterize the optimal value function using a system of equations form which an optimal stationary policy can be determined.

The study of stochastic systems endowed with the risk-sensitive average criterion can be traced back, at least, to the seminal papers by Howard and Matheson [16], Jacobson [17] and Jaquette [18, 19]. Recently, there has been an intensive work on (controlled) stochastic system endowed with the risk-sensitive average criterion; see, for instance, Flemming and McEneany [12], Di Masi and Stettner [9, 10, 11], Jaśkiewicz [20], Sladký and Montes-de-Oca [26], Sladký [25] and the references there in. A fundamental result

on the existence of solutions of the risk-sensitive optimality equation was obtained by Howard and Matheson [16], where controlled Markov chains with finite state and action spaces were studied, and it was shown that the optimal average cost is determined by a *single* equation whenever each stationary policy determines a communicating Markov chain. In such a case, the optimal average cost function is constant, say g, and the existence of a solution to the optimality equation was established using the Perron–Frobenius theory of nonnegative matrices (Gantmakher [13]). Other approaches have been used to obtain a solution to the optimality equation: the main result in Hernández-Hernández and Marcus [14] is based on game theoretical ideas, the approach in Cavazos-Cadena and Fernández-Gaucherand [4] relies on the risk-sensitive total cost criterion, and the discounted technique — involving contractive mappings — was employed in Di Masi and Stettner [9] and Cavazos-Cadena [5]. On the other hand, there is an interesting contrast between the risk-neutral and the risk-sensitive average cost criteria: Under strong recurrence conditions, like the simultaneous Doeblin condition — under which the Markov chain determined by each stationary policy has a single recurrent class — the risk-neutral optimality equation has a solution, but a similar conclusion is not valid in the risk-sensitive context, even if the optimal average cost is constant (Cavazos-Cadena and Fernández-Gaucherand [4], Cavazos-Cadena and Hernández-Hernandez [6]). Thus, the characterization of the optimal risk-sensitive average cost can not be based, in general, on a single equation, and the problem posed above is an interesting and natural one.

The characterization of a general (risk-sensitive) optimal average cost function was recently studied in Sladký [25] for models with finite state *and* action sets; in that paper, the analysis is based on Perron-Frobenius decompositions of a family of nonnegative matrices (Rothblum and Whittle [22], Sladký [23, 24], Whittle [27], Zijm [28]). On the other hand, the discounted approach has also been employed to study the case of a non necessarily constant optimal average index; see, for instance, Hernández-Hernández and Marcus [15] for models with denumerable state space, and Jaśkiewicz [20] and Cavazos-Cadena and Salem-Silva [8], which concern MDPs with Borel state space. Roughly, in those papers a characterization of the optimal average cost is obtained at *some* states where the optimal performance index attains its minimum. In the context of this work, the discounted technique will play a central role to obtain a *complete* characterization of the optimal average cost function.

The main results of this work involve the idea of *optimality system* introduced in Section 3, and can be briefly described as follows: An optimality system is determined by (i) a partition $S_1, \ldots, S_k$ of a state space, (ii) a sequence of pairs $\{(g_i, h_i(\cdot))\}_{i=1,\ldots,k}$, where g_i is a real number and h_i is a function defined on S_i, (iii) the specification of a (generally proper) subset of $B(x)$ of the original set admissible actions $A(x)$ at each state x. In terms of these objects, an equation — which is similar to the usual optimality equation — is stipulated for every $i = 1, 2, \ldots, k$, and the following conclusions, extending those in Cavazos-Cadena and Hernández-Hernández [7] for *uncontrolled* models, are obtained:

(1) An optimality system characterizes the optimal average cost function and renders an optimal stationary policy (the verification theorem);

(2) There exists and optimality system (the existence theorem).

The approach used below to establish these conclusions relies on basic probabilistic and dynamic programming ideas, which are used to establish the verification theorem, whereas the discounted method is employed to derive the existence result.

The organization of the paper is as follows: In Section 2 a brief description of the decision model is presented and, after introducing the notion of optimality system in Section 3, the verification and existence results are stated as Theorems 3.1 and 3.2, respectively. Then, in Section 4 a technical result on the inferior limit average criterion is presented, and it is used to establish the verification theorem in Section 5. Next in Section 6 the discounted approach is used to specify the components of an optimality system, and the exposition concludes in Section 7 with a proof of the existence theorem.

Notation. The set of all nonnegative integers is denoted by $\mathbb{N}$ and, for a given topological space $\mathbb{K}$, $\mathcal{B}(\mathbb{K})$ stands for the Banach space of all bounded functions $C : \mathbb{K} \to \mathbb{R}$ equipped with the supremum norm:

$$\|C\| := \sup_{x \in \mathbb{K}} |C(x)|.$$

On the other hand, for $x \in \mathbb{K}$, $\delta_x(\cdot)$ is the Dirac's measure concentrated at x, that is, for every Borel subset $D \subset \mathbb{K}$, $\delta_x(D) = 1$ if $x \in D$, and $\delta_x(D) = 0$ when $x \notin D$. If A is an event, the corresponding indicator function is denoted by $I[A]$ and, as usual, all relations involving conditional expectations are supposed to hold almost surely with respect to the underlying probability measure.

2. DECISION MODEL

Throughout the remainder $\mathcal{M} = (S, A, \{A(x)\}_{x \in S}, C, P)$ is an MDP, where the state space S is a finite set endowed with the discrete topology, and the action set A is a metric space. For each $x \in S$, $A(x) \subset A$ is the nonempty set of admissible actions at x, whereas $\mathbb{K} := \{(x, a) \,|\, a \in A(x), x \in S\}$ is the class of admissible pairs. On the other hand, $C \in B(\mathbb{K})$ is the cost function and $P = [p_{x\,y}(\cdot)]$ is the controlled transition law on S given $\mathbb{K}$, that is, for each $(x, a) \in \mathbb{K}$ and $z \in S$, $p_{x\,z}(a) \geq 0$ and $\sum_{y \in S} p_{x\,y}(a) = 1$. This model $\mathcal{M}$ is interpreted as follows: At each time $t \in \mathbb{N}$ the decision maker observes the state of a dynamical system, say $X_t = x \in S$, and selects the action (control) $A_t = a \in A(x)$. Then, a cost $C(x, a)$ is incurred and, regardless of the previous states and actions, the state of the system at time $t + 1$ will be $X_{t+1} = y \in S$ with probability $p_{x\,y}(a)$; this is the Markov property of the decision process.

Assumption 2.1. (i) For each $x \in S$, $A(x)$ is a compact subset of A.

(ii) For every $x, y \in S$, the mappings $a \mapsto C(x, a)$ and $a \mapsto p_{x\,y}(a)$ are continuous in $a \in A(x)$.

Policies. The space $\mathbb{H}_t$ of possible histories up to time $t \in \mathbb{N}$ is defined by $\mathbb{H}_0 := S$ and $\mathbb{H}_t := \mathbb{K}^t \times S$, $t \geq 1$, and $\mathbf{h}_t = (x_0, a_0, \ldots, x_i, a_i, \ldots, x_t)$ stands for a generic element of $\mathbb{H}_t$, where $a_i \in A(x_i)$. A policy $\pi = \{\pi_t\}$ is a special sequence of stochastic kernels: For each $t \in \mathbb{N}$ and $\mathbf{h}_t \in \mathbb{H}_t$, $\pi_t(\cdot|\mathbf{h}_t)$ is a probability measure on A concentrated on $A(x_t)$, and for each Borel subset $B \subset A$, the mapping $\mathbf{h}_t \mapsto \pi_t(B|\mathbf{h}_t)$, $\mathbf{h}_t \in \mathbb{H}_t$, is

Borel measurable; when the controller chooses actions according to π the control A_t applied at time t belongs to $B \subset A$ with probability $\pi_t(B|\mathbf{h}_t)$, where $\mathbf{h}_t$ is the observed history of the process up to time t. The class of all policies is denoted by $\mathcal{P}$. Given the policy π being used for choosing actions and the initial state $X_0 = x$, the distribution of the state-action process $\{(X_t, A_t)\}$ is uniquely determined (Araposthatis et al. [1], Puterman [21]), and such a distribution and the corresponding expectation operator are denoted by P_x^π and E_x^π, respectively. Next, define $\mathbb{F} := \prod_{x\in S} A(x)$ and notice that $\mathbb{F}$ is a compact metric space, which consists of all functions $f : S \to A$ such that $f(x) \in A(x)$ for each $x \in S$. A policy π is *stationary* if there exists a sequence $f \in \mathbb{F}$ such that the probability measure $\pi_t(\cdot|\mathbf{h}_t)$ is always concentrated at $f(x_t)$, and in this case π and f are naturally identified; with this convention, $\mathbb{F} \subset \mathcal{P}$.

