On Ornstein–Uhlenbeck driven by Ornstein–Uhlenbeck processes

Bernard Bercu [a,b], Frédéric Proïa [a,b,*], Nicolas Savy [c]

[a] Université Bordeaux 1, Institut de Mathématiques de Bordeaux, UMR 5251, 33405 Talence cedex, France
[b] INRIA Bordeaux Sud-Ouest, Team ALEA, 200 avenue de la Vieille Tour, 33405 Talence cedex, France
[c] Université Paul Sabatier, Institut de Mathématiques de Toulouse, UMR C5583, 31062 Toulouse Cedex 09, France

ARTICLE INFO

ABSTRACT

We investigate the asymptotic behavior of the maximum likelihood estimators of the unknown parameters of positive recurrent Ornstein–Uhlenbeck processes driven by Ornstein–Uhlenbeck processes.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction and motivation

Since the seminal work of Uhlenbeck and Ornstein (1930), a wide literature is available on Ornstein–Uhlenbeck processes driven by Brownian or fractional Brownian motions (Kutoyants, 2004; Liptser and Shiryaev, 2001). Many interesting papers are also available on Ornstein–Uhlenbeck processes driven by Lévy processes

$$\mathrm{d}X_t = \theta X_t \mathrm{d}t + \mathrm{d}L_t \tag{1.1}$$

where $\theta < 0$ and (L_t) is a continuous-time stochastic process starting from zero with stationary and independent increments. We refer the reader to Barndorff-Nielsen and Shephard (2001) for the mathematical foundation on Ornstein–Uhlenbeck processes driven by Lévy processes. Some recent extension on Ornstein–Uhlenbeck processes driven by fractional Lévy processes may be found in Barndorff-Nielsen and Basse-O'Connor (2011). More complex diffusions in which the volatility is itself given by an Ornstein–Uhlenbeck process are also available in Barndorff-Nielsen and Veraart (2013), whereas some continuous-time analogues of discrete-time ARMA models, based on general Ornstein–Uhlenbeck processes, can be found in Brockwell and Lindner (2012). Parametric estimation results for Ornstein–Uhlenbeck driven by α-stable Lévy processes are established in Hu and Long (2007) while nonparametric estimation results are given in Jongbloed et al. (2005). Two interesting applications related to money exchange rates and stock prices may be found in Barndorff-Nielsen and Shephard (2001) and Onalan (2009); see also the references therein. In short, actual research studies tend to treat volatility as more and more elaborate diffusions. We intend to transpose all correlation phenomena in the driving process, to lighten the investigation and conserve homoscedasticity.

---

* Corresponding author.
E-mail addresses: Bernard.Bercu@math.u-bordeaux1.fr (B. Bercu), Frederic.Proia@math.u-bordeaux1.fr (F. Proïa), Nicolas.Savy@math.univ-toulouse.fr (N. Savy).

To the best of our knowledge, no results are available on Ornstein–Uhlenbeck driven by Ornstein–Uhlenbeck processes defined, over the time interval $[0, T]$, by

$$\begin{cases} dX_t = \theta X_t dt + dV_t \\ dV_t = \rho V_t dt + dW_t \end{cases} \tag{1.2}$$

where $\theta < 0$, $\rho \leq 0$ and (W_t) is a standard Brownian motion. For the sake of simplicity, we choose the initial values $X_0 = 0$ and $V_0 = 0$.

Our motivation for studying (1.2) comes from two observations. On the one hand, Ornstein–Uhlenbeck driven by Ornstein–Uhlenbeck processes are clearly related to stochastic volatility models in financial mathematics (Barndorff-Nielsen and Veraart, 2013; Schoutens, 2000). On the other hand, (1.2) can be seen as a continuous-time version of the first-order stable autoregressive process driven by a first-order autoregressive process recently investigated in Bercu and Proïa (2013) and Proïa, 2013, such as Brockwell and Lindner (2012) does for ARMA processes. It could be interesting, as a future study, to compare the efficiency of our approach with dynamic volatility models on real financial data.

The paper is organized as follows. Section 2 is devoted to the maximum likelihood estimation for θ and ρ. A continuous-time equivalent of the Durbin–Watson statistic is also provided. In Section 3, we establish the almost sure convergence as well as the asymptotic normality of our estimates. One shall realize that there is a radically different behavior of the estimator of ρ in the two situations where $\rho < 0$ and $\rho = 0$. Our analysis relies on technical tools postponed to Section 4.

