Directional Regularity: Achieving faster rates of convergence in multivariate functional data

Omar Kassi, Sunny Wang
CREST-ENSAI

Brown Bag Seminar
Rennes

21st March, 2024

Outline

Recap on Functional Data

- The observation unit (entity), the datum, could be one or several curves, image(s), or several such objects

Recap on Functional Data

- The observation unit (entity), the datum, could be one or several curves, image(s), or several such objects
- *Related fields*: Signal Processing, Longitudinal Data...

Recap on Functional Data

- The observation unit (entity), the datum, could be one or several curves, image(s), or several such objects
- *Related fields*: Signal Processing, Longitudinal Data...
- Data are (in)dependent realizations of some variable

$$X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{F})$$

Recap on Functional Data

- The observation unit (entity), the datum, could be one or several curves, image(s), or several such objects
- *Related fields*: Signal Processing, Longitudinal Data...
- Data are (in)dependent realizations of some variable

$$X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{F})$$

- When $\mathcal{X}$ is a space of curves/images/signals
  - Functional Data problem

Recap on Functional Data

- The observation unit (entity), the datum, could be one or several curves, image(s), or several such objects
- *Related fields*: Signal Processing, Longitudinal Data...
- Data are (in)dependent realizations of some variable

$$X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{F})$$

- When $\mathcal{X}$ is a space of curves/images/signals
  - Functional Data problem
- Functional Data Analysis (FDA) deals with the statistical description and modeling of samples of random variable taking values in spaces of functions

Regularity in the univariate case (1/2)

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

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

Regularity in the 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 :

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

Regularity in the univariate case (2/2)

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

Regularity in the 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

Regularity in the 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 :

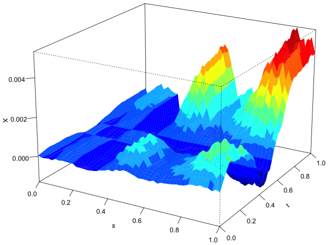
$$H(t_0) \approx \frac{\log(\theta(t_1,t_2)) - \log(\theta(t_1,t_3))}{2\log(2)}, \qquad t_0 \in [t_1,t_2] \subset [t_1,t_3]$$

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 functional data

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



Detour to non-parametric regression

- Let $(X_i, Y_i), i = 1, \ldots, n$ be data pairs observed under the model

$$Y_i = f(X_i) + \varepsilon_i, \qquad i = 1, \ldots, n,$$

where $f : [0,1]^d \to \mathbb{R}$ and the X_i's are i.i.d uniformly distributed on the hypercube

Detour to non-parametric regression

- Let $(X_i, Y_i), i = 1, \ldots, n$ be data pairs observed under the model

$$Y_i = f(X_i) + \varepsilon_i, \qquad i = 1, \ldots, n,$$

where $f : [0,1]^d \to \mathbb{R}$ and the X_i's are i.i.d uniformly distributed on the hypercube

- If f belongs to the anisotropic Hölder class, the minimax rate of estimation is $n^{-\beta/(2\beta+1)}$, where β is the effective smoothness:

$$\beta^{-1} = \sum_{i=1}^{d} \beta_i^{-1},$$

with β_i is the regularity along dimension $\mathbf{e}_i$

Are you really isotropic?

- Anisotropy / isotropy is often determined within the confines of the canonical basis as it is a notion of smoothness *along a dimension*

Are you really isotropic?

- Anisotropy / isotropy is often determined within the confines of the canonical basis as it is a notion of smoothness *along a dimension*
- But this is not the full story!

Are you really isotropic?

- Anisotropy / isotropy is often determined within the confines of the canonical basis as it is a notion of smoothness *along a dimension*
- But this is not the full story!
- Let $\mathcal{T}$ be an open subset of $\mathbb{R}^2$, and $f : \mathbb{R}^2 \to \mathbb{R}$, and $\{\mathbf{u_1}, \mathbf{u_2}\}$ be an orthonormal basis, where the function f is β_i-Hölder continuous along $\mathbf{u_i}$, for $i = 1, 2$

Are you really isotropic?

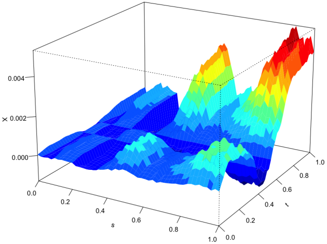
- Anisotropy / isotropy is often determined within the confines of the canonical basis as it is a notion of smoothness *along a dimension*
- But this is not the full story!
- Let $\mathcal{T}$ be an open subset of $\mathbb{R}^2$, and $f : \mathbb{R}^2 \to \mathbb{R}$, and $\{\mathbf{u_1}, \mathbf{u_2}\}$ be an orthonormal basis, where the function f is β_i-Hölder continuous along $\mathbf{u_i}$, for $i = 1, 2$
- Let $\mathbf{v} \in \mathbb{S}$ such that $\mathbf{v} = \alpha_1 \mathbf{u_1} + \alpha_2 \mathbf{u_2}$. Then we have

$$|f(\mathbf{t}) - f(\mathbf{t} - \Delta \mathbf{v})| \leq L_1 |\alpha_1 \Delta|^{\beta_1} + L_2 |\alpha_2 \Delta|^{\beta_2}$$

Are you really isotropic?

- Anisotropy / isotropy is often determined within the confines of the canonical basis as it is a notion of smoothness *along a dimension*
- But this is not the full story!
- Let $\mathcal{T}$ be an open subset of $\mathbb{R}^2$, and $f : \mathbb{R}^2 \to \mathbb{R}$, and $\{\mathbf{u_1}, \mathbf{u_2}\}$ be an orthonormal basis, where the function f is β_i-Hölder continuous along $\mathbf{u_i}$, for $i = 1, 2$
- Let $\mathbf{v} \in \mathbb{S}$ such that $\mathbf{v} = \alpha_1 \mathbf{u_1} + \alpha_2 \mathbf{u_2}$. Then we have

$$|f(\mathbf{t}) - f(\mathbf{t} - \Delta \mathbf{v})| \leq L_1 |\alpha_1 \Delta|^{\beta_1} + L_2 |\alpha_2 \Delta|^{\beta_2}$$



Let's define things

Directional Regularity

Let X be a continuous & non-differentiable stochastic process, $\mathbf{u} \in \mathbb{S}$ a unit vector and $H_{\mathbf{u}} : \mathcal{T} \to (0,1)$. X has local regularity $H_{\mathbf{u}}$ at $\mathbf{t} \in \mathcal{T}$ along the direction $\mathbf{u}$ if $L_u : \mathcal{T} \to \mathbb{R}_+$ exist such that :

$$\theta_{\mathbf{u}}(\mathbf{t}, \Delta) := \mathbb{E}\left[\left\{X\left(\mathbf{t} - \frac{\Delta}{2}\mathbf{u}\right) - X\left(\mathbf{t} + \frac{\Delta}{2}\mathbf{u}\right)\right\}^2\right] = L_{\mathbf{u}}(\mathbf{t})\Delta^{2H_{\mathbf{u}}(\mathbf{t})} + G(\mathbf{t}, \Delta),$$

where $G(\mathbf{t}, \Delta) \underset{\Delta \to 0}{=} o\left(\Delta^{2H_{\mathbf{u}}(\mathbf{t})}\right)$. We call the map $\mathbf{u} \mapsto H_{\mathbf{u}}$ directional regularity.

