J. Math. Anal. Appl. 420 (2014) 1337–1353

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Robust Markov control processes

Anna Jaśkiewicz [a,*], Andrzej S. Nowak [b]

[a] *Institute of Mathematics and Computer Science, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland*
[b] *Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Podgórna 50, 65-246 Zielona Góra, Poland*

ARTICLE INFO

Article history:
Received 2 October 2013
Available online 19 June 2014
Submitted by U. Stadtmueller

Keywords:
Robust control
Fatou's lemma
A generalised Tauberian relation
Optimality inequality

ABSTRACT

We examine the average minimax control problem on a general state space with an unbounded cost function. The controller is uncertain about his model in the sense that he deals with a range of alternative models as possibly true. The aim of this paper is to establish the optimality inequality and to provide a minimax strategy for the controller. A new feature in our approach is an application of a generalised Tauberian relation that, in turn, allows us to impose relatively weak assumptions.

1. Introduction

This paper addresses the problem of minimax control on a Borel state space with unbounded cost functions and weakly continuous transition probabilities. In contrast to the usual dynamic programming control problem with a single decision-maker, we assume here that there are two decision-makers: the controller himself and his opponent known also as nature. The task of nature is to aid the controller to design a strategy that alleviates its sensitivity to ambiguity of the underlying Markov process. The ambiguity is modelled by a sequence of decision rules of the opponent that influence both on the transition probabilities of the process and costs payed by the controller. Therefore, the controller is unaware of the probability model and he takes into account instead a class of approximate models carefully selecting a minimax strategy that assures the best performance in the worst achievable case. The minimax control problems on Borel state spaces have been already considered in [9,15,18,20,22]. They are sometimes referred to as *games against nature* or *robust Markov control processes*. Such models can also be recognised as a special class of two-person dynamic games, in which the second player obtains information on the current state and the action chosen by the controller.

---

[*] Corresponding author.
E-mail addresses: anna.jaskiewicz@pwr.edu.pl (A. Jaśkiewicz), a.nowak@wmie.uz.zgora.pl (A.S. Nowak).

http://dx.doi.org/10.1016/j.jmaa.2014.06.028
0022-247X/© 2014 Published by Elsevier Inc.

A standard application of the minimax approach is to control systems that depend on unknown parameters. Then, a common form of the evolution of the system's state process $\{x_t\}$ is given by

$$x_{t+1} = \Psi(x_t, a_t, b_t, \xi_t), \quad t = 0, 1, \ldots, \tag{1}$$

where $\{a_t\}$, $\{b_t\}$ are strategies for the controller and nature, $\{\xi_t\}$ is a sequence of random disturbances and $\Psi(\cdot)$ is usually a continuous function. To be more concrete, let us describe an approach suggested by Hansen and Sargent [11] in macroeconomics. Namely, they use entropy as a distance measure to calibrate the model uncertainty set. In other words, the model set consists of those models whose relative entropy from the nominal model is bounded by a specified value. Then, a Lagrange multiplier theorem is utilised to covert the entropy constraint into a penalty on perturbations from the model, which leads to the study of a system given in (1). The reader is also referred to Section 5 in [18], where a particular example is considered. Other models, e.g., with an application to the theory of optimal growth, are given in Examples 1 and 2 in this paper. Another problem concerns designing a replacement policy for a line of machines. Clearly, this problem can be well modelled by a Markov decision process with states describing ageing phases and actions representing either various repair or replacement alternatives. Although the parameters for this model can be estimated from the historical data, it is seldom done because of the following reasons. Firstly, one can rarely have an access to enough historical data. Secondly, there is inherent uncertainty in future fluctuation for the price (cost) of new equipment and thirdly, one cannot adequately assess the probability of a machine breaking down at the given age. Therefore, this example illustrates the need for developing criteria that address parameter uncertainty. Finally, it is worthy to emphasise that transition probabilities and cost functions estimated from historical data do not work in practice. This is because the optimal policy is pretty sensitive to any perturbations in the transition probability. Hence, ignorance of perturbation errors can lead to serious degradation in performance (see [7,15,25]).

In this work we study both the expected discounted cost criterion and the expected average cost criterion per unit time. The both aforementioned performance measures have been extensively considered in the literature. In the classical papers on stochastic dynamic programming (Markov decision processes) it is assumed that the payoffs are bounded or bounded from below, see for example [3,4,13,28,31]. Unbounded returns, on the other hand, are mainly treated with the aid of the weighted supremum norm [14]. Then, the expected discounted cost case is usually solved via the Banach contraction principle, whereas the expected average cost case is studied under some geometric ergodicity conditions and "vanishing discount factor" approach that lead to the optimality equation. Similar techniques were also applied to zero-sum stochastic games (see [17,18,21,26] and references cited therein) and robust Markov control processes with unbounded payoffs [9].

In this paper we propose a *new set* of assumptions that allow to establish the optimality inequality and to utilise a Tauberian relation. Although our requirements correspond somehow to the ones used by Sennott [33], Ritt and Sennott [29] and Schäl [32], they are weaker and the proofs contain novel elements. Firstly, we do not confine as in [29,32,33] to the non-negative costs and secondly, we apply a generalised Tauberian relation that allows us to consider a robust Markov control processes.

The remaining part of the paper is organised as follows. In the next section we formulate basic definitions and the Fatou lemma that is used later in the proof of Theorem 2. The robust Markov control model is described in Section 3. The next two sections are devoted to a formulation of main results for discounted and average cost cases. In Section 6, we compare our conditions to the ones used in the literature and give an example for the average cost case. Finally, Appendix A contains a crucial result for this paper, namely a generalised Tauberian relation.

2. Preliminaries

Let R be the set of all real numbers, $N = \{1, 2, ...\}$ and $N_0 = N \cup \{0\}$. A Borel space is a non-empty Borel subset of a complete separable metric space endowed with the Borel σ-algebra $\mathcal{B}(Y)$ and the relative