Performance Index. As already mentioned, the decision maker is supposed to be *risk-averse* with constant risk-sensitivity coefficient $\lambda > 0$, that is, the controller assesses a random cost Y using the expectation of $e^{\lambda Y}$; the certain equivalent of Y is the real number $\mathcal{E}[Y]$ determined by $e^{\lambda\mathcal{E}[Y]} = E[e^{\lambda Y}]$, so that the controller is indifferent between paying the certain equivalent $\mathcal{E}[Y]$ for sure, or incurring the random cost Y. It follows that

$$\mathcal{E}[Y] = \frac{1}{\lambda}\log\left(E[e^{\lambda Y}]\right),$$

whereas Jensen's inequality yields that if Y has finite expectation, then $\mathcal{E}[Y] \geq E[Y]$ and the strict inequality holds if Y is non constant. Suppose now that the controller is driving the system using policy $\pi \in \mathcal{P}$ starting at $x \in S$, and let $J_n(\lambda, \pi, x)$ be the certain equivalent of the total cost $\sum_{t=0}^{n-1} C(X_t, A_t)$ incurred before time n, that is,

$$J_n(\lambda, \pi, x) := \frac{1}{\lambda}\log\left(E_x^\pi\left[e^{\lambda\sum_{t=0}^{n-1} C(X_t,A_t)}\right]\right). \tag{2.1}$$

With this notation, the (long-run superior limit) λ-sensitive average cost at state x under policy π is given by

$$J(\lambda, \pi, x) := \limsup_{n\to\infty} \frac{1}{n} J_n(\lambda, \pi, x), \tag{2.2}$$

and

$$J^*(\lambda, x) := \inf_{\pi\in\mathcal{P}} J(\lambda, \pi, x), \quad x \in S, \tag{2.3}$$

is the optimal λ-sensitive average cost function; a policy $\pi^* \in \mathcal{P}$ is λ-optimal if $J(\lambda, \pi^*, x) = J^*(\lambda, x)$ for each $x \in S$.

Remark 2.1. When $X_0 = x$, the inferior limit λ-sensitive average criterion associated with $\pi \in \mathcal{P}$ and $x \in S$ is defined by

$$J_-(\lambda, \pi, x) := \liminf_{n\to\infty} \frac{1}{n} J_n(\lambda, \pi, x), \tag{2.4}$$

and the corresponding (inferior limit) λ-optimal value function is given by

$$J_*(\lambda, x) := \inf_{\pi\in\mathcal{P}} J_-(\lambda, \pi, x), \quad x \in S, \tag{2.5}$$

so that $J_*(\lambda, \cdot) \leq J^*(\lambda, \cdot)$; as it will be shown below, under Assumption 2.1 the optimal value functions $J_*(\lambda, \cdot)$ and $J^*(\lambda, \cdot)$ coincide.

The Problem. The optimality equation corresponding to the average criterion in (2.2) is given by

$$e^{\lambda(g+h(x))} = \inf_{a\in A(x)} \left[e^{\lambda C(x,a)} \sum_{y\in S} p_{x\,y}(a) e^{\lambda h(y)} \right], \quad x \in S, \tag{2.6}$$

where g is a real number and $h : S \to \mathbb{R}$ is a given function. When this equation is satisfied by the pair $(g, h(\cdot)) \in \mathbb{R} \times \mathcal{B}(S)$, the optimal average cost function $J^*(\lambda, \cdot)$ is constant and equal to g; moreover, Assumption 2.1 yields that there exists a policy $f^* \in \mathbb{F}$ such that

$$e^{\lambda(g+h(x))} = e^{\lambda C(x,f^*(x))} \sum_{y\in S} p_{x\,y}(f^*(x)) e^{\lambda h(y)}, \quad x \in S,$$

and such a policy f^* is λ-optimal. As already noted, a pair $(g, h(\cdot))$ satisfying (2.6) exists when the whole state space is a communicating class under the action of each stationary policy; however, it was shown un Cavazos-Cadena and Hernández-Hernández [7] that, if the Markov chain associated with some $f \in \mathbb{F}$ has two or more recurret classes, or if the set of transient states is nonempty, then (2.6) may not have a solution, even if the optimal average cost function is constant. On the other hand, for *uncontrolled* Markov chains it was recently shown in Cavazos-Cadena and Hernández-Hernández [7] that, in general, the average cost function is determined by a *system* of local Poisson equations, and *the main problem* considered in this note consists in extending such a conclusion to the present context of controlled models. The results in this direction involve the idea of *optimality system*, which is introduced in the following section.

3. OPTIMALITY SYSTEMS AND MAIN RESULTS

In this section the main conclusions of this note are stated as Theorems 3.1 and 3.2 below. These results involve the idea of *optimality system*, which extends the notion of optimality equation and allows to characterize the optimal value function in terms of a system of equations, as well as to obtain a λ-optimal stationary policy.

Definition 3.1. Let $\mathcal{M} = (S, A, \{A(x)\}_{x\in S}, C, P)$ be the MDP described in Section 2. An *optimality system* for $\mathcal{M}$ is a vector of triplets

$$\mathcal{O} = ((S_1, g_1, h_1), (S_2, g_2, h_2), \ldots, (S_k, g_k, h_k)) \tag{3.1}$$

satisfying the following conditions:

(i) $S_1, S_2, \ldots, S_k$ is a partition of S.

(ii) For each $i = 1, 2, \ldots, k$, $(g_i, h_i(\cdot)) \in \mathbb{R} \times \mathcal{B}(S_i)$ and

$$g_1 \le g_2 \le \cdots \le g_k. \tag{3.2}$$

(iii) For each $i = 1, 2, \ldots, k$,

$$B(x) := \{a \in A(x) \mid \textstyle\sum_{y\in S_1\cup S_2\cup\cdots\cup S_i} p_{x\,y}(a) = 1\}, \quad x \in S_i, \quad \text{is nonempty.} \tag{3.3}$$

(iv) For each $i = 1, 2, \ldots, k$,

$$e^{\lambda(g_i + h_i(x))} = \inf_{a \in B(x)} \left[e^{\lambda C(x,a)} \sum_{y \in S_i} p_{x\,y}(a) e^{\lambda h_i(y)} \right], \quad x \in S_i. \tag{3.4}$$

Remark 3.1. Notice that (3.4) implies that, for every $x \in S_i$, $\sum_{y \in S_i} p_{x\,y}(a) > 0$ for all $a \in B(x)$, since $e^{\lambda(g_i + h_i(x))} > 0$.

The number k of triplets in $\mathcal{O}$ will be referred to as *the order* of $\mathcal{O}$. The above idea is an extension of the notion of J-system used in Cavazos-Cadena and Hernández-Hernández [7] to characterize the average cost function for an uncontrolled Markov chain. In the present controlled context, the following result shows that an optimality system renders (i) the optimal value function, (ii) the equality of the superior and inferior limit optimal value functions, as well as (iii) a λ-optimal stationary policy.

Theorem 3.2. [Verification.] Let $\mathcal{M}$ be the model described in Section 2 and suppose that Assumption 2.1 holds. If $\mathcal{O} = ((S_1, g_1, h_1), (S_2, g_2, h_2), \ldots, (S_k, g_k, h_k))$ is an optimality system for $\mathcal{M}$, then the following assertions (i) – (iii) hold:
(i) For each $i = 1, 2, \ldots, k$, the optimal average cost at each state $x \in S_i$ is given by g_i:

$$J^*(\lambda, x) = g_i, \quad x \in S_i.$$

Moreover,
(ii) $J_*(\lambda, x) = J^*(\lambda, x)$ for all $x \in S$; see (2.3) and (2.5).
(iii) Suppose that the stationary policy $f \in \mathbb{F}$ satisfies that

$$f(x) \in B(x), \quad x \in S, \tag{3.5}$$

and

$$e^{\lambda(g_i + h_i(x))} = \left[e^{\lambda C(x, f(x))} \sum_{y \in S_i} p_{x\,y}(f(x)) e^{\lambda h_i(y)} \right], \quad x \in S_i, \quad i = 1, 2, \ldots, k. \tag{3.6}$$

In this case f is λ-optimal and

$$\lim_{n \to \infty} \frac{1}{n} J_n(\lambda, f, x) = J^*(\lambda, x), \quad x \in S.$$

Notice that Assumption 2.1 yields that the set $B(x)$ in (3.3) is always compact, a fact that using (3.4) implies the existence of a stationary policy f satisfying (3.5) and (3.6). The following result establishes the existence of an optimality system.

Theorem 3.3. [Existence.] Under Assumption 2.1, there exists an optimality system $\mathcal{O}$ for model $\mathcal{M}$.

The proof of Theorems 3.2 and 3.3 will be presented in Sections 5 and 7, respectively, after establishing the necessary preliminary results. The argument used to establish the verification result relies on standard probabilistic and dynamic programming arguments, whereas the existence of an optimality system will be obtained *via* the risk-sensitive discounted criterion.

4. A LOWER BOUND FOR THE INFERIOR LIMIT AVERAGE CRITERION

In this section a basic technical tool that will be used to prove Theorem 3.2 is established. The main objective is to show that if $\mathcal{O}$ is as optimality system for model $\mathcal{M}$, then a lower bound for the optimal inferior limit average cost function can be obtained, a result that is precisely stated in the following theorem.

Theorem 4.1. Let $\mathcal{O}$ in (3.1) be an optimality system for model $\mathcal{M}$. In this case, g_i is a lower bound for the inferior limit λ-sensitive average cost criterion at each state $x \in S_i$:

$$J_*(\lambda, x) \geq g_i, \quad x \in S_i, \quad i = 1, 2, \ldots, k; \tag{4.1}$$

see Remark 2.1.