Let's define things

Directional Regularity

Let X be a continuous & non-differentiable stochastic process, $\mathbf{u} \in \mathbb{S}$ a unit vector and $H_{\mathbf{u}} : \mathcal{T} \to (0,1)$. X has local regularity $H_{\mathbf{u}}$ at $\mathbf{t} \in \mathcal{T}$ along the direction $\mathbf{u}$ if $L_u : \mathcal{T} \to \mathbb{R}_+$ exist such that :

$$\theta_{\mathbf{u}}(\mathbf{t}, \Delta) := \mathbb{E}\left[\left\{X\left(\mathbf{t} - \frac{\Delta}{2}\mathbf{u}\right) - X\left(\mathbf{t} + \frac{\Delta}{2}\mathbf{u}\right)\right\}^2\right] = L_{\mathbf{u}}(\mathbf{t})\Delta^{2H_{\mathbf{u}}(\mathbf{t})} + G(\mathbf{t}, \Delta),$$

where $G(\mathbf{t}, \Delta) \underset{\Delta \to 0}{=} o\left(\Delta^{2H_{\mathbf{u}}(\mathbf{t})}\right)$. We call the map $\mathbf{u} \mapsto H_{\mathbf{u}}$ directional regularity.

- If $H_{\mathbf{u}}$ does not depend on the direction $\mathbf{u}$, we say X is isotropic

Let's define things

Directional Regularity

Let X be a continuous & non-differentiable stochastic process, $\mathbf{u} \in \mathbb{S}$ a unit vector and $H_{\mathbf{u}} : \mathcal{T} \to (0,1)$. X has local regularity $H_{\mathbf{u}}$ at $\mathbf{t} \in \mathcal{T}$ along the direction $\mathbf{u}$ if $L_u : \mathcal{T} \to \mathbb{R}_+$ exist such that :

$$\theta_{\mathbf{u}}(\mathbf{t}, \Delta) := \mathbb{E}\left[\left\{X\left(\mathbf{t} - \frac{\Delta}{2}\mathbf{u}\right) - X\left(\mathbf{t} + \frac{\Delta}{2}\mathbf{u}\right)\right\}^2\right] = L_{\mathbf{u}}(\mathbf{t})\Delta^{2H_{\mathbf{u}}(\mathbf{t})} + G(\mathbf{t}, \Delta),$$

where $G(\mathbf{t}, \Delta) \underset{\Delta \to 0}{=} o\left(\Delta^{2H_{\mathbf{u}}(\mathbf{t})}\right)$. We call the map $\mathbf{u} \mapsto H_{\mathbf{u}}$ directional regularity.

- If $H_{\mathbf{u}}$ does not depend on the direction $\mathbf{u}$, we say X is isotropic
- Otherwise, we call X anisotropic

Let's define things

Directional Regularity

Let X be a continuous & non-differentiable stochastic process, $\mathbf{u} \in \mathbb{S}$ a unit vector and $H_{\mathbf{u}} : \mathcal{T} \to (0,1)$. X has local regularity $H_{\mathbf{u}}$ at $\mathbf{t} \in \mathcal{T}$ along the direction $\mathbf{u}$ if $L_u : \mathcal{T} \to \mathbb{R}_+$ exist such that :

$$\theta_{\mathbf{u}}(\mathbf{t}, \Delta) := \mathbb{E}\left[\left\{X\left(\mathbf{t} - \frac{\Delta}{2}\mathbf{u}\right) - X\left(\mathbf{t} + \frac{\Delta}{2}\mathbf{u}\right)\right\}^2\right] = L_{\mathbf{u}}(\mathbf{t})\Delta^{2H_{\mathbf{u}}(\mathbf{t})} + G(\mathbf{t}, \Delta),$$

where $G(\mathbf{t}, \Delta) \underset{\Delta \to 0}{=} o\left(\Delta^{2H_{\mathbf{u}}(\mathbf{t})}\right)$. We call the map $\mathbf{u} \mapsto H_{\mathbf{u}}$ directional regularity.

- If $H_{\mathbf{u}}$ does not depend on the direction $\mathbf{u}$, we say X is isotropic
- Otherwise, we call X anisotropic
- Anisotropy is not just a notion of smoothness along a dimension, but also along a direction

Example 1: Sum of fBms (1/2)

- Let $\mathbf{u_1}, \mathbf{u_2}$ be two unit vectors that spans $\mathbb{R}^2$, and $H_1 < H_2 \in (0, 1)$.

Example 1: Sum of fBms (1/2)

- Let $\mathbf{u_1}, \mathbf{u_2}$ be two unit vectors that spans $\mathbb{R}^2$, and $H_1 < H_2 \in (0, 1)$.
- Consider
$$X(\mathbf{t}) = B^{H_1}(t_1) + B^{H_2}(t_2), \qquad \forall \mathbf{t} \in \mathcal{T},$$
where (t_1, t_2) are the coordinates of $\mathbf{t}$ in the $(\mathbf{u_1}, \mathbf{u_2})$

Example 1: Sum of fBms (1/2)

- Let $\mathbf{u_1}, \mathbf{u_2}$ be two unit vectors that spans $\mathbb{R}^2$, and $H_1 < H_2 \in (0, 1)$.
- Consider
$$X(\mathbf{t}) = B^{H_1}(t_1) + B^{H_2}(t_2), \qquad \forall \mathbf{t} \in \mathcal{T},$$
where (t_1, t_2) are the coordinates of $\mathbf{t}$ in the $(\mathbf{u_1}, \mathbf{u_2})$
- Independence of B_1 and B_2 implies:
$$\mathbb{E}\left[\{X(\mathbf{t} - (\Delta/2)\mathbf{u_i}) - X(\mathbf{t} + (\Delta/2)\mathbf{u_i})\}^2\right] = \Delta^{2H_i}.$$

Example 1: Sum of fBms (2/2)

- Take again the sum of two independent fBms, where $\mathbf{e_i} = a_{1,i}\mathbf{u_1} + a_{2,i}\mathbf{u_2}$

Example 1: Sum of fBms (2/2)

- Take again the sum of two independent fBms, where $\mathbf{e_i} = a_{1,i}\mathbf{u_1} + a_{2,i}\mathbf{u_2}$
- In the canonical basis, we have instead

$$\mathbb{E}\left[\left\{X(\mathbf{t} - (\Delta/2)\mathbf{e_i}) - X(\mathbf{t} + (\Delta/2)\mathbf{e_i})\right\}^2\right] = \left|a_{1,i}\Delta\right|^{2H_1} + \left|a_{2,i}\Delta\right|^{2H_2}.$$

Example 1: Sum of fBms (2/2)

- Take again the sum of two independent fBms, where $\mathbf{e_i} = a_{1,i}\mathbf{u_1} + a_{2,i}\mathbf{u_2}$
- In the canonical basis, we have instead

$$\mathbb{E}\left[\{X(\mathbf{t} - (\Delta/2)\mathbf{e_i}) - X(\mathbf{t} + (\Delta/2)\mathbf{e_i})\}^2\right] = |a_{1,i}\Delta|^{2H_1} + |a_{2,i}\Delta|^{2H_2}.$$

- In the “bad” basis (which in fact is almost all of them), the effective smoothness of X is instead given by $\beta^{-1} = 2H_1^{-1} < H_1^{-1} + H_2^{-1}$

Example 2: Product of fBms

- Define the product of two independent fBms, where $\{\mathbf{u_1}, \mathbf{u_2}\}$ is the orthonormal basis containing the maximising regularity:

$$X(\mathbf{t}) = B^{H_1}(t_1)B^{H_2}(t_2), \qquad \forall \mathbf{t} \in \mathcal{T}.$$

Example 2: Product of fBms

- Define the product of two independent fBms, where $\{\mathbf{u_1}, \mathbf{u_2}\}$ is the orthonormal basis containing the maximising regularity:

$$X(\mathbf{t}) = B^{H_1}(t_1)B^{H_2}(t_2), \qquad \forall \mathbf{t} \in \mathcal{T}.$$

