Learning the Smoothness of Weakly Dependent Functional Times Series

Hassan Maissoro[1,*], Valentin Patilea[2], Myriam Vimond[3]

[1] CREST, Datastorm & ENSAI; hassan.maissoro@datastorm.fr, [*] Presenting author
[2] CREST, ENSAI; valentin.patilea@ensai.fr
[3] CREST, ENSAI; myriam.vimond@ensai.fr

Motivation

We aim to study stationary functional time series (FTS) where the trajectory are measured with error at discretely, randomly sampled, domain points. Our goal is to estimate the local regularity parameters of the trajectories for FTS in the context of weak dependency, and to derive non-asymptotic bounds for the concentration of these estimators. Indeed, a majority of inference problems in FDA depends on the local regularity.

Weak dependency

Let $\boldsymbol{X} = (X_n)_{n\in\mathbb{Z}}$ be a stationary FTS, with continuous paths, defined on the interval $I = [0,1]$:

- $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$: space of square integrable functions ;
- $(\mathcal{C}, \|\cdot\|_\infty)$: space of continuous functions on I.

The space $\mathbb{L}^p_{\mathcal{C}}$ is the space of $\mathcal{C}$−valued random element X such that

$$\nu_p(X) = \left(\mathbb{E}\left[\|X\|^p_\infty\right]\right)^{1/p} < \infty.$$

The process $\{X_n\}_n$ is $\mathbb{L}^{\mathbf{p}}_{\mathcal{C}} - \mathbf{m}$**-approximable** if each $X_n \in \mathbb{L}^p_{\mathcal{C}}$ admits the MA representation:

$$X_n = f(\varepsilon_n, \varepsilon_{n-1}, \ldots),$$

where $\{\varepsilon_n\}$ are i.i.d. elements in a measurable space S, and $f : S^\infty \to \mathcal{H}$ is measurable. Moreover, we assume that if, for every $n \in \mathbb{Z}$, $\{\varepsilon^{(n)}_k\}_k$ is an independent copy of $\{\varepsilon_n\}_n$ defined on the same probability space, then letting

$$X^{(m)}_n = f(\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_{n-m-1}, \varepsilon^{(n)}_{n-m}, \varepsilon^{(n)}_{n-(m-1)} \ldots),$$

we have

$$\sum_{m\geq 1} \nu_p\left(X_m - X^{(m)}_m\right) < \infty.$$

Example. *FAR(1) is* $\mathbb{L}^p_{\mathcal{C}} - m$*-approximable:*

$$X_n(t) = \int_0^1 \beta(t,s) X_{n-1}(s) ds + \varepsilon_n(t)$$

$\{\varepsilon_n\}_{n\in\mathbb{Z}}$ *are i.i.d. fBm with Hurst exponent* H_ε.

The local regularity parameters

The process X, with non differentiable paths, admits a ***local regularity*** at $t \in I$, with

- local **exponent** $H_t \in (0,1)$,
- and local **Hölder constant** $L_t > 0$, if

$$\mathbb{E}\left[(X(u) - X(v))^2\right] \approx L_t^2 |u - v|^{2H_t},$$

for all $u, v \in [t - \Delta/2, t + \Delta/2]$ for some $\Delta > 0$.

The local regularity estimators

We introduce for any u, v close to t,

$$\widehat{\theta}(u,v) = \frac{1}{N}\sum_{n=1}^N \left\{\widetilde{X}_n(v) - \widetilde{X}_n(u)\right\}^2$$

Let $t_1 = t - \Delta/2$, $t_3 = t + \Delta/2$. The estimator of H_t is

$$\widehat{H}_t = \frac{\log(\widehat{\theta}(t_1,t_3)) - \log(\widehat{\theta}(t_1,t))}{2\log(2)}$$

A plug-in estimator for L_t^2 is

$$\widehat{L}_t^2 = \frac{\widehat{\theta}\,(t_1,t_3)}{\Delta^{2\widehat{H}_t}}.$$