This result will be proved below by induction. Since the argument is rather technical, to ease the presentation the simple auxiliary facts involved in the argument are established in the following three lemmas.

Lemma 4.2. If $\mathcal{O} = ((S_1, g_1, h_1), (S_2, g_2, h_2), \ldots, (S_k, g_k, h_k))$ is an optimality system for model $\mathcal{M}$, then the following assertions (i) and (ii) hold:
(i) For each positive integer n,

$$\frac{1}{n} J_n(\lambda, \pi, x) \geq g_k - \frac{2\|h_k\|}{n}, \quad x \in S_k, \quad \pi \in \mathcal{P}.$$

Consequently,
(ii) At each state $x \in S_k$, the constant g_k is a lower bound for the optimal inferior limit average cost function:

$$J_*(\lambda, x) \geq g_k, \quad x \in S_k;$$

see (2.1), (2.4) and (2.5).

P r o o f. Since $S_1 \cup \cdots \cup S_k = S$, from (3.3) it follows that $B(x) = A(x)$ when $x \in S_k$, and then the fourth part in Definition 3.1 yields that

$$e^{\lambda(g_k + h_k(x))} \leq e^{\lambda C(x,a)} \sum_{y \in S_k} p_{x\,y}(a) e^{\lambda h_k(y)}, \quad a \in A(x), \quad x \in S_k. \tag{4.2}$$

Now let $\pi \in \mathcal{P}$ be arbitrary. After integrating both sides of the above inequality with respect to $\pi_0(\cdot|x)$, it follows that

$$e^{\lambda(g_k + h_k(x))} \leq E_x^\pi \left[e^{\lambda C(X_0, A_0) + \lambda h_k(X_1)} I[X_1 \in S_k] \right], \quad x \in S_k, \quad \pi \in \mathcal{P}. \tag{4.3}$$

On the other hand, for every positive integer n, the Markov property yields that

$$\begin{aligned} E_x^\pi \left[e^{\lambda \sum_{t=0}^{n} C(X_t, A_t) + \lambda h_k(X_{n+1})} I[X_r \in S_k,\, 1 \leq r \leq n+1] \middle| (X_m, A_m), 1 \leq m \leq n \right] \\ = e^{\lambda \sum_{t=0}^{n-1} C(X_t, A_t)} I[X_r \in S_k,\, 1 \leq r \leq n] e^{\lambda C(X_n, A_n)} \sum_{y \in S_k} p_{X_n\,y}(A_n) e^{\lambda h_k(y)} \\ \geq e^{\lambda \sum_{t=0}^{n-1} C(X_t, A_t)} I[X_r \in S_k,\, 1 \leq r \leq n] e^{\lambda g_k + \lambda h_k(X_n)} \end{aligned}$$

where (4.2) was used to set the inequality. Therefore,

$$E_x^\pi\left[e^{\lambda\sum_{t=0}^{n}C(X_t,A_t)+\lambda h_k(X_{n+1})}I[X_r\in S_k,\,1\le r\le n+1]\right]$$
$$\ge e^{\lambda g_k}E_x^\pi\left[e^{\lambda\sum_{t=0}^{n-1}C(X_t,A_t)+\lambda h_k(X_n)}I[X_r\in S_k,\,1\le r\le n]\right].$$

Combining this last relation and (4.3), a simple induction argument yields that, for every positive integer n, $x\in S_k$ and $\pi\in\mathcal{P}$,

$$E_x^\pi\left[e^{\lambda\sum_{t=0}^{n-1}C(X_t,A_t)+\lambda h_k(X_n)}I[X_r\in S_k,\,1\le r\le n]\right]\ge e^{\lambda(ng_k+h_k(x))},$$

and then

$$\begin{aligned}e^{\lambda(J_n(\lambda,\pi,x)+\|h_k\|)} &\ge E_x^\pi\left[e^{\lambda\sum_{t=0}^{n-1}C(X_t,A_t)+\lambda h_k(X_n)}\right]\\ &\ge E_x^\pi\left[e^{\lambda\sum_{t=0}^{n-1}C(X_t,A_t)+\lambda h_k(X_n)}I[X_r\in S_k,\,1\le r\le n]\right]\\ &\ge e^{\lambda(ng_k+h_k(x))}\ge e^{\lambda(ng_k-\|h_k\|)}\end{aligned}$$

so that, for every positive integer n,

$$\frac{1}{n}J_n(\lambda,\pi,x)\ge g_k-\frac{2\|h_k\|}{n},\quad x\in S_k,\quad \pi\in\mathcal{P},$$

establishing part (i), and then the second assertion follows from (2.4) and (2.5). □

Now let the optimality system $\mathcal{O}$ be as in (3.1), suppose that $k>1$ and set

$$\hat{S}=S_1\cup\cdots\cup S_{k-1}. \tag{4.4}$$

Next, let $x\in\hat{S}$ be arbitrary, so that there exists $i<k$ such that $x\in S_i$ for some $i<k$; since

$$a\in B(x)\implies\sum_{y\in S_1\cup\cdots\cup S_i}p_{x\,y}(a)=1\implies\sum_{y\in\hat{S}}p_{x\,y}(a)=1,$$

it follows that

$$\hat{A}(x):=\left\{a\in A(x)\,|\,\textstyle\sum_{y\in\hat{S}}p_{x\,y}(a)=1\right\},\quad x\in\hat{S}, \tag{4.5}$$

is always nonempty. Set $\hat{\mathbb{K}}:=\{(x,a)|x\in\hat{S},a\in\hat{A}(x)\}$ and define the transition $\hat{P}=[\hat{p}_{x\,y}]$ and $\hat{C}:\hat{\mathbb{K}}\to\mathbb{R}$ by

$$\hat{p}_{x\,y}(a):=p_{x\,y}(a),\quad \hat{C}(x,a):=C(x,a),\quad (x,a)\in\hat{\mathbb{K}},\quad y\in\hat{S}. \tag{4.6}$$

Definition 4.1. Let $\mathcal{O}$ be an optimality system for model $\mathcal{M}$ as in Definition 3.1, and suppose that the order k of $\mathcal{O}$ is larger than 1. With the notation in (4.4) – (4.6), the reduced model $\hat{\mathcal{M}}$ is specified by

$$\hat{\mathcal{M}}=(\hat{S},A,\{\hat{A}(x)\}_{x\in\hat{S}},\hat{C},\hat{P}) \tag{4.7}$$

Combining Definitions 3.1 and 4.1 the following lemma follows immediately.

Lemma 4.3. If $\mathcal{O} = ((S_1, g_1, h_1), (S_2, g_2, h_2), \ldots, (S_k, g_k, h_k))$ is an optimality system for model model $\mathcal{M}$, where $k > 1$, then

$$\hat{\mathcal{O}} = ((S_1, g_1, h_1), (S_2, g_2, h_2), \ldots, (S_{k-1}, g_{k-1}, h_{k-1})), \tag{4.8}$$

is an optimality system for the reduced model $\hat{\mathcal{M}}$. Moreover, setting

$$\hat{B}(x) := \{x \in \hat{A}(x) \,|\, \sum_{y \in S_1 \cup \cdots \cup S_i} \hat{p}_{x\,y} = 1\}, \quad x \in S_i, \quad i = 1, 2, \ldots, k-1,$$

the equality $\hat{B}(x) = B(x)$ holds for every $x \in \hat{S}$; see (3.3).

Remark 4.4. The class of policies for model $\hat{\mathcal{M}}$ will be denoted by $\hat{\mathcal{P}}$. For $\Delta \in \hat{\mathcal{P}}$, $\hat{J}_-(\lambda, \Delta, \cdot)$ denotes the inferior limit λ-sensitive average cost criterion associated with Δ, and $\hat{J}_*(\lambda, \cdot) = \inf_{\Delta \in \hat{\mathcal{P}}} \hat{J}_-(\lambda, \Delta, \cdot)$ stands for the optimal inferior limit average cost function for model $\hat{\mathcal{M}}$.

The following lemma is the final step before the proof of Theorem 4.1. Write

$$H_n := (X_0, A_0, \ldots, X_{n-1}, A_{n-1}, X_n). \tag{4.9}$$

Lemma 4.5. Let $\mathcal{O}$ in (3.1) be an optimality system for model $\mathcal{M}$, where $k > 1$. Suppose that for some $r \in \{1, 2, \ldots, k-1\}$, the state $x \in S_r$ and $\pi \in \mathcal{P}$ satisfy

$$J_-(\lambda, \pi, x) < g_r. \tag{4.10}$$

In this case, the following assertions (i) – (iii) hold:

(i) With probability 1 with respect to P_x^π, the actions chosen by π after observing H_n always belong to $\hat{A}(X_n)$. More precisely,

$$1 = P_x^\pi[\pi_n(\hat{A}(X_n)|H_n) = 1], \quad n \in \mathbb{N}.$$

Now let $w : \hat{S} \to A$ be a stationary policy for model $\hat{\mathcal{M}}$, that is, $w(x) \in \hat{A}(x)$ for each $x \in \hat{S}$, and define the policy $\Delta \in \hat{\mathcal{P}}$ as follows: For each $n \in \mathbb{N}$ and $\mathbf{h}_n \in \hat{\mathbb{H}}_n$,

$$\Delta_n(D|\mathbf{h}_n) \colon = \pi_n(D \cap \hat{A}(x_n)|\mathbf{h}_n) + (1 - \pi_n(\hat{A}(x_n)|\mathbf{h}_n))\delta_{w(x_n)}(D), \quad D \in \mathcal{B}(A). \tag{4.11}$$

With this notation,

(ii) For every $n \in \mathbb{N}$,

$$P_x^\Delta[H_n \in D] = P_x^\pi[H_n \in D], \quad D \in \mathcal{B}(\hat{\mathbb{H}}_n), \tag{4.12}$$

and then,

(iii) $\hat{J}_-(\lambda, \Delta, x) = J_-(\lambda, \pi, x) < g_r$.