- Independence implies that along the $(\mathbf{u_1}, \mathbf{u_2})$ basis, we have

$$\mathbb{E}\left[\{X(\mathbf{t} - (\Delta/2)\mathbf{u_i}) - X(\mathbf{t} + (\Delta/2)\mathbf{u_i})\}^2\right] = t_j^{2H_j}\Delta^{2H_i}, \qquad j \neq i.$$

Example 2: Product of fBms

- Define the product of two independent fBms, where $\{\mathbf{u_1}, \mathbf{u_2}\}$ is the orthonormal basis containing the maximising regularity:

$$X(\mathbf{t}) = B^{H_1}(t_1)B^{H_2}(t_2), \qquad \forall \mathbf{t} \in \mathcal{T}.$$

- Independence implies that along the $(\mathbf{u_1}, \mathbf{u_2})$ basis, we have

$$\mathbb{E}\left[\{X(\mathbf{t} - (\Delta/2)\mathbf{u_i}) - X(\mathbf{t} + (\Delta/2)\mathbf{u_i})\}^2\right] = t_j^{2H_j}\Delta^{2H_i}, \qquad j \neq i.$$

- In contrast, along the canonical basis we have instead

$$\theta_{\mathbf{e_i}}(\mathbf{t}, \Delta) \approx t_j^{2H_j}|a_{1,i}\Delta|^{2H_i} + t_i^{2H_i}|a_{1,j}\Delta|^{2H_j}, \qquad j \neq i.$$

- Anisotropy of X is not invariant to the choice of basis!

The key lemma

Lemma

Let $(\mathbf{u_1}, \mathbf{u_2}) \in \mathbb{S}$ span $\mathbb{R}^2$ such that $H_{\mathbf{u_1}} < H_{\mathbf{u_2}}$. Suppose $L_{\mathbf{u_1}}$ and $L_{\mathbf{u_2}}$ are continuously differentiable. For any $\mathbf{v} \in \mathbb{S}$, we have the following dichotomy:

- If $\mathbf{v} \neq \pm\mathbf{u_2}$, then the regularity of X along $\mathbf{v}$ is $H_{\mathbf{u_1}}$.
- Otherwise, the local regularity along $\mathbf{v}$ is $H_{\mathbf{u_2}}$.

The key lemma

Lemma

Let $(\mathbf{u_1}, \mathbf{u_2}) \in \mathbb{S}$ span $\mathbb{R}^2$ such that $H_{\mathbf{u_1}} < H_{\mathbf{u_2}}$. Suppose $L_{\mathbf{u_1}}$ and $L_{\mathbf{u_2}}$ are continuously differentiable. For any $\mathbf{v} \in \mathbb{S}$, we have the following dichotomy:

- If $\mathbf{v} \neq \pm\mathbf{u_2}$, then the regularity of X along $\mathbf{v}$ is $H_{\mathbf{u_1}}$.
- Otherwise, the local regularity along $\mathbf{v}$ is $H_{\mathbf{u_2}}$.

- Map $\mathbf{v} \mapsto H_{\mathbf{v}}$ can only take at most two possible values

The key lemma

Lemma

Let $(\mathbf{u_1}, \mathbf{u_2}) \in \mathbb{S}$ span $\mathbb{R}^2$ such that $H_{\mathbf{u_1}} < H_{\mathbf{u_2}}$. Suppose $L_{\mathbf{u_1}}$ and $L_{\mathbf{u_2}}$ are continuously differentiable. For any $\mathbf{v} \in \mathbb{S}$, we have the following dichotomy:

- If $\mathbf{v} \neq \pm \mathbf{u_2}$, then the regularity of X along $\mathbf{v}$ is $H_{\mathbf{u_1}}$.
- Otherwise, the local regularity along $\mathbf{v}$ is $H_{\mathbf{u_2}}$.

- Map $\mathbf{v} \mapsto H_{\mathbf{v}}$ can only take at most two possible values
- Maximisation problem $\arg\max_{v \in \mathbb{S}} H_{\mathbf{v}}$ admits two solutions $\mathbf{u_2}$ and $-\mathbf{u_2}$

The key lemma

Lemma

Let $(\mathbf{u_1}, \mathbf{u_2}) \in \mathbb{S}$ span $\mathbb{R}^2$ such that $H_{\mathbf{u_1}} < H_{\mathbf{u_2}}$. Suppose $L_{\mathbf{u_1}}$ and $L_{\mathbf{u_2}}$ are continuously differentiable. For any $\mathbf{v} \in \mathbb{S}$, we have the following dichotomy:

- If $\mathbf{v} \neq \pm\mathbf{u_2}$, then the regularity of X along $\mathbf{v}$ is $H_{\mathbf{u_1}}$.
- Otherwise, the local regularity along $\mathbf{v}$ is $H_{\mathbf{u_2}}$.

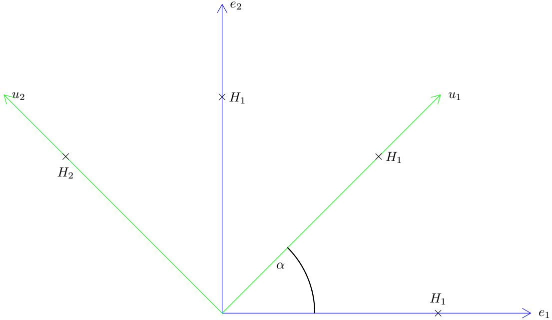
- Map $\mathbf{v} \mapsto H_{\mathbf{v}}$ can only take at most two possible values
- Maximisation problem $\arg\max_{v \in \mathbb{S}} H_{\mathbf{v}}$ admits two solutions $\mathbf{u_2}$ and $-\mathbf{u_2}$
- Finding the maximising direction $\mathbf{u_2}$ is equivalent to finding the angle $\alpha \in [0, \pi)$ between the two basis vectors $\mathbf{e_1}$ and $\mathbf{u_1}$:

$$\arg\max_{\mathbf{v} \in \mathbb{S}} H_{\mathbf{v}} = \arg\max_{\alpha \in [0,\pi)} H_{\mathbf{u}(\alpha)},$$

where $\mathbf{u}(\alpha) = \cos(\alpha)\mathbf{e_1} + \sin(\alpha)\mathbf{e_2}$.

The picture that says it all

Illustration of directional regularity



How to estimate α?

- Let H_1, H_2 denote the regularity of X along $\mathbf{u_1}$ and $\mathbf{u_2}$ respectively, where $\langle \mathbf{e_1}, \mathbf{u_1} \rangle = \cos(\alpha)$

Proposition

Suppose that $\mathbf{u}_1 \neq \pm \mathbf{e}_i$, for $i = 1, 2$. Then for a process X satisfying (8) and any fixed point $\mathbf{t} \in \mathcal{T}$, we have

$$|g(\alpha)| = \left(\frac{\theta_{\mathbf{e_2}}(\mathbf{t}, \Delta)}{\theta_{\mathbf{e_1}}(\mathbf{t}, \Delta)} \right)^{\frac{1}{2\underline{H}}} + O\left(\Delta^{\widetilde{\beta} \wedge |2H_1 - 2H_2|} \right),$$

where $g = \tan \mathbf{1}\{H_1 < H_2\} + \cot \mathbf{1}\{H_1 > H_2\}$, and $\underline{H} = \min\{H_1, H_2\}$.

- Angles can be computed, up to a reflection, by taking the ratios of mean-squared variations along the canonical basis

Data setting

- Observe an independent sample of random functions $X^{(1)}, \dots, X^{(N)}$ defined on $\mathcal{T}$, where $\mathcal{T}$ is an open subset of $\mathbb{R}_+^d$.
- Suppose that observations come in the form of $(Y^{(j)}(\mathbf{t}_m), \mathbf{t}_m)$, generated from

$$Y^{(j)}(\mathbf{t}_m) = X^{(j)}(\mathbf{t}_m) + \varepsilon_m^{(j)}, \qquad 1 \leq j \leq N, 1 \leq m \leq M_0, \mathbf{t}_m \in \mathcal{T},$$

where the homoscedastic errors are independent, centered random variables.

