Automatica 77 (2017) 219–229

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Optimal strategies for impulse control of piecewise deterministic Markov processes ☆

Benoîte de Saporta [a,c], François Dufour [b,c], Alizée Geeraert [b,c]

[a] Université de Montpellier IMAG CNRS UMR 5149, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France
[b] Université de Bordeaux, IMB CNRS UMR 5251, 351 cours de la Libération, 33405 Talence Cedex, France
[c] INRIA Bordeaux Sud Ouest team CQFD, 351 cours de la Libération, 33405 Talence Cedex, France

ARTICLE INFO

Article history:
Received 25 March 2016
Received in revised form
26 July 2016
Accepted 17 October 2016

Keywords:
Impulse control
Piecewise deterministic Markov processes
Optimal strategy
Long-run average cost
Dynamic programming

ABSTRACT

This paper deals with the general discounted impulse control problem of a piecewise deterministic Markov process. We investigate a new family of ϵ-optimal strategies. The construction of such strategies is explicit and only necessitates the previous knowledge of the cost of the no-impulse strategy. In particular, it does not require the resolution of auxiliary optimal stopping problem or the computation of the value function at each point of the state space. This approach is based on the iteration of a single-jump-or-intervention operator associated to the piecewise deterministic Markov process.

1. Introduction

The aim of this paper is to propose a new family of ϵ-optimal strategies for the impulse control problem of piecewise deterministic Markov processes (PDMPs) defined by O.L.V. Costa and M.H.A. Davis in Costa and Davis (1989). We consider the infinite horizon expected discounted impulse control problem where the controller instantaneously moves the process to a new point of the state space at some controller specified time.

Piecewise deterministic Markov processes have been introduced by M.H.A. Davis in Davis (1984) and Davis (1993) as a general class of stochastic hybrid models. These processes have two variables: a continuous one representing the physical parameters of the system and a discrete one which characterizes the regime of operation of the physical system and/or the environment. The process depends on three local characteristics: the flow, the jump intensity and the Markov kernel. The path of a PDMP consists of deterministic trajectories punctuated by random jumps. Starting from a point of the state space, the PDMP follows a deterministic trajectory determined by the flow, until the first jump time. This time is drawn either in a Poisson like fashion following the jump intensity or deterministically when the process hits the boundary of the state space. The new position and regime of the PDMP are selected by the Markov kernel. Then the process follows again a deterministic trajectory until the next jump time and so on. There are many and diverse applications of PDMPs for example in queuing or inventory systems, insurance, finance, maintenance models (Bäuerle & Rieder, 2011; Dassios & Embrechts, 1989; Davis, 1993) or in data transmission (Chafaï, Malrieu, & Paroux, 2010) and in biology (Doumic, Hoffmann, Krell, & Robert, 2015; Pakdaman, Thieullen, & Wainrib, 2010). The interested reader can also refer to de Saporta, Dufour, and Zhang (2015) for some applications in reliability.

Impulse control corresponds to the following situation: the process runs until a controller decides to intervene by instantaneously moving the process to some new point of the state space. Then, restarting at this new point, the process runs until the next intervention and so on. Many authors have considered impulse control for PDMPs, either by variational inequality (Dempster & Ye, 1995; Gątarek, 1992; Lenhart, 1989) or by value improvement (Costa & Davis, 1989). The simplest form of impulse control is optimal stopping, where the decision maker selects only one intervention time when the process is stopped. Optimal stopping for PDMPs has been

☆ This work was supported by the French National Agency of Research (ANR), grant PIECE, ANR-12-JS01-0006. The material in this paper was presented at the 28th European Conference on Operational Research, July 3–6 2016, Poznan, Poland. This paper was recommended for publication in revised form by Associate Editor Oswaldo Luiz V. Costa under the direction of Editor Richard Middleton.

E-mail addresses: benoite.de-saporta@umontpellier.fr (B. de Saporta), dufour@math.u-bordeaux1.fr (F. Dufour), alizee.geeraert@inria.fr (A. Geeraert).

http://dx.doi.org/10.1016/j.automatica.2016.11.039
0005-1098/

studied in Costa and Davis (1988), de Saporta, Dufour, and Gonzalez (2010) and Gugerli (1986).

