On Drift Parameter Estimation in Models with Fractional Brownian Motion by Discrete Observations

Yuliya Mishura
Taras Shevchenko Nat. Univ. of Kyiv

Kostiantyn Ralchenko
Taras Shevchenko Nat. Univ. of Kyiv

Abstract

We study a problem of an unknown drift parameter estimation in a stochastic differential equation driven by fractional Brownian motion. We represent the likelihood ratio as a function of the observable process. The form of this representation is in general rather complicated. However, in the simplest case it can be simplified and we can discretize it to establish the a. s. convergence of the discretized version of maximum likelihood estimator to the true value of parameter. We also investigate a non-standard estimator of the drift parameter showing further its strong consistency.

Keywords: fractional Brownian motion, parameter estimation, stochastic differential equation, strong consistency, discretization.

1. Introduction

The models with long-range dependence are very popular now because they correspond to various processes in economy, finances and tele-traffic. From the mathematical point of view, long-range dependence can be modeled with the help of fractional Brownian motion with Hurst parameter $H \in \left(\frac{1}{2}, 1\right)$. More promising are so called mixed models involving both the standard Wiener process and the fractional Brownian motion. Similarly to the standard semimartingale models, the problem of parameter estimation arises immediately when we want to adapt the model with long-range dependence to the specific situation. In particular, the problem of the drift parameter estimation in the diffusion model with fractional Brownian motion is rather important. The standard maximum likelihood estimator was considered by many authors, see, e.g., Mishura (2008) and Prakasa Rao (2010). It is constructed by continuous observations on the whole interval. Asymptotic properties when the interval of observations increases to the whole half-axis, were established. However, in practical considerations the observations are never continuous. So, the problem of the discretization of the estimate appears. Some papers are devoted to the parameter estimation for the models with fBm and discrete observations, see, e.g., Hu and Nualart (2010), Xiao, Zhang, and Xu (2011a), Xiao, Zhang, and Zhang (2011b), Bishwal (2011), Tanaka (2013), Hu and Song (2013), Zhang, Xiao, Zhang, and Niu (2014) but only restricted classes of models, basically linear models were considered. The situation is such that in the general case the maximum-likelihood estimator has a very

complicated representation via the observed process and the dicretized version does not allow reasonable form for calculations. Therefore, we have to propose some non-standard approach to construct strongly consistent drift parameter estimators for the discrete observations of the models with long-range dependence. One of such approaches was demonstrated in Mishura, Ral'chenko, Seleznev, and Shevchenko (2014), where some specific discretized estimators were proposed. In the present paper we propose two approaches. One of them consists in direct discretization of maximum-likelihood parameter estimator, however, only for the case when drift and diffusion coefficients coincide. It is one of the cases when the discretization leads to the reasonable form of the estimator. Another approach is to discretize the non-standard drift parameter estimator that was introduced in Kozachenko, Melnikov, and Mishura (2013). This also leads to the consistent estimator. Strong consistency is established for both estimators and illustrated with some simulations.

2. Maximum-likelihood estimation

2.1. Model description

Let $B^H = \{B_t^H, t \geq 0\}$ be a fractional Brownian motion with Hurst index $H \in (1/2, 1)$, defined on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Denote by $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by B^H. Consider the stochastic differential equation driven by fractional Brownian motion B^H:

$$dX_t = \theta a(t, X_t)dt + b(t, X_t)dB_t^H, \quad 0 \leq t \leq T, \quad T > 0,$$
$$X\big|_{t=0} = X_0 \in \mathbb{R}. \tag{1}$$

Here $\theta \in \mathbb{R}$ is unknown parameter to be estimated.

Suppose that the following assumptions hold:

(I) there exist positive constants C_1, C_2 such that for all $t \in [0, T]$, $x, y \in \mathbb{R}$

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq C_1 |x - y|,$$
$$|a(t, x)| + |b(t, x)| \leq C_2(1 + |x|);$$

(II) there exist constants $C_3 > 0$ and $\rho \in \left(\frac{1}{H} - 1, 1\right)$ such that for all $t \in [0, T]$, $x, y \in \mathbb{R}$

$$\left|b'_x(t, x) - b'_y(t, y)\right| \leq C_3 |x - y|^\rho;$$

(III) there exist constants $C_4 > 0$ and $\mu \in (1 - H, 1)$ such that for all $t, s \in [0, T]$, $x \in \mathbb{R}$

$$|b(t, x) - b(s, x)| + \left|b'_x(t, x) - b'_x(s, x)\right| \leq C_4 |t - s|^\mu.$$

According to (Nualart and Rascanu 2001, Theorem 2.1), under the conditions (I)–(III) there exists a unique solution X of the stochastic equation (1).

In addition, suppose that the following conditions hold:

(IV) $b(t, x) \neq 0$;

(V) $a, b \in C([0, \infty) \times \mathbb{R})$.

Denote $\alpha = H - \frac{1}{2}$, $\widetilde{\alpha} = (1 - 2\alpha)^{-1}$, $C_H = \left(\frac{\Gamma(2-2\alpha)}{2H\Gamma(1-\alpha)^3\Gamma(\alpha+1)}\right)^{\frac{1}{2}}$, $l_H(t, s) = C_H s^{-\alpha}(t - s)^{-\alpha}I_{\{0 < s < t\}}$, $\psi(t, x) = \frac{a(t,x)}{b(t,x)}$, $\varphi(t) = \psi(t, X_t)$, $I(t) = \int_0^t l_H(t, s)\varphi(s)ds$. Under the conditions (I), (III)–(V) $\varphi(t), t \in [0, T]$ is a continuous process with probability 1. Hence, it is Lebesgue integrable and for each $t \in [0, T]$ there exists an integral $\int_0^t l_H(t, s)\varphi(s)ds$.

