Is a Brownian motion skew?

Ernesto Mordecki

Sesión en honor a Mario Wschebor

Universidad de la República, Montevideo, Uruguay

XI CLAPEM - November 2009 - Venezuela[1]

[1] Joint work with Antoine Lejay and Soledad Torres

Outline

What is Skew Brownian motion

The Skew Brownian motion $X = \{X_t : 0 \leq t \leq T\}$ is the strong solution of the sde

$$X_t = x + B_t + \theta \ell_t^x,$$

- $B = \{B_t : 0 \leq t \leq T\}$ is a standard Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{P})$.
- $x \geq 0$ is the initial condition,
- $\theta \in [-1, 1]$ is the *skweness parameter*,
- $\ell^x = \{\ell_t^x : 0 \leq t \leq T\}$ is the local time at level zero of the (unknown) X.

Particular cases

- In case $\theta = 0$ we have the usual BM
- In case $\theta = 1$ we have the reflected BM, equal in distribution (by Skorohod's result) to $|x + B_t|$.
- In case $\theta = -1$ we have BM up to the first time the process hits $x = 0$, after, negative reflected BM.

Alternative construction of SBM (I)

We can define the SBm as the weak limit of scaled markov chain on $\mathbb{Z}$, with transition probabilities

$$p(i,j) = \begin{cases} 1/2, & \text{if } i \neq 0, j = i \pm 1 \\ p, & \text{if } i = 0, j = 1 \\ 1-p, & \text{if } i = 0, j = -1 \\ 0, & \text{in other cases} \end{cases}$$

with $p = (\theta + 1)/2$.

Remark

This construction shows that the process is Markovian

Alternative construction of SBM (II)

We can construct the SBm departing from the excursions of the original BM. In this case we choose with probability p whether the excursion is positive[2].

[2] Lejay, A., On the constructions of the Skew Brownian motion, *Probab. Surv.*, **3**, (2006), 413–466

Density

As a Markov process, the transition density of the SBM is given by

$$q_\theta(t,x,y) = p(t,y-x) + \mathrm{sgn}(y)\theta p(t,|x|+|y|),$$

where

$$p(t,x) = \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{x^2}{2t}\Big),$$

is the density of the Gaussian random variable with variance t and mean 0. Here

- the parameter θ is unknown,
- we estimate it through an observation of a trayectory on $[0,T]$, at times iT/n ($i = 0,\dots,n$),
- Our asymptotics is in the time discretization, n (T is fixed).

Maximum likelihood

Denote $X_i := X_{iT/n}$ $(i = 0, \ldots, n)$, our sample. With the previous formula for the density, we construct the likelihood:

$$Z_n(\theta) = \prod_{i=0}^{n-1} \frac{q_\theta(\Delta, X_i, X_{i+1})}{q_0(\Delta, X_i, X_{i+1})}, \qquad \text{with } \Delta = \frac{T}{n}.$$

It is also of interest the log-likelihood:

$$L_n(\theta) = \log \prod_{i=0}^{n-1} q_\theta(\Delta, X_i, X_{i+1})$$

Explicit form of the likelihood

Inserting the formula for the density, we obtain

$$Z_n(\theta) = \prod_{\substack{X_i>0 \\ X_{i+1}<0}} (1-\theta) \prod_{\substack{X_i<0 \\ X_{i+1}>0}} (1+\theta)$$

$$\times \prod_{\substack{X_i<0 \\ X_{i+1}<0}} \left(1-\theta e^{\frac{-2X_iX_{i+1}}{\Delta}}\right) \prod_{\substack{X_i>0 \\ X_{i+1}>0}} \left(1+\theta e^{\frac{-2X_iX_{i+1}}{\Delta}}\right)$$

$$= \prod_{i=0}^{n-1} \left(1 + h(\sqrt{n}X_i, \sqrt{n}\,(X_{i+1}-X_i)\right),$$

where

$$h(x,y) = \mathrm{sgn}(x+y)\exp\left(-(2/\Delta)(x(x+y))^+\right).$$

Main result

Definition (Stable convergence)

Consider $(\Omega, \mathcal{F}, \mathbf{P})$, and a σ-algebra $\mathcal{G} \subset \mathcal{F}$. The sequence of random variables $Y_1, Y_2 \ldots$ converge $\mathcal{G}$-stably in distribution to Y, denoted

$$Y_n \xrightarrow[n\to\infty]{\mathcal{G}\text{-stably in dist.}} Y$$

when

$$\mathbb{E}\left(Zf(Y_n)\right) \xrightarrow[n\to\infty]{} \mathbb{E}\left(Zf(Y)\right)$$

for any bounded $\mathcal{G}$ measurable random variable Z, and any bounded and continuous function f.