Plug-in estimators

- Natural plug-in estimator is

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

where $\widetilde{X}^{(j)}$ denotes the interpolation of $X^{(j)}$.

Plug-in estimators

- Natural plug-in estimator is

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

where $\widetilde{X}^{(j)}$ denotes the interpolation of $X^{(j)}$.

- Regularity $\underline{H}$ can be estimated with

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

Plug-in estimators

- Natural plug-in estimator is

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

where $\widetilde{X}^{(j)}$ denotes the interpolation of $X^{(j)}$.

- Regularity $\underline{H}$ can be estimated with

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

- Putting (1) and (2) together:

$$g^{-1}\widehat{|g(\alpha)|} = g^{-1}\left(\frac{\widehat{\theta}_{\mathbf{e_2}}(\mathbf{t}, \Delta)}{\widehat{\theta}_{\mathbf{e_1}}(\mathbf{t}, \Delta)}\right)^{\frac{1}{2\widehat{\underline{H}}}}.$$

Estimation Algorithm I

- Define $\texttt{mean}(A) := (\#\mathcal{T}_o)^{-1} \sum_{\mathbf{t} \in \mathcal{T}_o} A(\mathbf{t})$.

Require: Data $Y_i(\mathbf{t})$, Evaluation points $\mathcal{T}_o = \{\mathbf{t_1}, \ldots, \mathbf{t_k}\}$

Initialise $\widehat{\theta}_{\mathbf{e_1}}(\mathcal{T}_o) \leftarrow \emptyset$, $\widehat{\theta}_{\mathbf{e_2}}(\mathcal{T}_o) \leftarrow \emptyset$, $\underline{\widehat{H}}(\mathcal{T}_o) \leftarrow \emptyset$

for $\mathbf{t}$ $\texttt{in}$ $\mathcal{T}_o$ **do**

$\widehat{\theta}_{\mathbf{e_1}}(\mathbf{t}, \Delta) \leftarrow N^{-1} \sum_{j=1}^{N} \left\{ \widetilde{X}^{(j)}(\mathbf{t} - (\Delta/2)\mathbf{e_1}) - \widetilde{X}^{(j)}(\mathbf{t} + (\Delta/2)\mathbf{e_1}) \right\}^2$ ▷ fixed Δ

$\widehat{\theta}_{\mathbf{e_2}}(\mathbf{t}, \Delta) \leftarrow N^{-1} \sum_{j=1}^{N} \left\{ \widetilde{X}^{(j)}(\mathbf{t} - (\Delta/2)\mathbf{e_2}) - \widetilde{X}^{(j)}(\mathbf{t} + (\Delta/2)\mathbf{e_2}) \right\}^2$

if $\widehat{\theta}_{\mathbf{e_i}}(\mathbf{t}, \Delta) > 0$ **and** $\widehat{\theta}_{\mathbf{e_i}}(\mathbf{t}, \Delta) > 0$ **then**

$\underline{\widehat{H}}(\mathbf{t}, \Delta) \leftarrow \min_{i=1,2} \left\{ \left(\log(\widehat{\theta}_{\mathbf{e_i}}(\mathbf{t}, \mathbf{\Delta})) - \log(\widehat{\theta}_{\mathbf{e_i}}(\mathbf{t}, \mathbf{\Delta})) \right) / (2\log(2)) \right\}$

else

$\underline{\widehat{H}}(\mathbf{t}, \Delta) \leftarrow 1$

end if

Estimation Algorithm II

$\widehat{\theta}_{\mathbf{e_1}}(\mathcal{T}_o) \leftarrow \widehat{\theta}_{\mathbf{e_1}}(\mathcal{T}_o) \cup \widehat{\theta}_{\mathbf{e_1}}(\mathbf{t}, \boldsymbol{\Delta})$
$\widehat{\theta}_{\mathbf{e_2}}(\mathcal{T}_o) \leftarrow \widehat{\theta}_{\mathbf{e_2}}(\mathcal{T}_o) \cup \widehat{\theta}_{\mathbf{e_2}}(\mathbf{t}, \boldsymbol{\Delta})$
$\underline{\widehat{H}}(\mathcal{T}_o) \leftarrow \underline{\widehat{H}}(\mathcal{T}_o) \cup \underline{\widehat{H}}(\mathbf{t}, \Delta)$

end for

$\widehat{g(\alpha)} \leftarrow \left(\mathtt{mean}(\widehat{\theta}_{\mathbf{e_1}}(\mathcal{T}_o))/\mathtt{mean}(\widehat{\theta}_{\mathbf{e_2}}(\mathcal{T}_o))\right)^{1/(2*\mathtt{mean}(\underline{\widehat{H}})(\mathcal{T}_o))}$

$\widehat{\alpha}^{tan} \leftarrow \arctan \widehat{g(\alpha)}$

$\widehat{\alpha}^{cot} \leftarrow \text{arccot}\widehat{g(\alpha)}$

return $\widehat{\alpha}^{tan}, \widehat{\alpha}^{cot}$

Identification issues

- Two identification issues are present in (16)
- First is associated to g:

$$g = \tan \mathbf{1}\{H_1 < H_2\} + \cot \mathbf{1}\{H_1 > H_2\}.$$

Identification issues

- Two identification issues are present in (16)
- First is associated to g:

$$g = \tan \mathbf{1}\{H_1 < H_2\} + \cot \mathbf{1}\{H_1 > H_2\}.$$

- Second arises from the absolute value: we either have $tan(\alpha)$ or $tan(\pi - \alpha)$ if $g = \tan$, and similarly for the $\cot$ case

Identification issues

- Two identification issues are present in (16)
- First is associated to g:

$$g = \tan \mathbf{1}\{H_1 < H_2\} + \cot \mathbf{1}\{H_1 > H_2\}.$$

- Second arises from the absolute value: we either have $tan(\alpha)$ or $tan(\pi - \alpha)$ if $g = \tan$, and similarly for the $\cot$ case
- Basically we need to identify a unique angle amongst the four possible options

Resolving the identification problem

- Let's not forget what the angle gives us: the direction of the maximising regularity!

Resolving the identification problem

- Let's not forget what the angle gives us: the direction of the maximising regularity!
- Any unit vector $\mathbf{u} \in \mathbb{S}$ can be represented in the canonical basis:

$$\mathbf{u}(\beta) = \cos(\beta)\mathbf{e_1} + \sin(\beta)\mathbf{e_2}.$$

- Correct α between $\mathbf{u_1}$ and $\mathbf{e_1}$ is thus given by

$$\alpha = \arg\max_{\beta \in \{\gamma, \pi-\gamma, \pi/2-\gamma, \pi/2+\gamma\}} H_{\mathbf{u}(\beta)},$$

where $\gamma \approx$ arccot $\left((\theta_{\mathbf{e_1}}(t,\Delta)/\theta_{\mathbf{e_2}}(t,\Delta))^{1/(2\underline{H})} \right)$.