2. Maximum likelihood estimation

The maximum likelihood estimator of θ is given by

$$\widehat{\theta}_T = \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt} = \frac{X_T^2 - T}{2 \int_0^T X_t^2 dt}. \tag{2.1}$$

In the standard situation where $\rho = 0$, it is well-known that $\widehat{\theta}_T$ converges to θ almost surely. Moreover, as $\theta < 0$, the process (X_T) is positive recurrent and we have the asymptotic normality

$$\sqrt{T}\left(\widehat{\theta}_T - \theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, -2\theta).$$

We shall see in Section 3 that the almost sure limiting value of $\widehat{\theta}_T$ and its asymptotic variance will change as soon as $\rho < 0$. The estimation of ρ requires the evaluation of the residuals generated by the estimation of θ at stage T. For all $0 \leq t \leq T$, denote

$$\widehat{V}_t = X_t - \widehat{\theta}_T \Sigma_t \tag{2.2}$$

where

$$\Sigma_t = \int_0^t X_s ds. \tag{2.3}$$

By analogy with (2.1) and on the basis of the residuals (2.2), we estimate ρ by

$$\widehat{\rho}_T = \frac{\widehat{V}_T^2 - T}{2 \int_0^T \widehat{V}_t^2 dt}. \tag{2.4}$$

Therefore, we are in the position to define the continuous-time equivalent of the discrete-time Durbin–Watson statistic (Bercu and Proïa, 2013; Durbin and Watson, 1950, 1951, 1971),

$$\widehat{D}_T = \frac{2 \int_0^T \widehat{V}_t^2 dt - \widehat{V}_T^2 + T}{\int_0^T \widehat{V}_t^2 dt} = 2(1 - \widehat{\rho}_T). \tag{2.5}$$

3. Main results

The almost sure convergences of our estimates are as follows.

Theorem 3.1. *We have the almost sure convergences*

$$\lim_{T \to \infty} \widehat{\theta}_T = \theta^*, \qquad \lim_{T \to \infty} \widehat{\rho}_T = \rho^* \quad a.s. \tag{3.1}$$

where

$$\theta^* = \theta + \rho \quad and \quad \rho^* = \frac{\theta \rho (\theta + \rho)}{(\theta + \rho)^2 + \theta \rho}. \tag{3.2}$$

Proof. We immediately deduce from (1.2) that

$$\int_0^T X_t \, dX_t = \theta S_T + \rho P_T + M_T^X \tag{3.3}$$

where

$$S_T = \int_0^T X_t^2 \, dt, \qquad P_T = \int_0^T X_t V_t \, dt, \qquad M_T^X = \int_0^T X_t \, dW_t. \tag{3.4}$$

We shall see in Corollary 4.1 below that

$$\lim_{T \to \infty} \frac{1}{T} S_T = -\frac{1}{2(\theta + \rho)} \quad \text{a.s.} \tag{3.5}$$

and in the proof of Corollary 4.2 that

$$\lim_{T \to \infty} \frac{1}{T} P_T = -\frac{1}{2(\theta + \rho)} \quad \text{a.s.} \tag{3.6}$$

Moreover, if $(\mathcal{F}_t)$ stands for the natural filtration of the standard Brownian motion (W_t), then (M_t^X) is a continuous-time $(\mathcal{F}_t)$-martingale with the quadratic variation S_t. Hence, it follows from the strong law of large numbers for continuous-time martingales given *e.g.* in Feigin (1976) or Lépingle (1978), that $M_T^X = o(T)$ a.s. Consequently, we obtain from (3.3) that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T X_t \, dX_t = -\frac{\theta}{2(\theta + \rho)} - \frac{\rho}{2(\theta + \rho)} = -\frac{1}{2} \quad \text{a.s.} \tag{3.7}$$

which leads, via (2.1), to the first convergence in (3.1). The second convergence in (3.1) is more difficult to handle. We infer from (1.2) that

$$\int_0^T V_t \, dV_t = \rho \Lambda_T + M_T^V \tag{3.8}$$

where

$$\Lambda_T = \int_0^T V_t^2 \, dt \quad \text{and} \quad M_T^V = \int_0^T V_t \, dW_t. \tag{3.9}$$

On the one hand, if $\rho < 0$, it is well-known (see *e.g.* Feigin (1976), page 728) that

$$\lim_{T \to \infty} \frac{1}{T} \Lambda_T = -\frac{1}{2\rho} \quad \text{a.s.} \tag{3.10}$$

In addition, (M_t^V) is a continuous-time $(\mathcal{F}_t)$-martingale with the quadratic variation Λ_t. Consequently, $M_T^V = o(T)$ a.s. and we find from (3.8) that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T V_t \, dV_t = -\frac{1}{2} \quad \text{a.s.} \tag{3.11}$$

