Asymptotic analysis for a simple explicit estimator in Barndorff-Nielsen and Shephard stochastic volatility models

Petra Posedel
Department of Mathematics, Faculty of Economics and Business Zagreb, University of Zagreb, Croatia.

Keywords: Martingale estimating functions, stochastic volatility models with jumps, consistency and asymptotic normality
AMS: 60J75, 62F12

Abstract

We provide a simple explicit estimator for discretely observed Barndorff-Nielsen and Shephard models, prove rigorously consistency and asymptotic normality based on the single assumption that all moments of the stationary distribution of the variance process are finite, and give explicit expressions for the asymptotic covariance matrix.

We develop in detail the martingale estimating function approach for a bivariate model, that is not a diffusion, but admits jumps. We do not use ergodicity arguments.

We assume that both, logarithmic returns and instantaneous variance are observed on a discrete grid of fixed width, and the observation horizon tends to infinity. This analysis is a starting point and benchmark for further developments concerning optimal martingale estimating functions, and for theoretical and empirical investigations, that replace the (actually unobserved) variance process with a substitute, such as number or volume of trades or implied variance from option data. This is joint work with Friedrich Hubalek, Technische Universität Vienna.

1. Introduction

In [1] Barndorff-Nielsen and Shephard introduced a class of stochastic volatility models in continuous time, where the instantaneous variance follows an Ornstein-Uhlenbeck type process driven by an increasing Lévy process. Those models allow flexible modelling, capture many stylized facts of financial time series, and yet are of great analytical tractability. For further information see also [2]. BNS-models, as we will call them from now on, are affine models in the sense of [3] and [4], where the associated Riccati type equations can be solved up to quadrature in general. In several concrete cases the integration can be performed explicitly in closed form in terms of elementary functions, see [5].

BNS-models have been studied from various points of view in mathematical finance and related fields.
Strangely though, it seems that statistical estimation of the model is the most difficult problem, and most of the work in that area focused on computationally intensive methods.

The contributions of the present paper are as follows: first we develop a simple and explicit estimator for BNS models. Secondly, we give rigorous proofs of its consistency and asymptotic normality. In doing so we compute explicitly the asymptotic covariance matrix and develop to that purpose formulas for arbitrary bivariate integer moments of returns and variance.

2. The model

2.1. The continuous time model

2.1.1. The general setting

As in Barndorff-Nielsen and Shepard [1], we assume that the price process of an asset $S$ is defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ and is given by $S_t = S_0 \exp(X_t)$ with $S_0 > 0$ a constant. The process of logarithmic returns $X_t$ and the instantaneous variance process $V_t$ satisfy

\begin{align}
    dX(t) &= (\mu + \beta V(t-))dt + \sqrt{V(t-)}dW_\theta(t) + \rho dZ_{\lambda}(t), \quad X(0) = 0, \\
    dV(t) &= -\lambda V(t-)dt + dZ_{\lambda}(t), \quad V(0) = V_0,
\end{align}

where the parameters $\mu, \beta, \rho$ and $\lambda$ are real constants with $\lambda > 0$. The process $W$ is a standard Brownian motion, the process $Z$ is an increasing Lévy process, and we define $Z_\lambda(t) = Z(\lambda t)$ for notational simplicity. Adopting the terminology introduced by Barndorff-Nielsen and Shepard, we will refer to $Z$ as the background driving Lévy process (BDLP). The Brownian motion $W$ and the BDLP $Z$ are independent and $(\mathcal{F}_t)$ is assumed to be the usual augmentation of the filtration generated by the pair $(W, Z_{\lambda})$. The random variable $V_0$ has a self-decomposable distribution corresponding to the BDLP such that the process $V$ is strictly stationary and

\begin{align}
    E[V_0] &= \zeta, \quad \text{Var}[V_0] = \eta.
\end{align}

To shorten the notation we introduce the parameter vector

$\theta = (\lambda, \zeta, \eta, \mu, \beta, \rho)^T$, \quad (4)

and the bivariate process

$X = (X, V)$. \quad (5)$

If the distribution of $V_0$ is from a particular class $D$ then $X$ is called a BNS-DOU($\theta$) model.
2.1.2. The Γ-OU model

The Γ-OU model is obtained by constructing the BNS-model with stationary gamma distribution, \( V_0 \sim \Gamma(\nu, \alpha) \), where the parameters are \( \nu > 0 \) and \( \alpha > 0 \). The corresponding background driving Lévy process \( Z \) is a compound Poisson processes with intensity \( \nu \) and jumps from the exponential distribution with parameter \( \alpha \). Consequently both processes \( Z \) and \( V \) have a finite number of jumps in any finite time interval.

