Estimation non paramétrique dans un modèle de régression additif avec réponse et covariables fonctionnelles

Elodie BRUNEL
IMAG, CNRS
Université de Montpellier

travail en collaboration avec
Fabienne Comte (Univ. Paris-Cité) et Céline Duval (Univ. Paris Sorbonne)

Journées de Statistique et Optimisation en Occitanie

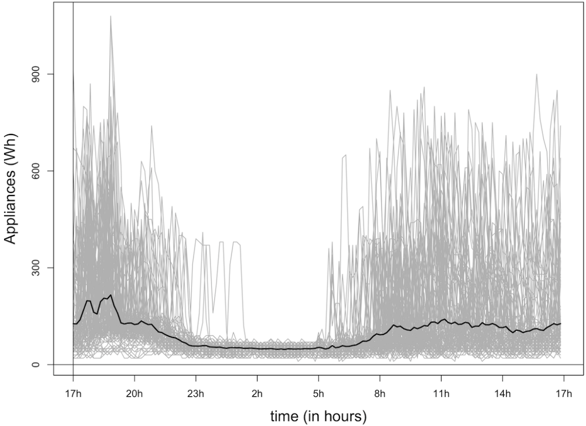
Céline

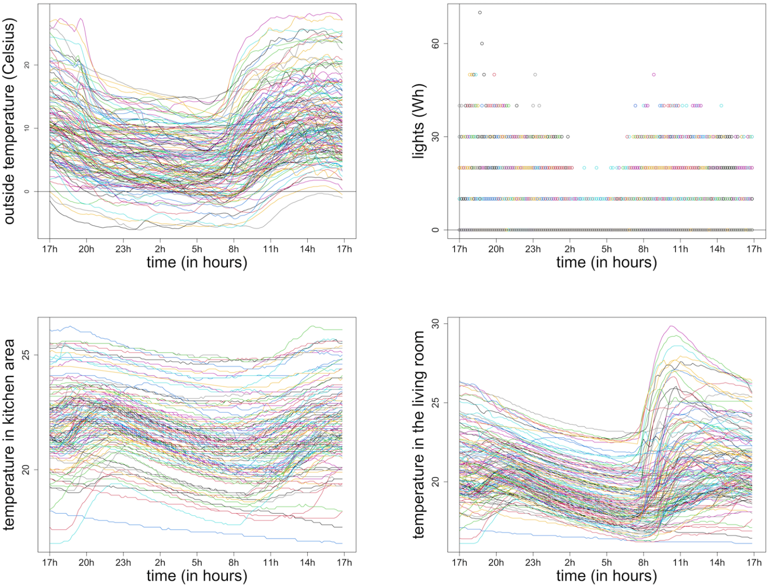
Fabienne

Previous works

- Functional data : Ramsay & Silverman, 2005.
- Important literature in the "functional linear model," meaning a real response variable with functional covariates : Cai et al. 2006 ; Cardot et al. 2007 ; Li and Hsing 2007 ; Ferraty et al. (2012), Hilgert et al. 2013 ; Cai and Yuan 2012 ; Brunel et al., 2016.
- The case where the response variable is also functional is less studied : "concurrent regression model" : Crambes & Mas 2013, Chagny et al. 2024.

We observe for $i = 1, \cdots, N$ individuals the N independent trajectories of a process $Y_i(t)$ (the functional response) and of K processes $(X_{i,1}(t), \cdots, X_{i,K}(t))$ (the functional explanatory variables) for $t \in [0, \tau]$, a fixed time interval.





outside temperature (Celsius)
0
10
20
17h
20h
23h
2h
5h
8h
11h
14h
17h
time (in hours)
lights (Wh)
0
30
60
17h
20h
23h
2h
5h
8h
11h
14h
17h
time (in hours)
temperature in kitchen area
20
25
17h
20h
23h
2h
5h
8h
11h
14h
17h
time (in hours)
temperature in the living room
20
25
30
17h
20h
23h
2h
5h
8h
11h
14h
17h
time (in hours)

Outline

We consider the regression model with functional responses and explanatory variables ("concurrent regression model") as follows :

$$Y_i(t) = \sum_{j=1}^{K} b_j(t) X_{i,j}(t) + \sigma(t, \mathbf{X}_i(t))\varepsilon_i(t), \quad t \in [0,\tau], \; i = 1, \ldots, N,$$

For $i = 1, \ldots, N$,

- The processes $(\varepsilon_i(t))_{t\in[0,\tau]}$ are centered, independent, and identically distributed (i.i.d.), with $\mathbb{E}(\varepsilon_1^2(t)) = 1$;
- The processes $(\mathbf{X}_i(t) = \begin{pmatrix} X_{i,1}(t) & \ldots & X_{i,K}(t) \end{pmatrix})_{t\in[0,\tau]}$ are i.i.d. and independent of $(\varepsilon_i(t))_{t\in[0,\tau]}$.
- The integer K is known and small compared to N : it is the number of explanatory processes.
- We may consider the homoscedastic case $\sigma(t, \mathbf{x}) = \sigma$.
  The model coefficients $b_1(t), \ldots, b_K(t)$ are deterministic functions, unknown and to be estimated.

