Learning the Smoothness of Weakly Dependent Functional Time Series

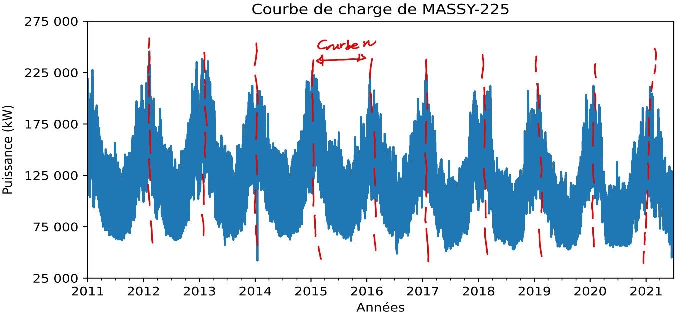
Hassan Maissoro, Valentin Patilea and Myriam Vimond

ENSAE-ENSAI stat/ML days, 12-13 September 2023

Introduction (1/3)

Example of a connection point for the extraction and injection of electricity

- A set of N time-dependent curves, $X_n : [0,1] \to \mathbb{R}$, $n = 1 \dots N$.



- The trajectories are irregular.
- We observe each curve every 10 mins + measurement errors.
- Regularity and final goal should be considered in reconstruction.

Introduction (2/3)

Observation scheme

For $n = 1, \dots, N$, X_n is measured with error at discrete, randomly sampled points :

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

- $\{X_n\}$ is a stationary process of $\mathcal{H} = \mathbb{L}^2[0,1]$,
- $M_1, \dots, M_N \overset{i.i.d.}{\sim} M$ with expectation λ,
- the observation times $T_{n,k} \sim T$ are i.i.d.,
- $\varepsilon_{n,k} \sim \epsilon$ are independent centered errors,
- $\{X_n\}$, $\{M_n\}$, $\{\varepsilon_{n,k}\}$, and $\{T_{n,k}\}$ are mutually independent.

Introduction (3/3)

Motivation

We aim to estimate the local regularity parameters of the trajectories for FTS in the context of weak dependency.

Using dependent curves measured with noise at random discrete points, our goal is to perform adaptive estimation of :

- mean and autocovariance kernel functions,
- depth functions, *etc*.

- The concept of local regularity was considered by Golovkine et al., (2022) for i.i.d. functional data.
- For FTS, mean and autocovariance estimators have already been considered by Rubìn and Panaretos (2020) under the hypothesis that these functions admit at least one derivative.
- We extend the results of Golovkine et al., (2022) to FTS to perform estimates that adapt to the local regularity.

Outline

Local regularity parameters (1/5)

Definition and estimation

Definition. The process X admits a *local regularity* at $t \in I$, with local exponent $H_t \in (0, 1)$ and 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 satisfying $t - \Delta/2 \leq u \leq t \leq v \leq t + \Delta/2$ for some $\Delta > 0$.

Local regularity parameters (1/5)

Definition and estimation

Definition. The process X admits a *local regularity* at $t \in I$, with local exponent $H_t \in (0,1)$ and 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 satisfying $t - \Delta/2 \leq u \leq t \leq v \leq t + \Delta/2$ for some $\Delta > 0$.

Estimation. We use some nonparametric estimates $\widetilde{X}_n$ to recover the X_n's. For any u, v close to t, let

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

Our estimators of H_t and L_t^2 are defined as empirical counterparts of their respective definition. Let $t_1 = t - \Delta/2$, $t_3 = t + \Delta/2$. The estimators of H_t and L_t^2 are

$$\widehat{H}_t = \frac{\log(\widehat{\theta}(t_1, t_3)) - \log(\widehat{\theta}(t_1, t))}{2\log(2)} \qquad \text{and} \qquad \widehat{L}_t^2 = \frac{\widehat{\theta}(t_1, t_3)}{\Delta^{2\widehat{H}_t}}.$$

Local regularity parameters (2/5)

Weak dependency assumption

Let $\{X_n\}_{n\in\mathbb{Z}}$ be a stationary FTS, with continuous paths, on $I = [0,1]$:

- $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$: 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

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

Weak dependency assumption : $\{X_n\}_n$ is $\mathbb{L}^{\mathbf{p}}_{\mathcal{C}}$ **– m-approximable**.