For the Γ-OU model it is more convenient to work with the parameters \( \nu \) and \( \alpha \). The connection to the generic parameters used in our general development is given by

\[
\zeta = \frac{\nu}{\alpha}, \quad \eta = \frac{\nu}{\alpha^2}.
\]

As the gamma distribution admits exponential moments we have integer moments of all orders and our Assumption 1 below is satisfied.

2.2. Discrete observations

We observe returns and variance process on a discrete grid of points in time,

\[
0 = t_0 < t_1 < \cdots < t_n.
\]

This implies

\[
V(t_i) = V(t_{i-1})e^{-\lambda(t_i-t_{i-1})} + \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i-s)} dZ_\lambda(s).
\]

Using

\[
V_i := V(t_i), \quad U_i := \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i-s)} dZ_\lambda(s)
\]

we have that \((U_i)_{i \geq 1}\) is a sequence of independent random variables, and it is independent of \( V_0 \). If the grid is equidistant, then \((U_i)_{i \geq 1}\) are iid. Observing the returns \( X \) on the grid we have

\[
X(t_i) - X(t_{i-1}) = \mu(t_i - t_{i-1}) + \beta(Y(t_i) - Y(t_{i-1})
+ \int_{t_{i-1}}^{t_i} \sqrt{V(s-)}dW(s) + \rho(Z_\lambda(t_i) - Z_\lambda(t_{i-1})).
\]

This suggests introducing the discrete time quantities

\[
X_i = X(t_i) - X(t_{i-1}), \quad Y_i = Y(t_i) - Y(t_{i-1}), \quad Z_i = Z_\lambda(t_i) - Z_\lambda(t_{i-1})
\]

and

\[
W_i = \frac{1}{\sqrt{Y_i}} \int_{t_{i-1}}^{t_i} \sqrt{V(s-)}dW(s).
\]
Furthermore, it is also convenient to introduce the discrete quantity

$$S_i = \frac{1}{\lambda}(Z_i - U_i).$$

(13)

It is not difficult to see (conditioning!) that \((W_i)_{i \geq 1}\) is an iid \(N(0, 1)\) sequence independent from all other discrete quantities. We note also that \((U_i, Z_i)_{i \geq 1}\) is a bivariate iid sequence, but \(U_i\) and \(Z_i\) are obviously dependent.

From now on, for notational simplicity, we consider the equidistant grid with

$$t_k = k\Delta,$$

(14)

where \(\Delta > 0\) is fixed. This implies

$$V_i = \gamma V_{i-1} + U_i$$

and

$$Y_i = \epsilon V_{i-1} + S_i,$$

(16)

where

$$\gamma = e^{-\lambda \Delta}, \quad \epsilon = \frac{1 - \gamma}{\lambda}.$$ 

(17)

Furthermore,

$$X_i = \mu \Delta + \beta Y_i + \sqrt{Y_i}W_i + \rho Z_i.$$ 

(18)

The sequence \((X_i, V_i)_{i \geq 0}\) is clearly Markovian. From now on we assume all moments of the stationary distribution of \(V_0\) exist.

**Assumption 1.**

$$E[V_0^n] < \infty \quad \forall n \in \mathbb{N}.$$ 

(19)

In the estimating context we assume all moments are finite with respect to all probability measures \(P_\theta, \theta \in \Theta\) under consideration, where \(\Theta\) is the parameter space.

No other assumptions are made, and all conditions required for consistency and asymptotic normality of our estimator will be proven rigorously from that assumption.

**Proposition 1.** We have for all \(n \in \mathbb{N}\) that

$$E[Z_1^n] < \infty, \quad E[U_1^n] < \infty, \quad E[S_1^n] < \infty,$$

(20)

and

$$E[Y_1^n] < \infty, \quad E[W_1^n] < \infty, \quad E[X_1^n] < \infty.$$ 

(21)

Consequently the expectation of any (multivariate) polynomial in \(Z_1, U_1, S_1, \sqrt{Y_1}, W_1, X_1\) exists under \(P_\theta\).
3. The simple explicit estimator

3.1. The simple estimating equations and their explicit solution

For estimation purposes we consider a probability space on which a parametrized family of probability measures is given:

\begin{equation}
(\Omega, \mathcal{F}, \{P_\theta : \theta \in \Theta\}),
\end{equation}

where \(\Theta = \{\theta \in \mathbb{R}^6 : \theta^1 > 0, \theta^2 > 0, \theta^3 > 0\}\). The data is generated under the true probability measure \(P_{\theta_0}\) with some \(\theta_0 \in \Theta\). The expectation with respect to \(P_\theta\) is denoted by \(E_\theta[\cdot]\) and with respect to \(P_{\theta_0}\) simply by \(E[\cdot]\).