Definition (Conditional Stable convergence)

Furthermore, consider sets $A, A_1, A_2, \ldots$. We say that $Y_1, Y_2 \ldots$ conditional on $A_1, A_2, \ldots$, converge $\mathcal{G}$-stably in distribution to Y conditional on A, denoted

$$Y_n \mid A_n \xrightarrow[n\to\infty]{\mathcal{G}\text{-stably}} Y \mid A,$$

when

$$\mathbb{E}\left(Zf(Y_n) \mid A_n\right) \xrightarrow[n\to\infty]{} \mathbb{E}\left(Zf(Y) \mid A\right)$$

for any bounded $\mathcal{G}$ measurable random variable Z, and any bounded and continuous function f.

Main result

Theorem

Under the null hypothesis ($\theta = 0$), the MLE θ_n satisfies

$$n^{1/4}\theta_n \mid A_n \xrightarrow[n\to\infty]{\mathcal{F}\text{-stably}} \frac{W(\ell_T^x)}{\ell_T^x} \mid A,$$

where $W = \{W_t : t \geq 0\}$ is a standard Brownian motion independent of B, and the sets are:

$$A_n = \{\omega : \inf_{1\leq i\leq n} X_i < 0\}, \qquad A = \{\omega : \inf_{0\leq t\leq T} X_t(\omega) < 0\}.$$

In particular, when $x = 0$, we have

$$n^{1/4}\theta_n \xrightarrow[n\to\infty]{\mathcal{F}\text{-stably}} \frac{W(\ell_T^x)}{\ell_T^x}.$$

Some comments on the result

- On the set A^c the limit is not defined, as $\ell_T^x = 0$. As the process does not hit the level $x = 0$, no statistical inference can be carried out.
- The limit is a mixture of normals, situation refered as *Local mixed asyptotic normality* (LAMN)in the terminology of statstical experiments.
- The speed of convergence is $n^{1/4}$, more slowly than the usual $n^{1/2}$, but typical in results involving the local time.

Computing the MLE by linearization

As the computation of the MLE is time consuming, consider the first order approximation of $L_n^{(1)}(\theta)$ at $\theta = 0$, that is

$$L_n^{(1)}\left(\frac{\theta}{n^{1/4}}\right) = L_n^{(1)}(0) + L_n^{(2)}(0)\frac{\theta}{n^{1/4}} + O\left(\frac{\theta^2}{n^{1/2}}\right),$$

that suggests that

$$\alpha_n = -n^{1/4}\frac{L_n^{(1)}(0)}{L_n^{(2)}(0)}$$

is a good approximation of $n^{1/4}\theta_n$ the MLE.

The proof in two steps

Step I: Prove that

$$n^{1/4}\theta_n - \alpha_n \to 0 \qquad \text{in probability.}$$

Step II: Prove that

$$\left(\frac{L_n^{(2)}(0)}{n^{1/2}}, \frac{L_n^{(1)}(0)}{n^{1/4}}, \mathbf{1}_{\mathbf{A_n}} \right) \xrightarrow[n\to\infty]{\mathcal{F}\text{-stably}} \left(c\ell_T^x, cW(\ell_T^x), \mathbf{1}_{\mathbf{A}}\right).$$

Here:

$$A_n = \{\omega \colon \inf_{1\le i\le n} X_i < 0\}, \qquad A = \{\omega \colon \inf_{0\le t\le T} X_t(\omega) < 0\}.$$

and from step II we obtain

$$\alpha_n \mid A_n = -n^{1/4}\frac{L_n^{(1)}(0)}{L_n^{(2)}(0)} \mid A_n \xrightarrow[n\to\infty]{\mathcal{F}\text{-stably}} \frac{W(\ell_T^x)}{\ell_T^x} \mid A.$$

Convergence of crossings of BM

In order to prove step II, for $k = 1, 2$, we compute

$$L_n^{(k)}(0) = (-1)^{k-1} \sum_{i=0}^{n-1} \left(h(\sqrt{n}X_i, \sqrt{n}(X_{i+1} - X_i)) \right)^2 .$$

The limits of this type of sums for $k = 2$ where obtained (for certain class of functions h) by Azaïs[3], and for $k = 1$ by Jacod[4]

[3] Azaïs, JM., Approximation des trajectoires et temps local des diffusions, *AIHP, PS*, **25**, 2, (1989), 175–194.

[4] Jacod, J., Rates of convergence to the local time of a diffusion, *AIHP, PS*, **34**, 4, (1998), 505–544.