Consider the new process $\widehat{B}_t^H := B_t^H + \theta \int_0^t \varphi(s)ds$. Suppose that the following assumptions hold.

(VI) There exists such function δ that belongs to $L_1[0, t]$ for all $t \in [0, T]$ a. s. and satisfies the equation

$$\theta \int_0^t l_H(t, s)\varphi(s)ds = (\widetilde{\alpha})^{-1/2} \int_0^t \delta_s ds;$$

(VII) $\mathsf{E} \int_0^t s^{2\alpha}\delta_s^2 ds < \infty$, $t \in [0, T]$;

(VIII) $\mathsf{E} \exp\left\{L_t - \frac{1}{2}\langle L\rangle_t\right\} = 1$, where $L_t = \int_0^t s^\alpha \delta_s d\widehat{B}_s$, and $\widehat{B}$ is Wiener process with respect to probability measure $\mathsf{P}_0(t)$ corresponding to the zero drift such that

$$\int_0^t l_H(t, s) \, d\widehat{B}_s^H = \widetilde{\alpha}^{-1/2} \int_0^t s^{-\alpha} \, d\widehat{B}_s.$$

(The existence of this Wiener process follows from the representation of fractional Brownian motion via Wiener process on a finite interval introduced in Norros, Valkeila, and Virtamo (1999).)

Then the likelihood ratio $\frac{d\mathsf{P}_\theta(t)}{d\mathsf{P}_0(t)}$ for the probability measure $\mathsf{P}_\theta(t)$ corresponding to our model and probability measure $\mathsf{P}_0(t)$ corresponding to the model with zero drift is equal to

$$\frac{d\mathsf{P}_\theta(t)}{d\mathsf{P}_0(t)} = \exp\left\{L_t - \frac{1}{2}\langle L\rangle_t\right\}.$$

Note that L_t is a square-integrable martingale. Now we present likelihood ratio as a function of the observed process X_t.

2.2. The explicit form for the likelihood ratio and a discretized version of MLE

We can present likelihood ratio as a function of the observed process X_t.

$$L_t = \int_0^t s^\alpha \delta_s d\widehat{B}_s = \int_0^t s^{2\alpha}\delta_s dY_s, \tag{2}$$

where

$$Y_s = \int_0^s u^{-\alpha} d\widehat{B}_u = \widetilde{\alpha}^{1/2} \int_0^s l_H(s, u)b^{-1}(u, X_u)dX_u, \tag{3}$$

$$\delta_s = \theta\widetilde{\alpha}^{1/2}\left(\int_0^s l_H(s, u)\varphi(u)du\right)' = C_H\theta\widetilde{\alpha}^{1/2}\left(\int_0^s (s-u)^{-\alpha}u^{-\alpha}\varphi(u)du\right)' \tag{4}$$

or

$$\delta_s = C_H\theta\widetilde{\alpha}^{1/2}\left(\frac{\varphi(s)}{s^{2\alpha}} + \alpha\int_0^s \frac{s^{-\alpha}\varphi(s) - u^{-\alpha}\varphi(u)}{(s-u)^{\alpha+1}}du\right). \tag{5}$$

According to (Mishura 2008, formula (6.3.13)), the maximum-likelihood estimator has the form

$$\hat{\theta}_t^{(1)} = \frac{\widetilde{\alpha}^{-1/2}\int_0^t s^\alpha I'(s)d\widehat{B}_s}{\int_0^t s^{2\alpha}(I'(s))^2 ds}.$$

Since $I'(s) = \delta_s\theta^{-1}\widetilde{\alpha}^{-1/2}$, we obtain

$$\hat{\theta}_t^{(1)} = \frac{\theta L_t}{\int_0^t s^{2\alpha}\delta_s^2 ds}. \tag{6}$$

Using (2), (3), (5) and the definition of the kernel $l_H(t,s)$ we can write

$$\hat{\theta}_t^{(1)} = \frac{\theta \int_0^t s^{2\alpha} \delta_s dY_s}{\int_0^t s^{2\alpha} \delta_s^2 ds}$$

$$= \frac{\int_0^t \left(\varphi(s) + \alpha s^{2\alpha} \int_0^s \frac{s^{-\alpha}\varphi(s) - u^{-\alpha}\varphi(u)}{(s-u)^{\alpha+1}} du \right) d\left(\int_0^s v^{-\alpha}(s-v)^{-\alpha} b^{-1}(v, X_v) dX_v \right)}{\int_0^t s^{2\alpha} \left(\frac{\varphi(s)}{s^{2\alpha}} + \alpha \int_0^s \frac{s^{-\alpha}\varphi(s) - u^{-\alpha}\varphi(u)}{(s-u)^{\alpha+1}} du \right)^2 ds}.$$

Remark 1. According to (Mishura 2008, Theorem 6.3.3), under assumptions (I)–(VIII) and

$$\int_0^\infty s^{2\alpha}(I'(s))^2 ds = \infty \quad \text{a.s.}$$

$\hat{\theta}_T^{(1)} \xrightarrow{\text{P1}} \theta$, $T \to \infty$.

