Exact asymptotic bias for estimators of the Ornstein-Uhlenbeck process

D. Bosq

LSTA, Université Pierre et Marie Curie - Paris 6

10ème Colloque Franco-Roumain de Mathématiques Appliquées

Outline

Definition

Consider a real stationary markov zero mean gaussian process $X = (X_t,\, t \in \mathbb{R})$ with a continuous nondegenerated autocorrelation $(\rho(h),\, h \geq 0)$, then, there exists $\theta > 0$ such that

$$\rho(h) = \exp(-\theta h),\ h \geq 0,$$

this is the so-called **Ornstein-Uhlenbeck process (OU)**.

One may also define OU as the unique stationary solution of the stochastic differential equation

$$dX_t = -\theta X_t \, dt + \sigma dW_t$$

where W is a bilateral standard Wiener process.
Interpretation: X is the speed of a particle submitted to brownian motion.

Finally, another simple form of OU is

$$X_t = \frac{e^{-\theta t}}{\sqrt{2\theta}} W_1(e^{2\theta t}), \; t \geq 0,$$

where W_1 is a standard Wiener process.

One may also define OU as the unique stationary solution of the stochastic differential equation

$$dX_t = -\theta X_t \, dt + \sigma dW_t$$

where W is a bilateral standard Wiener process.
Interpretation: X is the speed of a particle submitted to brownian motion.

Finally, another simple form of OU is

$$X_t = \frac{e^{-\theta t}}{\sqrt{2\theta}} W_1(e^{2\theta t}), \; t \geq 0,$$

where W_1 is a standard Wiener process.

The information inequality

In the following we suppose that $\sigma = 1$ and we intend to estimate θ and $g(\theta)$ from the observation of $X_{(T)} = (X_t, 0 \leq t \leq T)$.
The information inequality (or Fréchet-Darmois-Cramer-Rao inequality) is

$$E_\theta\left(g(\theta_T) - g(\theta)\right)^2 \geq \frac{\left(b_T^{'}(\theta) + g^{'}(\theta)\right)^2}{I_T(\theta)} + b_T^2(\theta),$$

where $\mathrm{I}_T(\theta)$ is the Fisher information and $\mathrm{b}_T(\theta)$ the bias of the estimator $g(\theta_T)$. Thus, in order to evaluate the above lower bound, it is necessary to study the bias and the bias derivative of $g(\theta_T)$.

A family of asymptotically efficient estimators

Consider the family $\mathcal{F}$ of estimators of the form

$$\theta_T = \theta_T(\alpha, \beta, \triangle_T) = \frac{T - \alpha X_0^2 - \beta X_T^2}{2B_T} + \triangle_T,$$

where $B_T = \int_0^T X_t^2\, dt$, $\alpha, \beta \in \mathbb{R}$ and $\triangle_T$ is a statistic satisfying

$$\textbf{(C)} \quad \Delta_T \xrightarrow{a.s.} 0, T^{\frac{p}{2}} E_\theta |\triangle_T|^p \to 0, p \geq 1, TE_\theta(\triangle_T) \to \delta_\theta, \quad T \to \infty$$

where δ_θ depends on $(\triangle_T)$ and θ

Examples

The empirical estimator (EE) is given by

$$\bar{\theta}_T = \frac{T}{2B_T}$$

The conditional likelihood of $X_{(T)}$ is

$$L = \exp\left(A_T\theta - B_T\frac{\theta^2}{2}\right)$$

where $A_T = \frac{T+X_0^2-X_T^2}{2}$, hence, the conditional maximum likelihood estimator (CMLE):

$$\hat{\theta}_T = \frac{A_T}{B_T}$$

Examples

The empirical estimator (EE) is given by

$$\bar{\theta}_T = \frac{T}{2B_T}$$

The conditional likelihood of $X_{(T)}$ is

$$L = \exp\left(A_T\theta - B_T\frac{\theta^2}{2}\right)$$

where $A_T = \frac{T+X_0^2-X_T^2}{2}$, hence, the conditional maximum likelihood estimator (CMLE):