However, we know from Itô's formula that

$$\frac{1}{T} \int_0^T X_t \, dX_t = \frac{1}{2} \left(\frac{X_T^2}{T} - 1 \right) \quad \text{and} \quad \frac{1}{T} \int_0^T V_t \, dV_t = \frac{1}{2} \left(\frac{V_T^2}{T} - 1 \right).$$

Then, we deduce from (3.7) and (3.11) that

$$\lim_{T \to \infty} \frac{X_T^2}{T} = 0 \quad \text{and} \quad \lim_{T \to \infty} \frac{V_T^2}{T} = 0 \quad \text{a.s.} \tag{3.12}$$

As $X_T = \theta \Sigma_T + V_T$, it clearly follows from (2.2) and (3.12) that

$$\lim_{T \to \infty} \frac{1}{2} \left(\frac{\widehat{V}_T^2}{T} - 1 \right) = -\frac{1}{2} \quad \text{a.s.} \tag{3.13}$$

Hereafter, we have from (2.4) the decomposition

$$\widehat{\rho}_T = \frac{T}{2\widehat{\Lambda}_T} \left(\frac{\widehat{V}_T^2}{T} - 1 \right) \tag{3.14}$$

where

$$\widehat{\Lambda}_T = \int_0^T \widehat{V}_t^2 \mathrm{d}t.$$

We shall see in Corollary 4.2 below that

$$\lim_{T \to \infty} \frac{1}{T}\widehat{\Lambda}_T = -\frac{1}{2\rho^*} \quad \text{a.s.} \tag{3.15}$$

Therefore, (3.14) together with (3.13) and (3.15) directly implies (3.1). On the other hand, if $\rho = 0$, it is clear from (1.2) that for all $t \geq 0$, $V_t = W_t$. Hence, we have from (2.2) and Itô's formula that

$$\widehat{V}_T^2 - T = 2M_T^W - 2W_T \Sigma_T(\widehat{\theta}_T - \theta) + \Sigma_T^2(\widehat{\theta}_T - \theta)^2 \tag{3.16}$$

and

$$\widehat{\Lambda}_T = \Lambda_T - 2(\widehat{\theta}_T - \theta)\int_0^T W_t \Sigma_t \mathrm{d}t + (\widehat{\theta}_T - \theta)^2 \int_0^T \Sigma_t^2 \mathrm{d}t \tag{3.17}$$

where

$$\Lambda_T = \int_0^T W_t^2 \mathrm{d}t \quad \text{and} \quad M_T^W = \int_0^T W_t \mathrm{d}W_t.$$

It is now necessary to investigate the almost sure asymptotic behavior of Λ_T. We deduce from the self-similarity of the Brownian motion (W_t) that

$$\Lambda_T = \int_0^T W_t^2 \mathrm{d}t \overset{\mathcal{L}}{=} T \int_0^T W_{t/T}^2 \mathrm{d}t = T^2 \int_0^1 W_s^2 \mathrm{d}s = T^2 \Lambda_1. \tag{3.18}$$

Consequently, it clearly follows from (3.18) that for any power $0 < a < 2$,

$$\lim_{T \to \infty} \frac{1}{T^a} \Lambda_T = +\infty \quad \text{a.s.} \tag{3.19}$$

As a matter of fact, since Λ_1 is almost surely positive, it is enough to show that

$$\lim_{T \to \infty} \mathbb{E}\left[\exp\left(-\frac{1}{T^a}\Lambda_T\right)\right] = 0. \tag{3.20}$$

However, we have from standard Gaussian calculations (see *e.g.* Liptser and Shiryaev (2001), page 232) that

$$\mathbb{E}\left[\exp\left(-\frac{1}{T^a}\Lambda_T\right)\right] = \mathbb{E}\left[\exp\left(-\frac{T^2}{T^a}\Lambda_1\right)\right] = \frac{1}{\sqrt{\cosh(v_T(a))}}$$

where $v_T(a) = \sqrt{2T^{2-a}}$ goes to infinity, which clearly leads to (3.20). Furthermore, (M_t^W) is a continuous-time $(\mathcal{F}_t)$-martingale with the quadratic variation Λ_t. We already saw that Λ_T which implies that $M_T^W = o(\Lambda_T)$ a.s. In addition, we obviously have $\Sigma_T^2 \leq TS_T$. One can observe that convergence (3.5) still holds when $\rho = 0$, which ensures that $\Sigma_T^2 \leq T^2$ a.s. Moreover, we deduce from the strong law of large numbers for continuous-time martingales that

$$(\widehat{\theta}_T - \theta)^2 = O\left(\frac{\log T}{T}\right) \quad \text{a.s.}$$