Regularity estimator

- Use the following noise-adapted estimator:

$$\widehat{H}_{\mathbf{v}} = \begin{cases} \frac{\log(\widehat{\theta}_{\mathbf{v}}(\mathbf{t},2\Delta)-2\widehat{\sigma}^2)-\log(\widehat{\theta}_{\mathbf{v}}(\mathbf{t},\Delta)-2\widehat{\sigma}^2)}{2\log(2)} & \text{if} \quad \widehat{\theta}_{\mathbf{v}}(\mathbf{t},2\Delta), \widehat{\theta}_{\mathbf{v}}(\mathbf{t},\Delta) > 2\widehat{\sigma}^2, \\ 1 & \text{otherwise.} \end{cases}$$

Regularity estimator

- Use the following noise-adapted estimator:

$$\widehat{H}_{\mathbf{v}} = \begin{cases} \frac{\log(\widehat{\theta}_{\mathbf{v}}(\mathbf{t},2\Delta)-2\widehat{\sigma}^2)-\log(\widehat{\theta}_{\mathbf{v}}(\mathbf{t},\Delta)-2\widehat{\sigma}^2)}{2\log(2)} & \text{if} \quad \widehat{\theta}_{\mathbf{v}}(\mathbf{t},2\Delta), \widehat{\theta}_{\mathbf{v}}(\mathbf{t},\Delta) > 2\widehat{\sigma}^2, \\ 1 & \text{otherwise.} \end{cases}$$

- Noise estimator is given by

$$\widehat{\sigma}_m^2 = \frac{1}{2N}\sum_{j=1}^{N}\left(Y^{(j)}(\mathbf{t}_m) - Y^{(j)}(\mathbf{t}_{m,1})\right)^2,$$

with $\mathbf{t}_{m,1}$ denoting the closest observed point to $\mathbf{t_m}$.

Regularity estimator

- Use the following noise-adapted estimator:

$$\widehat{H}_{\mathbf{v}} = \begin{cases} \frac{\log(\widehat{\theta}_{\mathbf{v}}(\mathbf{t},2\Delta)-2\widehat{\sigma}^2)-\log(\widehat{\theta}_{\mathbf{v}}(\mathbf{t},\Delta)-2\widehat{\sigma}^2)}{2\log(2)} & \text{if} \quad \widehat{\theta}_{\mathbf{v}}(\mathbf{t},2\Delta), \widehat{\theta}_{\mathbf{v}}(\mathbf{t},\Delta) > 2\widehat{\sigma}^2, \\ 1 & \text{otherwise.} \end{cases}$$

- Noise estimator is given by

$$\widehat{\sigma}_m^2 = \frac{1}{2N}\sum_{j=1}^{N}\left(Y^{(j)}(\mathbf{t}_m) - Y^{(j)}(\mathbf{t}_{m,1})\right)^2,$$

with $\mathbf{t}_{m,1}$ denoting the closest observed point to $\mathbf{t}_\mathbf{m}$.

- Compute $\widehat{H}_{\mathbf{v}}$ on a grid of spacings $\boldsymbol{\Delta}$:

$$\widehat{\alpha} = \arg\max_{\beta\in\{\widehat{\gamma},\pi-\widehat{\gamma},\pi/2-\widehat{\gamma},\pi/2+\widehat{\gamma}\}} \sum_{i=1}^{p} \widehat{H}_{\mathbf{u}(\beta)}(\Delta_i),$$

where $\widehat{\gamma} = \text{arccot}\left((\widehat{\theta}_{\mathbf{e_1}}(t,\Delta)/\widehat{\theta}_{\mathbf{e_2}}(t,\Delta))^{1/(2\underline{\widehat{H}}(\Delta))}\right)$.

Identification Algorithm

Require: $\widehat{\alpha}^{tan}$, $\widehat{\alpha}^{\cot}$, $Y_i(\mathbf{t})$, $\boldsymbol{\Delta}$, $\mathcal{T}_o$, $\widehat{\sigma}^2$

Identification Algorithm

Require: $\widehat{\alpha}^{tan}$, $\widehat{\alpha}^{\cot}$, $Y_i(\mathbf{t})$, $\mathbf{\Delta}$, $\mathcal{T}_o$, $\widehat{\sigma}^2$

Initialise $\widehat{H}_{\mathbf{v}(\beta)}(\Delta) \leftarrow \emptyset$, $\widehat{H}_{\mathbf{v}(\beta)} \leftarrow \emptyset$ $\triangleright\ \beta \in \{\widehat{\alpha}^{tan}, \widehat{\alpha}^{cot}, \pi - \widehat{\alpha}^{tan}, \pi - \widehat{\alpha}^{cot}\}$

Identification Algorithm

Require: $\widehat{\alpha}^{tan}$, $\widehat{\alpha}^{\cot}$, $Y_i(\mathbf{t})$, $\mathbf{\Delta}$, $\mathcal{T}_o$, $\widehat{\sigma}^2$

Initialise $\widehat{H}_{\mathbf{v}(\beta)}(\Delta) \leftarrow \emptyset$, $\widehat{H}_{\mathbf{v}(\beta)} \leftarrow \emptyset$ ▷ $\beta \in \{\widehat{\alpha}^{tan}, \widehat{\alpha}^{cot}, \pi - \widehat{\alpha}^{tan}, \pi - \widehat{\alpha}^{cot}\}$

$\mathbf{v}(\beta) \leftarrow (\cos(\beta), \sin(\beta))^\top$

for Δ `in` $\mathbf{\Delta}$ **do**

  for $\mathbf{t}$ `in` $\mathcal{T}_o$ **do**

    if $\widehat{\theta}_{\mathbf{v}(\beta)}(\mathbf{t}, 2\Delta) > 2\widehat{\sigma}^2$ `and` $\widehat{\theta}_{\mathbf{v}(\beta)}(\mathbf{t}, \Delta) > 2\widehat{\sigma}^2$ **then**

$$\widehat{H}_{\mathbf{v}(\beta)}(\mathbf{t}, \Delta) \leftarrow \left(\log(\widehat{\theta}_{\mathbf{v}(\beta)}(\mathbf{t}, \mathbf{2\Delta}) - 2\widehat{\sigma}^2) - \log(\widehat{\theta}_{\mathbf{v}(\beta)}(\mathbf{t}, \mathbf{\Delta}) - 2\widehat{\sigma}^2)\right) / (2\log(2))$$

    else

      $\widehat{H}_{\mathbf{v}(\beta)}(\mathbf{t}, \Delta) \leftarrow 1$

    end if

    $\widehat{H}_{\mathbf{v}(\beta)}(\Delta) \leftarrow \widehat{H}_{\mathbf{v}(\beta)}(\Delta) \cup \widehat{H}_{\mathbf{v}(\beta)}(\mathbf{t}, \Delta)$ ▷ $\widehat{H}$ now on grid of $\mathbf{t}$'s

  end for

  $\widehat{H}_{\mathbf{v}(\beta)} \leftarrow \widehat{H}_{\mathbf{v}(\beta)} \cup$ `mean`$(\widehat{H}_{\mathbf{v}(\beta)}(\Delta))$

end for

$\widehat{\alpha} \leftarrow \arg\max_\beta \sum_{\Delta \in \mathbf{\Delta}} \widehat{H}_{\mathbf{v}(\beta)}$

return $\widehat{\alpha}$

Theory - Assumptions

Assumptions.

- (H1) Let X be anisotropic process with the two regularities (H_1, H_2), and let $X^{(j)}$, $1 \leq j \leq N$, be independent realizations of X.

Theory - Assumptions

Assumptions.

- (H1) Let X be anisotropic process with the two regularities (H_1, H_2), and let $X^{(j)}$, $1 \leq j \leq N$, be independent realizations of X.

- (H2) Three positive constants $\mathfrak{a}$, $\mathfrak{A}$ and r exist such that, for any $\mathbf{t} \in \mathcal{T}$,

$$\mathbb{E}\left|X^{(j)}(\mathbf{t}) - X^{(j)}(\mathbf{s})\right|^{2p} \leq \frac{p!}{2}\mathfrak{a}\mathfrak{A}^{p-2}\|\mathbf{t} - \mathbf{s}\|^{2p\underline{H}(\mathbf{t})} \qquad \forall \mathbf{s} \in B(\mathbf{t}; r), \ \forall p \geq 1.$$

Theory - Assumptions

Assumptions.