$$\hat{\theta}_T = \frac{A_T}{B_T}$$

Now, the likelihood is

$$\tilde{L} = \sqrt{\frac{\theta}{\pi}} \exp(-\theta X_0^2) . L,$$

and the maximum likelihood estimator (MLE):

$$\tilde{\theta}_T = \frac{(A_T - X_0^2) + \sqrt{(A_T - X_0^2) + 2B_T}}{2B_T},$$

Finally, the reverse conditional maximum likelihood estimator(RCMLE) has the form

$$\check{\theta}_T = \frac{A_T'}{B_T},$$

where $A_T' = \frac{T + X_T^2 - X_0^2}{2}$.

Now, the likelihood is

$$\tilde{L} = \sqrt{\frac{\theta}{\pi}} \exp(-\theta X_0^2) . L,$$

and the maximum likelihood estimator (MLE):

$$\tilde{\theta}_T = \frac{(A_T - X_0^2) + \sqrt{(A_T - X_0^2) + 2B_T}}{2B_T},$$

Finally, the reverse conditional maximum likelihood estimator(RCMLE) has the form

$$\check{\theta}_T = \frac{A'_T}{B_T},$$

where $A'_T = \frac{T + X_T^2 - X_0^2}{2}$.

These estimators belong to $\mathcal{F}$:

$$\bar{\theta}_T = \theta_T(0,0,0)$$

$$\hat{\theta}_T = \theta_T(-1,1,0)$$

$$\check{\theta}_T = \theta_T(1,-1,0)$$

$$\widetilde{\theta}_T = (1,1,\triangle_T),$$

where $\triangle_T = \frac{T}{4B_T}\left[\left(\Gamma_T^2 + \frac{8B_T}{T^2}\right)^{\frac{1}{2}} - \Gamma_T\right]$, with $\Gamma_T = 1 - \frac{X_0^2+X_T^2}{T}$.

Note that $\mathcal{F}$ is a convex set.

Asymptotic efficiency

Proposition

For each $\theta_T \in \mathcal{F}$ one has

$$T^{\frac{p}{2}} E_\theta \left|\theta_T - \theta\right|^p \to (2\theta)^{\frac{p}{2}} E_\theta |N|^p, \; p \geq 1,$$

and

$$T^{\frac{1}{2}} (\theta_T - \theta) \Longrightarrow (2\theta)^{\frac{1}{2}} N,$$

where $N \sim \mathcal{N}(0,1)$.

Proof.

It is an easy consequence of Kutoyants (2004,2009). □

Bias of $\bar{\theta}_T$

First note that

$$\bar{\theta}_T = \frac{1}{2}\left(\hat{\theta}_T + \check{\theta}_T\right)$$

Then, since X is gaussian stationary, the three estimators have the same bias. Moreover this bias is positive:

$$E_\theta(\bar{\theta}_T) > \frac{1}{E_\theta(2T^{-1}B_T)} = \frac{1}{2(2\theta^{-1})} = \theta$$

In order to study this bias one may use the representation of X as the transform of a Wiener process for obtaining

$$b_T(\theta, X) = \theta b_{\theta T}(1, Y)$$

where $Y_t = \sqrt{\theta}\, X_{t/\theta}$.

Bias of $\bar{\theta}_T$

First note that

$$\bar{\theta}_T = \frac{1}{2}\left(\hat{\theta}_T + \check{\theta}_T\right)$$

Then, since X is gaussian stationary, the three estimators have the same bias. Moreover this bias is positive:

$$E_\theta(\bar{\theta}_T) > \frac{1}{E_\theta(2T^{-1}B_T)} = \frac{1}{2(2\theta^{-1})} = \theta$$

In order to study this bias one may use the representation of X as the transform of a Wiener process for obtaining

$$b_T(\theta, X) = \theta b_{\theta T}(1, Y)$$

where $Y_t = \sqrt{\theta}\, X_{t/\theta}$.

