Journal of Multivariate Analysis 102 (2011) 683–697

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Estimation of a multivariate stochastic volatility density by kernel deconvolution

Bert Van Es, Peter Spreij *

Korteweg–de Vries Institute for Mathematics, Universiteit van Amsterdam, PO Box 94248, 1090GE Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 19 November 2009
Available online 14 December 2010

AMS subject classifications:
62G07
62M07
62P20

Keywords:
Stochastic volatility models
Multivariate density estimation
Kernel estimator
Deconvolution
Mixing

ABSTRACT

We consider a continuous time stochastic volatility model. The model contains a stationary volatility process. We aim to estimate the multivariate density of the finite-dimensional distributions of this process. We assume that we observe the process at discrete equidistant instants of time. The distance between two consecutive sampling times is assumed to tend to zero.

A multivariate Fourier-type deconvolution kernel density estimator based on the logarithm of the squared processes is proposed to estimate the multivariate volatility density. An expansion of the bias and a bound on the variance are derived.

1. Introduction

Let S denote the log price process of some stock in a financial market. It is often assumed that S can be modeled as the solution of a stochastic differential equation or, more generally, as an Itô diffusion process. So we assume that we can write

$$dS_t = b_t\,dt + \sigma_t\,dW_t, \qquad S_0 = 0, \tag{1}$$

or, in integral form,

$$S_t = \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s, \tag{2}$$

where W is a standard Brownian motion and the processes b and σ are assumed to satisfy certain regularity conditions (see [16]) to have the integrals in (2) well defined. In a financial context, the process σ is called the volatility process. In the literature, the process σ is often taken to be independent of the Brownian motion W.

In this paper we adopt this independence assumption and we furthermore assume that σ is a strictly stationary positive process satisfying a mixing condition, for example an ergodic diffusion on $(0, \infty)$. We will assume that all p-dimensional marginal distributions of σ have invariant densities with respect to the Lebesgue measure on $(0, \infty)^p$. This is typically the case in virtually all stochastic volatility models that are proposed in the literature, where the evolution of σ is modeled by a stochastic differential equation, mostly in terms of σ^2, or $\log \sigma^2$ (cf. e.g. [27,14]).

Therefore, think of X as $X = \sigma^2$ or $X = \log \sigma^2$, as a motivation for nonparametric estimation procedures, consider stochastic differential equations of the type

$$dX_t = b(X_t)\,dt + a(X_t)\,dB_t,$$

* Corresponding author.
E-mail address: spreij@uva.nl (P. Spreij).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.12.003

with B equal to Brownian motion. Focusing on the invariant univariate density of X_t, we recall that it is up to a multiplicative constant equal to

$$x \mapsto \frac{1}{a^2(x)} \exp\left(2\int_{x_0}^{x} \frac{b(y)}{a^2(y)}\,\mathrm{d}y\right), \tag{3}$$

where x_0 is an arbitrary element of the state space (l, r); see e.g. [12] or [21]. From formula (3) one sees that the invariant distribution of X may take on many different forms, as is the case for the various models that have been proposed in the literature. Refraining from parametric assumptions on the functions a and b, nonparametric statistical procedures may be used to obtain information about the shape of the (one-dimensional) invariant distribution.

A phenomenon that is often observed in practice, is *volatility clustering*. This means that for different time instants $t_1, \ldots, t_p$ that are close, the corresponding values of $\sigma_{t_1}, \ldots, \sigma_{t_p}$ are close again. This can partly be explained by assumed continuity of the process σ, but it might also result from specific areas around the diagonal where the multivariate density of $(\sigma_{t_1}, \ldots, \sigma_{t_p})$ assumes high values. For instance, it is conceivable that for $p = 2$, the density of $(\sigma_{t_1}, \sigma_{t_2})$ has high concentrations around points (ℓ, ℓ) and (h, h), with $\ell < h$, a kind of bimodality on the diagonal of the joint distribution, with the interpretation that clustering occurs around a low value ℓ or around a high value h.