Let $t_k^n = \frac{k}{2^n}$, $k = 0, 1, 2, \ldots, 2^{2n}$. We can define a discretized version of the maximum-likelihood estimator

$$\hat{\theta}_n^{(2)} := \frac{\sum_{k=0}^{2^{2n}-1} \left(\varphi(t_k^n) + \alpha(t_k^n)^{2\alpha} \sum_{i=1}^{k-1} \frac{(t_k^n)^{-\alpha}\varphi(t_k^n) - (t_i^n)^{-\alpha}\varphi(t_i^n)}{(t_k^n - t_i^n)^{\alpha+1}} \frac{1}{2^n} \right) \left(\widetilde{Y}_{t_{k+1}^n} - \widetilde{Y}_{t_k^n} \right)}{\sum_{k=0}^{2^{2n}-1} (t_k^n)^{2\alpha} \left(\frac{\varphi(t_k^n)}{(t_k^n)^{2\alpha}} + \alpha \sum_{i=1}^{k-1} \frac{(t_k^n)^{-\alpha}\varphi(t_k^n) - t_i^{-\alpha}\varphi(t_i^n)}{(t_k^n - t_i^n)^{\alpha+1}} \frac{1}{2^n} \right)^2 \frac{1}{2^n}} \quad (7)$$

where

$$\widetilde{Y}_{t_k} = \sum_{i=1}^{k-1} (t_i^n)^{-\alpha}(t_k^n - t_i^n)^{-\alpha} b^{-1}(t_i^n, X_{t_i^n}) \left(X_{t_{i+1}^n} - X_{t_i^n} \right).$$

In the general case formula (7) is not suitable for applications because it involves a lot of weakly singular kernels and it is quite impossible to get its convergence to the true value of the parameter. But even if we get the convergence, the simulation error will be so great that annihilate our efforts in discretization. In order to avoid this technical difficulties, we start with the simplest case.

2.3. Estimation in the case $a = b$

Consider an equation

$$dX_t = \theta b(X_t)dt + b(X_t)dB_t^H. \quad (8)$$

In this case $\varphi \equiv 1$. So we get from (4) that

$$\delta_s = C_H \theta \widetilde{\alpha}^{1/2} \left(B(1-\alpha, 1-\alpha)s^{1-2\alpha} \right)' = C_H \theta \widetilde{\alpha}^{-1/2} B(1-\alpha, 1-\alpha)s^{-2\alpha}.$$

Then (2) and (3) imply

$$L_t = C_H \theta \widetilde{\alpha}^{-1/2} B(1-\alpha, 1-\alpha)Y_t = C_H \theta B(1-\alpha, 1-\alpha) \int_0^t l_H(t,s) b^{-1}(X_s) dX_s.$$

Therefore the maximum-likelihood estimator (6) can be written as follows:

$$\hat{\theta}_t^{(1)} = \frac{\int_0^t l_H(t,s) b^{-1}(X_s) dX_s}{C_H B(1-\alpha, 1-\alpha) t^{1-2\alpha}}. \quad (9)$$

It follows from (8) that

$$\hat{\theta}_t^{(1)} = \frac{\theta \int_0^t l_H(t,s) ds + \int_0^t l_H(t,s) dB_s^H}{C_H B(1-\alpha, 1-\alpha) t^{1-2\alpha}} = \theta + \frac{\int_0^t l_H(t,s) dB_s^H}{C_H B(1-\alpha, 1-\alpha) t^{1-2\alpha}}.$$

Since $\int_0^t l_H(t,s) dB_s^H$ is a square integrable martingale with angle bracket $t^{1-2\alpha} \to \infty$ we see that $\frac{\int_0^t l_H(t,s) dB_s^H}{t^{1-2\alpha}} \xrightarrow[n\to\infty]{\text{P1}} 0$. Hence the estimator $\hat{\theta}_t^{(1)}$ is strongly consistent.

Now we consider an estimator

$$\hat{\theta}_n^{(3)} = \frac{\sum_{k=1}^{2^{2n}-1}(t_k^n)^{-\alpha}(2^n - t_k^n)^{-\alpha}b^{-1}\left(X_{t_{k-1}^n}\right)\left(X_{t_k^n} - X_{t_{k-1}^n}\right)}{\mathrm{B}(1-\alpha,1-\alpha)2^{n(1-2\alpha)}},$$

where $t_k^n = \frac{k}{2^n}$, $k = 0, 1, \ldots, 2^{2n}$. This estimator is a discretized version of the estimator (9).

Theorem 1. *Suppose that there exist positive constants C_1, C_2, C_3, C_5 and $\rho \in (1/H - 1, 1]$, such that*

(a) $|b(x) - b(y)| \leq C_1 |x - y|$ *for all $x, y \in \mathbb{R}$,*

(b) $C_5 \leq |b(x)| \leq C_2(1 + |x|)$ *for all $x \in \mathbb{R}$,*

(c) $|b'(x) - b'(y)| \leq C_3 |x - y|^\rho$ *for all $x, y \in \mathbb{R}$*

Then $\hat{\theta}_n^{(3)} \xrightarrow{\mathrm{P1}} \theta, n \to \infty$. Moreover, for any $\beta \in (1/2, H)$ and $\gamma > 1/2$ there exists a random variable $\eta = \eta_{\beta,\gamma}$ with all finite moments such that $\left|\hat{\theta}_n^{(3)} - \theta\right| \leq \eta n^{\kappa+\gamma}2^{-\tau n}$, where $\kappa = \gamma/\beta$, $\tau = (1 - H) \wedge (2\beta - 1)$.