For a general infinite horizon expected discounted impulse control problem, a strategy consists in two sequences of controller-specified random variables defining the intervention times and new starting points of the process. Solving this problem involves finding a strategy that minimizes the expected sum of discounted running and intervention costs up to infinity. The minimal cost is called the value function. In general, optimal strategies do not exist. Instead, one considers ϵ-optimal strategies whose cost differs from the value function at most of ϵ.

The analysis of ϵ-optimal controls is interesting in its own right independently to the fact that the optimal control problem has an optimal solution. Indeed, from a theoretical point of view, PDMPs have been extensively studied. There exist several techniques to analyze such optimization problems that are known to be very efficient for establishing different mathematical properties (such as existence of optimal policies, smoothness of the value function, sufficiency of sub-classes of particular policies, etc.). However, the problem of solving explicitly an optimal control problem for a PDMP remains a critical issue. Except for very few specific models, the determination of an optimal policy and the value function is an extremely difficult problem to tackle. The standard approach for solving an optimal control problem for a PDMP is to develop numerical methods to get quasi-optimal solutions. This topic is, therefore, of crucial importance to demonstrate the practical interest of PDMP as a powerful modeling tool. Thus, it is important to obtain computable ϵ-optimal control. The crucial point is to derive an ϵ-optimal control that can be explicitly computed, generally in terms of the value function. The next step, that will be the objective of a next work, consists in a numerical approximation of such strategies.

There exists an extensive literature related to the study of the optimality equation associated to expected discounted control problems but few works are devoted to the characterization of ϵ-optimal strategies. The objective of the paper is to explicitly construct such strategies. An attempt in this direction has been proposed in Costa and Davis (1989, section 3.3). One step of their approach consists in solving an optimal stopping problem which makes this technique quite difficult to implement. Furthermore the knowledge of the optimal value function is required.

We propose a construction of an ϵ-optimal strategy which necessitates only the knowledge of the cost of the non-impulse strategy and without solving technical problems preliminary. This construction is based on the iteration of a single-jump-or-intervention operator associated to the PDMP. It builds on the explicit construction of ϵ-optimal stopping times developed in Gugerli (1986) for the optimal stopping problem. However, for the general impulse control problem, one must also optimally choose the new starting points of the process, which is a significant source of additional difficulties. It is important to emphasize that our method has the advantage of being constructive with regard to other works in the literature on impulse control problem.

This work is also the first step toward a tractable numerical approximation of ϵ-optimal strategies. A numerical method to compute the value function is proposed in de Saporta and Dufour (2012). It is based on the quantization of an underlying discrete-time Markov chain related to the continuous process and path-adapted time discretization grids. Discretization of ϵ-optimal strategies will be the object of a future work.

The paper is organized as follows. In Section 2 we recall the definition of a PDMP and state the impulse control problem under study. In Section 3, we construct a sequence of approximate value functions. In Section 4, we build an auxiliary process corresponding to an explicit family of strategies and we show that the cost of the controlled trajectories corresponds to the approximate value function built in Section 3. Technical details are gathered in the Appendix.

2. Impulse control problem of PDMP

We introduce first some standard notation before giving a precise definition of a piecewise deterministic Markov processes (PDMP) and of our impulse control problem.

For $a, b \in \mathbb{R}$, $a \wedge b = \min(a, b)$ is the minimum of a and b. By convention, set $\inf \varnothing = \infty$. Let $\mathcal{X}$ be a metric space with distance $d_{\mathcal{X}}$. For a subset A of $\mathcal{X}$, ∂A is the boundary of A and $\bar{A}$ its closure. We denote $\mathfrak{B}(\mathcal{X})$ the Borel σ-field of $\mathcal{X}$ and $\mathbf{B}(\mathcal{X})$ the set of real-valued, bounded and measurable functions defined on $\mathcal{X}$. For any function $w \in \mathbf{B}(\mathcal{X})$, we write C_w for the upper bound of $|w|$, that is $C_w = \sup_{x\in\mathcal{X}} |w(x)|$. For a Markov kernel P on $(\mathcal{X}, \mathfrak{B}(\mathcal{X}))$ and functions w in $\mathbf{B}(\mathcal{X})$, for any $x \in \mathcal{X}$, set $Pw(x) = \int_{\mathcal{X}} w(y)P(x, dy)$.

