

Regularity estimation in multivariate functional data analysis

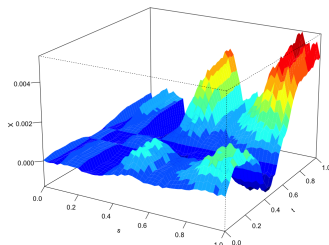
Omar KASSI & Nicolas KLUTCHNIKOFF & Valentin PATILEA

August 14, 2024
Bernoulli-IMS 11th World Congress



Multivariate functional data

- The realizations of the stochastic process X are surfaces
 - Satellite images
 - Measurements of temperature or salinity in oceanology



Data

- $\mathcal{T}$: open, bounded bi-dimensional rectangle, $\overline{\mathcal{T}} \subset (0, \infty)^2$

Data

- $\mathcal{T}$: open, bounded bi-dimensional rectangle, $\overline{\mathcal{T}} \subset (0, \infty)^2$
- $X^{(1)}, \dots, X^{(j)}, \dots, X^{(N)}$ are independent realizations of X

Data

- $\mathcal{T}$: open, bounded bi-dimensional rectangle, $\overline{\mathcal{T}} \subset (0, \infty)^2$
- $X^{(1)}, \dots, X^{(j)}, \dots, X^{(N)}$ are independent realizations of X
- The data associated to a sample path $X^{(j)}$ consist of the pairs $(Y_m^{(j)}, \mathbf{t}_m^{(j)}) \in \mathbb{R} \times \mathcal{T}$, where for $1 \leq j \leq N$ and $1 \leq m \leq M_j$

$$Y_m^{(j)} = X^{(j)}(\mathbf{t}_m^{(j)}) + \varepsilon_m^{(j)}, \quad \text{with} \quad \varepsilon_m^{(j)} = \sigma(\mathbf{t}_m^{(j)}, X^{(j)}(\mathbf{t}_m^{(j)}))e_m^{(j)}$$

Data

- $\mathcal{T}$: open, bounded bi-dimensional rectangle, $\overline{\mathcal{T}} \subset (0, \infty)^2$
- $X^{(1)}, \dots, X^{(j)}, \dots, X^{(N)}$ are independent realizations of X
- The data associated to a sample path $X^{(j)}$ consist of the pairs $(Y_m^{(j)}, \mathbf{t}_m^{(j)}) \in \mathbb{R} \times \mathcal{T}$, where for $1 \leq j \leq N$ and $1 \leq m \leq M_j$

$$Y_m^{(j)} = X^{(j)}(\mathbf{t}_m^{(j)}) + \varepsilon_m^{(j)}, \quad \text{with} \quad \varepsilon_m^{(j)} = \sigma(\mathbf{t}_m^{(j)}, X^{(j)}(\mathbf{t}_m^{(j)})) e_m^{(j)}$$

- $M_1, \dots, M_N$ be an independent sample of an integer-valued random variable M , $\mathbb{E}[M] = \mathfrak{m}$

Data

- $\mathcal{T}$: open, bounded bi-dimensional rectangle, $\overline{\mathcal{T}} \subset (0, \infty)^2$
- $X^{(1)}, \dots, X^{(j)}, \dots, X^{(N)}$ are independent realizations of X
- The data associated to a sample path $X^{(j)}$ consist of the pairs $(Y_m^{(j)}, \mathbf{t}_m^{(j)}) \in \mathbb{R} \times \mathcal{T}$, where for $1 \leq j \leq N$ and $1 \leq m \leq M_j$

$$Y_m^{(j)} = X^{(j)}(\mathbf{t}_m^{(j)}) + \varepsilon_m^{(j)}, \quad \text{with} \quad \varepsilon_m^{(j)} = \sigma(\mathbf{t}_m^{(j)}, X^{(j)}(\mathbf{t}_m^{(j)}))e_m^{(j)}$$

- $M_1, \dots, M_N$ be an independent sample of an integer-valued random variable M , $\mathbb{E}[M] = m$
- The $(\mathbf{t}_m^{(j)}, 1 \leq m \leq M_j)$ represent the observation points for the sample path $X^{(j)}$.

First steps : univariate case (1/2)

- For B^H a fBm with Hurst index $H \in (0, 1)$,

$$\mathbb{E} \left[\left\{ B^H(t) - B^H(s) \right\}^2 \right] = |t - s|^{2H}, \quad s, t \in \mathbb{R}_+$$

First steps : univariate case (1/2)

- For B^H a fBm with Hurst index $H \in (0, 1)$,

$$\mathbb{E} \left[\left\{ B^H(t) - B^H(s) \right\}^2 \right] = |t - s|^{2H}, \quad s, t \in \mathbb{R}_+$$

- Estimating equation for the Hurst parameter :

$$H = \frac{\log \left(\mathbb{E} \left[\left\{ B^H(t) - B^H(s) \right\}^2 \right] \right)}{2 \log |t - s|}$$

First steps : univariate case (2/2)

- Let X be a process defined on a subset of $\mathbb{R}$, with **non-differentiable sample paths**

First steps : univariate case (2/2)