- H1 Let X be anisotropic process with the two regularities (H_1, H_2), and let $X^{(j)}$, $1 \leq j \leq N$, be independent realizations of X.

- H2 Three positive constants $\mathfrak{a}$, $\mathfrak{A}$ and r exist such that, for any $\mathbf{t} \in \mathcal{T}$,

$$\mathbb{E}\left|X^{(j)}(\mathbf{t}) - X^{(j)}(\mathbf{s})\right|^{2p} \leq \frac{p!}{2}\mathfrak{a}\mathfrak{A}^{p-2}\|\mathbf{t}-\mathbf{s}\|^{2p\underline{H}(\mathbf{t})} \qquad \forall \mathbf{s} \in B(\mathbf{t}; r),\ \forall p \geq 1.$$

- H3 A constant $\mathfrak{G}$ exists such that

$$\mathbb{E}(\varepsilon^{2p}) \leq \frac{p!}{2}\mathfrak{G}^{p-2}\sigma^2, \qquad \forall p \geq 1. \tag{3}$$

Theoretical Properties

Theorem

Suppose that assumptions H1-H3 are satisfied. Then, three positive constants C_1, C_2 *and* $\mathfrak{u}$ *exist such that for any*

$$1 \geq \varepsilon \geq \mathfrak{u} \max\{\mathfrak{m}^{-2\underline{H}}, \Delta^{\widetilde{\beta} \wedge |2H_1 - 2H_2|}\},$$

$$\mathbb{P}\left(|\widehat{g(\alpha, \Delta)} - g(\alpha, \Delta)| \geq \varepsilon\right) \leq C_1 \exp\left(-C_2 \varepsilon^2 N \frac{\Delta^{6\underline{H}}}{\log^2(\Delta)}\right).$$

where g *is defined in Proposition 14.*

Theoretical Properties

Theorem

Suppose that assumptions H1-H3 are satisfied. Then, three positive constants C_1, C_2 and $\mathfrak{u}$ exist such that for any

$$1 \geq \varepsilon \geq \mathfrak{u} \max\{\mathfrak{m}^{-2\underline{H}}, \Delta^{\widetilde{\beta} \wedge |2H_1 - 2H_2|}\},$$

$$\mathbb{P}\left(|\widehat{g(\alpha, \Delta)} - g(\alpha, \Delta)| \geq \varepsilon\right) \leq C_1 \exp\left(-C_2 \varepsilon^2 N \frac{\Delta^{6\underline{H}}}{\log^2(\Delta)}\right).$$

where g is defined in Proposition 14.

Corollary

The following rates of convergence hold for $\widehat{\alpha}$:

$$|\widehat{\alpha}(\Delta) - \alpha| = O_{\mathbb{P}}\left(\max\left\{\frac{\#\boldsymbol{\Delta}}{\min\{\sqrt{N}, \mathfrak{m}^{\underline{H}}\}}, \frac{|\log \Delta|}{\sqrt{N}\Delta^{3\underline{H}}}, \mathfrak{m}^{-\underline{H}}\right\}\right).$$

Computational aspects of directional regularity

- Computational cost of identification dominates, due to the extra estimation of H's on a grid of spacings $\boldsymbol{\Delta}$
- Can restrict our analysis to the identification algorithm
- $O(M_0 \# \mathcal{T}_d)$ for interpolation in each surface, resulting in $O(NM_0 \# \mathcal{T}_d)$ for all surfaces, and thus $O(\# \boldsymbol{\Delta} NM_0 \# \mathcal{T}_d)$ on a grid of spacings

Simulation of Anisotropic Processes

- Need a fast way to simulate anisotropic processes to test our algorithms
- While many algorithms exist for simulation of processes such as fBm, they do not take into account anisotropy
- Based on circulant embedding method of Wood and Chan (1994), and exploiting the self-similarity and stationary increments of fBms

Simulator Idea

- Using basic trigonometry, can represent basis vectors as $\{\mathbf{u_1}, \mathbf{u_2}\}$ in the canonical basis:

$$\begin{aligned} \mathbf{u_1} &= \cos(\alpha)\mathbf{e_1} + \sin(\alpha)\mathbf{e_2}, \\ \mathbf{u_2} &= -\sin(\alpha)\mathbf{e_1} + \cos(\alpha)\mathbf{e_2}. \end{aligned} \tag{4}$$

Simulator Idea

- Using basic trigonometry, can represent basis vectors as $\{\mathbf{u_1}, \mathbf{u_2}\}$ in the canonical basis:

$$\begin{aligned} \mathbf{u_1} &= \cos(\alpha)\mathbf{e_1} + \sin(\alpha)\mathbf{e_2}, \\ \mathbf{u_2} &= -\sin(\alpha)\mathbf{e_1} + \cos(\alpha)\mathbf{e_2}. \end{aligned} \tag{4}$$

- Can try to take an equally spaced grid of points $\{\mathbf{u_{1k}}\}_{1\leq k\leq n}$,$\{\mathbf{u_{2k}}\}_{1\leq k\leq n}$ and simulate the fBms

Simulator Idea

- Using basic trigonometry, can represent basis vectors as $\{\mathbf{u_1}, \mathbf{u_2}\}$ in the canonical basis:

$$\begin{aligned} \mathbf{u_1} &= \cos(\alpha)\mathbf{e_1} + \sin(\alpha)\mathbf{e_2}, \\ \mathbf{u_2} &= -\sin(\alpha)\mathbf{e_1} + \cos(\alpha)\mathbf{e_2}. \end{aligned} \tag{4}$$

- Can try to take an equally spaced grid of points $\{\mathbf{u_{1k}}\}_{1\leq k\leq n}$,$\{\mathbf{u_{2k}}\}_{1\leq k\leq n}$ and simulate the fBms

- But this is not enough, since $-\sin(\alpha)$ can be negative, and $\cos(\alpha) < 0$ for $\alpha \in [\pi/2, 3\pi/2]$, while the fBm has a domain in $\mathbb{R}_+$!

Simulator Idea

- Using basic trigonometry, can represent basis vectors as $\{\mathbf{u_1}, \mathbf{u_2}\}$ in the canonical basis:

$$\begin{aligned} \mathbf{u_1} &= \cos(\alpha)\mathbf{e_1} + \sin(\alpha)\mathbf{e_2}, \\ \mathbf{u_2} &= -\sin(\alpha)\mathbf{e_1} + \cos(\alpha)\mathbf{e_2}. \end{aligned} \tag{4}$$

- Can try to take an equally spaced grid of points $\{\mathbf{u_{1k}}\}_{1\leq k\leq n}$,$\{\mathbf{u_{2k}}\}_{1\leq k\leq n}$ and simulate the fBms

- But this is not enough, since $-\sin(\alpha)$ can be negative, and $\cos(\alpha) < 0$ for $\alpha \in [\pi/2, 3\pi/2]$, while the fBm has a domain in $\mathbb{R}_+$!

- However, we can use the stationary increments to avoid the problem of negative values:

$$B(t) - B(s) \sim B(t-s),$$

  and take for example $t = 0$.