2.1. Definition of PDMP

Let M be the finite set of the possible regimes or modes of the system. For all modes m in M, let E_m be an open subset of $\mathbb{R}^d$ endowed with the usual Euclidean norm $|\cdot|$. Set $E = \{(m, \zeta), m \in M, \zeta \in E_m\}$. Define on E the following distance, for $x = (m, \zeta)$ and $x' = (m', \zeta') \in E$,

$$|x - x'| = |\zeta - \zeta'|\mathbb{1}_{\{m=m'\}} + \infty\mathbb{1}_{\{m\neq m'\}}.$$

A piecewise deterministic Markov process on the state space E is determined by three local characteristics:

- the flow $\Phi(x, t) = (m, \Phi_m(\zeta, t))$ for all $x = (m, \zeta)$ in E and $t \geq 0$, where $\Phi_m : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^d$ is continuous such that $\Phi_m(\cdot, t+s) = \Phi_m(\Phi_m(\cdot, t), s)$, for all $t, s \in \mathbb{R}^+$. It describes the deterministic trajectory between jumps. We set $t^*(x)$ the time the flow takes to reach the boundary of E when it starts from $x = (m, \zeta)$:
  $$t^*(x) = \inf\{t > 0 : \Phi_m(\zeta, t) \in \partial E_m\}.$$
- the jump intensity $\lambda : \bar{E} \to \mathbb{R}^+$ is a measurable function and has the following integrability property: for any $x = (m, \zeta)$ in E, there exists $\epsilon > 0$ such that
  $$\int_0^{\epsilon} \lambda(m, \Phi_m(\zeta, t))dt < +\infty.$$
  For all $x = (m, \zeta)$ in E and $t \in [0, t^*(x))$, we set
  $$\Lambda(m, \zeta, t) = \int_0^t \lambda(m, \Phi_m(\zeta, s))ds. \qquad (1)$$
- the Markov kernel Q on $(\bar{E}, \mathfrak{B}(\bar{E}))$ is the transition measure of the process. It selects the new location after a jump. It satisfies for all $x \in \bar{E}$, $Q(x, \{x\} \cup \partial E) = 0$: each jump is made in E and changes the location and/or the mode of the process.

It can be shown (Davis, 1993, section 25) that there exists a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \{\mathbb{P}_x\}_{x\in E})$ on which a process $\{X_t\}$ can be defined as a strong Markov process. The process $\{X_t\}$ has two components $X_t = (m_t, \zeta_t)$, the first one m_t is called the mode or the regime and the second one ζ_t is the so-called Euclidean variable. The motion of this process can be defined iteratively as follows. Starting at an initial point $X_0 = (m_0, \zeta_0) \in M \times E_{m_0}$, the first jump time T_1 is determined by

$$\mathbb{P}_{(m_0,\zeta_0)}(\{T_1 > t\}) = \mathrm{e}^{-\Lambda(m_0,\zeta_0,t)}\mathbb{1}_{\{t<t^*(m_0,\zeta_0)\}}. \qquad (2)$$

On $[0, T_1)$, the process $\{X_t\}$ follows the deterministic trajectory $m_t = m_0$, $\zeta_t = \Phi_{m_0}(\zeta_0, t)$. At the random time T_1, a jump occurs. It can produce either a discontinuity in the Euclidean variable ζ_t and/or change of mode. The process restarts at a new mode and/or position $X_{T_1} = (m_{T_1}, \zeta_{T_1})$, according to the distribution $Q((m_0, \Phi_{m_0}(\zeta_0, T_1)), \cdot)$. An inter jump time $T_2 - T_1$ is then selected in a similar way to Eq. (2), and on the interval $[T_1, T_2)$, the process