Proof. It follows from (8) that

$$
\begin{aligned}
X_{t_k^n} - X_{t_{k-1}^n} &= \theta \int_{t_{k-1}^n}^{t_k^n} b(X_v)dv + \int_{t_{k-1}^n}^{t_k^n} b(X_v)dB_v^H \\
&= \theta \int_{t_{k-1}^n}^{t_k^n} b\left(X_{t_{k-1}^n}\right) dv + \theta \int_{t_{k-1}^n}^{t_k^n} \left(b(X_v) - b\left(X_{t_{k-1}^n}\right)\right) dv \\
&\quad + \int_{t_{k-1}^n}^{t_k^n} \left(b(X_v) - b\left(X_{t_{k-1}^n}\right)\right) dB_v^H + \int_{t_{k-1}^n}^{t_k^n} b\left(X_{t_{k-1}^n}\right) dB_v^H.
\end{aligned}
\tag{10}
$$

Then

$$\hat{\theta}_n^{(3)} = \frac{\theta A_n + \theta B_n + D_n + E_n}{\mathrm{B}(1-\alpha,1-\alpha)}, \tag{11}$$

where

$$A_n = 2^{n(2\alpha-2)} \sum_{k=1}^{2^{2n}-1} (t_k^n)^{-\alpha}(2^n - t_k^n)^{-\alpha},$$

$$B_n = 2^{n(2\alpha-1)} \sum_{k=1}^{2^{2n}-1} (t_k^n)^{-\alpha}(2^n - t_k^n)^{-\alpha}b^{-1}\left(X_{t_{k-1}^n}\right) \int_{t_{k-1}^n}^{t_k^n} \left(b(X_v) - b\left(X_{t_{k-1}^n}\right)\right) dv,$$

$$D_n = 2^{n(2\alpha-1)} \sum_{k=1}^{2^{2n}-1} (t_k^n)^{-\alpha}(2^n - t_k^n)^{-\alpha}b^{-1}\left(X_{t_{k-1}^n}\right) \int_{t_{k-1}^n}^{t_k^n} \left(b(X_v) - b\left(X_{t_{k-1}^n}\right)\right) dB_v^H,$$

$$E_n = 2^{n(2\alpha-1)} \sum_{k=1}^{2^{2n}-1} (t_k^n)^{-\alpha}(2^n - t_k^n)^{-\alpha}\left(B_{t_k^n}^H - B_{t_{k-1}^n}^H\right).$$

It is not hard to show that the sequence

$$A_n = \sum_{k=1}^{2^{2n}-1} \left(\frac{k}{2^{2n}}\right)^{-\alpha} \left(1 - \frac{k}{2^{2n}}\right)^{-\alpha} \frac{1}{2^{2n}}$$

converges to $\int_0^1 x^{-\alpha}(1-x)^{-\alpha}dx = \mathrm{B}(1-\alpha,1-\alpha)$, moreover,

$$|A_n - \mathrm{B}(1-\alpha,1-\alpha)| \leq c_1 2^{-2n(1-\alpha)} \tag{12}$$

where c_1 is a constant. Indeed, $h(x) = x^{-\alpha}(1-x)^{-\alpha}$ is a decreasing function when $x \in \left(0, \frac{1}{2}\right]$, then

$$\int_0^{\frac{1}{2}} h(x)dx = \sum_{k=0}^{2^{2n-1}-1} \int_{\frac{k}{2^{2n}}}^{\frac{k+1}{2^{2n}}} h(x)dx < \int_0^{\frac{1}{2^{2n}}} h(x)dx + \sum_{k=1}^{2^{2n-1}} h\left(\frac{k}{2^{2n}}\right) \frac{1}{2^{2n}}.$$

On the other hand,

$$\int_0^{\frac{1}{2}} h(x)dx = \sum_{k=1}^{2^{2n-1}} \int_{\frac{k-1}{2^{2n}}}^{\frac{k}{2^{2n}}} h(x)dx > \sum_{k=1}^{2^{2n-1}} h\left(\frac{k}{2^{2n}}\right) \frac{1}{2^{2n}}.$$

So

$$0 < \int_0^{\frac{1}{2}} h(x)dx - \sum_{k=1}^{2^{2n-1}} h\left(\frac{k}{2^{2n}}\right) \frac{1}{2^{2n}} < \int_0^{\frac{1}{2^{2n}}} h(x)dx \leq \left(1 - \frac{1}{2^{2n}}\right)^{-\alpha} \frac{2^{-2n(1-\alpha)}}{1-\alpha}. \qquad (13)$$

Similarly one can show that

$$0 < \int_{\frac{1}{2}}^1 h(x)dx - \sum_{k=2^{2n-1}+1}^{2^{2n}-1} h\left(\frac{k}{2^{2n}}\right) \frac{1}{2^{2n}} < \left(1 - \frac{1}{2^{2n}}\right)^{-\alpha} \frac{2^{-2n(1-\alpha)}}{1-\alpha}. \qquad (14)$$

Combining (13) and (14), we obtain (12).