Concentration bounds

Let $\{X_n\}$ be $\mathbb{L}^4_{\mathcal{C}} - m$-approximable. Assume that the $\mathbb{L}^2$-risk of smoothing is suitably bounded. Then, for any $\mu \geq \mu_0$, for some μ_0, and for $\Delta > 0$ and $\varphi > 0$ depending on μ, we have

$$\mathbb{P}\left(|\widehat{H}_t - H_t| > \varphi\right) \leq \frac{4\mathfrak{f}_1}{N\varphi^2\Delta^{4H_t}} + 4\mathfrak{b}\exp\left(-\mathfrak{f}_2 N\varphi^2\Delta^{4H_t}\right),$$

$$\mathbb{P}\left(\left|\widehat{L}_t^2 - L_t^2\right| > \varphi\right) \leq \frac{5\mathfrak{l}_1}{N\varphi^2\Delta^{4H_t+4\varphi}} + 5\mathfrak{b}\exp\left(-\mathfrak{l}_2 N\varphi^2\Delta^{4H_t+4\varphi}\right),$$

where $\mathfrak{b} > 0$ is a constant and $\mathfrak{f}_1, \mathfrak{f}_2, \mathfrak{l}_1, \mathfrak{l}_2 > 0$ are also constants depending on the dependence measure.

Data observation Framework

For $n = 1 \ldots N$, the trajectory X_n is measured with error at discretely, randomly sampled, domain points:

$$Y_{n,k} = X_n(T_{n,k}) + \varepsilon_{n,k}, \quad 1 \leq k \leq M_n,$$

where

- $M_1, \ldots, M_N \overset{i.i.d.}{\sim} M$ with expectation μ,
- the observation times $T_{n,k} \sim T$ are independent,
- $\varepsilon_{n,k} \sim \epsilon$ are independent centered errors,
- $\mathbf{X}$, M, ϵ, and T are mutually independent.

For recovering the trajectories, we use the nonparametric estimation to construct an estimator $\widetilde{X}_n$ for each X_n, using its sampled points $(Y_{n,k}, T_{n,k})_k$.

Simulation

We simulate a FAR(1) where $\{\varepsilon_n\}$ are i.i.d. 'tied-down' *multifractional Brownian motion* (see [1]) paths with :

- a logistic H_t function and $L_t^2 = 4$,
- and a kernel $\beta(s,t) = \alpha st$, with $\alpha = 9/4$.

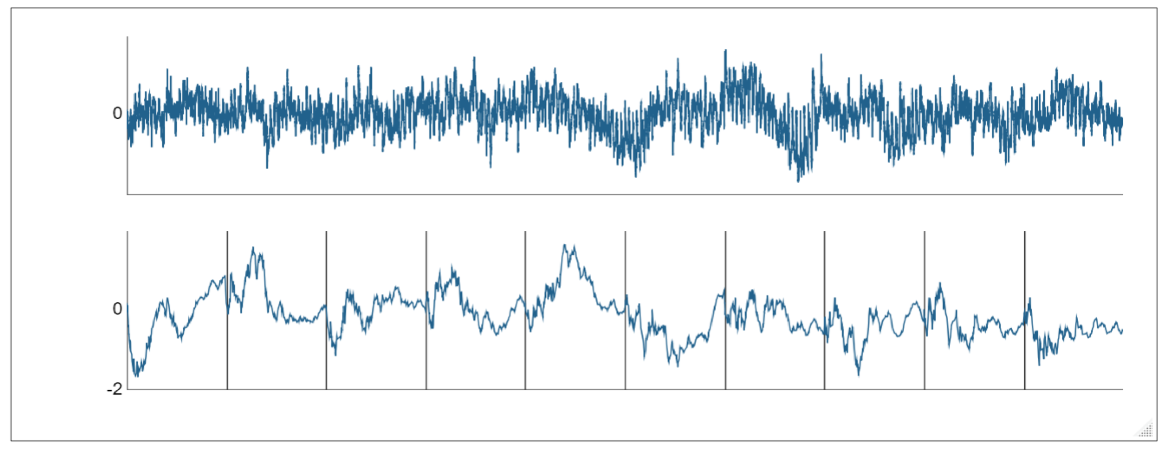

Figure: Time series of $N = 250$ observations of a simulated FAR(1) without error. The last ten functions are shown in the bottom graph.