which implies that $\Sigma_T^2(\widehat{\theta}_T - \theta)^2 = O(T \log T) = o(\Lambda_T)$ a.s. By the same token, as $X_T^2 = o(T)$ and $W_T^2 = o(T \log T)$ a.s., we find that

$$W_T \Sigma_T(\widehat{\theta}_T - \theta) = o(\Lambda_T) \quad \text{a.s.}$$

Consequently, we obtain from (3.16) that

$$\widehat{V}_T^2 - T = o(\Lambda_T) \quad \text{a.s.} \tag{3.21}$$

It remains to study the almost sure asymptotic behavior of $\widehat{\Lambda}_T$. One can easily see that

$$\int_0^T \Sigma_t^2 \mathrm{d}t \leq \frac{2}{\theta^2}(S_T + \Lambda_T).$$

However, it follows from (3.5) and (3.19) that $S_T = o(\Lambda_T)$ a.s. which ensures that

$$(\widehat{\theta}_T - \theta)^2 \int_0^T \Sigma_t^2 \mathrm{d}t = o(\Lambda_T) \quad \text{a.s.} \tag{3.22}$$

Via the same arguments,

$$(\widehat{\theta}_T - \theta) \int_0^T W_t \Sigma_t \mathrm{d}t = o(\Lambda_T) \quad \text{a.s.} \tag{3.23}$$

Then, we find from (3.17), (3.22) and (3.23) that

$$\widehat{\Lambda}_T = \Lambda_T(1 + o(1)) \quad \text{a.s.} \tag{3.24}$$

Finally, the second convergence in (3.1) follows from (3.21) and (3.24), which achieves the proof of Theorem 3.1. □

Our second result deals with the asymptotic normality of our estimates.

Theorem 3.2. *If* $\rho < 0$, *we have the joint asymptotic normality*

$$\sqrt{T} \begin{pmatrix} \widehat{\theta}_T - \theta^* \\ \widehat{\rho}_T - \rho^* \end{pmatrix} \xrightarrow{\mathscr{L}} \mathscr{N}(0, \Gamma) \tag{3.25}$$

where the asymptotic covariance matrix is given by

$$\Gamma = \begin{pmatrix} \sigma_\theta^2 & \ell \\ \ell & \sigma_\rho^2 \end{pmatrix} \tag{3.26}$$

with $\sigma_\theta^2 = -2\theta^*$,

$$\ell = \frac{2\rho^* \left((\theta^*)^2 - \theta\rho\right)}{(\theta^*)^2 + \theta\rho} \quad \text{and} \quad \sigma_\rho^2 = -\frac{2\rho^* \left((\theta^*)^6 + \theta\rho \left((\theta^*)^4 - \theta\rho \left(2(\theta^*)^2 - \theta\rho\right)\right)\right)}{\left((\theta^*)^2 + \theta\rho\right)^3}.$$

In particular, we have

$$\sqrt{T} \left(\widehat{\theta}_T - \theta^*\right) \xrightarrow{\mathscr{L}} \mathscr{N}(0, \sigma_\theta^2) \tag{3.27}$$

and

$$\sqrt{T} \left(\widehat{\rho}_T - \rho^*\right) \xrightarrow{\mathscr{L}} \mathscr{N}(0, \sigma_\rho^2). \tag{3.28}$$

Proof. We obtain from (2.1) the decomposition

$$\widehat{\theta}_T - \theta^* = \frac{M_T^X}{S_T} + \frac{R_T^X}{S_T} \tag{3.29}$$

where

$$R_T^X = \rho \int_0^T X_t(V_t - X_t)\mathrm{d}t = -\theta\rho \int_0^T \Sigma_t \mathrm{d}\Sigma_t = -\frac{\theta\rho}{2} \Sigma_T^2.$$

We shall now establish a similar decomposition for $\widehat{\rho}_T - \rho^*$. It follows from (2.2) that for all $0 \le t \le T$,

$$\widehat{V}_t = X_t - \widehat{\theta}_T \Sigma_t = V_t - (\widehat{\theta}_T - \theta)\Sigma_t = V_t - (\widehat{\theta}_T - \theta^*)\Sigma_t - \rho\Sigma_t$$
$$= V_t - \frac{\rho}{\theta}(X_t - V_t) - \frac{1}{\theta}(\widehat{\theta}_T - \theta^*)(X_t - V_t) = \frac{\theta^*}{\theta}V_t - \frac{\rho}{\theta}X_t - \frac{1}{\theta}(\widehat{\theta}_T - \theta^*)(X_t - V_t),$$