It follows that

$$b_T'(\theta) = \mathcal{O}\left(\frac{\ln T}{T}\right)$$

and

$$T\, b_T(\theta) \to 2$$

The general case

For the general θ_T one obtains

$$\frac{\partial}{\partial \theta} E_\theta(\theta_T - \theta) \to 0,$$

and

$$T.E_\theta\left(\theta_T(\alpha, \beta, \Delta_T) - \theta\right) \to 2 - \frac{\alpha + \beta}{2} + \delta_\theta.$$

For the MLE one has again

$$T.E_\theta(\widetilde{\theta}_T - \theta) \to 2.$$

The general case

For the general θ_T one obtains

$$\frac{\partial}{\partial\theta} E_\theta(\theta_T - \theta) \to 0,$$

and

$$T.E_\theta\left(\theta_T(\alpha,\beta,\Delta_T) - \theta\right) \to 2 - \frac{\alpha+\beta}{2} + \delta_\theta.$$

For the MLE one has again

$$T.E_\theta(\widetilde{\theta}_T - \theta) \to 2.$$

Asymptotic efficiency of g(θ_T^*)

Let $g : \mathbb{R}_+^* \mapsto \mathbb{R}$, in order to estimate $g(\theta)$, one sets

$$\theta_T^* = \max\left(\theta_T, e^{-T}\right), \theta_T \in \mathscr{F}, \; T > 0.$$

Clearly θ_T^* and θ_T have the same asymptotic behaviour and, under mild conditions, $g(\theta_T^*)$ is asymptotically efficient.

For example, if g is derivable, one has

$$T^{1/2}\left(g(\theta_T^*) - g(\theta)\right) \Rightarrow (2\theta)^{1/2} \left|g'(\theta)\right| N,$$

and if, in addition $\left|g'(\theta)\right| \leq c_\theta \, \theta^m$, $m \geq 0$, then

$$E_\theta\left(T^{p/2} \left|g(\theta_T^*) - g(\theta)\right|^p\right) \to (2\theta)^{p/2} \left|g'(\theta)\right|^p E_\theta\left[|N|^p\right] \quad p \geq 1.$$

Asymptotic efficiency of $g(\theta_T^*)$

Let $g : \mathbb{R}_+^* \mapsto \mathbb{R}$, in order to estimate $g(\theta)$, one sets

$$\theta_T^* = \max\left(\theta_T, e^{-T}\right), \theta_T \in \mathscr{F}, \; T > 0.$$

Clearly θ_T^* and θ_T have the same asymptotic behaviour and, under mild conditions, $g(\theta_T^*)$ is asymptotically efficient.
For example, if g is derivable, one has

$$T^{1/2}\left(g(\theta_T^*) - g(\theta)\right) \Rightarrow (2\theta)^{1/2} \left|g^{'}(\theta)\right| N,$$

and if, in addition $\left|g^{'}(\theta)\right| \leq c_\theta\, \theta^m$, $m \geq 0$, then

$$E_\theta\left(T^{p/2} \left|g(\theta_T^*) - g(\theta)\right|^p\right) \to (2\theta)^{p/2} \left|g^{'}(\theta)\right|^p E_\theta\left[|N|^p\right] \quad p \geq 1.$$

Asymptotic bias for $g(\theta)$

Proposition

If g has three continuous derivatives with

$$|g'''(\theta)| \leq c\,\theta^m,\ \theta > 0\ \ (c > 0, m \geq 0)$$

and $\theta_T = \theta_T(\alpha, \beta, \Delta_T)$ *then*

$$T.E_\theta\left(g(\theta_T^*) - g(\theta)\right) \to \left(2 - \frac{\alpha+\beta}{2} + \delta_\theta\right) g'(\theta) + \theta g''(\theta).$$

Again, the four "classical" estimators have the same asymptotic bias: $2\,g'(\theta) + \theta\, g''(\theta)$.