- Let X be a process defined on a subset of $\mathbb{R}$, with **non-differentiable sample paths**
- GKP (2022) : $H(t_0) \in (0, 1)$ and $L(t_0) > 0$ exist such that

$$\mathbb{E} \left[\{X(t) - X(s)\}^2 \right] \approx L(t_0)^2 |t - s|^{2H(t_0)}, \quad \forall s \leq t_0 \leq t$$

for t and s close to t_0

First steps : univariate case (2/2)

- Let X be a process defined on a subset of $\mathbb{R}$, with **non-differentiable sample paths**
- GKP (2022) : $H(t_0) \in (0, 1)$ and $L(t_0) > 0$ exist such that

$$\mathbb{E} \left[\{X(t) - X(s)\}^2 \right] \approx L(t_0)^2 |t - s|^{2H(t_0)}, \quad \forall s \leq t_0 \leq t$$

for t and s close to t_0

- Estimating equation :

$$H(t_0) \approx \frac{\log(\theta(t_1, t_2)) - \log(\theta(t_1, t_3))}{2 \log(2)}, \quad t_0 \in [t_1, t_2] \subset [t_1, t_3]$$

where

$$\theta(t, s) = \mathbb{E} \left[\{X(t) - X(s)\}^2 \right] \quad \text{and} \quad |t_1 - t_2| = 2|t_1 - t_3|.$$

Multivariate case : notation

- $H_1, H_2 : \mathcal{T} \rightarrow (0, 1)$ are continuously differentiable functions.
Let

$$\overline{H} = \max\{H_1, H_2\}$$

Multivariate case : notation

- $H_1, H_2 : \mathcal{T} \rightarrow (0, 1)$ are continuously differentiable functions.
Let

$$\overline{H} = \max\{H_1, H_2\}$$

- $L_1^{(1)}, L_2^{(1)}, L_1^{(2)}, L_2^{(2)}$: Non negative Lipschitz continuous functions defined on $\mathcal{T}$ such that

$$L_k^{(1)}(\mathbf{t}) + L_k^{(2)}(\mathbf{t}) > 0, \quad \forall \mathbf{t} \in \mathcal{T} \subset \mathbb{R}^2, \quad k = 1, 2.$$

Multivariate case : notation

- $H_1, H_2 : \mathcal{T} \rightarrow (0, 1)$ are continuously differentiable functions.
Let

$$\overline{H} = \max\{H_1, H_2\}$$

- $L_1^{(1)}, L_2^{(1)}, L_1^{(2)}, L_2^{(2)}$: Non negative Lipschitz continuous functions defined on $\mathcal{T}$ such that

$$L_k^{(1)}(\mathbf{t}) + L_k^{(2)}(\mathbf{t}) > 0, \quad \forall \mathbf{t} \in \mathcal{T} \subset \mathbb{R}^2, \quad k = 1, 2.$$

- For $X \in \mathcal{L}^2$, we denote for sufficiently small scalars Δ

$$\theta_{\mathbf{t}}^{(i)}(\Delta) = \mathbb{E} \left[\left\{ X \left(\mathbf{t} + \frac{\Delta}{2} \mathbf{e}_i \right) - X \left(\mathbf{t} - \frac{\Delta}{2} \mathbf{e}_i \right) \right\}^2 \right], \quad i = 1, 2,$$

where $(\mathbf{e}_1, \mathbf{e}_2)$ is canonical basis of $\mathbb{R}^2$

A class of multivariate processes

Definition

We say $X \in \mathcal{H}^{H_1, H_2}(\mathbf{L}, \mathcal{T})$ if three constants $\Delta_0, C, \beta > 0$ exist such that for any $\mathbf{t} \in \mathcal{T}$ and $0 < \Delta \leq \Delta_0$,

$$\left| \theta_{\mathbf{t}}^{(i)}(\Delta) - L_1^{(i)}(\mathbf{t})\Delta^{2H_1(\mathbf{t})} - L_2^{(i)}(\mathbf{t})\Delta^{2H_2(\mathbf{t})} \right| \leq C\Delta^{2\bar{H}(\mathbf{t})+\beta}, \quad i = 1, 2.$$

Let

$$\mathcal{H}^{H_1, H_2} = \bigcup_{\mathbf{L}} \mathcal{H}^{H_1, H_2}(\mathbf{L}, \mathcal{T}),$$

where $\mathbf{L} = (L_1^{(1)}, L_2^{(1)}, L_1^{(2)}, L_2^{(2)})$.

The functions H_1, H_2 define the local regularity of the process, while $\mathbf{L}$ represent the local Hölder constants.

Example : Sum of two fractional Brownian motion

- Let $B_1^{H_1}$ and $B_2^{H_2}$ be two independent fBm with Hurst index H_1 and H_2 .
- Let

$$\mathbf{X}_1(\mathbf{t}) = B_1^{H_1}(t_1) + B_2^{H_2}(t_2), \quad \forall \mathbf{t} = (t_1, t_2) \in \mathbb{R}^2.$$

Then $\mathbf{X}_1 \in \mathcal{H}^{H_1, H_2}$ where $\mathbf{L} = (1, 0, 0, 1)$.

