Time-Average Optimality for Semi-Markov Control Processes with Feller Transition Probabilities

Anna Jaśkiewicz[1] and Andrzej S. Nowak[2]

[1] Instytut Matematyki
Politechnika Wrocławska
50-370 Wrocław, Poland
(e-mail: ajaskiew@im.pwr.wroc.pl)
[2] Wydział Matematyki, Informatyki i Ekonometrii
Uniwersytet Zielonogórski
65-516 Zielona Góra, Poland
(e-mail: a.nowak@wmie.uz.zgora.pl)

Abstract. Semi-Markov control processes with Borel state space and Feller transition probabilities are considered. We prove that under fairly general conditions the two expected average costs: the time-average and the ratio-average coincide for stationary policies. Moreover, the optimal stationary policy for the ratio-average cost criterion is also optimal for the time-average cost criterion.
Keywords: semi-Markov control models, average cost optimality equation.

1 The model

Let X and A be Borel spaces, the state and the action space, respectively. By $A(x)$ we denote the compact set of actions available in state x. Define

$$K := \{(x, a) : x \in X, a \in A(x)\},$$

the set of admissible pairs as a Borel subset of $X \times A$.

If the current state is x and an action $a \in A(x)$ is selected, then the immediate cost of $c_1(x, a)$ is incurred and the system remains in state $x_0 = x$ for a

random time T with the cumulative distribution $G(\cdot|x, a)$ depending only on x and a. The cost of $c_2(x, a)$ per unit time is incurred until the next transition occurs. Afterwards the system jumps to the state $x_1 = y$ according to the probability distribution (*transition law*) $q(\cdot|x, a)$. This procedure repeats itself and yields a trajectory $(x_0, a_0, t_1, x_1, a_1, t_2, \ldots)$ of some stochastic process, where x_n is the state, a_n is the control variable and t_n is the time of the nth transition, $n \geq 0$.

A *control policy* $\pi = \{\pi_n\}$ and a stationary policy $\pi = \{f, f, \ldots\}$ are defined in a usual way. By Π and F we denote the set of all policies and the set of all stationary policies, respectively. Further, we will identify any stationary policy $\pi = \{f, f, \ldots\}$ with $f \in F$.

Let $(\Omega, \mathcal{F})$ be the measurable space consisiting of the sample space $\Omega := (X \times A \times [0, +\infty))^\infty$ and the corresponding product σ-algebra $\mathcal{F}$. Obviously, any policy π, the transition law q, and the conditional cumulative distribution function G of the differences $\{T_{n+1} - T_n\}$ generate the stochastic process $\{x_n, a_n, T_n\}$, $n \geq 0$ on $(\Omega, \mathcal{F})$.

Let E_x^π be the expectation operator with respect to the probability measure P_x^π defined on the product space Ω.

Let $\pi \in \Pi$, $x \in X$ and $t \geq 0$ be fixed. Put

$$N(t) := \max\{n \geq 0 : T_n \leq t\}$$

as the counting process, and

$$\tau(x, a) := \int_0^\infty t P_x^a(dt) = \int_0^\infty t G(dt|x, a) = E_x^a T$$

as the mean holding (sojourn) time. By our assumptions $P_x^\pi(N(t) < \infty) = 1$

We shall consider the two average expected costs:

- the *ratio-average cost*

$$J(x, \pi) := \limsup_{n \to \infty} \frac{E_x^\pi \left(\sum_{k=0}^{n-1} c(x_k, a_k) \right)}{E_x^\pi T_n},$$

- the *time-average cost*

$$j(x, \pi) := \limsup_{t \to \infty} \frac{E_x^\pi \left(\sum_{k=0}^{N(t)} c(x_k, a_k) \right)}{t},$$

where

$$c(x, a) := c_1(x, a) + \tau(x, a) c_2(x, a)$$

for each $(x, a) \in K$.
We impose the following assumptions on the model.

(**B**) *Basic assumptions*:
(i) for each $x \in X$, $A(x)$ is a compact metric space and, moreover, the set-valued mapping $x \mapsto A(x)$ is upper semicontinuous, i.e. $\{x \in X : A(x) \cap B \neq \emptyset\}$ is closed for every closed set B in A;
(ii) the cost function c is lower semicontinuous on K;
(iii) the transition law q is weakly continuous on K, i.e.,

$$\int_X u(y) q(dy|x, a)$$

is continuous function of (x, a) for every bounded continuous function u on X;

(iv) the mean holding time τ is continuous on K, and there exist positive constants b and B such that

$$b \leq \tau(x,a) \leq B$$

for all $(x,a) \in K$;
(v) there exist a constant $L > 0$ and a continuous function $V : X \mapsto [1,\infty)$ such that $|c(x,a)| \leq LV(x)$ for every $(x,a) \in K$;
(vi) the function

$$\int_X V(y)(dy|x,a)$$

is continuous on K.
(**GE**) *Geometric ergodicity assumptions*:
(i) there exists a Borel set $C \subset X$ such that for some $\lambda \in (0,1)$ and $\eta > 0$, we have

$$\int_X V(y)q(dy|x,a) \leq \lambda V(x) + \eta 1_C(x)$$

for each $(x,a) \in K$; V is the function introduced in (**B**, v);
(ii) the function V is bounded on C, i.e.,

$$v_C := \sup_{x \in C} V(x) < \infty;$$

(iii) there exist some $\delta \in (0,1)$ and a probability measure μ concentrated on the Borel set C with the property that

$$q(D|x,a) \geq \delta\mu(D)$$

for each Borel set $D \subset C$, $x \in C$ and $a \in A(x)$.