Bias derivative

Proposition

If g has one continuous derivative such that

$$\left|g'(u)\right| \leq c\,|u|^m \quad u \in \mathbb{R},$$

for some $c > 0$ and $m \geq 0$, and if $E\left(g(\theta.\bar{\theta}_{\theta T}(Y)\right)$ is differentiable under expectation, then

$$\frac{\partial}{\partial \theta} E_\theta\left(g(\bar{\theta}_T) - g(\theta)\right) \xrightarrow[T \to \infty]{} 0,$$

and the same property holds for each θ_T in F.

Examples

If $g(\theta)$ is a polynomial the result applies. In particular

$$T.E_\theta(\bar{\theta}_T^2 - \theta^2) \to 6\theta.$$

If $g(\theta) = \exp(\text{-}\theta h) = \rho(h)$, $(h > 0)$, one obtains

$$TE_\theta\left(\exp(-\theta_T^* h) - \exp(-\theta h)\right) \to \left(\theta h - 2 + \frac{\alpha + \beta}{2} - \delta_\theta\right) h \exp(-\theta h).$$

Examples

If $g(\theta)$ is a polynomial the result applies. In particular

$$T.E_\theta(\bar{\theta}_T^2 - \theta^2) \to 6\theta.$$

If $g(\theta) = \exp(\text{-}\theta h) = \rho(h)$, $(h > 0)$, one obtains

$$TE_\theta\left(\exp(-\theta_T^* h) - \exp(-\theta h)\right) \to \left(\theta h - 2 + \frac{\alpha+\beta}{2} - \delta_\theta\right) h \exp(-\theta h).$$

Examples

If $g(\theta) = \frac{c}{\theta} + d$ (c and d constants), assumption in the previous Proposition is not satisfied and we have

$$2g'(\theta) + \theta g''(\theta) = 0.$$

Actually, a slight modification of the proof gives

$$T.E_\theta(g(\bar{\theta}_T) - g(\theta)) \to 0,$$

which is natural since $g(\bar{\theta}_T)$ is an unbiased estimator of $g(\theta)$!

Reducing the bias for θ

If $\delta_\theta = \delta$ does not depends on θ, one may reduce the bias of θ_T by putting

$$\theta_T^{(1)} = \theta_T - \frac{2 - \frac{\alpha+\beta}{2} + \delta}{T},$$

then, clearly, $\theta_T^{(1)}$ remains asymptotically efficient and

$$T.E_\theta(\theta_T^{(1)} - \theta) \to 0.$$

Note that $\theta_T^{(1)} \in \mathscr{F}$, actually $\theta_T^{(1)} = \theta_T\left(\alpha, \beta, \Delta_T - \frac{2-\frac{\alpha+\beta}{2}+\delta}{T}\right)$. In particular, for $\hat{\theta}_T, \check{\theta}_T, \tilde{\theta}_T$ and$\bar{\theta}_T$, $\theta_T^{(1)}$ is obtained by substracting $\frac{2}{T}$.

Reducing the bias for g(θ)

The situation is somewhat different for $g(\theta)$; putting

$$\tilde{\bar{\theta}}_T = \max\left(\bar{\theta}_T - \frac{2}{T}, \exp(-T)\right)$$

one obtains

$$TE_\theta\left(g(\tilde{\bar{\theta}}_T) - g(\theta)\right) \to \theta g''(\theta).$$

If $g'.g''$ is positive, the absolute value of the asymptotic bias is reduced, but it is not the case in a general situation (cf $g(\theta)$=exp(-θ) at $\theta = 3$).

References

Billingsley, P.
Convergence of probability measures.
Wiley, Probability and Statistics, New York, second edition, 1999.

Bosq, D.
Exact asymptotic bias for estimators of the Ornstein-Uhlenbeck process.
Stat. Infer. Stoch. Process., 13, 133-145, 2010.

Bosq, D. and Blanke, D.
Inference and prediction in large dimensions.
Wiley-Dunod, Chichester, 2007.

Kutoyants, Yu. A.
Statistical Inference for Ergodic Diffusion Processes.
Springer Series In Statistics, 2004.

Kutoyants, Yu. A., Personnal communication, 2009.

Liptser, R. S. and Shiryaev, A. N.
Statistics of random processes I, II.
Springer, New-York, 2nd edition, 2001.