- Let $\beta > 0$ and define

$$\mathbf{X}_2(\mathbf{t}) = \mathbf{X}_1 \left(\begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \mathbf{t} \right), \quad \forall \mathbf{t} \in \mathbb{R}^2.$$

Then $\mathbf{X}_2 \in \mathcal{H}^{H_1, H_2}$ with

$$\mathbf{L} = (|\cos \beta|^{2H_1}, |\sin \beta|^{2H_2}, |\sin \beta|^{2H_1}, |\cos \beta|^{2H_2}).$$

Identification issues

- Let $H_1, H_2, \tilde{H}_1$ and $\tilde{H}_2$ be some continuously differentiable functions taking values in $(0, 1)$

Identification issues

- Let $H_1, H_2, \tilde{H}_1$ and $\tilde{H}_2$ be some continuously differentiable functions taking values in $(0, 1)$
- Assume $X \in \mathcal{H}^{H_1, H_2}$ and $X \in \mathcal{H}^{\tilde{H}_1, \tilde{H}_2}$

Identification issues

- Let $H_1, H_2, \tilde{H}_1$ and $\tilde{H}_2$ be some continuously differentiable functions taking values in $(0, 1)$
- Assume $X \in \mathcal{H}^{H_1, H_2}$ and $X \in \mathcal{H}^{\tilde{H}_1, \tilde{H}_2}$
- We then necessarily have

$$\min\{H_1(\mathbf{t}), H_2(\mathbf{t})\} = \min\{\tilde{H}_1(\mathbf{t}), \tilde{H}_2(\mathbf{t})\}$$

and

$$\max\{H_1(\mathbf{t}), H_2(\mathbf{t})\} = \max\{\tilde{H}_1(\mathbf{t}), \tilde{H}_2(\mathbf{t})\}$$

Identification issues

- Let $H_1, H_2, \tilde{H}_1$ and $\tilde{H}_2$ be some continuously differentiable functions taking values in $(0, 1)$
- Assume $X \in \mathcal{H}^{H_1, H_2}$ and $X \in \mathcal{H}^{\tilde{H}_1, \tilde{H}_2}$
- We then necessarily have

$$\min\{H_1(\mathbf{t}), H_2(\mathbf{t})\} = \min\{\tilde{H}_1(\mathbf{t}), \tilde{H}_2(\mathbf{t})\}$$

and

$$\max\{H_1(\mathbf{t}), H_2(\mathbf{t})\} = \max\{\tilde{H}_1(\mathbf{t}), \tilde{H}_2(\mathbf{t})\}$$

- Notation :

$$\underline{H}(\mathbf{t}) = \min\{H_1(\mathbf{t}), H_2(\mathbf{t})\}, \quad \overline{H}(\mathbf{t}) = \max\{H_1(\mathbf{t}), H_2(\mathbf{t})\}.$$

Estimating equations for $\underline{H}$ and $\overline{H}$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) = \mathbb{E} \left[\{X(\mathbf{t} - \Delta \mathbf{e}_i/2) - X(\mathbf{t} + \Delta \mathbf{e}_i/2)\}^2 \right], \quad i = 1, 2,$$

Estimating equations for $\underline{H}$ and $\overline{H}$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) = \mathbb{E} \left[\{X(\mathbf{t} - \Delta \mathbf{e}_i/2) - X(\mathbf{t} + \Delta \mathbf{e}_i/2)\}^2 \right], \quad i = 1, 2,$$

- Denote for any $\mathbf{t} \in \mathcal{T}$

$$\gamma_{\mathbf{t}}(\Delta) = \theta_{\mathbf{t}}^{(1)}(\Delta) + \theta_{\mathbf{t}}^{(2)}(\Delta)$$

Estimating equations for $\underline{H}$ and $\overline{H}$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) = \mathbb{E} \left[\{X(\mathbf{t} - \Delta \mathbf{e}_i/2) - X(\mathbf{t} + \Delta \mathbf{e}_i/2)\}^2 \right], \quad i = 1, 2,$$

- Denote for any $\mathbf{t} \in \mathcal{T}$

$$\gamma_{\mathbf{t}}(\Delta) = \theta_{\mathbf{t}}^{(1)}(\Delta) + \theta_{\mathbf{t}}^{(2)}(\Delta)$$

- Then

$$\underline{H}(\mathbf{t}) \approx \frac{\log(\gamma_{\mathbf{t}}(2\Delta)) - \log(\gamma_{\mathbf{t}}(\Delta))}{2 \log(2)}.$$

Estimating equations for $\underline{H}$ and $\overline{H}$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) = \mathbb{E} \left[\{X(\mathbf{t} - \Delta \mathbf{e}_i/2) - X(\mathbf{t} + \Delta \mathbf{e}_i/2)\}^2 \right], \quad i = 1, 2,$$

- Denote for any $\mathbf{t} \in \mathcal{T}$

$$\gamma_{\mathbf{t}}(\Delta) = \theta_{\mathbf{t}}^{(1)}(\Delta) + \theta_{\mathbf{t}}^{(2)}(\Delta)$$

- Then

$$\underline{H}(\mathbf{t}) \approx \frac{\log(\gamma_{\mathbf{t}}(2\Delta)) - \log(\gamma_{\mathbf{t}}(\Delta))}{2 \log(2)}.$$

- Let

$$\alpha_{\mathbf{t}}(\Delta) = \left| \frac{\gamma_{\mathbf{t}}(2\Delta)}{(2\Delta)^{2\underline{H}(\mathbf{t})}} - \frac{\gamma_{\mathbf{t}}(\Delta)}{\Delta^{2\underline{H}(\mathbf{t})}} \right|.$$