Estimation results

Estimation of H_t and L_t^2 at $t = 1/2$ using the previous FAR(1) and taking $\epsilon \sim \mathcal{N}(0,1)$.

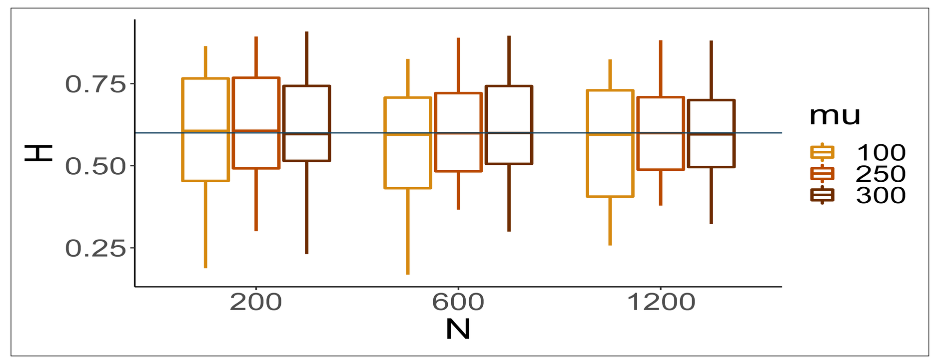

Figure: Estimates of $\widehat{H}_t$. The line is the true $H_t = 0.6$.

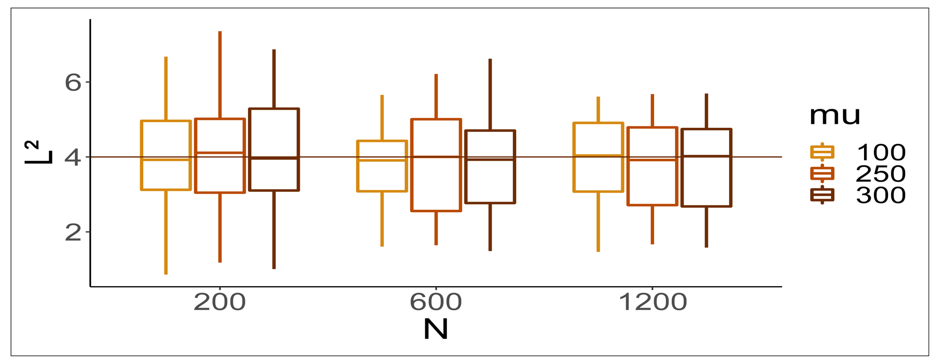

Figure: Estimates of $\widehat{L}_t^2$. The line is the true $L_t^2 = 4$.

Perspectives

Build adaptive estimation of :

- mean and covariance functions,
- auto-covariance function,
- dynamic functional principal component,
- depth functions, etc.

References

[1] Stilian A. Stoev and Murad S. Taqqu. How rich is the class of multifractional brownian motions? *Stochastic Processes and their Applications*, 116(2):200–221, 2006.

[2] Siegfried Hörmann and Piotr Kokoszka. Weakly dependent functional data. *The Annals of Statistics*, 38(3):1845–1884, 2010.

[3] Steven Golovkine, Nicolas Klutchnikoff, and Valentin Patilea. Learning the smoothness of noisy curves with application to online curve estimation. *Electronic Journal of Statistics*, 16(1):1485–1560, 2022.

[4] Weidong Liu, Han Xiao, and Wei Biao Wu. Probabilities and moment inequalities under dependence. *Statistica Sinica*, 23:1257–1272, 2013.