J. Math. Anal. Appl. 464 (2018) 152–163

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Solutions of the average cost optimality equation for Markov decision processes with weakly continuous kernel: The fixed-point approach revisited

Óscar Vega-Amaya [1]

Departamento de Matemáticas, Universidad de Sonora, Sonora, Mexico

A R T I C L E I N F O

Article history:
Received 18 April 2017
Available online 4 April 2018
Submitted by J.A. Filar

Keywords:
Markov decision processes
Average cost optimality equation
Lyapunov stability condition
Banach fixed-point theorem

A B S T R A C T

This paper shows the existence of lower semicontinuous solutions of the average cost optimality equation for Markov decision processes with Borel spaces, possible unbounded cost function and weakly continuous transition kernel. This is done imposing a growth condition on the cost function, a Lyapunov stability condition on the transition kernel and a set of standard compactness-continuity conditions. The solution of the average cost optimality equation is obtained by means of the Banach fixed-point theorem.

1. Introduction

The study of the average cost (AC) criterion goes back to the beginning of the Markov decision processes theory. In fact, Bellman [4] himself studied this criterion but as an auxilary tool to analize the growth rate of the optimal costs in finite horizon. Since then, the AC optimal control problem has been extensively studied, and it has shown to be a difficult problem with a number of counter-examples [37, Ch. 4]. For a detailed survey on approaches and early results see reference [3] and the monographs [22,23] for the progress made up to the nineties for models with Borel spaces and unbounded costs; for more recent contribution see references [10,11,14,16,17,20,26,27,30,40,42].

The different approaches devised to study the AC optimal control problem need of some "regularity" conditions to work, and most of them guarantee an AC optimal stationary policy by establishing either the average cost optimality equation (ACOE) or the average cost optimality inequality (ACOI). In particular, the validity of the ACOE is tied to certain recurrence/ergodicity properties and some compacteness-continuity conditions [3,5–7,9,25]. Concerning to the ACOI see the comments below.

E-mail address: ovega@mat.uson.mx.

[1] *Postal address:* Luis Encinas y Rosales s/n, Col. Centro, C. P. 83000, Hermosillo, Sonora, Mexico.

https://doi.org/10.1016/j.jmaa.2018.03.077

For discrete state spaces, the mentioned recurrence conditions ask that the mean return times to a distinguished state satisfy a boundness condition. (This property can be stated in several equivalent forms, for instance, as a Lyapunov stability condition [5,9].) However, this approach does not work for the uncountable state space case since it tipically involves events of probability zero, thus requiring stronger ergodicity conditions than the discrete counterpart needs; see, for instance, references [3,25] for bounded cost functions, and references [19,30,40] for unbounded cost functions in the state variable that does not grow faster than a function satisfying a Lyapunov stability condition.

The paper by Gordienko and Hernández-Lerma [19] is, to the best author's knowledge, the first work dealing with the existence of solutions of the ACOE for models with Borel spaces and possibly unbounded costs. The stability condition used in this latter reference is an adaptation to Markov decision processes of a stability condition introduced by Karthashov [32,33] to prove weighted geometric ergodicity for (non-controlled) Markov chains with general spaces. Gordienko and Hernández-Lerma [19] first show that their stability condition guarantees a uniform convergence rate on the class of stationary policies–in a weighted norm–and then, after adding a strong continuity assumption, they show the existence of a solution of the ACOE using the vanishing discount factor approach.

It was noticed in reference [40] that the core of the stability condition in [19] is a contraction property, and then it was used to obtain solutions of the ACOE by means of the classical Banach fixed-point theorem but applied to an operator closely related to the standard dynamic programming operator. This fixed-point approach rids of a number of assumptions of [19] and at the same time allows simplified proofs; in fact, reference [40] does not use, at least explicitly, the weighted geometric ergodicity and its proofs are practically self-contained. See references [18,35] for recent applications in wireless control systems of this fixed-point approach and [41] for an extension to zero-sum semi-Markov games.

The present paper extends the fixed-point approach introduced in reference [40] for models with a strongly continuous transition kernel to models with a weakly continuous transition kernel. More precisely, it is shown the existence of an AC optimal stationary policy by proving the existence of a lower semicontinuous solution of the ACOE (Theorem 3.7). The above is done in a setting provided by three kinds of assumptions: (i) a growth condition on the cost function (Assumption 3.1); (ii) a Lyapunov stability condition on the transition kernel (Assumption 3.2); and (iii) a standard set of compactness and continuity conditions, which includes the weak continuity of the transition kernel (Assumption 3.5). After that, strengthening the compacteness-continuity Assumption 3.5, it is shown the existence of a continuous solution to the ACOE (Corollary 3.8).

It is important to mention that Jaśkiewicz [28] have already extended the fixed-point approach of reference [40] to semi-Markov decision processes under conditions similar to Assumptions 3.1, 3.2 and 3.5, but the analysis in the present paper relies on more simpler arguments and it uses weaker assumptions; in fact, the proofs of the main results, namely, Theorem 3.7 and Corollary 3.8 mentioned previously, follow from truly elementary arguments. Moreover, the approach of the present paper directly extends to semi-Markov decision processes, and perhaps also to zero-sum semi-Markov games [29]. For a comprehensive account of zero-sum stochastic games see the recent survey [31].

It is also worth mentioning that "fixed-point arguments" were previously used in average cost Markov decision processes but restricted to bounded cost functions and under a very strong ergodicity assumption–see, for instance, references [3,25] or [21, Lemma 3.5 and Comments 3.7, pp. 59 and 61]. On the other hand, references [12,13] developed a quite different approach that could be qualified as a "fixed-point approach" to the ACOE. These latter references show the validity of the ACOE reducing the average cost optimal control problem into a discounted one. The mentioned reduction is accomplished under assumptions that turn out to be very restrictive, even for denumerable state spaces.

As mentioned paragraphs above, the ACOI also yields AC optimal stationary policies. This was noticed by Sennott [39] and she proved the validity of such inequality using the so-called vanishing discount factor approach under weaker conditions than those required for the ACOE [8]. This renewed the interest on the