P r o o f. (i) The argument is by contradiction. Suppose that, for some $n \in \mathbb{N}$,

$$0 < P_x^\pi[\pi_n(A(X_n) \setminus \hat{A}(X_n)|H_n) > 0]. \tag{4.13}$$

Notice now that

$$\begin{aligned} P_x^\pi[X_{n+1} \in S_k|H_n] &= \int_{A(X_n)} \sum_{y\in S_k} p_{X_n\, y}(a)\pi_n(\mathrm{d}a|H_n) \\ &\geq \int_{A(X_n)\setminus \hat{A}(X_n)} \sum_{y\in S_k} p_{X_n\, y}(a)\pi_n(\mathrm{d}a|H_n) \\ &= \int_{A(X_n)\setminus \hat{A}(X_n)} \sum_{y\in S\setminus \hat{S}} p_{X_n\, y}(a)\pi_n(\mathrm{d}a|H_n). \end{aligned}$$

For $a \in A(X_n) \setminus \hat{A}(X_n)$ the summation inside the integral is positive, by (4.5), and then the integral is larger that zero on the event $[\pi_n(A(X_n) \setminus \hat{A}(X_n)|H_n) > 0]$. It follows from (4.13) that $P_x^\pi[X_{n+1} \in S_k|H_n] > 0$ with positive P_x^π-probability, so that

$$P_x^\pi[X_{n+1} \in S_k] > 0. \tag{4.14}$$

Next, given $\tilde{\mathbf{h}}_n \in \mathbb{H}_n$ and $\tilde{a} \in A(x_n)$, define the (shifted) policy $\pi^{\tilde{\mathbf{h}}_n,\tilde{a}}$ as follows:

$$\pi_t^{\tilde{\mathbf{h}}_n,\tilde{a}}(\cdot|\mathbf{h}_t) \colon = \pi_{n+1+t}(\cdot|\tilde{\mathbf{h}}_n, \tilde{a}, \mathbf{h}_t), \quad \mathbf{h}_t \in \mathbb{H}_t, \quad t \in \mathbb{N}.$$

With this specification, the Markov property yields that for every $m > n + 1$

$$\begin{aligned} &E_x^\pi[e^{\lambda\sum_{t=0}^{m-1} C(X_t,A_t)} I[X_{n+1} \in S_k]|H_n, A_n, X_{n+1}] \\ &\quad = e^{\lambda\sum_{t=0}^{n} C(X_t,A_t)} I[X_{n+1} \in S_k] E_{X_{n+1}}^{\pi^{H_n,A_n}}[e^{\lambda\sum_{t=0}^{m-n-2} C(X_t,A_t)}] \\ &\quad \geq e^{-\lambda(n+1)\|C\|} I[X_{n+1} \in S_k] E_{X_{n+1}}^{\pi^{H_n,A_n}}[e^{\lambda\sum_{t=0}^{m-n-2} C(X_t,A_t)}] \\ &\quad \geq e^{-\lambda(n+1)\|C\|} I[X_{n+1} \in S_k] e^{\lambda J_{m-n-1}(\lambda,\pi^{H_n,A_n},X_{n+1})}; \end{aligned}$$

see (2.1). From this point, Lemma 4.2(i) yields that

$$\begin{aligned} &E_x^\pi[e^{\lambda\sum_{t=0}^{m-1} C(X_t,A_t)} I[X_{n+1} \in S_k]|H_n, A_n, X_{n+1}] \\ &\qquad \geq e^{-\lambda(n+1)\|C\|} I[X_{n+1} \in S_k] e^{\lambda(m-n-1)g_k - 2\lambda\|h_k\|} \end{aligned}$$

and then

$$\begin{aligned} e^{\lambda J_m(\lambda,\pi,x)} &= E_x^\pi[e^{\lambda\sum_{t=0}^{m-1} C(X_t,A_t)}] \\ &\geq E_x^\pi[e^{\lambda\sum_{t=0}^{m-1} C(X_t,A_t)} I[X_{n+1} \in S_k]] \\ &\geq e^{-\lambda(n+1)\|C\|} P_x^\pi[X_{n+1} \in S_k] e^{\lambda(m-n-1)g_k - 2\lambda\|h_k\|}. \end{aligned}$$

Using (4.14), this inequality immediately yields that

$$J_-(\lambda,\pi,x) = \liminf_{m\to\infty} \frac{1}{m} J_m(\lambda,\pi,x) \geq g_k$$

and then (4.10) implies that $g_k < g_r$, an inequality that, recalling that $r < k$, contradicts (3.2). Therefore, (4.13) does not hold and it follows that

$$0 = P_x^\pi[\pi_n(A(X_n) \setminus \hat{A}(X_n)|H_n) > 0],$$

that is, $1 = P_x^\pi[\pi_n(\hat{A}(X_n)|H_n) = 1]$.

(ii) The argument is by induction. For $n = 0$, both sides of (4.12) are equal to $\delta_x(D)$. Assume now that (4.12) holds for certain nonnegative integer n, and let $D \in \mathcal{B}(\hat{\mathbb{H}}_n)$, $D_1 \in \mathcal{B}(\hat{A})$ and $D_2 \subset \hat{S}$ be arbitrary. Next observe that

$$P_x^\pi[H_n \in D, A_n \in D_1, X_{n+1} \in D_2|H_n] = I[H_n \in D] \int_{a\in D_1} \sum_{y\in D_2} p_{X_n\, y}(a)\pi_n(\mathrm{d}a|H_n);$$

since the equality $\Delta_n(\cdot|H_n) = \pi_n(\cdot|H_n)$ holds P_x^π-a.s., by part (i) and (4.11), it follows that

$$\begin{aligned} &P_x^\pi[H_n \in D, A_n \in D_1, X_{n+1} \in D_2|H_n] \\ &\qquad = I[H_n \in D] \int_{a\in D_1} \sum_{y\in D_2} p_{X_n\, y}(a)\Delta_n(\mathrm{d}a|H_n) \quad P_x^\pi\text{-a.s.}, \end{aligned}$$

and then

$$\begin{aligned} &P_x^\pi[H_n \in D, A_n \in D_1, X_{n+1} \in D_2] \\ &\qquad = \int_{\mathbf{h}_n\in D} \left[\int_{a\in D_1} \sum_{y\in D_2} p_{x_n\, y}(a)\Delta_n(\mathrm{d}a|\mathbf{h}_n) \right] P_x^\pi[\mathrm{d}\mathbf{h}_n]. \end{aligned}$$

By the induction hypothesis the distribution of H_n is the same under P_x^π and P_x^Δ, so that

$$\begin{aligned} &P_x^\pi[H_n \in D, A_n \in D_1, X_{n+1} \in D_2] \\ &\qquad = \int_{\mathbf{h}_n\in D} \left[\int_{a\in D_1} \sum_{y\in D_2} p_{x_n\, y}(a)\Delta_n(\mathrm{d}a|\mathbf{h}_n) \right] P_x^\Delta[\mathrm{d}\mathbf{h}_n], \end{aligned}$$

that is,

$$P_x^\pi[H_n \in D, A_n \in D_1, X_{n+1} \in D_2] = P_x^\Delta[H_n \in D, A_n \in D_1, X_{n+1} \in D_2].$$

Since $D \in \mathcal{B}(\hat{H}_n), D_1 \in \mathcal{B}(\hat{A})$ and $D_2 \subset \hat{S}$ are arbitrary, Theorem 10.4 in Billingsley [2] yields that (4.12) holds with $n + 1$ instead of n.

(iii) The previous part yields that $E_x^\Delta[e^{\lambda \sum_{t=0}^{n-1} C(X_t,A_t)}] = E_x^\pi[e^{\lambda \sum_{t=0}^{n-1} C(X_t,A_t)}]$ for every positive integer n. Therefore, $J_n(\lambda, \Delta, x)/n = J_n(\lambda, \pi, x)/n$, and the conclusion follows after taking the inferior limit as n goes to ∞ in both sides of this equality. □

After the above preliminaries, the proof of the main result of this section is presented below.