Estimating equations for $\underline{H}$ and $\overline{H}$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) = \mathbb{E} \left[\{X(\mathbf{t} - \Delta \mathbf{e}_i/2) - X(\mathbf{t} + \Delta \mathbf{e}_i/2)\}^2 \right], \quad i = 1, 2,$$

- Denote for any $\mathbf{t} \in \mathcal{T}$

$$\gamma_{\mathbf{t}}(\Delta) = \theta_{\mathbf{t}}^{(1)}(\Delta) + \theta_{\mathbf{t}}^{(2)}(\Delta)$$

- Then

$$\underline{H}(\mathbf{t}) \approx \frac{\log(\gamma_{\mathbf{t}}(2\Delta)) - \log(\gamma_{\mathbf{t}}(\Delta))}{2 \log(2)}.$$

- Let

$$\alpha_{\mathbf{t}}(\Delta) = \left| \frac{\gamma_{\mathbf{t}}(2\Delta)}{(2\Delta)^{2\underline{H}(\mathbf{t})}} - \frac{\gamma_{\mathbf{t}}(\Delta)}{\Delta^{2\underline{H}(\mathbf{t})}} \right|.$$

- Then

$$\overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t}) \approx \frac{\log(\alpha_{\mathbf{t}}(2\Delta)) - \log(\alpha_{\mathbf{t}}(\Delta))}{2 \log(2)}.$$

Estimators for $\underline{H}$ and $\overline{H}$: presmoothing

- In general, the sheets $X^{(j)}$, $j \in \{1, \dots, N\}$, are not available

Estimators for $\underline{H}$ and $\overline{H}$: presmoothing

- In general, the sheets $X^{(j)}$, $j \in \{1, \dots, N\}$, are not available
- Let $\tilde{X}^{(j)}$ be an observable approximation of $X^{(j)}$.

Estimators for $\underline{H}$ and $\overline{H}$: presmoothing

- In general, the sheets $X^{(j)}$, $j \in \{1, \dots, N\}$, are not available
- Let $\tilde{X}^{(j)}$ be an observable approximation of $X^{(j)}$.
 - If X is observed everywhere and without noise, then

$$\tilde{X}^{(j)}(\mathbf{t}) = X^{(j)}(\mathbf{t}), \quad \forall \mathbf{t} \in \mathcal{T}$$

Estimators for $\underline{H}$ and $\overline{H}$: presmoothing

- In general, the sheets $X^{(j)}$, $j \in \{1, \dots, N\}$, are not available
- Let $\tilde{X}^{(j)}$ be an observable approximation of $X^{(j)}$.
 - If X is observed everywhere and without noise, then

$$\tilde{X}^{(j)}(\mathbf{t}) = X^{(j)}(\mathbf{t}), \quad \forall \mathbf{t} \in \mathcal{T}$$

- If X is observed with noise or/and on a discrete grid, then $\tilde{X}^{(j)}$ is an estimator of $X^{(j)}$ (local polynomial, splines, interpolation...)

Estimators for $\underline{H}$ and $\overline{H}$ (1/2)

- The observable approximation allows to build estimates :

$$\widehat{\theta}_{\mathbf{t}}^{(i)}(\Delta) = \frac{1}{N} \sum_{j=1}^N \left\{ \widetilde{X}^{(j)}(\mathbf{t} - (\Delta/2)\mathbf{e}_i) - \widetilde{X}^{(j)}(\mathbf{t} + (\Delta/2)\mathbf{e}_i) \right\}^2,$$

Estimators for $\underline{H}$ and $\overline{H}$ (1/2)

- The observable approximation allows to build estimates :

$$\widehat{\theta}_{\mathbf{t}}^{(i)}(\Delta) = \frac{1}{N} \sum_{j=1}^N \left\{ \widetilde{X}^{(j)}(\mathbf{t} - (\Delta/2)\mathbf{e}_i) - \widetilde{X}^{(j)}(\mathbf{t} + (\Delta/2)\mathbf{e}_i) \right\}^2,$$

$$\widehat{\gamma}_{\mathbf{t}}(\Delta) = \widehat{\theta}_{\mathbf{t}}^{(1)}(\Delta) + \widehat{\theta}_{\mathbf{t}}^{(2)}(\Delta).$$

Estimators for $\underline{H}$ and $\overline{H}$ (1/2)

- The observable approximation allows to build estimates :

$$\widehat{\theta}_{\mathbf{t}}^{(i)}(\Delta) = \frac{1}{N} \sum_{j=1}^N \left\{ \widetilde{X}^{(j)}(\mathbf{t} - (\Delta/2)\mathbf{e}_i) - \widetilde{X}^{(j)}(\mathbf{t} + (\Delta/2)\mathbf{e}_i) \right\}^2,$$

$$\widehat{\gamma}_{\mathbf{t}}(\Delta) = \widehat{\theta}_{\mathbf{t}}^{(1)}(\Delta) + \widehat{\theta}_{\mathbf{t}}^{(2)}(\Delta).$$

- The first estimator follows :

$$\underline{\widehat{H}}(\mathbf{t}) = \begin{cases} \frac{\log(\widehat{\gamma}_{\mathbf{t}}(2\Delta)) - \log(\widehat{\gamma}_{\mathbf{t}}(\Delta))}{2 \log(2)} & \text{if } \widehat{\gamma}_{\mathbf{t}}(2\Delta), \widehat{\gamma}_{\mathbf{t}}(\Delta) > 0 \\ 1 & \text{otherwise} \end{cases}.$$