We assume there is a process \(X\) that is BNS-DOU(\(\theta\)) under \(P_\theta\). We want to find an estimator for \(\theta_0\) using observations \(X_1, \ldots, X_n, V_1, \ldots, V_n\). We are interested in asymptotics as \(n \to \infty\). To that purpose let us consider the following martingale estimating functions:

\begin{align}
G^1_n(\theta) &= \sum_{k=1}^n [V_k - \mathbb{E}_\theta[V_k|V_0 = v]], & f^1(v, \theta) &= \mathbb{E}_\theta[V_1|V_0 = v] \\
G^2_n(\theta) &= \sum_{k=1}^n [V_k V_{k-1} - \mathbb{E}_\theta[V_k V_{k-1}|V_0 = v]], & f^2(v, \theta) &= \mathbb{E}_\theta[V_1 V_0|V_0 = v] \\
G^3_n(\theta) &= \sum_{k=1}^n [V_k^2 - \mathbb{E}_\theta[V_k^2|V_0 = v]], & f^3(v, \theta) &= \mathbb{E}_\theta[V_1^2|V_0 = v] \\
G^4_n(\theta) &= \sum_{k=1}^n [X_k - \mathbb{E}_\theta[X_k|V_0 = v]], & f^4(v, \theta) &= \mathbb{E}_\theta[X_1|V_0 = v] \\
G^5_n(\theta) &= \sum_{k=1}^n [X_k V_{k-1} - \mathbb{E}_\theta[X_k V_{k-1}|V_0 = v]], & f^5(v, \theta) &= \mathbb{E}_\theta[X_1 V_0|V_0 = v] \\
G^6_n(\theta) &= \sum_{k=1}^n [X_k V_k - \mathbb{E}_\theta[X_k V_k|V_0 = v]], & f^6(v, \theta) &= \mathbb{E}_\theta[X_1 V_1|V_0 = v]
\end{align}

The estimator \(\hat{\theta}_n\) is obtained by solving the estimating equation \(G_n(\hat{\theta}_n) = 0\) and it turns out that this equation has a simple explicit solution.

**Proposition 2.** The estimating equation \(G_n(\hat{\theta}_n) = 0\) admits for every \(n \geq 2\) on the event

\begin{equation}
C_n = \{\xi_n^2 - \xi_n^1 v_n^1 > 0, v_n^2 - (v_n^1)^2 > 0\}
\end{equation}
a unique solution \( \hat{\theta}_n = (\lambda_n, \zeta_n, \eta_n, \beta_n, \rho_n, \mu_n) \) that is given by

\[
\begin{align*}
\gamma_n &= \left(\xi_2^2 - \xi_1^1 v_1^1\right)/\left(v_n^2 - (v_1^1)^2\right); \\
\zeta_n &= \left(\xi_1^1 - \gamma_n v_1^1\right)/(1 - \gamma_n); \\
\eta_n &= \left((\xi_3^1 - (\xi_1^1)^2) - \gamma_n^2 (v_n^2 - (v_1^1)^2)\right)/(1 - \gamma_n^2); \\
\lambda_n &= -\log(\gamma_n)/\Delta; \\
\epsilon_n &= (1 - \gamma_n)/\lambda_n; \\
\beta_n &= \left(\xi_5^1 - \xi_4^1 \xi_1^1\right)/(\epsilon_n (v_n^2 - (v_1^1)^2)); \\
\rho_n &= \left(\xi_6^1 - \xi_4^1 \xi_1^1\right) - \beta_n \epsilon_n (\eta_n (1 - \gamma_n) + \gamma_n (v_n^2 - (v_1^1)^2))/\left(2(1 - \gamma_n)\eta_n\right); \\
\mu_n &= \left(\xi_4^1 - \beta_n \epsilon_n (v_1^1 - \zeta_n)\right)/\Delta - (\beta_n + \lambda_n \rho_n) \zeta_n;
\end{align*}
\]

where

\[
\begin{align*}
\xi_1^1 &= \frac{1}{n} \sum_{i=1}^{n} V_i, & \xi_2^1 &= \frac{1}{n} \sum_{i=1}^{n} V_i V_{i-1}, & \xi_3^1 &= \frac{1}{n} \sum_{i=1}^{n} V_i^2, \\
\xi_4^1 &= \frac{1}{n} \sum_{i=1}^{n} X_i, & \xi_5^1 &= \frac{1}{n} \sum_{i=1}^{n} X_i V_{i-1}, & \xi_6^1 &= \frac{1}{n} \sum_{i=1}^{n} X_i^2, \\
\end{align*}
\]

and

\[
\begin{align*}
v_1^1 &= \frac{1}{n} \sum_{i=1}^{n} V_{i-1}, & v_2^1 &= \frac{1}{n} \sum_{i=1}^{n} V_i V_{i-1}.
\end{align*}
\]