By (Mishura *et al.* 2014, Lemma 2), there exist random variables ξ_1 and ξ_2 with all finite moments such that for all $n \geq 1$ and $k = 1, 2, \ldots, 2^{2n}$

$$\left| \int_{t_{k-1}^n}^{t_k^n} \left(b(X_u) - b(X_{t_{k-1}^n}) \right) du \right| \leq \xi_1 n^\kappa 2^{-n(\beta+1)}$$

and

$$\left| \int_{t_{k-1}^n}^{t_k^n} \left(b(X_u) - b(X_{t_{k-1}^n}) \right) dB_u^H \right| \leq \xi_2 n^{\gamma+\kappa} 2^{-2n\beta},$$

Then

$$|B_n| \leq C_5^{-1} \xi_1 n^\kappa 2^{n(2\alpha-\beta-2)} \sum_{k=1}^{2^{2n}-1} (t_k^n)^{-\alpha}(2^n - t_k^n)^{-\alpha} = C_5^{-1} \xi_1 n^\kappa 2^{-n\beta} A_n \leq c_2 \xi_1 n^\kappa 2^{-n\beta}; \qquad (15)$$

$$|D_n| \leq C_5^{-1} \xi_2 n^{\gamma+\kappa} 2^{n(2\alpha-1-2\beta)} \sum_{k=1}^{2^{2n}-1} (t_k^n)^{-\alpha}(2^n - t_k^n)^{-\alpha} \leq c_2 \xi_2 n^{\gamma+\kappa} 2^{-n(2\beta-1)}. \qquad (16)$$

Finally we estimate E_n. Start by writing

$$\mathsf{E}\left[E_n^2 \right] = 2^{2n(2\alpha-1)} \mathsf{E}\left[\left(\sum_{k=1}^{2^{2n}-1} \int_{t_{k-1}^n}^{t_k^n} (t_k^n)^{-\alpha} (2^n - t_k^n)^{-\alpha} dB_s^H \right)^2 \right].$$

According to (Mishura 2008, Corollary 1.9.4), for $f \in L_{1/H}[0,t]$ there exist a constant $C_H > 0$ such that

$$\mathsf{E}\left[\left(\int_0^t f(s)dB_s^H \right)^2 \right] \leq C_H \left(\int_0^t |f(s)|^{1/H} ds \right)^{2H}.$$

Hence,

$$\mathsf{E}\left[E_n^2 \right] \leq c_3 2^{2n(2\alpha-1)} \left(\sum_{k=1}^{2^{2n}-1} \int_{t_{k-1}^n}^{t_k^n} (t_k^n)^{-\alpha/H} (2^n - t_k^n)^{-\alpha/H} ds \right)^{2H}$$

$$= c_3 2^{2n(H-1)} \left(\sum_{k=1}^{2^{2n}-1} \left(\frac{k}{2^{2n}} \right)^{-\alpha/H} \left(1 - \frac{k}{2^{2n}} \right)^{-\alpha/H} \frac{1}{2^{2n}} \right)^{2H}.$$

As above,

$$\sum_{k=1}^{2^{2n}-1} \left(\frac{k}{2^{2n}}\right)^{-\alpha/H} \left(1-\frac{k}{2^{2n}}\right)^{-\alpha/H} \frac{1}{2^{2n}} \to \mathrm{B}\left(1-\alpha/H, 1-\alpha/H\right), \ n\to\infty,$$

which implies that $\mathsf{E}\left[E_n^2\right] \leq c_4 2^{2n(H-1)}$. Since E_n is Gaussian, we have $\mathsf{E}\left[|E_n|^p\right] \leq c_5(p)2^{pn(H-1)}$ for any $p \geq 1$. Therefore, for any $\nu > 1$

$$\mathsf{E}\left[\sum_{n=1}^{\infty}\frac{|E_n|^p}{n^{\nu}2^{pn(H-1)}}\right] = \sum_{n=1}^{\infty}\frac{\mathsf{E}\left[|E_n|^p\right]}{n^{\nu}2^{pn(H-1)}} \leq c_5(p)\sum_{n=1}^{\infty}n^{-\nu} < \infty.$$

Consequently,

$$\xi_3 := \sup_{n\geq 1}\frac{|E_n|}{n^{\nu/p}2^{n(H-1)}} < \infty$$

almost surely, moreover, by Fernique's theorem, all moments of ξ_3 are finite. Therefore,

$$|E_n| \leq \xi_3 n^{\delta}2^{-n(1-H)}, \tag{17}$$

where $\delta > 0$ can be taken arbitrarily small.

Combining (11), (12) and (15)–(17) we obtain

$$\left|\hat{\theta}_n^{(3)} - \theta\right| \leq \frac{\theta\left|A_n - \mathrm{B}(1-\alpha, 1-\alpha)\right| + \theta\left|B_n\right| + |D_n| + |E_n|}{\mathrm{B}(1-\alpha, 1-\alpha)}$$

$$\leq \frac{\theta c_1 2^{-2n(1-\alpha)} + \theta c_2 \xi_1 n^{\kappa}2^{-n\beta} + c_2\xi_2 n^{\gamma+\kappa}2^{-n(2\beta-1)} + \xi_3 n^{\delta}2^{-n(1-H)}}{\mathrm{B}(1-\alpha, 1-\alpha)}.$$

Note that $2(1-\alpha) = 3 - 2H > 1 - H \geq \tau$ and $\beta > 1/2 > 1 - H \geq \tau$. Then,

$$\left|\hat{\theta}_n^{(3)} - \theta\right| \leq \eta n^{\kappa+\gamma}2^{-\tau n},$$

where $\eta \leq c_6(\theta)(1 + \xi_1 + \xi_2 + \xi_3)$. □

3. Non-standard estimators

In the paper Kozachenko et al. (2013) the following non-standard estimator for θ in the equation (1) was considered:

$$\hat{\theta}_t^{(4)} = \frac{\int_0^t a(s, X_s)b^{-2}(s, X_s)dX_s}{\int_0^t a^2(s, X_s)b^{-2}(s, X_s)ds}.$$

