Available online at www.sciencedirect.com
ScienceDirect
automatica
Automatica 44 (2008) 177–184
www.elsevier.com/locate/automatica

Brief paper

Discrete-time drift counteraction stochastic optimal control: Theory and application-motivated examples☆

I.V. Kolmanovsky[a,*], L. Lezhnev[b], T.L. Maizenberg[c]

[a] Powertrain Control Research and Advanced Engineering, Ford Motor Company, Dearborn, MI, USA
[b] Russian Central Automotive and Engine Research Institute (NAMI), Moscow, Russia
[c] Moscow State Mining University, Moscow, Russia

Received 27 October 2006; received in revised form 28 April 2007; accepted 17 June 2007
Available online 5 September 2007

Abstract

We develop stochastic optimal control results for nonlinear discrete-time systems driven by disturbances modeled by a Markov chain. A characterization and a computational procedure for a control law which maximizes a cost functional, related to expected time-to-violate specified constraints or to expected total yield before constraint violation occurs, are discussed. Such an optimal control law may be viewed as providing drift counteraction and is, therefore, referred to as drift counteraction stochastic optimal control. Two simulation examples highlight opportunities for applications of these results to hybrid electric vehicle (HEV) powertrain management and to oil extraction.

Keywords: Stochastic control; Dynamic programming; Constrained systems; Automotive applications; Hybrid electric vehicles

1. Introduction

Consider a discrete-time system

$$x(t+1) = f(x(t), v(t), w(t)), \tag{1}$$

where $x(t)$ is the state vector, $v(t)$ is the control vector, $w(t)$ is the vector of measured disturbances, and t is an integer, $t \in Z^+$. The system has control constraints which are expressed in the form $v(t) \in U$, where U is a given set.

The behavior of $w(t)$ is modeled by a Markov chain (Dynkin & Yushkevich, 1967) with a finite number of states $w(t) \in W = \{w^j, j \in J\}$. The transition probability from $w(t) = w^i \in W$ to $w(t+1) = w^j \in W$ is denoted by $P(w^j|w^i, \bar{x})$, and in our treatment of the problem we allow this transition probability to depend on the state $x(t) = \bar{x}$.

Our objective is to determine a control function $u(x, w)$ such that, with $v(t) = u(x(t), w(t))$, a cost functional of the form

$$J^{x_0,w_0,u} = E_{x_0,w_0} \sum_{t=0}^{\tau^{x_0,w_0,u}(G)-1} g(x(t), v(t), w(t)) \tag{2}$$

is maximized. Here, $\tau^{x_0,w_0,u}(G) \in Z^+$ denotes the first time instant the trajectory of $x(t)$ and $w(t)$, denoted by $\{x^u, w^u\}$, resulting from the application of the control $v(t) = u(x(t), w(t))$, exits a prescribed compact set G. See Fig. 1. The specification of the set G reflects the constraints existing in the system. Note that $\{x^u, w^u\}$ is a random process, $\tau^{x_0,w_0,u}(G)$ is a random variable, and $E_{x_0,w_0}[\cdot]$ denotes the expectation conditional to initial values of x and w, i.e., $x(0) = x_0$, $w(0) = w_0$. When clear from the context, we omit the subscript and square brackets after E.

Note that if $g = 1$, then $J^{x_0,w_0,u} = E[\tau^{x_0,w_0,u}(G)]$, and our objective is to find a control law which maximizes the expected time of exiting G (or, equivalently, the expected time-to-violate specified constraints). For instance, for a vehicle using an adaptive cruise control to follow another, randomly accelerating and decelerating vehicle, the control objective may be to keep the distance to the lead vehicle within specified limits for as long as possible with only very gradual (small) and smooth accelerations and decelerations.

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor George Yin under the direction of Editor Ian Petersen.

* Corresponding author. Tel.: +1 313 845 10 40.
E-mail addresses: ikolmano@ford.com (I.V. Kolmanovsky), levalezhnev@hotmail.com (L. Lezhnev).

0005-1098/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2007.06.002

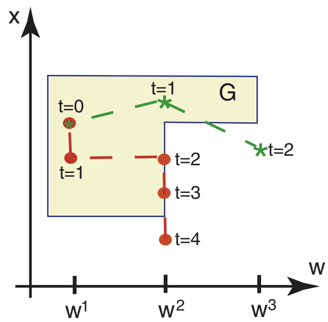

Fig. 1. The set G and two trajectories, $\{x^u, w^u\}$, exiting G at random time instants due to a random realization of $w(t)$. Here $W = \{w^1, w^2, w^3\}$.

For a more general g, our objective may be interpreted as maximizing expected cumulative yield before the combined trajectory, $\{x^u(t), w^u(t)\}$, is forced outside of G. The regions where larger instantaneous yield is attained may also be regions where exiting G is more likely due to the properties of the transition probabilities. As one example, operating a machine harder may yield larger instantaneous yield, but it may also cause earlier failure of the machine (wherein 'working' and 'failed' states of the machine may be modeled as two states in the Markov chain).