- $\mathbb{L}^{\mathbf{p}}$ **– m**-approximation for $\mathcal{H}-$valued functional data was introduced by HÖRMANN and KOKOSZKA (2010).
- We need a dependency type of $\{X_n\}$ that can be inherited by $\{X_n(t)\}$ because we are studying $\{X_n\}$ locally at $t \in I$ and such we use $\|\cdot\|_\infty$ instead of $\|\cdot\|_{\mathcal{H}}$.

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

Local regularity parameters (3/5)

Concentration bounds

- Let $\{X_n\}$ be $\mathbb{L}^{\mathbf{4}}_{\mathcal{C}}$ − **m-approximable**.
- Assume that the $\mathbb{L}^2$-risk of smoothing is suitably bounded.

Then, for some $\varphi, \psi \in (0,1)$ such that

$$6L_t^2 \Delta^{-2\varphi} \varphi |log \Delta| < \psi,$$

and for λ large enough, we have :

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

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

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

Local regularity parameters (4/5)

Application : sample paths of a FAR(1)

We simulate a FAR(1) where $\{\xi_n\}$ are i.i.d. 'tied-down' *multifractional Brownian motion* (see STOEV and TAQQU (2006)) paths with :

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

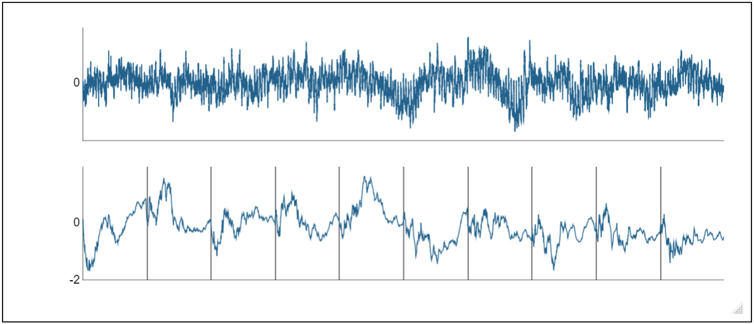


Figure – Time series of $N = 250$ observations of a simulated FAR(1) without error. The last ten functions are shown in the bottom graph.

Local regularity parameters (5/5)

Application : estimation of local regularity parameters

Estimation of H_t and L_t^2 at $t = 1/2$ using the previous FAR(1) and taking $\epsilon \sim \mathcal{N}(0, \sigma^2 = 0.04)$.

- Obtained reasonably good results :

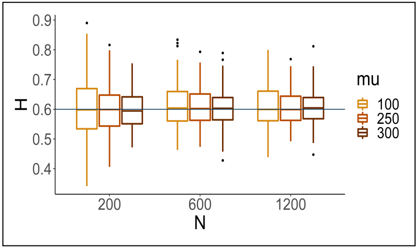


Figure – Estimates of $\widehat{H}_t$. The line is the true $H_t = 0.6$.

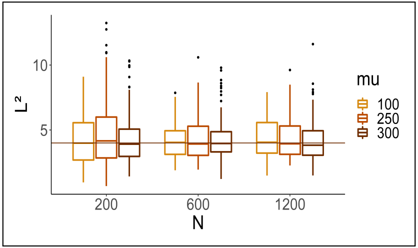


Figure – Estimates of $\widehat{L_t^2}$. The line is the true $L_t^2 = 4$.

Adaptive mean function estimation

Adaptive mean function estimation

Let $\mu(t) = \mathbb{E}(X_n(t))$ be the mean function of the stationary process $\{X_n\}$.