Statistical application: testing hypothesis

The limit distribution

Lemma

For $x = 0$ and $T = 1$, the limit distribution of $W(\ell_1^0)/\ell_1^0 =: \Upsilon$ is symmetric, with density

$$f_\Upsilon(x) = \int_0^{+\infty} \mathrm{d}y \int_0^1 \frac{\sqrt{y}}{2\pi\sqrt{t^3}} \exp\left(\frac{-xy}{2} - \frac{y^2}{2t}\right) \mathrm{d}t\,,$$

equal in distribution to

$$\Upsilon = \frac{G(H)}{H} \text{ with } H = \frac{1}{2}(U + \sqrt{V + U^2}),$$

where $G(H)$, U and V are independent random variables, $G(H)$, $G(H) \sim \mathcal{N}(0, H)$, $U \sim \mathcal{N}(0, t)$ and $V \sim \exp(1/2)$.

Density of Υ

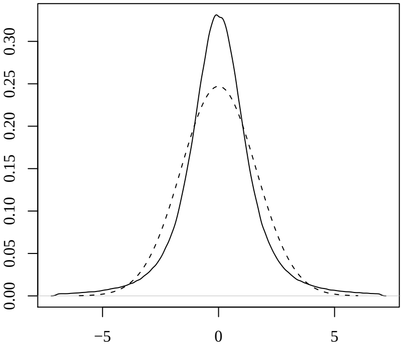

Figure: Density of Υ (solid) and density of the normal distribution with variance $\mathrm{Var}(\Upsilon)$ (dashed).

Testing hypothesis

It is then possible to develop a hypothesis test of $\theta = 0$ against $\theta \neq 0$. For this, let us compute

$$\mathbf{P}\left[|\theta_n| \geq \frac{K}{n^{1/4}}\right] \approx \mathbf{P}\left[|\Upsilon| \geq \frac{K}{cn^{1/4}}\right].$$

Of course, the second type error cannot be computed, and we do not know the asymptotic behavior of $L_n(\theta)$ when $\theta \neq 0$. However, it is rather easy to perform simulation and thus to get some numerical informations about $\Lambda_n(\theta)$ and α_n. For example, we see in Figure 2 an approximation of the density of $\alpha_n/n^{1/4}$ for $\theta = 0.5$ compared to an approximation of the density of $\alpha_n/n^{1/4}$ for the Brownian motion with $n = 1000$. We can note that the histogram of $\alpha_n/n^{1/4}$ has its peak on 0.5.

Some numerical considerations: θ_n against $\alpha_n/n^{1/4}$

n	$n^{-1/2}$	mean	std dev	quant. 90 %
100	0.100	0.026	0.057	0.082
200	0.070	0.028	0.083	0.057
500	0.044	0.013	0.055	0.026
1000	0.031	0.013	0.040	0.033
2000	0.022	0.006	0.025	0.015
5000	0.014	0.006	0.041	0.006
10000	0.010	0.002	0.005	0.003

Table: Statistics of $|\theta_n - \alpha_n/n^{1/4}|$ over 100 paths.

This table suggests that the error of $|\theta_n - \alpha_n/n^{1/4}|$ is of order $1/n^{1/2}$, so that $\alpha_n/n^{1/4}$ is a pretty good approximation of θ_n, and is much more faster to compute.

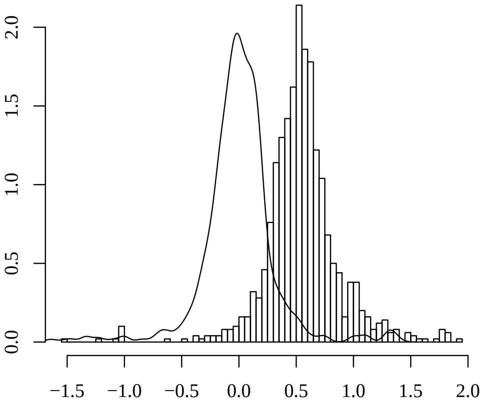

Figure: Histogram of $\alpha_n/n^{1/4}$ ($n = 1000$) with $\theta = 0.5$ against the an approximation of the density of $\alpha_n/n^{1/4}$ for $\theta = 0$.

One open question

- The convergence for $\theta \neq 0$ is open. This a Theorem on crossings under the alternative hypothesis, then for the SBm instead of the BM. Some natural questions arise:
  - Whether the speed the same ($n^{1/4}$),
  - Whether it is still possible to get the explicit limit distribution,
  - Can we merge this results in Le Cam / Ibraguimov - Hasminsikii Theory of convergence of statistical experiments.

Agradecimientos

- A Mario Wschebor, por
  - sus contribuciones científicas nacionales e internacionales
  - sus contribuciones a la organización de los científicos en Uruguay, Latinoamérica, y en el mundo
  - sus enseñanzas y su solidaridad
  - su ejemplo permanente como ciudadano del mundo
  - ...
- A Jean Marc Azaïs
- A los organizadores del IX CLAPEM, en especial, J.R. Chichi León y Stella Brassesco