Estimators for $\underline{H}$ and $\overline{H}$ (2/2)

- Moreover

$$\hat{\alpha}_{\mathbf{t}}(\Delta) = \begin{cases} \left| \frac{\hat{\gamma}_{\mathbf{t}}(2\Delta)}{(2\Delta)^{2\hat{H}(\mathbf{t})}} - \frac{\hat{\gamma}_{\mathbf{t}}(\Delta)}{\Delta^{2\hat{H}(\mathbf{t})}} \right| & \text{if } \frac{\hat{\gamma}_{\mathbf{t}}(2\Delta)}{(2\Delta)^{2\hat{H}(\mathbf{t})}} \neq \frac{\hat{\gamma}_{\mathbf{t}}(\Delta)}{\Delta^{2\hat{H}(\mathbf{t})}} \\ 1 & \text{otherwise} \end{cases}.$$

Estimators for $\underline{H}$ and $\overline{H}$ (2/2)

- Moreover

$$\hat{\alpha}_t(\Delta) = \begin{cases} \left| \frac{\hat{\gamma}_t(2\Delta)}{(2\Delta)^{2\hat{H}(t)}} - \frac{\hat{\gamma}_t(\Delta)}{\Delta^{2\hat{H}(t)}} \right| & \text{if } \frac{\hat{\gamma}_t(2\Delta)}{(2\Delta)^{2\hat{H}(t)}} \neq \frac{\hat{\gamma}_t(\Delta)}{\Delta^{2\hat{H}(t)}} \\ 1 & \text{otherwise} \end{cases}.$$

- Hence

$$\widehat{(\overline{H} - \underline{H})}(t) = \frac{\log(\hat{\alpha}_t(2\Delta)) - \log(\hat{\alpha}_t(\Delta))}{2 \log(2)}$$

Estimators for $\underline{H}$ and $\overline{H}$ (2/2)

- Moreover

$$\hat{\alpha}_{\mathbf{t}}(\Delta) = \begin{cases} \left| \frac{\hat{\gamma}_{\mathbf{t}}(2\Delta)}{(2\Delta)^{2\hat{H}(\mathbf{t})}} - \frac{\hat{\gamma}_{\mathbf{t}}(\Delta)}{\Delta^{2\hat{H}(\mathbf{t})}} \right| & \text{if } \frac{\hat{\gamma}_{\mathbf{t}}(2\Delta)}{(2\Delta)^{2\hat{H}(\mathbf{t})}} \neq \frac{\hat{\gamma}_{\mathbf{t}}(\Delta)}{\Delta^{2\hat{H}(\mathbf{t})}} \\ 1 & \text{otherwise} \end{cases}.$$

- Hence

$$(\widehat{\overline{H} - \underline{H}})(\mathbf{t}) = \frac{\log(\hat{\alpha}_{\mathbf{t}}(2\Delta)) - \log(\hat{\alpha}_{\mathbf{t}}(\Delta))}{2 \log(2)}$$

- We then set

$$A_N(\tau) = \left\{ (\widehat{\overline{H} - \underline{H}})(\mathbf{t}) \geq \tau \right\},$$

and define

$$\widehat{\overline{H}}(\mathbf{t}) = \underline{\hat{H}}(\mathbf{t}) + (\widehat{\overline{H} - \underline{H}})(\mathbf{t}) 1_{A_N(\tau)}.$$

Estimating equations for $L_1^{(i)}(\mathbf{t})$ and $L_2^{(i)}(\mathbf{t})$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) \approx L_1^{(i)}(\mathbf{t})\Delta^{2H_1(\mathbf{t})} + L_2^{(i)}(\mathbf{t})\Delta^{2H_2(\mathbf{t})}, \quad i = 1, 2$$

Estimating equations for $L_1^{(i)}(\mathbf{t})$ and $L_2^{(i)}(\mathbf{t})$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) \approx L_1^{(i)}(\mathbf{t})\Delta^{2H_1(\mathbf{t})} + L_2^{(i)}(\mathbf{t})\Delta^{2H_2(\mathbf{t})}, \quad i = 1, 2$$

- Assume

$$H_1(\mathbf{t}) = \underline{H}(\mathbf{t}) < \overline{H}(\mathbf{t}) = H_2(\mathbf{t})$$

Estimating equations for $L_1^{(i)}(\mathbf{t})$ and $L_2^{(i)}(\mathbf{t})$

- Recall

$$\theta_{\mathbf{t}}^{(i)}(\Delta) \approx L_1^{(i)}(\mathbf{t})\Delta^{2H_1(\mathbf{t})} + L_2^{(i)}(\mathbf{t})\Delta^{2H_2(\mathbf{t})}, \quad i = 1, 2$$

- Assume

$$H_1(\mathbf{t}) = \underline{H}(\mathbf{t}) < \overline{H}(\mathbf{t}) = H_2(\mathbf{t})$$