Simulation Algorithm I

Require: $\alpha \in [0, 2\pi]$, $N \in \mathbb{N}$, $H_1, H_2 \in (0,1)$,$n \in \mathbb{N}$, f, $\mathbf{v} \in \{(i/n, j/n)\}_{0\le i,j\le n}$

Initialise $Y(\mathbf{v}) \leftarrow \emptyset$

if $\alpha > \pi$ **then**

$\alpha \leftarrow \alpha - \pi$

end if

$\mathbf{u_1} \leftarrow (\cos(\alpha), \sin(\alpha))^\top$

$\mathbf{u_2} \leftarrow (-\sin(\alpha), \cos(\alpha))^\top$

$\mathbf{t} \leftarrow \{n^{-1}(|\cos(\alpha)| + \sin(\alpha))k\}_{0\le k\le n}$

for i `from` 1 `to` N **do**

$\widetilde{B}_1 \leftarrow \texttt{fbm}(H_1, n, |\cos(\alpha)| + \sin(\alpha))$

$\widetilde{B}_2 \leftarrow \texttt{fbm}(H_2, n, |\cos(\alpha)| + \sin(\alpha))$

if $\alpha \le \pi/2$ **then**

$B_1 \leftarrow \widetilde{B}_1$

$\mathbf{s} \leftarrow \{-\sin(\alpha) + (k/n)(\cos(\alpha) + \sin(\alpha))\}_{0\le k\le n}$

else

$\mathbf{t}^{proj} \leftarrow \{\cos(\alpha) + (k/n)(\sin(\alpha) - \cos(\alpha)\}_{0\le k\le n}$

$B_1^- \leftarrow -\widetilde{B}_1(-\mathbf{t}^{proj}\mathbf{1}\{\mathbf{t_k}^{proj} < 0\})$

$B_1^+ \leftarrow \widetilde{B}_1(\mathbf{t}^{proj}\mathbf{1}\{\mathbf{t_k}^{proj} \ge 0\})$

Simulation Algorithm II

$B_1 \leftarrow B_1^- \cup B_1^+$
$\mathbf{s} \leftarrow \{(\cos(\alpha) + \sin(\alpha)) + (k/n)(\sin(\alpha) - \cos(\alpha)\}_{0 \leq k \leq n}$
end if
$\underline{s}_k \leftarrow \arg\min_{x \in \mathbf{t}} |\mathbf{s_k} - x| \mathbf{1}\{\mathbf{s_k} < 0\}$
$\overline{s}_k \leftarrow \arg\min_{x \in \mathbf{t}} |\mathbf{s_k} - x| \mathbf{1}\{\mathbf{s_k} \geq 0\}$
$B_2^- \leftarrow -\widetilde{B}_2(-\underline{s}_k)$
$B_2^+ \leftarrow \widetilde{B}_2(\overline{s}_k)$
$B_2 \leftarrow B_2^- \cup B_2^+$
$X^{(i)}(\mathbf{v}) \leftarrow f\left(B_1\left(\langle \mathbf{v}, \mathbf{u_1} \rangle\right), B_2\left(\langle \mathbf{v}, \mathbf{u_2} \rangle\right)\right)$ ▷ f is some composition function
$Y^{(i)}(\mathbf{v}) \leftarrow X^{(i)}(\mathbf{v}) + \epsilon^{(i)}(\mathbf{v})$
$Y(\mathbf{v}) \leftarrow Y(\mathbf{v}) \cup Y^{(i)}(\mathbf{v})$
end for
return $Y(\mathbf{v})$

Computational aspects of simulator

- fBm simulator on the canonical basis runs in $O(n \log n)$ for each sample path, where n is the number of points of the grid
- For our anisotropic simulator, the complexity is thus $O(Nn \log n)$, where N is the number of surfaces
- Because the complexity of searching for the right coordinates is negligible ($O(n)$)

Simulation setup

- Consider the sum and product of two fBms $f_1(B_1, B_2) = B_1 + B_2$, $f_2(B_1, B_2) = B_1 B_2$
- surfaces $N \in \{100, 200\}$, $M_0 = 51 \times 51$ points, noise $\sigma \in \{0, 0.01, 0.05, 0.1\}$, Angles $\alpha \in \{\pi/3, \pi/6, 5\pi/6\}$, $H_1 = 0.8, H_2 = 0.5$
- $\Delta = M_0^{-1/4}(1 + \Delta_c)$, where $\Delta_c = 0.25$ for estimation of α
- $\boldsymbol{\Delta} = \{M_0^{-1/4}, \Delta_1, \ldots, \Delta_{k-1}, 0.4\}$, where $\#\boldsymbol{\Delta} = 15$ for identification
- Risk measure

$$\mathcal{R}_\alpha = |\widehat{\alpha} - \alpha|$$

Simulation Results - Sum

Figure 1: Boxplots for $M = 51$ (sum)

N100_M51_sum.pdf

N200_M51_sum.pdf

Simulation Results - Product

Figure 2: Boxplots for $M = 51$ (product)

N100_M51_prod.pdf   N200_M51_prod.pdf

So... what's the point?

- Rates of convergence of various quantities in fda depend crucially on the regularity of the process

So... what's the point?

- Rates of convergence of various quantities in fda depend crucially on the regularity of the process
- By considering the *directional regularity*, one can exploit the inherent anisotropy of the process and possibly obtain faster rates

So... what's the point?

- Rates of convergence of various quantities in fda depend crucially on the regularity of the process
- By considering the *directional regularity*, one can exploit the inherent anisotropy of the process and possibly obtain faster rates
- Done by simply applying a transformation to the data, of the form

$$Z(\mathbf{t}) := X(\mathbf{R}_\alpha^{-1} \cdot \mathbf{t}), \qquad \forall \mathbf{t} \in \mathcal{T},$$

where

$$\mathbf{R}_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

and α can be estimated using our methodology

Smoothing Application: Setup

- Let $\{X(\mathbf{t}), \mathbf{t} \in \mathcal{T}\}$ be a bi-variate stochastic process with maximising direction $\mathbf{u_1}$

Smoothing Application: Setup

- Let $\{X(\mathbf{t}), \mathbf{t} \in \mathcal{T}\}$ be a bi-variate stochastic process with maximising direction $\mathbf{u_1}$
- Observations associated to $\{X(\mathbf{t}), \mathbf{t} \in \mathcal{T}\}$ come in the form of pairs $(Y_m^{(j)}, \mathbf{t}_m)$, such that

$$Y_m^{(j)} = X^{(j)}(\mathbf{t}_m) + \varepsilon_m^{(j)}, \qquad 1 \leq m \leq M_0, 1 \leq j \leq N,$$

where $(Y_m^{(j)}, \mathbf{t}_m)$ is the learning set.

Smoothing Application: Setup

- Let $\{X(\mathbf{t}), \mathbf{t} \in \mathcal{T}\}$ be a bi-variate stochastic process with maximising direction $\mathbf{u_1}$

- Observations associated to $\{X(\mathbf{t}), \mathbf{t} \in \mathcal{T}\}$ come in the form of pairs $(Y_m^{(j)}, \mathbf{t}_m)$, such that
$$Y_m^{(j)} = X^{(j)}(\mathbf{t}_m) + \varepsilon_m^{(j)}, \qquad 1 \leq m \leq M_0, 1 \leq j \leq N,$$
where $(Y_m^{(j)}, \mathbf{t}_m)$ is the learning set.

- Consider a new realisation X^{new} of X, where pairs $(Y_m^{new}, \mathbf{t}_m)$ are observed such that
$$Y_m^{new} = X^{new}(\mathbf{t}_m) + \varepsilon_m^{new}, \qquad 1 \leq m \leq M_1,$$
where $(Y_m^{new}, \mathbf{t}_m)$ is the online set.