which leads to

$$\widehat{\Lambda}_T = I_T + (\widehat{\theta}_T - \theta^*)\left(J_T + (\widehat{\theta}_T - \theta^*)K_T\right), \tag{3.30}$$

where

$$I_T = \frac{1}{\theta^2} \left(\rho^2 S_T + (\theta^*)^2 \Lambda_T - 2\theta^* \rho P_T\right),$$

$$J_T = \frac{1}{\theta^2} \left(2\rho S_T + 2\theta^* \Lambda_T - 2(\theta + 2\rho)P_T\right),$$

$$K_T = \frac{1}{\theta^2} \left(S_T + \Lambda_T - 2P_T\right).$$

Then, we deduce from (2.4) and (3.30) that

$$\widehat{\Lambda}_T(\widehat{\rho}_T - \rho^*) = \frac{I_T^V}{2} + \frac{1}{2}(\widehat{\theta}_T - \theta^*)\left(J_T^V + (\widehat{\theta}_T - \theta^*)K_T^V\right) \tag{3.31}$$

in which $I_T^V = \widehat{V}_T^2 - T - 2\rho^* I_T, J_T^V = -2\rho^* J_T$, and $K_T^V = -2\rho^* K_T$. At this stage, in order to simplify the complicated expression (3.31), we repeatedly make use of Itô's formula. For all $0 \le t \le T$, we have

$$\Lambda_t = \frac{1}{2\rho}V_t^2 - \frac{1}{\rho}M_t^V - \frac{t}{2\rho},$$

$$P_t = \frac{1}{\theta^*}X_t V_t - \frac{1}{2\theta^*}V_t^2 - \frac{1}{\theta^*}M_t^X - \frac{t}{2\theta^*},$$

$$S_t = \frac{1}{2\theta}X_t^2 + \frac{\rho}{2\theta^*\theta}V_t^2 - \frac{\rho}{\theta^*\theta}X_t V_t - \frac{1}{\theta^*}M_t^X - \frac{t}{2\theta^*},$$

where the continuous-time martingales M_t^X and M_t^V were previously defined in (3.4) and (3.9). Therefore, it follows from tedious but straightforward calculations that

$$\widehat{\Lambda}_T(\widehat{\rho}_T - \rho^*) = C_X M_T^X + C_V M_T^V + \frac{J_T^V}{2}(\widehat{\theta}_T - \theta^*) + R_T^V \tag{3.32}$$

where

$$C_V = \frac{(\theta^*)^2 \rho^*}{\theta^2 \rho} \quad \text{and} \quad C_X = -\frac{\rho(2\theta + \rho)\rho^*}{\theta^2 \theta^*}.$$

The remainder R_T^V is similar to R_T^X and they play a negligible role. The combination of (3.29) and (3.32) leads to the vectorial expression

$$\sqrt{T}\begin{pmatrix} \widehat{\theta}_T - \theta^* \\ \widehat{\rho}_T - \rho^* \end{pmatrix} = \frac{1}{\sqrt{T}}A_T Z_T + \sqrt{T}R_T \tag{3.33}$$

where

$$A_T = \begin{pmatrix} S_T^{-1}T & 0 \\ B_T \widehat{\Lambda}_T^{-1}T & C_V \widehat{\Lambda}_T^{-1}T \end{pmatrix}, \qquad R_T = \begin{pmatrix} S_T^{-1}R_T^X \\ \widehat{\Lambda}_T^{-1}D_T \end{pmatrix}$$

with $B_T = C_X + J_T^V(2S_T)^{-1}$ and $D_T = R_T^V + J_T^V(2S_T)^{-1}R_T^X$. The leading term in (3.33) is the continuous-time vector $(\mathcal{F}_t)$-martingale (Z_t) with the predictable quadratic variation $\langle Z \rangle_t$ given by

$$Z_t = \begin{pmatrix} M_t^X \\ M_t^V \end{pmatrix} \quad \text{and} \quad \langle Z \rangle_t = \begin{pmatrix} S_t & P_t \\ P_t & \Lambda_t \end{pmatrix}.$$

We deduce from (3.5), (3.6) and (3.10) that

$$\lim_{T \to \infty} A_T = A \quad \text{a.s.} \tag{3.34}$$

where A is the limiting matrix given by

$$A = \begin{pmatrix} -2\theta^* & 0 \\ -2\rho^*(C_X - 2(\theta\rho)^{-1}\theta^*\rho^*) & -2\rho^* C_V \end{pmatrix}.$$

By the same token, we immediately have from (3.5), (3.6) and (3.10) that

$$\lim_{T \to \infty} \frac{\langle Z \rangle_T}{T} = \Delta = -\frac{1}{2\theta^*}\begin{pmatrix} 1 & 1 \\ 1 & \theta^*\rho^{-1} \end{pmatrix} \quad \text{a.s.} \tag{3.35}$$