Here is an example where this happens. We consider a regime switching volatility process. Assume that for $i = 0, 1$ we have two stationary processes X^i, each of them having multivariate invariant distributions having densities. Call these $f^i_{t_1,\ldots,t_p}(x_1, \ldots, x_p)$, whereas for $p = 1$ we simply write f^i. We assume these two processes to be independent, and also independent of a two-state homogeneous Markov chain U with states 0, 1. Let $Q(t)$ be the matrix of transition probabilities $q_{ij}(t) = P(X_t = i|X_0 = j)$. Let A be the matrix of transition intensities and write

$$A = \begin{pmatrix} -a_0 & a_1 \\ a_0 & -a_1 \end{pmatrix},$$

with $a_0, a_1 > 0$. Then $\dot{Q}(t) = AQ(t)$, and

$$Q(t) = \frac{1}{a_0 + a_1}\begin{pmatrix} a_1 + a_0 \mathrm{e}^{-(a_0+a_1)t} & a_1 - a_1 \mathrm{e}^{-(a_0+a_1)t} \\ a_0 - a_0 \mathrm{e}^{-(a_0+a_1)t} & a_0 + a_1 \mathrm{e}^{-(a_0+a_1)t} \end{pmatrix}.$$

The stationary distribution of U is given by $\pi_i := P(U_t = i) = \frac{a_{1-i}}{a_0+a_1}$ and we assume that U_0 has this distribution. We finally define the process ξ by

$$\xi_t = U_t X_t^1 + (1 - U_t) X_t^0.$$

Then ξ is stationary too and it has a bivariate stationary distribution with a density, related by $P(\xi_s \in \mathrm{d}x, \xi_t \in \mathrm{d}y) = f_{s,t}(x, y)\,\mathrm{d}x\,\mathrm{d}y$. Elementary calculations lead to the following expression for $f_{s,t}$ for $0 < s < t$

$$f_{s,t}(x, y) = q_{11}(t-s)\pi_1 f^1_{s,t}(x, y) + q_{10}(t-s)\pi_0 f^0(x) f^1(y) + q_{01}(t-s)\pi_1 f^1(x) f^0(y) + q_{00}(t-s)\pi_0 f^0_{s,t}(x, y).$$

Suppose that the volatility process is defined by $\sigma_t = \exp(\xi_t)$ and that the X^i are both Ornstein–Uhlenbeck processes given by

$$\mathrm{d}X_t^i = -a(X_t^i - \mu_i)\,\mathrm{d}t + b\,\mathrm{d}W_t^i,$$

with W^1, W^2 independent Brownian motions, $\mu_1 \neq \mu_2$ and $a > 0$. Suppose that the X^i start in their stationary $N(\mu_i, \frac{b^2}{2a})$ distributions. Then the center of the distribution of (X_s^i, X_t^i) is (μ_i, μ_i), whereas the center of the distribution of (X_s^0, X_t^1) is (μ_0, μ_1). Hence the density $f_{s,t}$ is a mixture of four hump shaped contours, each of them having a different center of location. If $t - s$ is small, this effectively reduces to mixture of distributions with centers (μ_1, μ_1) and (μ_2, μ_2). Similar qualitative observations can be made for models that switch between different GARCH regimes.

Nonparametric procedures are able to detect such a property of a bivariate distribution, and are consequently appropriate tools to get some partial insight into the behavior of the volatility. In the present paper we propose a nonparametric estimator for the multivariate density of the volatility process. Using ideas from deconvolution theory, we will propose a procedure for the estimation of this density at a number of fixed time instants. Related work on estimating a stationary univariate density of the volatility process has been done by Van Es et al. [22], Comte and Genon-Catalot [3], Van Zanten and Zareba [24], whereas a deconvolution approach has also been adopted to estimate a regression function for a discrete time stochastic volatility model by Franke et al. [8], Comte [1] and Comte et al. [2]. In [4] it is assumed that the volatility process solves a stochastic differential equation and nonparametric estimators for the drift and diffusion coefficients of that equation are studied. In most of these papers, one works with a simplified model with detrended log prices, which amounts to modeling the process S by Eq. (1) with a zero drift coefficient. In the present paper we adhere to this approach, which has become the tradition, as well.

The observations of the log-asset price S process are assumed to take place at the time instants $\Delta, 2\Delta, \ldots, n\Delta$, where the time gap satisfies $\Delta = \Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. This means that we base our estimator on the so-called *high frequency data*.