- For $i = 1, 2$,

$$L_1^{(i)}(\mathbf{t}) \approx \frac{\theta_{\mathbf{t}}^{(i)}(\Delta)}{\Delta^{2H_1(\mathbf{t})}}$$

and

$$L_2^{(i)}(\mathbf{t}) \approx \frac{1}{(4^{D(\mathbf{t})} - 1)\Delta^{2D(\mathbf{t})}} \left| \frac{\theta_{\mathbf{t}}^{(i)}(2\Delta)}{(2\Delta)^{2H_1(\mathbf{t})}} - \frac{\theta_{\mathbf{t}}^{(i)}(\Delta)}{\Delta^{2H_1(\mathbf{t})}} \right|$$

with $D(\mathbf{t}) = H_2(\mathbf{t}) - H_1(\mathbf{t})$

Estimators for $L_1^{(i)}(\mathbf{t})$ and $L_2^{(i)}(\mathbf{t})$

- Plug into the estimating equations for $L_j^{(i)}(\mathbf{t})$ the estimators of the unknown quantities, as defined above
- Special attention requires the case $\underline{H}(\mathbf{t}) = \overline{H}(\mathbf{t})$
 - A diagnostic tool is provided

Non-asymptotic results

Proposition 1 : Constants $C_1, \dots, C_5$ exist such that,

$$\forall \varepsilon, \tau \in (0, 1) \quad \max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq \varepsilon \leq 2\tau,$$

Non-asymptotic results

Proposition 1 : Constants $C_1, \dots, C_5$ exist such that,

$$\forall \varepsilon, \tau \in (0, 1) \quad \max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq \varepsilon \leq 2\tau,$$

$$\mathbb{P} \left[|\hat{\underline{H}}(\mathbf{t}) - \underline{H}(\mathbf{t})| \geq \varepsilon \right] \leq p_1,$$

Non-asymptotic results

Proposition 1 : Constants $C_1, \dots, C_5$ exist such that,

$$\forall \varepsilon, \tau \in (0, 1) \quad \max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq \varepsilon \leq 2\tau,$$

$$\mathbb{P} \left[|\hat{\underline{H}}(\mathbf{t}) - \underline{H}(\mathbf{t})| \geq \varepsilon \right] \leq p_1,$$

and

$$\mathbb{P} \left[\left| \hat{\overline{H}}(\mathbf{t}) - \overline{H}(\mathbf{t}) \right| \geq \varepsilon \right] \leq C_3 \{p_1 + p_2 + p_3\},$$

Non-asymptotic results

Proposition 1 : Constants $C_1, \dots, C_5$ exist such that,

$$\forall \varepsilon, \tau \in (0, 1) \quad \max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq \varepsilon \leq 2\tau,$$

$$\mathbb{P} \left[|\hat{\underline{H}}(\mathbf{t}) - \underline{H}(\mathbf{t})| \geq \varepsilon \right] \leq p_1,$$

and

$$\mathbb{P} \left[\left| \hat{\overline{H}}(\mathbf{t}) - \overline{H}(\mathbf{t}) \right| \geq \varepsilon \right] \leq C_3 \{p_1 + p_2 + p_3\},$$

with

- $p_1 = C_1 \exp \left(-C_2 N \times \varepsilon^2 \times \Delta^{4\underline{H}(\mathbf{t})} \varrho(\Delta, \mathbf{m}) \right),$

Non-asymptotic results

Proposition 1 : Constants $C_1, \dots, C_5$ exist such that,

$$\forall \varepsilon, \tau \in (0, 1) \quad \max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq \varepsilon \leq 2\tau,$$

$$\mathbb{P} \left[|\hat{\underline{H}}(\mathbf{t}) - \underline{H}(\mathbf{t})| \geq \varepsilon \right] \leq p_1,$$

and

$$\mathbb{P} \left[|\hat{\overline{H}}(\mathbf{t}) - \overline{H}(\mathbf{t})| \geq \varepsilon \right] \leq C_3 \{p_1 + p_2 + p_3\},$$

with

- $p_1 = C_1 \exp \left(-C_2 N \times \varepsilon^2 \times \Delta^{4\underline{H}(\mathbf{t})} \varrho(\Delta, \mathbf{m}) \right),$
- $p_2 = \exp \left[-C_4 N \times \varepsilon^2 \times \frac{\Delta^{4\overline{H}(\mathbf{t})} \varrho(\Delta, \mathbf{m})}{\log^2(\Delta)} \Delta^{4D(\mathbf{t})} \right] 1_{\{\underline{H}(\mathbf{t}) < \overline{H}(\mathbf{t})\}},$

Non-asymptotic results

Proposition 1 : Constants $C_1, \dots, C_5$ exist such that,

$$\forall \varepsilon, \tau \in (0, 1) \quad \max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq \varepsilon \leq 2\tau,$$

$$\mathbb{P} \left[|\widehat{\underline{H}}(\mathbf{t}) - \underline{H}(\mathbf{t})| \geq \varepsilon \right] \leq p_1,$$

and

$$\mathbb{P} \left[\left| \widehat{\overline{H}}(\mathbf{t}) - \overline{H}(\mathbf{t}) \right| \geq \varepsilon \right] \leq C_3 \{p_1 + p_2 + p_3\},$$

with

- $p_1 = C_1 \exp \left(-C_2 N \times \varepsilon^2 \times \Delta^{4\underline{H}(\mathbf{t})} \varrho(\Delta, \mathbf{m}) \right),$
- $p_2 = \exp \left[-C_4 N \times \varepsilon^2 \times \frac{\Delta^{4\overline{H}(\mathbf{t})} \varrho(\Delta, \mathbf{m})}{\log^2(\Delta)} \Delta^{4D(\mathbf{t})} \right] 1_{\{\underline{H}(\mathbf{t}) < \overline{H}(\mathbf{t})\}},$
- $p_3 = \exp \left[-C_5 N \times \tau^2 \times \frac{\Delta^{4\overline{H}(\mathbf{t})} \varrho(\Delta, \mathbf{m})}{\log^2(\Delta)} \Delta^{4D(\mathbf{t})} \right],$

where

$$\varrho(\Delta, \mathbf{m}) = \max\{\Delta^{2\underline{H}(\mathbf{t})}, \rho(\mathbf{m})^2\}^{-1}.$$