Proof: The first three equations \( G_j(\theta) = 0 \), for \( j = 1, 2, 3 \) contain only the unknowns \( \zeta, \eta, \lambda \) and are easily solved. In fact we get a familiar estimator for the first two moments and the autocorrelation coefficient of an AR(1) process. The last three equations \( G_k(\theta) = 0 \), for \( j = 4, 5, 6 \) can be seen as a linear system for the unknowns \( \mu, \beta, \rho \), once the other parameters have been determined.

### 3.2 Consistency

Let us investigate the consistency of the estimator from the previous section.

**Theorem 1.** We have \( P(C_n) \to 1 \) when \( n \to \infty \) and the estimator \( \hat{\theta}_n \) is consistent on \( C_n \), namely

\[
\hat{\theta}_n \xrightarrow{a.s.} \theta_0
\]

on \( C_n \) as \( n \to \infty \).
3.3. Asymptotic normality

For a concise vector notation we introduce
\[ \Xi_k = (V_k, V_k V_{k-1}, V_k^2, X_k, X_k V_{k-1}, X_k V_k)^\top, \]  
and write the estimating equations in the form
\[ G_n^i(\theta) = \sum_{k=1}^{n} \left[ \Xi_k^i - f^i(V_{k-1}, \theta) \right], \quad i = 1, \ldots, 6 \]
where
\[ f^i(v, \theta) = \sum_{\ell=v_i}^{p_i+r_i+q_i} \phi_i^\ell(\theta)v^\ell. \]

We will use, that \( f^i(v, \theta) \) is a polynomial in \( v \), and that its coefficients \( \phi \) are smooth functions in \( \theta \). We shall first prove the central limit theorem for the estimating functions.

**Proposition 3.** We have
\[ \frac{1}{\sqrt{n}} G_n(\theta_0) \xrightarrow{D} N(0, \Upsilon), \]
where
\[ \Upsilon_{ij} = E[Cov(\Xi_{i1}^1, \Xi_{j1}^1|V_0)]. \]
Lemma 1. We have
\[ \frac{1}{\sqrt{n}}[\xi_n - \xi] \overset{D}{\longrightarrow} N(0, \Sigma), \] (35)
where
\[ \Sigma = P^{-1} \Upsilon (P^{-1})^T \] (36)
and
\[ P_{ij} = \delta_{ij} - \phi_1^1 \delta_{1j} - \phi_2^2 \delta_{3j} \] (37)
with \( \delta_{ij} \) denoting the Kronecker delta.

Finally, we have all the ingredients for proving the following result.

Theorem 2. The estimator
\[ \hat{\theta}_n = (\lambda_n, \zeta_n, \eta_n, \beta_n, \rho_n, \mu_n) \] (38)
is asymptotically normal, namely
\[ \sqrt{n}[\hat{\theta}_n - \theta_0] \overset{D}{\longrightarrow} N(0, T), \] (39)
as \( n \to \infty \), where
\[ T = D\Sigma D^T \] (40)
and \( D \) can be calculated explicitly.

4. Further and alternative developments

4.1. Intra-day observations

Our approach is based on the explicit calculation of conditional and unconditional moments. Those calculations can be done for BNS-models on arbitrary time intervals. Hence our analysis is not restricted to a fixed time grid with the number of observation intervals tending to infinity, but could be performed also on a fixed horizon, with the number of intra-day observations increasing to infinity. The resulting estimators should then be compared to power-variation methods, cf. [6].

4.2. Comparison to the generalized method of moments

We would be interested in a comparison of our results to the related generalized methods of moments. For a rigorous treatment of the latter, a precise specification of the weighting matrix is required, see [7] and the references therein.
4.3. Unobserved volatility and substitutes for volatility

Finally, perhaps the biggest issue is, that the instantaneous variance is not observed in discrete time. In [8] it is reported, that the number of trades is an excellent substitute for statistical purposes. This is certainly a promising starting point for an empirical analysis. For a theoretical analysis a joint model for the number prices and number of trades has to be specified.

Acknowledgements: Financial support from the Austrian Science Fund (FWF) under grant P15889 is gratefully acknowledged.

5. Bibliography