P r o o f o f Theorem 4.1. The argument is by induction in the order k of the optimality system $\mathcal{O}$. If $k = 1$ then (4.1) follows from Lemma 4.2(ii). Suppose now that (4.1) holds when $k = m - 1$ for certain integer $m \geq 2$, and let $\mathcal{O}$ be an optimality system for $\mathcal{M}$

with order m. The reduced optimality system $\hat{\mathcal{O}}$ in Definition 4.1 has order $m-1$, and then the optimal inferior limit average cost corresponding to $\hat{\mathcal{M}}$ satisfies

$$\hat{J}_*(\lambda, x) \geq g_i, \quad x \in S_i, \quad i = 1, 2, \ldots, m-1, \tag{4.15}$$

by the induction hypothesis, a fact that will be used to verify that

$$J_*(\lambda, x) \geq g_i, \quad x \in S_i, \quad i = 1, 2, \ldots, m-1. \tag{4.16}$$

Indeed, if this relation fails, there exist $r < m$ and a state $x \in S_r$ such that $J_*(\lambda, x) < g_r$, and then $J_-(\lambda, \pi, x) < g_r$ for some policy $\pi \in \mathcal{P}$. Using Lemma 4.5(iii), there exists a policy $\Delta \in \hat{\mathcal{P}}$ such that $\hat{J}_-(\lambda, \Delta, x) = J_-(\lambda, \pi, x) < g_r$, and then $\hat{J}_*(\lambda, x) < g_r$, contradicting (4.15). Thus, (4.16) holds, whereas an application of Lemma 4.2(i) to the present optimality system of order m yields that $J^*(\lambda, x) \geq g_m$ for all $x \in S_m$, a fact that together with (4.16) yields that (4.1) holds when $k = m$, concluding the argument. □

5. PROOF OF THE VERIFICATION THEOREM

In this section Theorem 3.2 will be established. The argument combines Theorem 4.1 with the following result, which provides an upper bound for the (superior limit) average cost function associated with a stationary policy f satisfying (3.6). Although such a result can be obtained from Cavazos-Cadena and Hernández-Hernández [7], for the sake of completeness a different proof is presented, which uses simple probabilistic arguments. The following notation is involved in the argument: For each set $W \subset S$, the corresponding hitting time is given by

$$T_W := \min\{n > 0 \,|\, X_n \in W\}, \tag{5.1}$$

where the minimum of the empty set is ∞.

Theorem 5.1. (i) Let f be a stationary policy as in the statement of Theorem 3.2(ii). In this case,

$$J(\lambda, f, x) \leq g_i, \quad x \in S_i, \quad i = 1, 2, \ldots, k. \tag{5.2}$$

Consequently,

(ii) For each $i \in \{1, 2, \ldots, k\}$ and $x \in S_i$, $J^*(\lambda, x) \leq g_i$.

P r o o f . To begin with, notice that (3.3) and (3.5) together imply that the set $S_1 \cup \cdots \cup S_i$ is closed under the action of policy f, that is, for each $i = 1, 2, \ldots, k$,

$$x \in S_1 \cup \cdots \cup S_i \quad \text{and} \quad p_{x\,y}(f(x)) > 0 \implies y \in S_1 \cup \cdots \cup S_i. \tag{5.3}$$

On the other hand, starting from (3.6), a standard induction argument using the Markov property yields that, for every $n = 1, 2, 3, \ldots$ and $i = 1, 2, \ldots k$,

$$e^{\lambda(ng_i + h_i(x))} = E_x^f[e^{\lambda \sum_{t=0}^{n-1} C(X_t, A_t) + \lambda h_i(X_n)} I[X_t \in S_i,\, t < n]], \quad x \in S_i. \tag{5.4}$$

Since (5.3) implies that $1 = P_x^f[X_t \in S_1]$ for every $x \in S_1$ and $t \in \mathbb{N}$, it follows that if the initial state x belongs to S_1, the equality

$$e^{\lambda(ng_1+h_1(x))} = E_x^f\left[e^{\lambda\sum_{t=0}^{n-1}C(X_t,A_t)}e^{\lambda h_1(X_n)}\right]$$

holds for each $n > 0$, and in this case $E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}\right] \le e^{\lambda(ng_1+2\|h_1\|)}$, that is,

$$J_n(\lambda, x, f) \le ng_1 + 2\|h_1\|, \quad x \in S_1, \quad n = 1, 2, 3, \ldots \tag{5.5}$$

see (2.1). Next, for $i \in \{1, 2, \ldots, k\}$ consider the following claim. $\mathcal{C}_i : J(\lambda, f, x) \le g_i$ for every $x \in S_i$. It will be proved, by induction, that $\mathcal{C}_i$ is valid for every $i = 1, 2, \ldots, k$. To achieve this goal, observe that (5.5) implies that

$$J(\lambda, f, x) = \limsup_{n\to\infty}\frac{1}{n}J_n(\lambda, x, f) \le g_1, \quad x \in S_1,$$

so that $\mathcal{C}_1$ is valid. Now, suppose that $\mathcal{C}_j$ holds for $j = 1, 2, \ldots, i-1$, where $i \in \{2, 3, \ldots, k\}$. In this case, given $\varepsilon > 0$, for each $x \in S_j$ with $1 \le j \le i-1$, there exists a positive integer $N(x)$ such that $J_n(\lambda, f, x)/n \le g_j + \varepsilon$ for $n \ge N(x)$, a relation that *via* (2.1) is equivalent to

$$E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}\right] \le e^{\lambda n(g_j+\varepsilon)}, \quad n \ge N(x);$$

since $g_j \le g_i$ for $j < i$, by (3.2), it follows that

$$E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}\right] \le D(\varepsilon)e^{\lambda n(g_i+\varepsilon)}, \quad x \in S_1 \cup \cdots \cup S_{i-1}, \quad n = 1, 2, 3, \ldots, \tag{5.6}$$

where, setting

$$\tilde{D}(\varepsilon) := \max\left\{e^{-\lambda n(g_i+\varepsilon)}E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}\right] \;\middle|\; 1 \le n < N(x),\ x \in \bigcup_{j<i}S_j\right\},$$

$D(\varepsilon) \ge 1$ is given by

$$D(\varepsilon) := \max\{\tilde{D}(\varepsilon), 1\}.$$

Next, let $x \in S_i$ be arbitrary but fixed, and observe that (5.3) yields that

$$P_x^f[X_t \in S_1 \cup \cdots \cup S_i,\ t = 1, 2, 3, \ldots] = 1. \tag{5.7}$$

Combining this relation with the specification of the hitting time T_W in (5.1), it follows that for every positive integers n and r the following equalities occur with probability 1 with respect to P_x^f:

$$\begin{aligned} I[T_{S_1\cup\cdots\cup S_{i-1}} = r] &= I[X_m \in S_i, 1 \le m < r]I[X_r \in S_1 \cup \cdots \cup S_{i-1}] \\ I[T_{S_1\cup\cdots\cup S_{i-1}} > n-1] &= I[X_m \in S_i, 1 \le m \le n-1]. \end{aligned}$$

Therefore, for a positive integer n,

$$\begin{aligned}
&E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}\right]\\
&\quad=\sum_{r=1}^{n-1}E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}I[T_{S_1\cup\cdots\cup S_{i-1}}=r]\right]\\
&\qquad+E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}I[T_{S_1\cup\cdots\cup S_{i-1}}>n-1]\right] \qquad (5.8)\\
&\quad=\sum_{r=1}^{n-1}E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}I[X_m\in S_i,1\le m<r]I[X_r\in\cup_{j=1}^{i-1}S_j]\right]\\
&\qquad+E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}I[X_t\in S_i,\,t=1,2,\ldots,n-1]\right].
\end{aligned}$$

To continue, each one of the terms in this last equality will be analyzed. First, recalling that $x\in S_i$, notice that (5.4) immediately implies that

$$\begin{aligned}
&E_x^f\left[e^{\sum_{t=0}^{r-1}C(X_t,A_t)}I[X_t\in S_i,\,t=1,2,\ldots,r-1]\right]\\
&\qquad\qquad\le e^{\lambda(rg_i+2\|h_i\|)},\quad r=1,2,3,\ldots. \qquad (5.9)
\end{aligned}$$

Next, for $r\in\{1,2,\ldots,n-1\}$, the Markov property yields

$$\begin{aligned}
&E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}I[X_m\in S_i,1\le m<r]I[X_r\in\cup_{j=1}^{i-1}S_j]\,\middle|\,H_r\right]\\
&\quad=I[X_m\in S_i,1\le m<r]I[X_r\in\cup_{j=1}^{i-1}S_j]e^{\sum_{t=0}^{r-1}C(X_t,A_t)}\\
&\qquad\qquad\times E_{X_r}^f\left[e^{\sum_{t=0}^{n-r-1}C(X_t,A_t)}\,\middle|\,H_r\right]\\
&\quad\le I[X_m\in S_i,1\le m<r]I[X_r\in\cup_{j=1}^{i-1}S_j]e^{\sum_{t=0}^{r-1}C(X_t,A_t)}D(\varepsilon)e^{\lambda(n-r)(g_i+\varepsilon)}\\
&\quad\le I[X_m\in S_i,1\le m<r]e^{\sum_{t=0}^{r-1}C(X_t,A_t)}D(\varepsilon)e^{\lambda(n-r)(g_i+\varepsilon)}
\end{aligned}$$