Smoothing Application: Setup

- Let $\{X(\mathbf{t}), \mathbf{t} \in \mathcal{T}\}$ be a bi-variate stochastic process with maximising direction $\mathbf{u_1}$
- Observations associated to $\{X(\mathbf{t}), \mathbf{t} \in \mathcal{T}\}$ come in the form of pairs $(Y_m^{(j)}, \mathbf{t}_m)$, such that
$$Y_m^{(j)} = X^{(j)}(\mathbf{t}_m) + \varepsilon_m^{(j)}, \qquad 1 \leq m \leq M_0, 1 \leq j \leq N,$$
where $(Y_m^{(j)}, \mathbf{t}_m)$ is the learning set.
- Consider a new realisation X^{new} of X, where pairs $(Y_m^{new}, \mathbf{t}_m)$ are observed such that
$$Y_m^{new} = X^{new}(\mathbf{t}_m) + \varepsilon_m^{new}, \qquad 1 \leq m \leq M_1,$$
where $(Y_m^{new}, \mathbf{t}_m)$ is the online set.
- Goal: recovery of the online set $X^{new}(\mathbf{t}_m)$ with the$(Y_m^{new}, \mathbf{t}_m)$ by using some estimator $\widehat{X}^{new}(\mathbf{t}_m)$

Smoothing Application: Methodology

- With the transformation, observed data is $(Y_m^{new}, \mathbf{R}_\alpha \mathbf{t}_m), 1 \leq m \leq M_1$, from

$$Y_m^{new} = Z^{new}(\mathbf{R}_\alpha \mathbf{t}_m) + \varepsilon_m^{new}, \qquad 1 \leq m \leq M_1$$

Smoothing Application: Methodology

- With the transformation, observed data is $(Y_m^{new}, \mathbf{R}_\alpha \mathbf{t}_m), 1 \leq m \leq M_1$, from

$$Y_m^{new} = Z^{new}(\mathbf{R}_\alpha \mathbf{t}_m) + \varepsilon_m^{new}, \qquad 1 \leq m \leq M_1$$

- Consider the Nadaraya-Watson estimator of the form

$$\widehat{Z}^{new}(\mathbf{t}; \mathbf{B}) = \sum_{m=1}^{M_1} Y_m^{new} \frac{K\left(\mathbf{B}(R_\alpha \mathbf{t}_m - \mathbf{t})\right)}{\sum_{m=1}^{M_0} K\left(\mathbf{B}(R_\alpha \mathbf{t}_m - \mathbf{t})\right)}.$$

Smoothing Application: Methodology

- With the transformation, observed data is $(Y_m^{new}, \mathbf{R}_\alpha \mathbf{t}_m), 1 \leq m \leq M_1$, from

$$Y_m^{new} = Z^{new}(\mathbf{R}_\alpha \mathbf{t}_m) + \varepsilon_m^{new}, \qquad 1 \leq m \leq M_1$$

- Consider the Nadaraya-Watson estimator of the form

$$\widehat{Z}^{new}(\mathbf{t}; \mathbf{B}) = \sum_{m=1}^{M_1} Y_m^{new} \frac{K\left(\mathbf{B}(R_\alpha \mathbf{t}_m - \mathbf{t})\right)}{\sum_{m=1}^{M_0} K\left(\mathbf{B}(R_\alpha \mathbf{t}_m - \mathbf{t})\right)}.$$

- This is equivalent to

$$\widehat{X}^{new}(\mathbf{t}; \mathbf{B}) = \sum_{m=1}^{M_1} Y_m^{new} \frac{K\left(\mathbf{B} R_\alpha(\mathbf{t}_m - \mathbf{t})\right)}{\sum_{m=1}^{M_1} K\left(\mathbf{B} R_\alpha(\mathbf{t}_m - \mathbf{t})\right)}.$$

Smoothing Application: Theory

- Consider the risk

$$\mathcal{R}\left(\mathbf{B}, M_1\right) = \mathbb{E}\left[\|\widehat{Z}(\mathbf{B}, M_1) - Z\|_2^2\right].$$

Smoothing Application: Theory

- Consider the risk

$$\mathcal{R}\left(\mathbf{B}, M_1\right) = \mathbb{E}\left[\|\widehat{Z}(\mathbf{B}, M_1) - Z\|_2^2\right].$$

- Effective smoothness is

$$\omega = \{H_1^{-1} + H_2^{-1}\}^{-1}$$

Smoothing Application: Theory

- Consider the risk
$$\mathcal{R}\left(\mathbf{B}, M_1\right) = \mathbb{E}\left[\|\widehat{Z}(\mathbf{B}, M_1) - Z\|_2^2\right].$$

- Effective smoothness is
$$\omega = \{H_1^{-1} + H_2^{-1}\}^{-1}$$

- Optimal bandwidth is given by
$$h_i \asymp M_1^{-\frac{\omega}{(2\omega+1)H_i}}, \qquad i = 1, 2,$$
which gives us the following rate of convergence:
$$\mathcal{R}\left(\mathbf{B}, M_1\right) \lesssim M_1^{-\frac{2\omega}{2\omega+1}}.$$

Simulation Setup

- DGP: Sum of two fBms

Simulation Setup

- DGP: Sum of two fBms
- Parameters for learning set: $\alpha \in \{\pi/3, 5\pi/6\}, N = 150, M_0 = 101, \sigma = 0.05, H_1 = 0.8, H_2 = 0.5$ with the same Δ settings as before

Simulation Setup

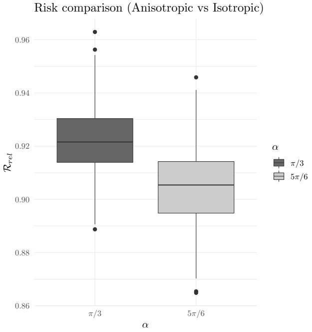
- DGP: Sum of two fBms
- Parameters for learning set: $\alpha \in \{\pi/3, 5\pi/6\}, N = 150, M_0 = 101, \sigma = 0.05, H_1 = 0.8, H_2 = 0.5$ with the same Δ settings as before
- Parameters for online set: $M_1^{true} = 201$, $M_1 = 121, \sigma = 0.05$

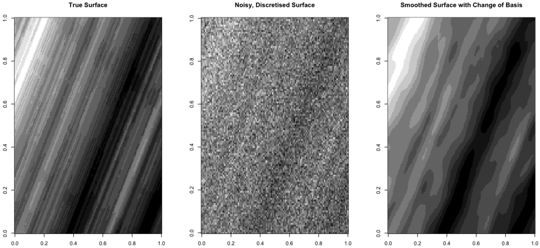
Simulation Setup

- DGP: Sum of two fBms
- Parameters for learning set: $\alpha \in \{\pi/3, 5\pi/6\}, N = 150, M_0 = 101, \sigma = 0.05, H_1 = 0.8, H_2 = 0.5$ with the same Δ settings as before
- Parameters for online set: $M_1^{true} = 201$, $M_1 = 121, \sigma = 0.05$
- Risk measure for comparison:

$$\mathcal{R}_{rel} = \frac{\mathcal{R}^{ani}(\mathbf{B}, M_1)}{\mathcal{R}^{iso}(\mathbf{B}, M_1)}$$

Simulation results





Conclusion

- Anisotropy depends not only on the dimension, but the direction

Conclusion

- Anisotropy depends not only on the dimension, but the direction
- Taking into account the directional regularity can allow one to obtain faster rates of convergence, even if isotropic on the canonical basis

Conclusion

- Anisotropy depends not only on the dimension, but the direction
- Taking into account the directional regularity can allow one to obtain faster rates of convergence, even if isotropic on the canonical basis
- Algorithms for the estimation and identification of the directional regularity that works well in practice are constructed

Conclusion

- Anisotropy depends not only on the dimension, but the direction
- Taking into account the directional regularity can allow one to obtain faster rates of convergence, even if isotropic on the canonical basis
- Algorithms for the estimation and identification of the directional regularity that works well in practice are constructed
- One application is the improved rate in smoothing surfaces

Conclusion

- Anisotropy depends not only on the dimension, but the direction
- Taking into account the directional regularity can allow one to obtain faster rates of convergence, even if isotropic on the canonical basis
- Algorithms for the estimation and identification of the directional regularity that works well in practice are constructed
- One application is the improved rate in smoothing surfaces
- But the consequences are not limited to smoothing! Thus recommend it as a standard pre-processing step in multivariate fda

THANK YOU!