Furthermore, it clearly follows from Corollary 4.3 below that

$$\frac{X_T^2}{\sqrt{T}} \xrightarrow{\mathscr{P}} 0 \quad \text{and} \quad \frac{V_T^2}{\sqrt{T}} \xrightarrow{\mathscr{P}} 0. \tag{3.36}$$

Finally, as $\Gamma = A\Delta A'$, the joint asymptotic normality (3.25) follows from the conjunction of (3.33)–(3.36) together with Slutsky's lemma and the central limit theorem for continuous-time vector martingales given *e.g.* in Feigin (1976), which achieves the proof of Theorem 3.2. □

Theorem 3.3. *If $\rho = 0$, we have the convergence in distribution*

$$T \widehat{\rho}_T \xrightarrow{\mathscr{L}} \mathcal{W} \tag{3.37}$$

where the limiting distribution $\mathcal{W}$ is given by

$$\mathcal{W} = \frac{\int_0^1 B_s dB_s}{\int_0^1 B_s^2 ds} = \frac{B_1^2 - 1}{2 \int_0^1 B_s^2 ds} \tag{3.38}$$

and (B_t) is a standard Brownian motion.

Proof. Via the same reasoning as in Section 2 of Feigin (1979), it follows from the self-similarity of the Brownian motion (W_t) that

$$\left(\int_0^T W_t^2 dt, \frac{1}{2} \left(W_T^2 - T \right) \right) \stackrel{\mathscr{L}}{=} \left(T \int_0^T W_{t/T}^2 dt, \frac{T}{2} \left(W_1^2 - 1 \right) \right)$$

$$= \left(T^2 \int_0^1 W_s^2 ds, \frac{T}{2} \left(W_1^2 - 1 \right) \right). \tag{3.39}$$

Moreover, we obtain from (3.30) that

$$\widehat{\Lambda}_T = \alpha_T S_T + \beta_T P_T + \gamma_T \Lambda_T \tag{3.40}$$

where

$$\alpha_T = \frac{1}{\theta^2} (\widehat{\theta}_T - \theta)^2,$$

$$\beta_T = -\frac{2}{\theta} (\widehat{\theta}_T - \theta) - \frac{2}{\theta^2} (\widehat{\theta}_T - \theta)^2,$$

$$\gamma_T = 1 + \frac{2}{\theta} (\widehat{\theta}_T - \theta) + \frac{1}{\theta^2} (\widehat{\theta}_T - \theta)^2.$$

By Theorem 3.1, $\widehat{\theta}_T$ converges almost surely to θ which implies that α_T, β_T, and γ_T converge almost surely to 0, 0 and 1. Hence, we deduce from (3.5), (3.6) and (3.40) that

$$\widehat{\Lambda}_T = \Lambda_T (1 + o(1)) \quad \text{a.s.} \tag{3.41}$$

Furthermore, one can observe that $\widehat{V}_T^2 / T$ shares the same asymptotic distribution as W_T^2 / T. Finally, (3.37) follows from (3.39) and (3.41) together with the continuous mapping theorem. $\square$

Remark 3.1. The asymptotic behavior of $\widehat{\rho}_T$ when $\rho < 0$ and $\rho = 0$ is closely related to the results previously established for the unstable discrete-time autoregressive process (Chan and Wei, 1988; Feigin, 1979; White, 1958). According to Corollary 3.1.3 of Chan and Wei (1988), we can express

$$\mathcal{W} = \frac{\mathcal{T}^2 - 1}{2 \mathcal{S}}$$

where $\mathcal{T}$ and $\mathcal{S}$ are given by the Karhunen–Loeve expansions

$$\mathcal{T} = \sqrt{2} \sum_{n=1}^{\infty} \gamma_n Z_n \quad \text{and} \quad \mathcal{S} = \sum_{n=1}^{\infty} \gamma_n^2 Z_n^2$$

with $\gamma_n = 2(-1)^n / ((2n - 1)\pi)$ and (Z_n) being a sequence of independent random variables with $\mathcal{N}(0, 1)$ distribution.

Remark 3.2. It immediately follows from our previous results that $\widehat{D}_T$ converges almost surely to $D^* = 2 (1 - \rho^*)$. In addition, if $\rho < 0$, we have the asymptotic normality

$$\sqrt{T} \left(\widehat{D}_T - D^* \right) \xrightarrow{\mathscr{L}} \mathcal{N}(0, \sigma_D^2)$$

where $\sigma_D^2 = 4 \sigma_\rho^2$ whereas, if $\rho = 0$,

$$T \left(\widehat{D}_T - 2 \right) \xrightarrow{\mathscr{L}} -2 \mathcal{W}.$$

4. Some technical tools

First of all, most of our results rely on the following keystone lemma.