A risk bound for the anisotropy detection

Proposition 3 : Let

$$A_N(\tau) = \left\{ \widehat{(\overline{H} - \underline{H})}(\mathbf{t}) \geq \tau \right\}.$$

A risk bound for the anisotropy detection

Proposition 3 : Let

$$A_N(\tau) = \left\{ \widehat{(\overline{H} - \underline{H})}(\mathbf{t}) \geq \tau \right\}.$$

If

$$\max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq 2\tau,$$

A risk bound for the anisotropy detection

Proposition 3 : Let

$$A_N(\tau) = \left\{ \widehat{(\overline{H} - \underline{H})}(\mathbf{t}) \geq \tau \right\}.$$

If

$$\max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq 2\tau,$$

and

$$2\tau \leq \{\overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t})\} + 1_{\{\underline{H}(\mathbf{t}) = \overline{H}(\mathbf{t})\}}.$$

A risk bound for the anisotropy detection

Proposition 3 : Let

$$A_N(\tau) = \left\{ \widehat{(\overline{H} - \underline{H})}(\mathbf{t}) \geq \tau \right\}.$$

If

$$\max\{|\log(\Delta)| |R(\underline{H})(\mathbf{t})|, |R(\overline{H} - \underline{H})(\mathbf{t})|\} \leq 2\tau,$$

and

$$2\tau \leq \{\overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t})\} + 1_{\{\underline{H}(\mathbf{t}) = \overline{H}(\mathbf{t})\}}.$$

Then

$$\mathbb{P} \left(1_{A_N(\tau)} \neq 1_{\{\underline{H}(\mathbf{t}) < \overline{H}(\mathbf{t})\}} \right) \leq C_3 \exp \left[-C_5 N \times \tau^2 \times \frac{\Delta^{4\overline{H}(\mathbf{t})} \varrho(\Delta, \mathbf{m})}{\log^2(\Delta)} \Delta^{4D(\mathbf{t})} \right],$$

where C_3 and C_5 are the positive constants from Proposition 1.

MfB sheet with domain deformation (1/2)

- Let W be a MfB sheet with Hurst functional parameter η

MfB sheet with domain deformation (1/2)

- Let W be a MfB sheet with Hurst functional parameter η
- Let $A = (A_1, A_2)$ be a continuously differentiable deformation of a domain in the plane, satisfying some mild conditions

MfB sheet with domain deformation (1/2)

- Let W be a MfB sheet with Hurst functional parameter η
- Let $A = (A_1, A_2)$ be a continuously differentiable deformation of a domain in the plane, satisfying some mild conditions
- Let

$$X = W \circ A$$

MfB sheet with domain deformation (1/2)

- Let W be a MfB sheet with Hurst functional parameter η
- Let $A = (A_1, A_2)$ be a continuously differentiable deformation of a domain in the plane, satisfying some mild conditions
- Let

$$X = W \circ A$$

- Then for $\mathbf{t}, \mathbf{s} \in \mathcal{T}$, we have

$$\begin{aligned}\theta(\mathbf{t}, \mathbf{s}) &= \mathbb{E} [\{X(\mathbf{t}) - X(\mathbf{s})\}^2] \\ &\approx |A_1(\mathbf{t})|^{2H_1(\mathbf{t})} |\partial_1 A_2(\mathbf{t})(t_1 - s_1) + \partial_2 A_2(\mathbf{t})(t_2 - s_2)|^{2H_2(\mathbf{t})} \\ &\quad + |A_2(\mathbf{t})|^{2H_2(\mathbf{t})} |\partial_1 A_1(\mathbf{t})(t_1 - s_1) + \partial_2 A_1(\mathbf{t})(t_2 - s_2)|^{2H_1(\mathbf{t})},\end{aligned}$$

where

$$H_1 = \eta_1 \circ A \quad \text{and} \quad H_2 = \eta_2 \circ A.$$

MfB sheet with domain deformation (2/2)

- Assume that there exist $\rho \in (0, 1)$ such that

$$0 \leq \overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t}) \leq \frac{1 - \rho}{2}$$

MfB sheet with domain deformation (2/2)

- Assume that there exist $\rho \in (0, 1)$ such that

$$0 \leq \overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t}) \leq \frac{1 - \rho}{2}$$

- Then $X = W \circ A \in \mathcal{H}^{H_1, H_2}(\mathbf{L}, \mathcal{T})$, with $\mathbf{L}$ given by :

MfB sheet with domain deformation (2/2)

- Assume that there exist $\rho \in (0, 1)$ such that

$$0 \leq \overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t}) \leq \frac{1 - \rho}{2}$$