where (5.6) was used to set the first inequality. Thus,

$$\begin{aligned}
&E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}I[X_m\in S_i,1\le m<r]I[X_r\in\cup_{j=1}^{i-1}S_j]\right]\\
&\quad\le D(\varepsilon)e^{\lambda(n-r)(g_i+\varepsilon)}E_x^f\left[e^{\sum_{t=0}^{r-1}C(X_t,A_t)}I[X_m\in S_i,1\le m<r]\right]\\
&\quad\le D(\varepsilon)e^{\lambda(n-r)(g_i+\varepsilon)}e^{\lambda(rg_i+2\|h_i\|)}\\
&\quad\le e^{2\lambda\|h_i\|}D(\varepsilon)e^{\lambda n(g_i+\varepsilon)}
\end{aligned}$$

where (5.9) was used to set the second inequality. Combining this last display and (5.9) and recalling that $D(\varepsilon)\ge 1$, from (5.8) it follows that

$$\begin{aligned}
e^{\lambda J_n(\lambda,f,x)} &= E_x^f\left[e^{\sum_{t=0}^{n-1}C(X_t,A_t)}\right]\\
&\le \sum_{r=1}^{n-1}e^{2\lambda\|h_i\|}D(\varepsilon)e^{\lambda n(g_i+\varepsilon)}+e^{2\lambda\|h_i\|}e^{\lambda ng_i}\\
&\le D(\varepsilon)ne^{2\lambda\|h_i\|}e^{\lambda n(g_i+\varepsilon)},
\end{aligned}$$

that is,

$$J_n(\lambda, f, x) \leq \frac{\log(n) + 2\lambda\|h_i\| + \log(D(\varepsilon))}{\lambda} + n(g_i + \varepsilon),$$

a relation that leads to

$$J(\lambda, f, x) = \limsup_{n\to\infty} \frac{1}{n} J_n(\lambda, f, x) \leq g_i + \varepsilon;$$

since $x \in S_i$ and $\varepsilon > 0$ are arbitrary, it follows that $\mathcal{C}_i$ holds, concluding the induction argument. Therefore $\mathcal{C}_j$ occurs for every $j = 1, 2, \ldots, k$, a fact that is equivalent to (5.2). □

P r o o f o f Theorem 3.2. Since $J_*(\lambda, \cdot) \leq J^*(\lambda, \cdot)$, Theorems 4.1 and 5.1(ii) together yield that

$$g_i \leq J_*(\lambda, x) \leq J^*(\lambda, x) \leq g_i, \quad x \in S_i, \quad i = 1, 2, \ldots, k,$$

a relation that immediately implies parts (i) and (ii). Now, let $f \in \mathbb{F}$ be as in (3.5) and (3.6). Using that $J(\lambda, f, \cdot) \geq J_-(\lambda, f, \cdot) \geq J_*(\lambda, \cdot)$, by (2.2), (2.4) and (2.5), the above displayed relation and Theorem 5.1(i) lead to

$$J(\lambda, f, x) = J_-(\lambda, f, x) = g_i = J^*(x), \quad x \in S_i, \quad i = 1, 2, \ldots, k,$$

where part (i) was used to set the last equality. Therefore, f is λ-optimal and, *via* (2.2) and (2.4), $\lim_{n\to\infty} J_n(\lambda, f, x)/n = J^*(\lambda, x)$ for all $x \in S$, completing the proof. □

6. DISCOUNTED APPROACH

This section presents the necessary technical tools that will be used to establish the existence of an optimality system for model $\mathcal{M}$. The approach relies on the discounted operators introduced below which, when λ is small enough and appropriate communication conditions are satisfied by the transition law, have been used to construct solutions of the optimality equation (2.6) (Di Masi and Stettner [9], Cavazos-Cadena [5]).

Definition 6.1. Given $\alpha \in (0, 1)$ define the operator $T_\alpha : \mathcal{B}(S) \to \mathcal{B}(S)$ as follows: For each $V \in \mathcal{B}(S)$ and $x \in S$, $T_\alpha[V](x)$ is determined by

$$e^{\lambda T_\alpha[V](x)} = \inf_{a\in A(x)} \left[e^{\lambda C(x,a)} \sum_{y\in S} p_{x\,y}(a) e^{\lambda\alpha V(y)} \right], \quad x \in S. \tag{6.1}$$

According to this specification, $T_\alpha[V](x)$ is the minimum certain equivalent of the random cost $C(X_0, A_0) + \alpha V(X_1)$ that can be achieved when the initial state is $X_0 = x$. On the other hand, it is not difficult to see that T is a monotone and α-homogeneous operator, that is, for $V, W \in \mathcal{B}(S)$ (i) $V \geq W$ implies that $T[V] \geq T[W]$, and (ii) $T[V + r] = T[V] + \alpha r$ for every $r \in \mathbb{R}$. Combining these properties with the relation $W - \|W - V\| \leq V \leq W + \|W - V\|$, it follows that

$$T[W] - \alpha\|W - V\| \leq T[V] \leq T[W] + \alpha\|W - V\|, \quad V, W \in \mathcal{B}(S), \tag{6.2}$$

so that $\|T[W] - T[V]\| \leq \alpha\|V - W\|$, showing that T_α is a contractive operator on the space $\mathcal{B}(S)$ endowed with the maximum norm. Consequently, by Banach's fixed point theorem, there exists a unique function $V_\alpha \in \mathcal{B}(S)$ satisfying $T_\alpha[V_\alpha] = V_\alpha$, that is,

$$e^{\lambda V_\alpha(x)} = \inf_{a\in A(x)} \left[e^{\lambda C(x,a)} \sum_{y\in S} p_{x\,y}(a) e^{\lambda\alpha V_\alpha(y)} \right], \quad x \in S, \quad \alpha \in (0,1). \tag{6.3}$$

Notice now that (6.1) yields that $T_\alpha[0](x) = \inf_{a\in A(x)} C(x,a)$, so that $\|T_\alpha[0]\| \leq \|C\|$. Using (6.2) with V_α and 0 instead of W and V, respectively, it follows that

$$(1-\alpha)\|V_\alpha\| \leq \|C\|. \tag{6.4}$$

In the remainder of the section, the family $\{V_\alpha\}_{\alpha\in(0,1)}$ of fixed points will be used to construct the components of an optimality system, and the idea in the following definition is the essential step in that direction. Throughout the remainder, $\{\alpha_m\} \subset (0,1)$ is a fixed sequence satisfying the following requirements:

$$\alpha_m \nearrow 1 \quad \text{as} \quad m \nearrow \infty \tag{6.5}$$

and

$$\begin{aligned} &\text{for every } x, y \in S, \text{ the following limits exist:} \\ &\lim_{m\to\infty} [V_{\alpha_m}(x) - V_{\alpha_m}(y)] \in [-\infty, \infty] \\ &\lim_{m\to\infty} (1-\alpha_m) V_{\alpha_m}(x) \in [-\|C\|,\ \|C\|], \end{aligned} \tag{6.6}$$

where the last inclusion follows from (6.4).

Definition 6.2. The relation '$\sim$' in the state space S is specified as follows:

$$x \sim y \iff \lim_{m\to\infty} [V_{\alpha_m}(x) - V_{\alpha_m}(y)] \in (-\infty, \infty). \tag{6.7}$$

From this definition it is not difficult to see that '$\sim$' is an equivalence relation, and then it induces a partition of S into equivalence classes. Notice that for $x, y \in S$, (6.6) and Definition 6.2 yield that

$$x \not\sim y \iff \lim_{m\to\infty} [V_{\alpha_m}(x) - V_{\alpha_m}(y)] = \infty \quad \text{or} \quad \lim_{m\to\infty} [V_{\alpha_m}(x) - V_{\alpha_m}(y)] = -\infty; \tag{6.8}$$

moreover,

$$\begin{aligned} &\text{if } x \sim x_1 \text{ and } y \sim y_1 \text{ and } \lim_{m\to\infty} [V_{\alpha_m}(x) - V_{\alpha_m}(y)] = \infty, \\ &\qquad \text{then } \lim_{m\to\infty} [V_{\alpha_m}(x_1) - V_{\alpha_m}(y_1)] = \infty. \end{aligned} \tag{6.9}$$

Definition 6.3. The relation '$\prec$' in the family of equivalence classes determined by the equivalence relation in (6.7) is defined as follows: If $\mathcal{E}$ and $\mathcal{E}'$ are two different equivalence classes, then

$$\mathcal{E} \prec \mathcal{E}' \iff \lim_{m\to\infty} [V_{\alpha_m}(x) - V_{\alpha_m}(y)] = \infty \text{ for some } x \in \mathcal{E}' \text{ and some } y \in \mathcal{E}.$$

By (6.9) this relation is well-defined, whereas (6.8) implies that $\prec$ is a (strict) total order, that is, if $\mathcal{E}$ and $\mathcal{E}'$ are two different equivalences classes, then either $\mathcal{E} \prec \mathcal{E}'$ or $\mathcal{E}' \prec \mathcal{E}$. Moreover, combining the above definition and (6.9), it follows that

$$\mathcal{E} \prec \mathcal{E}' \iff \lim_{m\to\infty}[V_{\alpha_m}(x) - V_{\alpha_m}(y)] = \infty \quad \text{for all } x \in \mathcal{E}' \text{ and all } y \in \mathcal{E}. \tag{6.10}$$

Throughout the remainder,

$$S_1^*, \ldots, S_k^* \text{ are the different equivalence clasess of } S \text{ with respect to '}\sim\text{'} \tag{6.11}$$

where, without loss of generality, the labelling of the equivalence classes is such that

$$S_i^* \prec S_{i+1}^* \quad 1 \le i < k; \tag{6.12}$$

also, the states $x_1, \ldots, x_k$ are fixed and satisfy

$$x_i \in S_i^*, \quad i = 1, 2, \ldots, k. \tag{6.13}$$

Now, for $i \in \{1, 2, \ldots, k\}$, define

$$g_i^* := \lim_{m\to\infty}(1 - \alpha_m)V_{\alpha_m}(x_i), \tag{6.14}$$

and

$$h_i^*(x) = \lim_{m\to\infty}[V_{\alpha_m}(x) - V_{\alpha_m}(x_i)], \quad x \in S_i^*. \tag{6.15}$$

Notice that $g_i^* \in [-\|C\|, \|C\|]$, by (6.4) whereas, observing that $x_i \sim x$ for every $x \in S_i$, from Definition 6.2 it follows that $h_i(x)$ is finite for every $x \in S_i^*$; the above objects S_i^*, g_i^* and $h_i^*(\cdot)$ will be used to build an optimality system for model $\mathcal{M}$.