Lemma 4.1. *The process* (X_t) *is geometrically ergodic.*

Proof. It follows from (1.2) that

$$\mathrm{d}X_t = (\theta + \rho)X_t\mathrm{d}t - \theta\rho\Sigma_t\mathrm{d}t + \mathrm{d}W_t \qquad (4.1)$$

where we recall that

$$\Sigma_t = \int_0^t X_s\mathrm{d}s.$$

Consequently, if

$$\Phi_t = \begin{pmatrix} X_t \\ \Sigma_t \end{pmatrix},$$

we clearly deduce from (4.1) that

$$\mathrm{d}\Phi_t = A\Phi_t\mathrm{d}t + \mathrm{d}B_t$$

where

$$A = \begin{pmatrix} \theta + \rho & -\theta\rho \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B_t = \begin{pmatrix} W_t \\ 0 \end{pmatrix}.$$

The geometric ergodicity of (Φ_t) only depends on the sign of $\lambda_{\max}(A)$, *i.e.* the largest eigenvalue of A, which has to be negative. An immediate calculation shows that

$$\lambda_{\max}(A) = \max(\theta, \ \rho)$$

which ensures that $\lambda_{\max}(A) < 0$ as soon as $\rho < 0$. Moreover, if $\rho = 0$, (X_t) is an ergodic Ornstein–Uhlenbeck process since $\theta < 0$, which completes the proof of Lemma 4.1. $\quad\square$

Corollary 4.1. *We have the almost sure convergence*

$$\lim_{T \to \infty} \frac{1}{T}S_T = -\frac{1}{2(\theta + \rho)} \quad a.s. \qquad (4.2)$$

Proof. According to Lemma 4.1, it is necessary to establish the asymptotic behavior of $\mathbb{E}[X_t^2]$. Denote $\alpha_t = \mathbb{E}[X_t^2]$, $\beta_t = \mathbb{E}[\Sigma_t^2]$ and $\gamma_t = \mathbb{E}[X_t\Sigma_t]$. One obtains from Itô's formula that

$$\frac{\partial U_t}{\partial t} = CU_t + I$$

where

$$U_t = \begin{pmatrix} \alpha_t \\ \beta_t \\ \gamma_t \end{pmatrix}, \qquad C = \begin{pmatrix} 2(\theta + \rho) & 0 & -2\theta\rho \\ 0 & 0 & 2 \\ 1 & -\theta\rho & \theta + \rho \end{pmatrix}, \qquad I = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

It is not hard to see that $\lambda_{\max}(C) = \max(\theta + \rho, \ 2\theta, \ 2\rho)$. On the one hand, if $\rho < 0$, $\lambda_{\max}(C) < 0$ which implies that

$$\lim_{t \to \infty} U_t = -C^{-1}I.$$

It means that

$$\lim_{t \to \infty} \alpha_t = -\frac{1}{2(\theta + \rho)}, \qquad \lim_{t \to \infty} \beta_t = -\frac{1}{2\theta\rho(\theta + \rho)}, \qquad \lim_{t \to \infty} \gamma_t = 0.$$

Hence, (4.2) follows from Lemma 4.1 together with the ergodic theorem. On the other hand, if $\rho = 0$, (X_t) is a positive recurrent Ornstein–Uhlenbeck process and convergence (4.2) is well-known. $\quad\square$

Corollary 4.2. *If* $\rho < 0$, *we have the almost sure convergence*

$$\lim_{T \to \infty} \frac{1}{T}\widehat{\Lambda}_T = -\frac{(\theta + \rho)^2 + \theta\rho}{2\theta\rho(\theta + \rho)} \quad a.s.$$

Proof. If $\rho < 0$, (V_t) is a positive recurrent Ornstein–Uhlenbeck process and it is well-known that

$$\lim_{T \to \infty} \frac{1}{T}\Lambda_T = -\frac{1}{2\rho} \quad a.s.$$

In addition, as $X_t = \theta \Sigma_t + V_t$,

$$\int_0^T X_t \Sigma_t \mathrm{d}t = \frac{1}{\theta}(S_T - P_T).$$

However, we already saw in the proof of Corollary 4.1 that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T X_t \Sigma_t \mathrm{d}t = 0 \quad \text{a.s.}$$

which leads, via (4.2), to the almost sure convergence

$$\lim_{T \to \infty} \frac{P_T}{T} = -\frac{1}{2(\theta + \rho)} \quad \text{a.s.}$$