- Then $X = W \circ A \in \mathcal{H}^{H_1, H_2}(\mathbf{L}, \mathcal{T})$, with $\mathbf{L}$ given by :

$$L_1^{(1)}(\mathbf{t}) = |A_2(\mathbf{t})|^{2H_2(\mathbf{t})} |\partial_1 A_1(\mathbf{t})|^{2H_1(\mathbf{t})},$$

$$L_2^{(1)}(\mathbf{t}) = |A_1(\mathbf{t})|^{2H_1(\mathbf{t})} |\partial_1 A_2(\mathbf{t})|^{2H_2(\mathbf{t})},$$

$$L_1^{(2)}(\mathbf{t}) = |A_2(\mathbf{t})|^{2H_2(\mathbf{t})} |\partial_2 A_1(\mathbf{t})|^{2H_1(\mathbf{t})},$$

$$L_2^{(2)}(\mathbf{t}) = |A_1(\mathbf{t})|^{2H_1(\mathbf{t})} |\partial_2 A_2(\mathbf{t})|^{2H_2(\mathbf{t})}$$

MfB sheet with domain deformation (2/2)

- Assume that there exist $\rho \in (0, 1)$ such that

$$0 \leq \overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t}) \leq \frac{1 - \rho}{2}$$

- Then $X = W \circ A \in \mathcal{H}^{H_1, H_2}(\mathbf{L}, \mathcal{T})$, with $\mathbf{L}$ given by :

$$L_1^{(1)}(\mathbf{t}) = |A_2(\mathbf{t})|^{2H_2(\mathbf{t})} |\partial_1 A_1(\mathbf{t})|^{2H_1(\mathbf{t})},$$

$$L_2^{(1)}(\mathbf{t}) = |A_1(\mathbf{t})|^{2H_1(\mathbf{t})} |\partial_1 A_2(\mathbf{t})|^{2H_2(\mathbf{t})},$$

$$L_1^{(2)}(\mathbf{t}) = |A_2(\mathbf{t})|^{2H_2(\mathbf{t})} |\partial_2 A_1(\mathbf{t})|^{2H_1(\mathbf{t})},$$

$$L_2^{(2)}(\mathbf{t}) = |A_1(\mathbf{t})|^{2H_1(\mathbf{t})} |\partial_2 A_2(\mathbf{t})|^{2H_2(\mathbf{t})}$$

- Deduce estimating equations for the components of the deformation, depending on $H_1, H_2, \mathbf{L}$ and the variance of X

MfB sheet with domain deformation (2/2)

- Assume that there exist $\rho \in (0, 1)$ such that

$$0 \leq \overline{H}(\mathbf{t}) - \underline{H}(\mathbf{t}) \leq \frac{1 - \rho}{2}$$

- Then $X = W \circ A \in \mathcal{H}^{H_1, H_2}(\mathbf{L}, \mathcal{T})$, with $\mathbf{L}$ given by :

$$L_1^{(1)}(\mathbf{t}) = |A_2(\mathbf{t})|^{2H_2(\mathbf{t})} |\partial_1 A_1(\mathbf{t})|^{2H_1(\mathbf{t})},$$

$$L_2^{(1)}(\mathbf{t}) = |A_1(\mathbf{t})|^{2H_1(\mathbf{t})} |\partial_1 A_2(\mathbf{t})|^{2H_2(\mathbf{t})},$$

$$L_1^{(2)}(\mathbf{t}) = |A_2(\mathbf{t})|^{2H_2(\mathbf{t})} |\partial_2 A_1(\mathbf{t})|^{2H_1(\mathbf{t})},$$

$$L_2^{(2)}(\mathbf{t}) = |A_1(\mathbf{t})|^{2H_1(\mathbf{t})} |\partial_2 A_2(\mathbf{t})|^{2H_2(\mathbf{t})}$$

- Deduce estimating equations for the components of the deformation, depending on $H_1, H_2, \mathbf{L}$ and the variance of X
- Estimates of A are easily obtained by plug-in

Take away

- Functional data are noisy, discretely observed realizations of a stochastic process

Take away

- Functional data are noisy, discretely observed realizations of a stochastic process
- A general class of stochastic processes defined on the plane is introduced ; the MfB sheet is an example

Take away

- Functional data are noisy, discretely observed realizations of a stochastic process
- A general class of stochastic processes defined on the plane is introduced ; the MfB sheet is an example
- The characteristics of the process are estimated nonparametrically, exploiting the replication feature of the functional data

Take away

- Functional data are noisy, discretely observed realizations of a stochastic process
- A general class of stochastic processes defined on the plane is introduced ; the MfB sheet is an example
- The characteristics of the process are estimated nonparametrically, exploiting the replication feature of the functional data
- Non-asymptotic exponential bounds of the estimators are derived

Take away

- Functional data are noisy, discretely observed realizations of a stochastic process
- A general class of stochastic processes defined on the plane is introduced ; the MfB sheet is an example
- The characteristics of the process are estimated nonparametrically, exploiting the replication feature of the functional data
- Non-asymptotic exponential bounds of the estimators are derived
- Two applications are proposed
 - Multifractional Brownian sheet with domain deformation
 - Optimal smoothing for reconstructing the sheets

QR code to the paper on arxiv



THANK YOU