7. PROOF OF THE EXISTENCE RESULT

In this section it will be verified that an optimality system for model $\mathcal{M}$ exists. With the notation in (6.11) – (6.15), define the sequence of triplets $\mathcal{O}^*$ as follows:

$$\mathcal{O}^* := ((S_1^*, g_1^*, h_1^*), \ldots, (S_k^*, g_k^*, h_k^*)). \tag{7.1}$$

P r o o f o f Theorem 3.3. It will be shown that $\mathcal{O}^*$ specified above is an optimality system for model $\mathcal{M}$. To achieve this goal, the four conditions in Definition 3.1 will be verified.

(i) Since $S_1^*, \ldots, S_k^*$ are the different equivalence classes of S with respect to the equivalence relation in Definition 6.2, those sets S_i^* form a partition of S.

(ii) As already noted, g_i^* is a finite number and $h_i^* \in \mathcal{B}(S_i^*)$. Now let $i < j$ be arbitrary in $\{1, 2, \ldots, k\}$. Recall now that $x_i \in S_i$ and $x_j \in S_j$, by (6.13), and combine Definition 6.3 with (6.10) and (6.12) to obtain that $\lim_{m\to\infty}[V_{\alpha_m}(x_j) - V_{\alpha_m}(x_i)] = \infty$, so that $V_{\alpha_m}(x_j) > V_{\alpha_m}(x_i)$ for m large enough, a fact that leads to

$$g_j^* = \lim_{m\to\infty}(1 - \alpha_m)V_{\alpha_m}(x_j) \ge \lim_{m\to\infty}(1 - \alpha_m)V_{\alpha_m}(x_i) = g_i^*,$$

and then $g_1^* \leq \cdots \leq g_k^*$.

(iii) Setting

$$B^*(x) = \{a \in A(x) \,|\, \sum_{y \in S_1^* \cup \cdots \cup S_i^*} p_{x\,y}(a) = 1\}, \quad x \in S_i^*, \quad i = 1, 2, \ldots, k, \tag{7.2}$$

it will be shown below that $B^*(x)$ is always a nonempty set. To achieve this goal, notice that Assumption 2.1 yields that, for each $\alpha \in (0,1)$, there exists a policy $f_\alpha \in \mathbb{F}$ such that, for every $x \in S$,

$$e^{\lambda V_\alpha(x)} = e^{\lambda C(x, f_\alpha(x))} \sum_{y \in S} p_{x\,y}(f_\alpha(x)) e^{\lambda \alpha V_\alpha(y)}. \tag{7.3}$$

Now, let the sequence $\{\alpha_m\}$ be as in (6.5) and (6.6), and consider the sequence $\{f_{\alpha_m}\} \subset \mathbb{F}$. Recalling that $\mathbb{F}$ is a compact metric space, taking a subsequence (if necessary), without loss of generality it can be assumed that there exists $f^* \in \mathbb{F}$ such that

$$\lim_{m \to \infty} f_{\alpha_m}(x) = f^*(x). \tag{7.4}$$

Next, it will be shown that $f^*(x)$ always belongs to $B^*(x)$, an assertion that will be verified by contradiction. Let $i \in \{1, 2, \ldots, k\}$ and $x \in S_i^*$ be arbitrary but fixed, and *suppose* that

$$p_{x\,z}(f^*(x)) > 0 \quad \text{for some } z \in S_j^* \text{ where } j > i. \tag{7.5}$$

Replacing α by α_m in (7.3) and multiplying both sides of the resulting equality by $e^{-\lambda V_{\alpha_m}(x_i)}$, where x_i is the fixed state in (6.13), direct calculations yield that

$$\begin{aligned} &e^{\lambda(1-\alpha_m)V_{\alpha_m}(x_i) + \lambda[V_{\alpha_m}(x) - V_{\alpha_m}(x_i)]} \\ &\qquad = e^{\lambda C(x, f_{\alpha_m}(x))} \sum_{y \in S} p_{x\,y}(f_{\alpha_m}(x)) e^{\lambda \alpha_m [V_{\alpha_m}(y) - V_{\alpha_m}(x_i)]}, \end{aligned} \tag{7.6}$$

and then

$$\begin{aligned} &e^{\lambda(1-\alpha_m)V_{\alpha_m}(x_i) + \lambda[V_{\alpha_m}(x) - V_{\alpha_m}(x_i)]} \\ &\qquad \geq e^{\lambda C(x, f_{\alpha_m}(x))} p_{x\,z}(f_{\alpha_m}(x)) e^{\lambda \alpha_m [V_{\alpha_m}(z) - V_{\alpha_m}(x_i)]}. \end{aligned} \tag{7.7}$$

Since $x, x_i \in S_i^*$, taking the limit as m goes to ∞ in both sides of this inequality, the continuity of the transition law and the cost function together with (6.6), (6.14), (6.15) and (7.4), lead to

$$e^{\lambda g_i^* + \lambda h_i^*(x)} \geq e^{\lambda C(x, f^*(x))} p_{x\,z}(f^*(x)) e^{\lambda \lim_{m \to \infty} [V_{\alpha_m}(z) - V_{\alpha_m}(x_i)]};$$

since $z \in S_j^*$ and $x_i \in S_i^*$ with $j > i$, *via* (6.10) and (6.12) it follows that

$$\lim_{m \to \infty} [V_{\alpha_m}(z) - V_{\alpha_m}(x_i)] = \infty,$$

and recalling that λ and $p_{x\,z}(f^*(x))$ are positive, the above display yields that $e^{\lambda g_i^* + \lambda h_i^*(x)} \geq \infty$, a contradiction that stems from (7.5). Therefore, $p_{x\,z}(f^*(x)) = 0$ when $z \in S_j^*$ with $j > i$, and it follows that

$$\sum_{y \in S_1^* \cup \cdots \cup S_i^*} p_{x,y}(f^*(x)) = 1,$$

that is,

$$f^*(x) \in B^*(x); \tag{7.8}$$

since $x \in S_i^*$ and $i \in \{1, 2, \ldots, k\}$ were arbitrary in this argument, it follows that $B^*(x)$ is always a nonempty set.

(iv) It will be verified that

$$e^{\lambda(g_i^* + h_i^*(x))} = \inf_{a \in B^*(x)} \left[e^{\lambda C(x,a)} \sum_{y \in S_i^*} p_{x\,y}(a) e^{\lambda h_i^*(y)} \right], \quad x \in S_i^*. \tag{7.9}$$

Let $i \in \{1, 2, \ldots, k\}$ and $x \in S_i^*$ be arbitrary but fixed. Now take an arbitrary action $a \in B^*(x) \subset A(x)$ and notice that (7.2) yields that $p_{x\,y}(a) = 0$ when $y \notin S_1^* \cup \cdots \cup S_i^*$. Using this fact (6.3) implies that, for every positive integer m,

$$e^{\lambda V_{\alpha_m}(x)} \leq e^{\lambda C(x,a)} \sum_{y \in S_1^* \cup \cdots \cup S_i^*} p_{x\,y}(a) e^{\lambda \alpha_m V_{\alpha_m}(y)},$$

and multiplying both sides of this inequality by $e^{-\lambda V_{\alpha_m}(x_i)}$ it follows that

$$e^{\lambda(1-\alpha_m)V_{\alpha_m}(x_i) + \lambda[V_{\alpha_m}(x) - V_{\alpha_m}(x_i)]} \leq e^{\lambda C(x,a)} \sum_{y \in S_1^* \cup \cdots \cup S_i^*} p_{x\,y}(a) e^{\lambda \alpha_m [V_{\alpha_m}(y) - V_{\alpha_m}(x_i)]};$$

recalling that $x_i \in S_i^*$ and using (6.6), (6.14) and (6.15), taking the limit as m goes to ∞ in both sides of the above inequality the following relation is obtained:

$$\begin{aligned} e^{\lambda g_i^* + \lambda h_i^*(x)} \quad &\leq \quad e^{\lambda C(x,a)} \sum_{y \in S_i^*} p_{x\,y}(a) e^{\lambda h_i^*(y)} \\ &\quad + e^{\lambda C(x,a)} \sum_{y \in \cup_{1 \leq j < i} S_j^*} p_{x\,y}(a) e^{\lambda \lim_{m \to \infty} [V_{\alpha_m}(y) - V_{\alpha_m}(x_i)]}. \end{aligned} \tag{7.10}$$