For continuous-time systems, under an assumption that $w(t)$ is a Wiener or a Poisson process, it can be shown (Afanas'ev, Kolmanovskii, & Nosov, 1996) that determining an optimal control in this kind of a problem reduces to solving a non-smooth partial-differential equation (PDE). For instance, for a first order stochastic system, $\mathrm{d}x = (v - w_0)\,\mathrm{d}t + \sigma \cdot \mathrm{d}w$, where w_0 is a constant, w is a standard Wiener process, the control v satisfies $|v| \leqslant \bar{v}$, and $g = 1$, this PDE has the form

$$\frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + \frac{\partial V}{\partial x}(-w_0) + \left|\frac{\partial V}{\partial x}\right| \bar{v} + 1 = 0.$$

The boundary conditions for this PDE are $V(x) = 0$ for $x \in \partial G$, where ∂G denotes the boundary of G. The optimal control has the form

$$v = \bar{v}\,\mathrm{sign}\left(\frac{\partial V}{\partial x}\right).$$

As compared to solving the above PDE numerically, the discrete-time treatment of the problem, which is the focus of the present paper, appears to provide a more computationally tractable approach to determining the optimal control. In what follows, we will treat this discrete-time problem within the framework of optimal stopping (Dynkin & Yushkevich, 1967) and drift counteraction (Kolmanovsky & Maizenberg, 2002) stochastic optimal control, and we will discuss opportunities for applications of our results to hybrid electric vehicle (HEV) management and to oil extraction.

The paper is organized as follows. In Section 2 we develop theoretical results to characterize optimal control in problem (1), (2). In Section 3 we discuss computational procedures to approximately compute this optimal control law. Since one of our motivations is to develop theory for more effective control of automotive systems, in Section 4 we consider an example which highlights opportunities for application of our results to HEV powertrain management. In Section 5 another application-oriented example is discussed, where the objective is to maximize the cumulative yield of oil extracted from a well before the well is lost. Finally, concluding remarks are made in Section 6.

2. Theoretical results

Given a state vector, x^-, and disturbance vectors, $w^-, w^+ \in W$, we define

$$\begin{aligned} L^u V(x^-, w^-) &\triangleq E_{x^-,w^-}[V(f(x^-, u(x^-, w^-), w^-), w^+)] - V(x^-, w^-) \\ &= \sum_{j \in J} V(f(x^-, u(x^-, w^-), w^-), w^j) \cdot P(w^j | w^-, x^-) \\ &\quad - V(x^-, w^-). \end{aligned} \tag{3}$$

The following theorem provides sufficient conditions for the optimal control law, $u_*(x, w)$:

Theorem 1. *Suppose that $g(x, u, w) \geqslant \varepsilon$ for some $\varepsilon > 0$ and that there exists a control function $u_*(x, w)$ and a continuous, non-negative function $V(x, w)$ such that*

$$\begin{aligned} &L^{u_*} V(x, w) + g(x, u_*(x, w), w) = 0 \quad \text{if } (x, w) \in G, \\ &L^u V(x, w) + g(x, u(x, w), w) \leqslant 0 \quad \text{if } (x, w) \in G, \quad u \neq u_*, \\ &V(x, w) = 0 \quad \text{if } (x, w) \notin G. \end{aligned} \tag{4}$$

Then, u_* maximizes (2), *and, for all $(x_0, w_0) \in G$, $V(x_0, w_0) = J^{x_0, w_0, u_*}$. Furthermore, $J^{x_0, w_0, u}$, $E[\tau^{x_0, w_0, u}(G)]$ are finite for any policy u, and the function V, satisfying* (4), *if exists, is unique.*

Proof. We define $T(t) = \min\{t, \tau^{x_0, w_0, u}(G)\}$, where $x_0 = x^u(0)$, $w_0 = w^u(0)$. Following the same arguments as in the proof of Theorem 5.1 in Dynkin (1963) and taking advantage of $T(t)$ being a Markov moment (Dynkin, 1963), of finiteness of $E[T(t)]$, and of boundedness of V, Dynkin's formula for the discrete-time system (1) holds in the form

$$\begin{aligned} &E[V(x^u(T(t)), w^u(T(t)))] - V(x^u(0), w^u(0)) \\ &\quad = E \sum_{k=0}^{T(t)-1} L^u V(x^u(k), w^u(k)). \end{aligned} \tag{5}$$

Using (4) and $g \geqslant \varepsilon$ we obtain

$$\begin{aligned} &E[V(x^u(T(t)), w^u(T(t)))] - V(x^u(0), w^u(0)) \\ &\quad \leqslant -E \sum_{k=0}^{T(t)-1} g(x^u(k), w^u(k)) \leqslant -\varepsilon \cdot E[T(t)]. \end{aligned}$$

Thus $(2/\varepsilon)\max_{(x, w^i) \in G} V(x, w^i) \geqslant E[T(t)]$, where the left-hand side is finite because V is continuous and G is compact. Thus $E[T(t)]$ is bounded by the same upper bound for any t. Viewing $T(t) = T(t; \omega)$ as a random variable (a function of elementary event, ω) we note that it is monotonically non-decreasing with t. Considering expectation as a Lebesgue integral and applying Levi's theorem