Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 3092–3102
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Likelihood-based inference for a class of multivariate diffusions with unobserved paths

Konstantinos Kalogeropoulos*

Department of Engineering, Signal Processing Group, University of Cambridge, Trumpington street, CB2 1PZ, Cambridge, UK

Received 11 October 2005; accepted 10 May 2006
Available online 15 March 2007

Abstract

This paper presents a Markov chain Monte Carlo algorithm for a class of multivariate diffusion models with unobserved paths. This class is of high practical interest as it includes most diffusion driven stochastic volatility models. The algorithm is based on a data augmentation scheme where the paths are treated as missing data. However, unless these paths are transformed so that the dominating measure is independent of any parameters, the algorithm becomes reducible. The methodology developed in Roberts and Stramer [2001a. On inference for partial observed nonlinear diffusion models using the metropolis-hastings algorithm. Biometrika 88(3); 603–621] circumvents the problem for scalar diffusions. We extend this framework to the class of models of this paper by introducing an appropriate reparametrisation of the likelihood that can be used to construct an irreducible data augmentation scheme. Practical implementation issues are considered and the methodology is applied to simulated data from the Heston model.

MSC: 62M99; 60J60; 91B84

Keywords: Data augmentation; Markov chain Monte Carlo; Stochastic volatility; Heston model

1. Introduction

Diffusion processes constitute a natural and useful tool for modelling phenomena evolving continuously in time. They find applications in many different fields including finance, biology, physics, engineering, etc. A diffusion process V_t is defined through a stochastic differential equation (SDE):

$$dV_t = \mu(V_t, \theta)\, dt + \sigma(V_t, \theta)\, dB_t, \quad V_0 = v_0, \quad 0 \leqslant t \leqslant T, \quad V_t \in \Re^d, \tag{1}$$

where B_t is a standard Brownian motion. The drift μ and volatility σ of the diffusion should satisfy some regularity conditions (locally Lipschitz with a growth bound) to ensure that the SDE will have a weakly unique solution; see of Kloeden and Platen (1995, Chapter 4). Throughout this paper we assume that their functional form is known and we focus on the problem of parametric inference.

In practice we can only observe a discrete skeleton of the diffusion V. Depending on what kind of data we observe, we can further classify the diffusion models into two categories. The first category includes models where we have

* Tel.: +44 1223 322766.
E-mail address: kk384@cam.ac.uk.

0378-3758/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.05.017

observations on every coordinate of the vector V. In the second category, the process is divided into two components, say X and α, and we only observe points of X. In this paper we focus on a subclass of the second category, henceforth denoted by $\mathbb{C}$, considering multivariate diffusions with unobserved paths that satisfy the following SDE ($0 \leqslant t \leqslant T$):

$$\begin{pmatrix} \mathrm{d}X_t \\ \mathrm{d}\alpha_t \end{pmatrix} = \begin{pmatrix} \mu_x(\alpha_t, \theta) \\ \mu_\alpha(\alpha_t, \theta) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \sigma_x(\alpha_t, \theta) & 0 \\ 0 & \sigma_\alpha(\alpha_t, \theta) \end{pmatrix} \begin{pmatrix} \mathrm{d}B_t \\ \mathrm{d}W_t \end{pmatrix}, \tag{2}$$

where B and W denote standard Brownian motions that can potentially be correlated. The class $\mathbb{C}$ includes many interesting diffusions. For instance it contains most diffusion driven stochastic volatility models which are particularly useful in financial applications; see for instance Ghysels et al. (1996) and Shephard (2005). Famous examples of stochastic volatility models in $\mathbb{C}$ are the models introduced in Heston (1993), Stein and Stein (1991) and Hull and White (1987).[1] Section 4 provides an example based on the model of Heston (1993).

A crucial difference between the models of these two categories is the availability of the Markov property. Suppose that we observe V (or X accordingly) at times $\{t_k, 0 \leqslant t_k \leqslant T, k = 0, 1, \ldots, n\}$ and let $Y = \{Y_k = V_{t_k}\}$. In the first category we can write the likelihood using the transition density of the diffusion (conditional on the initial point Y_0):

$$p_\theta(Y) = \prod_{k=1}^{n} p_\theta(Y_k | Y_{k-1}). \tag{3}$$

However, this is not always true for the models of the second category and consequently for the diffusions in $\mathbb{C}$.

In both cases the likelihood is not generally available in closed form and the problem of inference is quite complicated. As a result of this, the literature contains various methodologies that may or not be based on the likelihood; see Sørensen (2004) for an extensive review. Likelihood based approaches are either analytical (Aït-Sahalia, 2002, 2005), or they use simulations (Pedersen, 1995; Durham and Gallant, 2002). They usually approximate the likelihood in a way so that the discretisation error can become arbitrarily small, although the methodology developed in Beskos et al. (2006) succeeds exact inference in the sense that it allows only for Monte Carlo error. Unfortunately, all of the above rely on the Markov property and therefore become hard to generalise to the non-Markovian case.

A natural way to proceed is via data augmentation, a methodology introduced by Tanner and Wong (1987). The idea is based on the fact that the likelihood can always be well approximated given the entire path of V or a sufficiently fine partition of it. Therefore, the unobserved paths of V are treated as missing data and a finite number of points, large enough to make the approximation error arbitrarily small, is imputed. Elerian et al. (2001), Eraker (2001) and Jones (2003) use Markov chain Monte Carlo (MCMC) approaches and approximate the posterior through the Euler–Maruyama approximation of the transition density (3). As noted in Roberts and Stramer (2001a,b), however, there exists a strong dependence between the imputed sample paths and the volatility coefficients. In fact the algorithm becomes reducible as the number of imputed points increases.

Roberts and Stramer (2001a) tackle the problem for scalar diffusions by a reparametrisation on the paths of V and Kalogeropoulos et al. (2006a) offer an extension for some multivariate diffusions. As explained in Section 2.1, however, this framework does not cover the models in $\mathbb{C}$. This paper focuses on this class and we introduce a novel reparametrisation of the likelihood that may serve as the basis for data augmentation schemes. Alternative approaches to this problem can be found in Chib et al. (2005) and Golightly and Wilkinson (2005).

The paper is organised as follows. Next section elaborates on the need for a reparametrisation and provides a likelihood that may serve as the basis for inference purposes. Section 3 presents the details of a data augmentation scheme that can handle models in $\mathbb{C}$. In Section 4, the methodology is applied to simulated data. We highlight the necessity of the reparametrisation in a simple stochastic volatility model and we perform a simulation study based on the model of Heston (1993) to illustrate the proposed methodology. Finally we conclude in Section 5 with some relevant discussion.

[1] It is easier to see this if we use $S = \exp(X)$ instead of X.