Since

$$\lim_{m \to \infty} [V_{\alpha_m}(y) - V_{\alpha_m}(x_i)] = -\infty \text{ when } y \in S_j^* \text{ with } j < i, \tag{7.11}$$

by (6.10) and (6.12), the positivity of λ yields that the second summation in the above display vanishes, so that

$$e^{\lambda g_i^* + \lambda h_i^*(x)} \leq e^{\lambda C(x,a)} \sum_{y \in S_i^*} p_{x\,y}(a) e^{\lambda h_i^*(y)}$$

and then, since $a \in B^*(x)$ was arbitrary in this argument,

$$e^{\lambda g_i^* + \lambda h_i^*(x)} \leq \inf_{a \in B^*(x)} \left[e^{\lambda C(x,a)} \sum_{y \in S_i^*} p_{x\,y}(a) e^{\lambda h_i^*(y)} \right]. \tag{7.12}$$

To establish the reverse inequality, notice that (7.6) yields that

$$\begin{aligned} &e^{\lambda(1-\alpha_m)V_{\alpha_m}(x_i) + \lambda[V_{\alpha_m}(x) - V_{\alpha_m}(x_i)]} \\ &\qquad \geq e^{\lambda C(x, f_{\alpha_m}(x))} \sum_{y \in S_i^* \cup \cdots \cup S_i^*} p_{x\,y}(f_{\alpha_m}(x)) e^{\lambda \alpha_m [V_{\alpha_m}(y) - V_{\alpha_m}(x_i)]}. \end{aligned}$$

Taking the limit as m goes to ∞, the specifications of g_I^* and $h_i^*(\cdot)$ together with Assumption 2.1 and (7.4) lead to

$$\begin{aligned} e^{\lambda g_i^* + \lambda h_i^*(x)} &\geq e^{\lambda C(x, f^*(x))} \sum_{y \in S_i^*} p_{x\,y}(f^*(x)) e^{\lambda h_i^*(y)} \\ &\quad + e^{\lambda C(x, f_{\alpha_m}(x))} \sum_{y \in \cup_{1 \leq j < i} S_j} p_{x\,y}(f^*(x)) e^{\lambda \lim_{m \to \infty} [V_{\alpha_m}(y) - V_{\alpha_m}(x_i)]} \end{aligned}$$

and then (7.11) and the positivity of λ yield that

$$\begin{aligned} e^{\lambda g_i^* + \lambda h_i^*(x)} &\geq e^{\lambda C(x, f^*(x))} \sum_{y \in S_i^*} p_{x\,y}(f^*(x)) e^{\lambda h_i^*(y)} \\ &\geq \inf_{a \in B^*(x)} \left[e^{\lambda C(x,a)} \sum_{y \in S_i^*} p_{x\,y}(a) e^{\lambda h_i^*(y)} \right] \end{aligned}$$

where the second inequality follows from the inclusion in (7.8). This display and (7.12) together imply that

$$e^{\lambda g_i^* + \lambda h_i^*(x)} = \inf_{a \in B^*(x)} \left[e^{\lambda C(x,a)} \sum_{y \in S_i^*} p_{x\,y}(a) e^{\lambda h_i^*(y)} \right];$$

since $i \in \{1, 2, \ldots, k\}$ and $x \in S_i^*$ are arbitrary, (7.9) follows.

In short, it has been verified that $\mathcal{O}^*$ in (7.1) is an optimality system for $\mathcal{M}$, establishing the conclusion of Theorem 3.3. □

ACKNOWLEDGEMENT

The authors are grateful to the reviewers for their careful reading of the original manuscript, and for their helpful suggestions to improve the content and presentation of the paper.
This work was partially supported by CONACYT under Grant No. 105657.

(Received March 8, 2011)

REFERENCES

[1] A. Arapstathis, V. K. Borkar, E. Fernández-Gaucherand, M. K. Gosh, and S. I. Marcus: Discrete-time controlled Markov processes with average cost criteria: a survey. SIAM J. Control Optim. *31* (1993), 282–334.

[2] P. Billingsley: Probability and Measure. Third edition. Wiley, New York 1995.

[3] R. Cavazos–Cadena and E. Fernández-Gaucherand: Controlled Markov chains with risk-sensitive criteria: average cost, optimality equations and optimal solutions. Math. Method Optim. Res. *43* (1999), 121–139.

[4] R. Cavazos–Cadena and E. Fernández–Gaucherand: Risk-sensitive control in communicating average Markov decision chains. In: Modelling Uncertainty: An examination of Stochastic Theory, Methods and Applications (M. Dror, P. L'Ecuyer and F. Szidarovsky, eds.), Kluwer, Boston 2002, pp. 525–544.

[5] R. Cavazos–Cadena: Solution to the risk-sensitive average cost optimality equation in a class of Markov decision processes with finite state space. Math. Method Optim. Res. *57* (2003), 263–285.

[6] R. Cavazos–Cadena and D. Hernández-Hernández: A characterization of the optimal risk-sensitive average cost in finite controlled Markov chains. Ann. App. Probab., *15* (2005), 175–212.

[7] R. Cavazos–Cadena and D. Hernández-Hernández: A system of Poisson equations for a non-constant Varadhan functional on a finite state space. Appl. Math. Optim. *53* (2006), 101–119.

[8] R. Cavazos–Cadena and F. Salem-Silva: The discounted method and equivalence of average criteria for risk-sensitive Markov decision processes on Borel spaces. Appl. Math. Optim. *61* (2009), 167–190.

[9] G. B. Di Masi and L. Stettner: Risk-sensitive control of discrete time Markov processes with infinite horizon. SIAM J. Control Optim. *38* 1999, 61–78.

[10] G. B. Di Masi and L. Stettner: Infinite horizon risk sensitive control of discrete time Markov processes with small risk. Syst. Control Lett. *40* (2000), 15–20.

[11] G. B. Di Masi and L. Stettner: Infinite horizon risk sensitive control of discrete time Markov processes under minorization property. SIAM J. Control Optim. *46* (2007), 231–252.

[12] W. H. Fleming and W. M. McEneany: Risk-sensitive control on an infinite horizon. SIAM J. Control Optim. *33* (1995), 1881–1915.

[13] F. R. Gantmakher: The Theory of Matrices. Chelsea, London 1959.

[14] D. Hernández-Hernández and S. I. Marcus: Risk-sensitive control of Markov processes in countable state space. Syst. Control Lett. *29* (1996), 147–155.

[15] D. Hernández-Hernández and S. I. Marcus: Existence of risk sensitive optimal stationary policies for controlled Markov processes. Appl. Math. Optim. *40* (1999), 273–285.

[16] A. R. Howard and J. E. Matheson: Risk-sensitive Markov decision processes. Management Sci. *18* (1972), 356–369.

[17] D. H. Jacobson: Optimal stochastic linear systems with exponential performance criteria and their relation to stochastic differential games. IEEE Trans. Automat. Control *18* (1973), 124–131.

[18] S. C. Jaquette: Markov decison processes with a new optimality criterion: discrete time. Ann. Statist. *1* (1973), 496–505.

[19] S. C. Jaquette: A utility criterion for Markov decision processes. Management Sci. *23* (1976), 43–49.

[20] A. Jaśkiewicz: Average optimality for risk sensitive control with general state space. Ann. App. Probab. *17* (2007), 654–675.

[21] M. L. Puterman: Markov Decision Processes. Wiley, New York 1994.

[22] U. G. Rothblum and P. Whittle: Growth optimality for branching Markov decision chains. Math. Oper. Res. *7* (1982), 582–601.

[23] K. Sladký: Successive approximation methods for dynamic programming models. In: Proc. Third Formator Symposium on the Analysis of Large-Scale Systems (J. Beneš and L. Bakule, eds.), Academia, Prague 1979, pp. 171–189.

[24] K. Sladký: Bounds on discrete dynamic programming recursions I. Kybernetika *16* (1980), 526-547.

[25] K. Sladký: Growth rates and average optimality in risk-sensitive Markov decision chains. Kybernetika *44* (2008), 205–226.

[26] K. Sladký and R. Montes-de-Oca: Risk-sensitive average optimality in Markov decision chains. In: Operations Research Proceedings, Vol. 2007, Part III (2008), pp. 69–74.

[27] P. Whittle: Optimization Over Time–Dynamic Programming and Stochastic Control. Wiley, Chichester 1983.

[28] W. H. M. Zijm: Nonnegative Matrices in Dynamic Programming. Mathematical Centre Tract, Amsterdam 1983.

Alfredo Alanís-Durán, Facultad de Ciencias Físico-Matemáticas, Universidad Autónoma de Nuevo León, Avenida Universidad s/n, Ciudad Universitaria, San Nicolás de los Garza NL 66451. México.

e-mail: alfredo.alanisdr@uanl.edu.mx

Rolando Cavazos-Cadena, Departamento de Estadística y Cálculo, Universidad Autónoma Agraria Antonio Narro, Boulevard Antonio Narro 1923, Buenavista, Saltillo COAH 25315. México.

e-mail: rcavazos@uaaan.mx