Stationary policies with Markov partition property

John E. Goulionis *

Dimitrios I. Stengos

George Tzavelas

Department of Statistics & Insurance Science
University of Piraeus
80 Karaoli & Dimitriou Street
18534 Piraeus
Greece

Abstract

This paper treats the infinite horizon discounted cost control problem for *partially observable Markov decision processes* (POMDPs). Sondik [14] studied the class of finitely transient policies and showed that their value functions over an infinite time horizon are piecewise linear (p.w.l) and can be computed exactly by solving a system of linear equations. However the condition for finite transience is stronger than is needed to ensure p.w.l value functions. In this paper we introduce alternatively the class of periodic policies whose value functions turn out to be also p.w.l. Moreover we examine a more general condition than finite transience and periodicity. We implement these ideas in a replacement problem under Markovian deterioration, investigate for periodic policies and give numerical examples.

1.   Introduction

Partially Observable Markov Decision Processes (POMDPs) are models which provide a powerful framework for decision-theoretic planning of actions. Unfortunately, the computational burden associated with solving POMDPs is overwhelming, precluding their application to problems of practical size. Sondik [9] reviews a policy iteration algorithm for infinite horizon POMDP problems. The two main steps in policy iteration is value determination and policy improvement. For an arbitrary infinite horizon

E-mail: `sondi@otenet.gr`