According to (Kozachenko et al. 2013, Theorem 4), if the assumptions (I)–(IV), (VI)–(VII) hold and there exist such $\beta > 1 - H$ and $p > 1$ that

$$\frac{T^{H+\beta-1}(\log T)^p \int_0^T \left|(D_{0+}^{\beta}\varphi)(s)\right| ds}{\int_0^T \varphi_s^2 ds} \to 0 \ \text{ a.s. as } T\to\infty,$$

then the estimator $\hat{\theta}_T^{(4)}$ is well-defined and strongly consistent as $T \to \infty$.

We define a discretized version of $\hat{\theta}_T^{(4)}$ for the equation

$$X_t = X_0 + \theta\int_0^t a(X_s)ds + \int_0^t b(X_s)dB_s^H. \tag{18}$$

Put

$$\hat{\theta}_n^{(5)} := \frac{\sum_{k=1}^{2^{2n}} a\left(X_{t_{k-1}^n}\right) b^{-2}\left(X_{t_{k-1}^n}\right)\left(X_{t_k^n} - X_{t_{k-1}^n}\right)}{\sum_{k=1}^{2^{2n}} a^2\left(X_{t_{k-1}^n}\right) b^{-2}\left(X_{t_{k-1}^n}\right)\frac{1}{2^n}},$$

$$\widehat{\varphi}_n(t) := \sum_{k=0}^{2^{2n}-1} \varphi(t_k^n) I_{[t_k^n, t_{k+1}^n)}(t).$$

Theorem 2. *Suppose that there exist positive constants C_1, C_3, C_5, C_6 $\rho \in (1/H - 1, 1]$, $\beta > 1 - H$, $p > 1$ such that*

(a) $|a(x) - a(y)| + |b(x) - b(y)| \leq C_1 |x - y| \quad$ *for all $x, y \in \mathbb{R}$,*

(b) $C_5 \leq |a(x)| \leq C_6$, $C_5 \leq |b(x)| \leq C_6 \quad$ *for all $x \in \mathbb{R}$,*

(c) $|b'(x) - b'(y)| \leq C_3 |x - y|^\rho \quad$ *for all $x, y \in \mathbb{R}$,*

(d) $\dfrac{2^{n(H+\beta)} n^p \int_0^{2^n} \left|\left(D_{0+}^\beta \widehat{\varphi}_n\right)(s)\right| ds}{\sum_{k=1}^{2^{2n}} \varphi^2(t_{k-1}^n)} \to 0$ *a. s. as $n \to \infty$.*

Then with probability one, $\hat{\theta}_n^{(5)} \to \theta, n \to \infty$.

Proof. It follows from (18) that

$$\begin{aligned}
X_{t_k^n} - X_{t_{k-1}^n} &= \theta \int_{t_{k-1}^n}^{t_k^n} a(X_v) dv + \int_{t_{k-1}^n}^{t_k^n} b(X_v) dB_v^H \\
&= \theta \int_{t_{k-1}^n}^{t_k^n} a\left(X_{t_{k-1}^n}\right) dv + \theta \int_{t_{k-1}^n}^{t_k^n} \left(a(X_v) - a\left(X_{t_{k-1}^n}\right)\right) dv \qquad (19) \\
&\quad + \int_{t_{k-1}^n}^{t_k^n} \left(b(X_v) - b\left(X_{t_{k-1}^n}\right)\right) dB_v^H + \int_{t_{k-1}^n}^{t_k^n} b\left(X_{t_{k-1}^n}\right) dB_v^H.
\end{aligned}$$

Then

$$\hat{\theta}_n^{(5)} = \theta + \frac{\theta B_n + E_n + D_n}{A_n},$$

where

$$A_n = 2^{-n} \sum_{k=1}^{2^{2n}} \varphi^2(t_{k-1}^n),$$

$$B_n = \sum_{k=1}^{2^{2n}} a\left(X_{t_{k-1}^n}\right) b^{-2}\left(X_{t_{k-1}^n}\right) \int_{t_{k-1}^n}^{t_k^n} \left(a(X_v) - a\left(X_{t_{k-1}^n}\right)\right) dv,$$

$$E_n = \sum_{k=1}^{2^{2n}} a\left(X_{t_{k-1}^n}\right) b^{-2}\left(X_{t_{k-1}^n}\right) \int_{t_{k-1}^n}^{t_k^n} \left(b(X_v) - b\left(X_{t_{k-1}^n}\right)\right) dB_v^H,$$

$$D_n = \sum_{k=1}^{2^{2n}} \varphi(t_{k-1}^n) \left(B_{t_k^n}^H - B_{t_{k-1}^n}^H\right).$$

D_n can be represented in the form

$$D_n = \int_0^{2^n} \widehat{\varphi}_n(s) dB_s^H.$$

Applying (Kozachenko *et al.* 2013, Theorem 3) we can estimate

$$\sup_{0 \leq t \leq 2^n} \left|\left(D_{2^{n-}}^{1-\beta} B_{2^{n-}}^H\right)(t)\right| \leq \xi(p) 2^{n(H+\beta-1)} n^p (\log 2)^p.$$