- A naive estimator of $\mu(t)$: $\widehat{\mu}_N(t,h) = N^{-1}(\widehat{X}_1(t,h) + \cdots + \widehat{X}_N(t,h))$, where $\widehat{X}_n(t,h)$ is a nonparametric estimator of X_n, and h a bandwidth.

Adaptive mean function estimation

Let $\mu(t) = \mathbb{E}(X_n(t))$ be the mean function of the stationary process $\{X_n\}$.

- A naive estimator of $\mu(t)$: $\widehat{\mu}_N(t, h) = N^{-1}(\widehat{X}_1(t, h) + \cdots + \widehat{X}_N(t, h))$, where $\widehat{X}_n(t, h)$ is a nonparametric estimator of X_n, and h a bandwidth.
- **The objective :** estimation of $\mu(t)$ by selection of h according to the local regularity of $\{X_n\}$ at time t and selection of the relevant curves of the sample.
- The proposed estimator is $\widehat{\mu}_P(t; h^*_\mu)$, with

$$\widehat{\mu}_P(t; h) = P_N(t; h)^{-1}\left(\pi_1(t; h)\widehat{X}_1(t; h) + \cdots + \pi_N(t; h)\widehat{X}_N(t; h)\right)^{-1}$$

$$\pi_n(t; h) = \begin{cases} 1 & \text{if } \sum_{k=1}^{M_n} \mathbb{1}\{|T_{n,k} - t| \leq h\} \geq 1 \\ 0 & \text{otherwise} \end{cases}, \quad P_N(t; h) = \sum_{n=1}^{N} \pi_n(t; h).$$

- h^*_μ minimises a sharp upper bound of the quadratic risk of $\mu(t)$.

Adaptive mean function estimation

Let $\mu(t) = \mathbb{E}(X_n(t))$ be the mean function of the stationary process $\{X_n\}$.

- A naive estimator of $\mu(t)$: $\widehat{\mu}_N(t,h) = N^{-1}(\widehat{X}_1(t,h) + \cdots + \widehat{X}_N(t,h))$, where $\widehat{X}_n(t,h)$ is a nonparametric estimator of X_n, and h a bandwidth.
- **The objective :** estimation of $\mu(t)$ by selection of h according to the local regularity of $\{X_n\}$ at time t and selection of the relevant curves of the sample.
- The proposed estimator is $\widehat{\mu}_P(t; h^*_\mu)$, with

$$\widehat{\mu}_P(t;h) = P_N(t;h)^{-1}\left(\pi_1(t;h)\widehat{X}_1(t;h) + \cdots + \pi_N(t;h)\widehat{X}_N(t;h)\right)^{-1}$$

$$\pi_n(t;h) = \begin{cases} 1 & \text{if } \sum_{k=1}^{M_n} \mathbb{1}\{|T_{n,k} - t| \leq h\} \geq 1 \\ 0 & \text{otherwise} \end{cases}, \quad P_N(t;h) = \sum_{n=1}^{N} \pi_n(t;h).$$

- h^*_μ minimises a sharp upper bound of the quadratic risk of $\mu(t)$.

Adaptive autocovariance kernel estimation (work in progress).

- **The objective :** The same methodology is under development for the autocovariance kernel for lag-ℓ, $\ell > 0$.

Take home message

1. Estimation of local regularity for FTS.
   - Local regularity parameters are : exponent H_t and Hölder constant L_t^2.
   - Exponential bound for the concentration of the estimators of H_t and L_t^2 under $\mathbb{L}_{\mathcal{C}}^{\mathbf{4}}$ − **m-approximation**.
   - The simulations show that $\widehat{H}_t$ and $\widehat{L_t^2}$ give satisfactory results.

2. Adaptive estimation of the mean and autocovariance kernel functions.
   - Optimal smoothing parameter used to reconstruct curves depends on the final goal.
   - Simulations show satisfactory results for the mean function.

- Perspectives :
  - Adaptive estimators for anomaly detection,
  - Robust prediction model, etc.

Thanks for your attention !