For any function $u : X \mapsto R$ define the V-norm

$$\|u\|_V := \sup_{x \in X} \frac{|u(x)|}{V(x)}.$$

By L_V^∞ we denote the Banach space of all Borel measurable functions u for which $\|u\|_V$ is finite.

Let L_V denote the subset of L_V^∞ consisting of all lower semicontinuous functions.

Under (**GE**) the embedded state process $\{x_n\}$ governed by a stationary policy is a positive recurrent aperiodic Markov chain and for each stationary policy f, there exists a unique invariant probability measure, denoted by π_f (see Theorem 11.3.4 and page 116 in [Meyn and Tweedie, 1993]). Moreover, by Theorem 2.3 in [Meyn and Tweedie, 1994], $\{x_n\}$ is V-uniformly ergodic. Thi results in the following

$$J(f) := J(x,f) = \frac{\int_X c(x,f(x))\pi_f(dx)}{\int_X \tau(x,f(x))\pi_f(dx)}$$

for every $f \in F$.

We also make two additional assumptions on the sojourn time T.

(**R**) *Regularity condition*:
there exist $\epsilon > 0$ and $\beta < 1$ such that

$$P_x^a(T \leq \epsilon) \leq \beta$$

for all $x \in C$ and $a \in A(x)$.
(**I**) *Uniform integrability condition*:

$$\lim_{t\to\infty} \sup_{x\in C} \sup_{a\in A(x)} P_x^a(T > t) = 0.$$

For further and broad discussion of the assumptions the reader is referred to [Jaśkiewicz, 2001] and [Ross, 1970].

2 Main results

In this section we present two new theorems on SMCPs with Borel state spaces. Theorem 1 concerns the existence of the optimal stationary policy for the ratio-average criterion. The proof combines some ideas and tools used in [Jaśkiewicz, 2001].

For the ε-perturbed SMCPs, we prove that the associated with them the average cost optimality equation has a solution.

Next, taking into account slightly modified solutions, we obtain a certain optimality inequality, which is enough to obtain an average optimal policy. It is worth pointing out that compared with previous work [Jaśkiewicz, 2001] in the limit passage we need to use of Fatou's lemma for weakly convergent measures [Serfozo, 1982].

Theorem 1. Assume (**B**, **GE**). There exist a constatant g^*, a function $h_* \in L_V$ and $f^* \in F$ such that

$$\begin{aligned} h_*(x) &\geq \min_{a\in A(x)} \left[c(x,a) - g^*\tau(x,a) + \int_X h_*(y)q(dy|x,a) \right] \\ &= c(x,f^*(x)) - g^*\tau(x,f^*(x)) + \int_X h_*(y)q(dy|x,f^*(x)) \end{aligned} \tag{1}$$

for all $x \in X$. Moreover, f^* is an average optimal policy and g^* is optimal cost with respect to the ratio-average criterion, i.e.,

$$g^* = \inf_{\pi\in\Pi} J(x,\pi) = J(f^*)$$

for every $x \in X$.

Theorem 2 deals with the equivalence of the two expected average cost criteria for SMCPs with Feller transition probabilities. Related result under the strong continuity of $q(\cdot|x,a)$ in $a \in A(x)$ is given in [Jaśkiewicz, 2004].

To obtain the mentioned equivalence we use two inequalities as the point of departure. Using them we define a supermartingale and submartingale, and then by Doob's theorem we obtain the equality of the two optimal costs according to the ratio-average and time-average cost criteria. To apply the optional sampling theorem we have to prove the uniform integrability of the supermartingale and submartingale involved. This issue is studied in [Jaśkiewicz, 2004]. The whole analysis relies on dealing with the consecutive returns of the process (induced by q, an arbitrary π, and the cumulative distribution G) to the small set C.

Theorem 2. Assume (**B**, **GE**, **R**, **I**). Then
(a) $g^* = \inf_{\pi \in \Pi} j(x, \pi)$;
(b) $j(x, f) = J(x, f)$ for any $f \in F$.

References

[Jaśkiewicz, 2001]A. Jaśkiewicz. An approximation approach to ergodic semi-markov control processes. *Math. Methods Oper. Res.*, pages 1–19, 2001.

[Jaśkiewicz, 2004]A. Jaśkiewicz. On the equivalence of two expected average cost criteria for semi-markov control processes. *Math. Oper. Res.*, pages 326–338, 2004.

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

[Meyn and Tweedie, 1994]S.P. Meyn and R.L. Tweedie. Computable bounds for geometric convergence rates of markov chains. *Ann. Appl. Prob.*, pages 981–1011, 1994.

[Ross, 1970]S.M. Ross. *Applied Probability Models with Optimization Applications.* Holden-Day, San Francisco, 1970.

[Serfozo, 1982]R. Serfozo. Convergence of lebesgue integrals with varying measures. *Sankhya, Ser. A*, pages 380–402, 1982.