Therefore

$$|D_n| \leq \sup_{0 \leq t \leq 2^n} \left| \left(D_{2^n-}^{1-\beta} B_{2^n-}^H \right)(s) \right| \cdot \int_0^{2^n} \left| \left(D_{0+}^{\beta} \widehat{\varphi}_n \right)(s) \right| ds$$

$$\leq \xi(p)(\log 2)^p 2^{n(H+\beta-1)} n^p \int_0^{2^n} \left| \left(D_{0+}^{\beta} \widehat{\varphi}_n \right)(s) \right| ds.$$

Then

$$\left| \frac{D_n}{A_n} \right| \leq \xi(p)(\log 2)^p \frac{2^{n(H+\beta)} n^p \int_0^{2^n} \left| \left(D_{0+}^{\beta} \widehat{\varphi}_n \right)(s) \right| ds}{\sum_{k=1}^{2^{2n}} \varphi^2(t_{k-1}^n)} \to 0 \text{ a.s. as } n \to \infty.$$

Using the condition (b) we can write

$$\left| \frac{B_n}{A_n} \right| \leq C_6^{-1} 2^{-n} \sum_{k=1}^{2^{2n}} \int_{t_{k-1}^n}^{t_k^n} \left| a(X_v) - a\left(X_{t_{k-1}^n} \right) \right| dv,$$

$$\left| \frac{E_n}{A_n} \right| \leq C_6^{-1} 2^{-n} \sum_{k=1}^{2^{2n}} \int_{t_{k-1}^n}^{t_k^n} \left| b(X_v) - b\left(X_{t_{k-1}^n} \right) \right| dB_v^H.$$

It now follows from (Mishura *et al.* 2014, Lemma 2) that $\left| \frac{B_n}{A_n} \right| \to 0$, $\left| \frac{E_n}{A_n} \right| \to 0$ as $n \to \infty$. $\square$

Example 1. Consider the model (8):

$$dX_t = \theta b(X_t)dt + b(X_t)dB_t^H.$$

Suppose that there exist positive constants C_1, C_3, C_5, C_6 $\rho \in (1/H - 1, 1]$, $\beta > 1 - H$, $p > 1$ such that

(a) $|b(x) - b(y)| \leq C_1 |x - y|$ for all $x, y \in \mathbb{R}$,

(b) $C_5 \leq |b(x)| \leq C_6$ for all $x \in \mathbb{R}$,

(c) $|b'(x) - b'(y)| \leq C_3 |x - y|^\rho$ for all $x, y \in \mathbb{R}$.

In this case the non-standard estimator $\hat{\theta}_n^{(5)}$ has the form

$$\hat{\theta}_n^{(6)} = 2^{-n} \sum_{k=1}^{2^{2n}} b^{-1}\left(X_{t_{k-1}^n} \right) \left(X_{t_k^n} - X_{t_{k-1}^n} \right), \tag{20}$$

$\widehat{\varphi}_n(t) = 1$. Then $\left(D_{0+}^{\beta} \widehat{\varphi}_n \right)(s) = \frac{1}{\Gamma(1-\beta)} \cdot s^{-\beta}$ and

$$\frac{2^{n(H+\beta)} n^p \int_0^{2^n} \left| \left(D_{0+}^{\beta} \widehat{\varphi}_n \right)(s) \right| ds}{\sum_{k=1}^{2^{2n}} \varphi^2(t_{k-1}^n)} = \frac{n^p}{\Gamma(2-\beta) \cdot 2^{n(1-H)}} \to 0, \quad n \to \infty.$$

Consequently the conditions of Theorem 2 are satisfied and the estimator (20) is strongly consistent.

4. Simulations

In this section we illustrate quality of the estimators with the help of simulation experiments. We consider the equation (18) with $X_0 = 1$, $\theta = 1$. For each set of parameters, we simulate 100 trajectories of the solution. In the case $a = b$ we compute the average relative error

Table 1: $a(x) = b(x) = \sin x + 2$.

n	$H = 0.6$		$H = 0.7$		$H = 0.8$		$H = 0.9$	
	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$
3	0.0929	0.0967	0.0935	0.1015	0.0926	0.0956	0.1144	0.0935
4	0.0512	0.0512	0.0510	0.0509	0.0497	0.0471	0.0522	0.0495
5	0.0262	0.0258	0.0264	0.0258	0.0251	0.0244	0.0287	0.0261
6	0.0122	0.0121	0.0120	0.0120	0.0123	0.0127	0.0125	0.0107

Table 2: $a(x) = b(x) = \cos x + 2$.

n	$H = 0.6$		$H = 0.7$		$H = 0.8$		$H = 0.9$	
	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$
3	0.0898	0.0891	0.1006	0.0964	0.1021	0.0971	0.1088	0.0935
4	0.0567	0.0568	0.0501	0.0474	0.0544	0.0517	0.0611	0.0521
5	0.0228	0.0227	0.0276	0.0277	0.0255	0.0245	0.0280	0.0244
6	0.0139	0.0138	0.0122	0.0123	0.0130	0.0133	0.0140	0.0137

Table 3: $a(x) = b(x) = \frac{1}{1+x^2}$.

n	$H = 0.6$		$H = 0.7$		$H = 0.8$		$H = 0.9$	
	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$
3	0.0969	0.0950	0.1027	0.1003	0.1079	0.0963	0.1198	0.0922
4	0.0467	0.0457	0.0473	0.0477	0.0485	0.0444	0.0489	0.0437
5	0.0257	0.0259	0.0289	0.0282	0.0235	0.0245	0.0307	0.0263
6	0.0123	0.0123	0.0129	0.0128	0.0120	0.0116	0.0130	0.0114