Consequently, we deduce from (3.1) together with (3.30) that

$$\lim_{T \to \infty} \frac{1}{T} \widehat{\Lambda}_T = \lim_{T \to \infty} \frac{1}{T} I_T = -\frac{(\theta + \rho)^2 + \theta \rho}{2 \theta \rho (\theta + \rho)} \quad \text{a.s.}$$

which achieves the proof of Corollary 4.2. □

Corollary 4.3. *If $\rho < 0$, we have the asymptotic normalities*

$$X_T \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, -\frac{1}{2(\theta + \rho)}\right) \quad \text{and} \quad V_T \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, -\frac{1}{2\rho}\right).$$

The asymptotic normality of X_T still holds in the particular case where $\rho = 0$.

Proof. This asymptotic normality is a well-known result for the Ornstein–Uhlenbeck process (V_t) with $\rho < 0$. In addition, one can observe that for all $t \geq 0$, $\mathbb{E}[X_t] = 0$. The end of the proof is a direct consequence of the Gaussianity of (X_t) together with Lemma 4.1 and Corollary 4.1. □

Acknowledgments

References

Barndorff-Nielsen, O.E., Basse-O'Connor, A., 2011. Quasi Ornstein–Uhlenbeck processes. Bernoulli 17 (3), 916–941.
Barndorff-Nielsen, O.E., Shephard, N., 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (2), 167–241.
Barndorff-Nielsen, O.E., Veraart, A., 2013. Stochastic volatility of volatility and variance risk premia. J. Financ. Econ. 11 (1), 1–46.
Bercu, B., Proïa, F., 2013. A sharp analysis on the asymptotic behavior of the Durbin–Watson statistic for the first-order autoregressive process. ESAIM Probab. Stat. 17 (1), 500–530.
Brockwell, P.J., Lindner, A., 2012. Ornstein–Uhlenbeck related models driven by Lévy processes. In: Statistical Methods for Stochastic Differential Equations. In: Monogr. Statist. Appl. Probab., vol. 124. CRC Press, pp. 383–427.
Chan, N.H., Wei, C., 1988. Limiting distributions of least squares estimates of unstable autoregressive processes. Ann. Statist. 16 (1), 367–401.
Durbin, J., Watson, G.S., 1950. Testing for serial correlation in least squares regression. I. Biometrika 37, 409–428.
Durbin, J., Watson, G.S., 1951. Testing for serial correlation in least squares regression. II. Biometrika 38, 159–178.
Durbin, J., Watson, G.S., 1971. Testing for serial correlation in least squares regession. III. Biometrika 58, 1–19.
Feigin, P.D., 1976. Maximum likelihood estimation for continuous-time stochastic processes. Adv. Appl. Probab. 8 (4), 712–736.
Feigin, P.D., 1979. Some comments concerning a curious singularity. J. Appl. Probab. 16 (2), 440–444.
Hu, Y., Long, H., 2007. Parameter estimation for Ornstein–Uhlenbeck processes driven by α-stable Lévy motions. Commun. Stoch. Anal. 1 (2), 175–192.
Jongbloed, G., Van der Meulen, F.H., Van der Vaart, A.W., 2005. Nonparametric inference for Lévy-driven Ornstein–Uhlenbeck processes. Bernoulli 11 (5), 759–791.
Kutoyants, Y.A., 2004. Statistical inference for ergodic diffusion processes. In: Springer Series in Statistics. Springer-Verlag London Ltd., London.
Lépingle, D., 1978. Sur le comportement asymptotique des martingales locales. In: Séminaire de Probabilités, XII. In: Lecture Notes in Math, vol. 649. Springer, Berlin, pp. 148–161.
Liptser, R.S., Shiryaev, A.N., 2001. Statistics of Random Processes. II. In: Applications of Mathematics (New York), vol. 6. Springer-Verlag, Berlin.
Onalan, O., 2009. Financial modelling with Ornstein–Uhlenbeck processes driven by Lévy process. In: Proceedings of the World Congress Engineering, vol. 2. pp. 1–6.
Proïa, F., 2013. Further results on the H-test of Durbin for stable autoregressive processes. J. Multivariate Anal. 118, 77–101.
Schoutens, W., 2000. Stochastic Processes and Orthogonal Polynomials. In: Lecture Notes in Statistics, vol. 146. Springer-Verlag, New York.
Uhlenbeck, G.E., Ornstein, L.S., 1930. On the theory of brownian motion. Phys. Rev. 36, 823–841.
White, J.S., 1958. The limiting distribution of the serial correlation coefficient in the explosive case. Ann. Math. Statist. 29, 1188–1197.