Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 124 (2014) 51–80

stochastic processes and their applications

www.elsevier.com/locate/spa

Parametric inference for discretely observed multidimensional diffusions with small diffusion coefficient

Romain Guy[a,b,*], Catherine Larédo[a,b], Elisabeta Vergu[a]

[a] *UR 341 Mathématiques et Informatique Appliquées, INRA, Jouy-en-Josas, France*
[b] *UMR 7599 Laboratoire de Probabilités et Modèles aléatoires, Université Denis Diderot Paris 7 and CNRS, Paris, France*

Received 6 June 2012; received in revised form 3 May 2013; accepted 23 July 2013
Available online 6 August 2013

Abstract

We consider a multidimensional diffusion X with drift coefficient $b(\alpha, X_t)$ and diffusion coefficient $\epsilon\sigma(\beta, X_t)$. The diffusion sample path is discretely observed at times $t_k = k\Delta$ for $k = 1 \dots n$ on a fixed interval $[0, T]$. We study minimum contrast estimators derived from the Gaussian process approximating X for small ϵ. We obtain consistent and asymptotically normal estimators of α for fixed Δ and $\epsilon \to 0$ and of (α, β) for $\Delta \to 0$ and $\epsilon \to 0$ without any condition linking ϵ and Δ. We compare the estimators obtained with various methods and for various magnitudes of Δ and ϵ based on simulation studies. Finally, we investigate the interest of using such methods in an epidemiological framework.

Keywords: Minimum contrast estimators; Low frequency data; High frequency data; Epidemic data

1. Introduction

In this study we focus on the parametric inference in the drift coefficient $b(\alpha, X_t^\epsilon)$ and in the diffusion coefficient $\epsilon\sigma(\beta, X_t^\epsilon)$ of a multidimensional diffusion model $\left(X_t^\epsilon\right)_{t\geq 0}$ with small diffusion coefficient, when it is observed at discrete times on a fixed time interval in the

* Corresponding author at: UR 341 Mathématiques et Informatique Appliquées, INRA, Jouy-en-Josas, France. Tel.: +33 0 1 34 65 22 58.
E-mail address: romain.guy@jouy.inra.fr (R. Guy).

0304-4149/$ - see front matter
http://dx.doi.org/10.1016/j.spa.2013.07.009

asymptotics $\epsilon \to 0$. This asymptotics has been widely studied and has proved fruitful in applied problems, see e.g. [6]. Our interest in considering this kind of diffusions is motivated by the fact that they are natural approximations of epidemic processes. Indeed, the classical stochastic SIR model in a closed population, describing variations over time in Susceptible (S), Infectious (I) and Removed (R) individuals, is a bi-dimensional continuous-time Markovian jump process. The population size (N) based normalization of this process asymptotically leads to an ODE system. Before passing to the limit, the forward Kolmogorov diffusion equation allows describing the epidemic dynamics through a bi-dimensional diffusion, with diffusion coefficient proportional to $1/\sqrt{N}$. Moreover, epidemics are discretely observed and therefore we are interested in the statistical setting defined by discrete data sampled at times $t_k = k\Delta$ on a fixed interval $[0, T]$ with $T = n\Delta$. The number of data points is n and Δ, the sampling interval, is not necessarily small.

Historically, statistics for diffusions were developed for continuously observed processes leading to explicit formulations of the likelihood [16,18]. In this context, two asymptotics exist for estimating α for a diffusion continuously observed on a time interval $[0, T]$: $T \to \infty$ for recurrent diffusions and T fixed and the diffusion coefficient tends to 0. In practice, however, observations are not continuous but partial, with various mechanisms underlying the missingness, which leads to intractable likelihoods. One classical case consists in sample paths discretely observed with a sampling interval Δ. This adds another asymptotic framework $\Delta \to 0$ and raises the question of estimating parameters in the diffusion coefficient (see [8,24] for T fixed and [12,15,23] for $T \to \infty$).

Since nineties, statistical methods associated to discrete data have been developed in the asymptotics of a small diffusion coefficient (e.g. [17,7,22]). Considering a discretely observed diffusion on $\mathbb{R}$ with constant ($=\epsilon$) diffusion coefficient, Genon-Catalot (1990) obtained, using the Gaussian approximating process [2], a consistent and ϵ^{-1}-normal and efficient estimator of α under the condition $\left\{\epsilon \to 0, \Delta \to 0, \epsilon/\sqrt{\Delta} = O(1)\right\}$. The author additionally proved that this estimator possessed good properties also for Δ fixed. Uchida [22] obtained similar results using approximate martingale estimating equations. Then, Sørensen [20] obtained, as $\epsilon \to 0$, consistent and ϵ^{-1}-normal estimators of a parameter θ present in both the drift and diffusion coefficients, with no assumption on Δ, but under additional conditions not verified in the case of distinct parameters in the drift and diffusion coefficients. For this latter case, Sørensen and Uchida [21] obtained consistent and ϵ^{-1}-normal estimators of α and consistent and $\sqrt{n}$-normal estimators of β under the condition $\Delta/\epsilon \to 0$ and $\sqrt{\Delta}/\epsilon$ bounded. This result was later extended by Gloter and Sørensen [10] to the case where $\epsilon^{-1}\Delta^{\rho}$ is bounded for some $\rho > 0$. Their results rely on a class of contrast processes based on the expansion of the infinitesimal generator of the diffusion, the order of the expansion being driven by the respective magnitude of ϵ and Δ and requiring this knowledge (value of ρ), which might be a drawback when applying the method. Moreover, this contrast becomes difficult to handle for values of Δ that are not very small with respect to ϵ.

To overcome this drawback, we consider a simple contrast based on the Gaussian approximation of the diffusion process X^{ϵ} [2,6]. Contrary to Gloter and Sørensen [10], our contrast has generic formulation, regardless of the ratio between Δ and ϵ. Thus, the standard balance condition between ϵ and Δ of previous works is here removed. Our study extends the results of [7] to the case of multidimensional diffusion processes with parameters in both the drift and diffusion coefficients. We consider successively the cases Δ fixed and $\Delta \to 0$. We obtain consistent and ϵ^{-1}-normal estimators of α (when β is unknown or equal to a known function of α) for fixed Δ. For high frequency data, we obtain results similar to [10], but without