Table 4: $a(x) = b(x) = 1$.

n	$H = 0.6$		$H = 0.7$		$H = 0.8$		$H = 0.9$	
	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$	$\delta_n^{(3)}$	$\delta_n^{(5)}$
3	0.0947	0.0925	0.1037	0.0998	0.1129	0.1106	0.1177	0.1052
4	0.0498	0.0513	0.0510	0.0504	0.0556	0.0522	0.0578	0.0504
5	0.0275	0.0271	0.0248	0.0255	0.0262	0.0260	0.0248	0.0236
6	0.0124	0.0125	0.0116	0.0118	0.0137	0.0138	0.0125	0.0120

$\delta_n^{(i)} = \left| \hat{\theta}_n^{(i)} - \theta \right| / \theta$ for each of estimators $\hat{\theta}_n^{(i)}$, $i = 3, 5$ (Tables 1–4). In the case $a \neq b$ we compute the average relative error only for the estimator $\hat{\theta}_n^{(5)}$ (Table 5).

In the case of equal coefficients we see that the estimators $\hat{\theta}_n^{(3)}$ and $\hat{\theta}_n^{(5)}$ have similar performance. The advantage of $\hat{\theta}_n^{(5)}$ is its independence of the parameter H (which might be unknown). But in the case of known H the estimator $\hat{\theta}_n^{(3)}$ is preferable because it is com-

Table 5: $a(x) = \sin x + 2$, $b(x) = \cos x + 2$.

n	$H = 0.6$ $\delta_n^{(5)}$	$H = 0.7$ $\delta_n^{(5)}$	$H = 0.8$ $\delta_n^{(5)}$	$H = 0.9$ $\delta_n^{(5)}$
3	0.0756	0.0792	0.0757	0.0751
4	0.0411	0.0361	0.0453	0.0459
5	0.0200	0.0199	0.0159	0.0200
6	0.0099	0.0113	0.0094	0.0100

putable faster.

Also the simulation results show that the rate on convergence probably does not depend on H. Moreover, it seems that it is around 2^{-n}, so the bound in Theorem 1 is not optimal.

References

Bishwal J (2011). "Minimum Contrast Estimation in Fractional Ornstein-Uhlenbeck Process: Continuous and Discrete Sampling." *Fractional Calculus and Applied Analysis*, **14**(3), 375–410. ISSN 1311-0454.

Hu Y, Nualart D (2010). "Parameter Estimation for Fractional Ornstein-Uhlenbeck Processes." *Statistics & Probability Letters*, **80**(11-12), 1030–1038.

Hu Y, Song J (2013). "Parameter Estimation for Fractional Ornstein-Uhlenbeck Processes with Discrete Observations." In F Viens, J Feng, Y Hu, E Nualart (eds.), *Malliavin Calculus and Stochastic Analysis*, volume 34 of *Springer Proceedings in Mathematics & Statistics*, pp. 427–442. Springer US. ISBN 978-1-4614-5905-7. doi:10.1007/978-1-4614-5906-4_19.

Kozachenko Y, Melnikov A, Mishura Y (2013). "On Drift Parameter Estimation in Models with Fractional Brownian Motion." Accepted by Statistics.

Mishura Y (2008). *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, volume 1929 of *Lecture Notes in Mathematics*. Springer, Berlin.

Mishura Y, Ral'chenko K, Seleznev O, Shevchenko G (2014). "Asymptotic Properties of Drift Parameter Estimator Based on Discrete Observations of Stochastic Differential Equation Driven by Fractional Brownian Motion." Submitted to Modern Trends in Stochastics.

Norros I, Valkeila E, Virtamo J (1999). "An Elementary Approach to a Girsanov Formula and Other Analytical Results on Fractional Brownian Motions." *Bernoulli*, **5**(4), 571–587.

Nualart D, Rascanu A (2001). "Differential Equations Driven by Fractional Brownian Motion." *Collectanea Mathematica*, **53**, 55–81.

Prakasa Rao B (2010). *Statistical Inference for Fractional Diffusion Processes*. Wiley Online Library: Books. Wiley. ISBN 9780470667132.

Tanaka K (2013). "Distributions of the Maximum Likelihood and Minimum Contrast Estimators Associated with the Fractional Ornstein-Uhlenbeck Process." *Statistical Inference for Stochastic Processes*, **16**(3), 173–192. ISSN 1387-0874.

Xiao W, Zhang W, Xu W (2011a). "Parameter Estimation for Fractional Ornstein-Uhlenbeck Processes at Discrete Observation." *Applied Mathematical Modelling*, **35**, 4196–4207.

Xiao W, Zhang W, Zhang XL (2011b). "Maximum-likelihood Estimators in the Mixed Fractional Brownian Motion." *Statistics*, **45**(1), 73–85.

Zhang P, Xiao W, Zhang XL, Niu P (2014). "Parameter Identification for Fractional Ornstein-Uhlenbeck Processes Based on Discrete Observation." *Economic Modelling*, **36**(C), 198–203.

Affiliation:

Yuliya Mishura and Kostiantyn Ralchenko
Department of Probability Theory, Statistics and Actuarial Mathematics
Taras Shevchenko National University of Kyiv
64 Volodymyrska
01601 Kyiv
Ukraine
E-mail: myus@univ.kiev.ua, k.ralchenko@gmail.com