Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 127 (2017) 2482–2507

stochastic processes and their applications

www.elsevier.com/locate/spa

Approximating a diffusion by a finite-state hidden Markov model

I. Kontoyiannis[a,*], S.P. Meyn[b]

[a] *Department of Informatics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece*
[b] *Department of Electrical and Computer Engineering, University of Florida, Gainesville, USA*

Received 3 July 2015; received in revised form 25 April 2016; accepted 27 November 2016
Available online 10 December 2016

Abstract

For a wide class of continuous-time Markov processes evolving on an open, connected subset of $\mathbb{R}^d$, the following are shown to be equivalent:

(i) The process satisfies (a slightly weaker version of) the classical Donsker–Varadhan conditions;
(ii) The transition semigroup of the process can be approximated by a finite-state hidden Markov model, in a strong sense in terms of an associated operator norm;
(iii) The resolvent kernel of the process is 'v-separable', that is, it can be approximated arbitrarily well in operator norm by finite-rank kernels.

Under any (hence all) of the above conditions, the Markov process is shown to have a purely discrete spectrum on a naturally associated weighted L_∞ space.

Keywords: Markov process; Hidden Markov model; Hypoelliptic diffusion; Stochastic Lyapunov function; Discrete spectrum

1. Introduction

Consider a continuous-time Markov process $\mathbf{\Phi} = \{\Phi(t) : t \geq 0\}$ taking values in an open, connected subset X of $\mathbb{R}^d$, equipped with its associated Borel σ-field $\mathcal{B}$. We begin by assuming

* Corresponding author.
E-mail addresses: yiannis@aueb.gr (I. Kontoyiannis), meyn@ece.ufl.edu (S.P. Meyn).

http://dx.doi.org/10.1016/j.spa.2016.11.004

that $\mathbf{\Phi}$ is a diffusion; that is, it is the solution of the stochastic differential equation,

$$d\Phi(t) = u(\Phi(t))dt + M(\Phi(t))dB(t), \quad t \geq 0, \;\; \Phi(0) = x, \tag{1}$$

where $u = (u_1, u_2, \dots, u_d)^T : \mathsf{X} \to \mathbb{R}^d$ and $M : \mathsf{X} \to \mathbb{R}^d \times \mathbb{R}^k$ are locally Lipschitz, and $\boldsymbol{B} = \{B(t) : t \geq 0\}$ is k-dimensional standard Brownian motion. [Extensions to more general Markov processes are briefly discussed in Section 1.4.] Unless explicitly stated otherwise, throughout the paper we assume that:

$$\left.\begin{array}{l}\text{The strong Markov process } \mathbf{\Phi} \text{ is the unique strong solution of (1)}\\ \text{with continuous sample paths.}\end{array}\right\} \tag{A1}$$

The distribution of the process $\mathbf{\Phi}$ is described by the initial condition $\Phi(0) = x \in \mathsf{X}$ and the transition semigroup $\{P^t\}$: For any $t \geq 0$, $x \in \mathsf{X}$, $A \in \mathcal{B}$,

$$P^t(x, A) := \mathsf{P}_x\{\Phi(t) \in A\} := \Pr\{\Phi(t) \in A \mid \Phi(0) = x\}.$$

Recall that the kernel P^t acts as a linear operator on functions $f : \mathsf{X} \to \mathbb{R}$ on the right and on signed measures ν on $(\mathsf{X}, \mathcal{B})$ on the left, respectively, as,

$$P^t f(x) = \int f(y)P^t(x, dy), \qquad \nu P^t(A) = \int \nu(dx)P^t(x, A), \quad x \in \mathsf{X}, \; A \in \mathcal{B},$$

whenever the above integrals exist. Also, for any signed measure ν on $(\mathsf{X}, \mathcal{B})$ and any function $f : \mathsf{X} \to \mathbb{R}$ we write $\nu(f) := \int f d\nu$, whenever the integral exists. In this paper we will constrain the domain of functions f to a Banach space defined with respect to a weighted L_∞ norm.

One of the central assumptions we make throughout the paper is the following regularity condition on the semigroup:

$$\left.\begin{array}{l}\text{The transition semigroup admits a continuous density: There is a}\\ \text{continuous function } p \text{ on } (0, \infty) \times \mathsf{X} \times \mathsf{X} \text{ such that,}\\ P^t(x, A) = \displaystyle\int_A p(t, x, y)\, dy, \quad x \in \mathsf{X}, \;\; A \in \mathcal{B}.\end{array}\right\} \tag{A2}$$

Hörmander's theorem [29, Thm. 38.16] gives sufficient conditions for (A2). Explicit bounds on the density are also available; see [26] and its references.

1.1. Irreducibility, drift, and semigroup approximations

The ergodic theory of continuous-time Markov processes is often most easily addressed by translating results from the discrete-time domain. This is achieved, e.g., in [8,22,23,21] through consideration of the Markov chain whose transition kernel is defined by one of the *resolvent kernels* of $\mathbf{\Phi}$, defined as,

$$R_\alpha := \int_0^\infty e^{-\alpha t} P^t \, dt, \quad \alpha > 0. \tag{2}$$

In the case $\alpha = 1$ we simply write $R := R_1 = \int_0^\infty e^{-t} P^t \, dt$, and call R "the" resolvent kernel of the process $\mathbf{\Phi}$.

The family of resolvent kernels $\{R_\alpha\}$ is simply the Laplace transform of the semigroup, so that each R_α admits a density under (A2). This density will not be continuous in general, so we