Outline

Least Squares Estimators

Let $(\varphi_j)_{j\geq 1}$ be a basis of $\mathbb{L}^2_\tau := \mathbb{L}^2([0,\tau])$ of measurable functions such that $\int_0^\tau \varphi_j^2(x)dx \leq 1$, and let S_m be the subspace spanned by $(\varphi_j)_{1\leq j\leq m}$. Our estimator of $\mathbf{b}(t) = (b_1(t),\ldots,b_K(t))^T$ is constructed on the product space $S_{\mathbf{m}} := S_{m_1} \times \cdots \times S_{m_K}$, $\mathbf{m} = (m_1,\ldots,m_K)$, i.e.,

$$\widehat{\mathbf{b}}_{\mathbf{m}}(t) = (\widehat{b}_1(t),\ldots,\widehat{b}_K(t))^T \quad \text{where} \quad \widehat{b}_k(t) = \sum_{j=1}^{m_k} \widehat{\beta}_{k,j}\varphi_j(t),$$

For $i = 1,\cdots,N$, $(Y_i(t),\mathbf{X}_i(t))_{0\leq t\leq\tau}$ and $\mathbf{X}_i(t) = (X_{i,1}(t),\cdots,X_{i,K}(t))^T$ and $\mathbf{h} = (h_1,\cdots,h_K)^T \in (\mathbb{L}^2_\tau)^K$

$$\widehat{\mathbf{b}}_{\mathbf{m}}(t) = \arg\min_{\mathbf{h}\in S_{\mathbf{m}}} \frac{1}{N}\sum_{i=1}^{N}\int_0^\tau (Y_i(t) - \mathbf{h}(t)^T\mathbf{X}_i(t))^2 dt$$

Minimizing with respect to $\mathbf{h}$ the contrast $\frac{1}{N}\sum_{i=1}^{N}\int_0^\tau (Y_i(t) - \mathbf{h}(t)^T\mathbf{X}_i(t))^2 dt$ is equivalent to minimizing :

$$U_N(\mathbf{h}) = \frac{1}{N}\sum_{i=1}^{N}\int_0^\tau (\mathbf{h}(t)^T\mathbf{X}_i(t))^2 dt - \frac{2}{N}\sum_{i=1}^{N}\int_0^\tau Y_i(t)\mathbf{h}(t)^T\mathbf{X}_i(t)dt$$

and taking the expectation :

$$\begin{aligned}\mathbb{E}[U_N(\mathbf{h})] &= \mathbb{E}\left\{\int_0^\tau (\mathbf{h}(t)^T\mathbf{X}_1(t))^2 dt\right\} - 2\mathbb{E}\left[\int_0^\tau Y_1(t)\mathbf{h}(t)^T\mathbf{X}_1(t)dt\right] \\ &= \int_0^\tau \mathbf{h}(t)^T\underbrace{\mathbb{E}[\mathbf{X}_1(t)\mathbf{X}_1(t)^T]}_{:=\Gamma(t)}\mathbf{h}(t)dt - 2\int_0^\tau \mathbf{b}(t)^T\mathbb{E}[\mathbf{X}_1(t)\mathbf{X}_1(t)^T]\mathbf{h}(t)dt \\ &:= \|\mathbf{h}\|_\Gamma^2 - 2\langle \mathbf{b};\mathbf{h}\rangle_\Gamma = \|\mathbf{h}-\mathbf{b}\|_\Gamma^2 - \|\mathbf{b}\|_\Gamma^2\end{aligned}$$

Thus, $\mathbb{E}[U_N(\mathbf{h})]$ is minimized for $\mathbf{h}=\mathbf{b}$ and $\|\cdot\|_\Gamma$ appears as the natural semi-norm (or norm if Γ is invertible) of the problem.

Minimizing with respect to **h** the contrast $\frac{1}{N}\sum_{i=1}^{N}\int_0^\tau (Y_i(t)-\mathbf{h}(t)^T\mathbf{X}_i(t))^2dt$ is equivalent to minimizing :

$$U_N(\mathbf{h}) = \frac{1}{N}\sum_{i=1}^{N}\int_0^\tau (\mathbf{h}(t)^T\mathbf{X}_i(t))^2dt - \frac{2}{N}\sum_{i=1}^{N}\int_0^\tau Y_i(t)\mathbf{h}(t)^T\mathbf{X}_i(t)dt$$

and taking the expectation :

$$\begin{aligned}\mathbb{E}[U_N(\mathbf{h})] &= \mathbb{E}\left\{\int_0^\tau (\mathbf{h}(t)^T\mathbf{X}_1(t))^2dt\right\} - 2\mathbb{E}\left[\int_0^\tau Y_1(t)\mathbf{h}(t)^T\mathbf{X}_1(t)dt\right] \\ &= \int_0^\tau \mathbf{h}(t)^T\underbrace{\mathbb{E}[\mathbf{X}_1(t)\mathbf{X}_1(t)^T]}_{:=\Gamma(t)}\mathbf{h}(t)dt - 2\int_0^\tau \mathbf{b}(t)^T\mathbb{E}[\mathbf{X}_1(t)\mathbf{X}_1(t)^T]\mathbf{h}(t)dt \\ &:= \|\mathbf{h}\|_\Gamma^2 - 2\langle\mathbf{b};\mathbf{h}\rangle_\Gamma = \|\mathbf{h}-\mathbf{b}\|_\Gamma^2 - \|\mathbf{b}\|_\Gamma^2\end{aligned}$$

Thus, $\mathbb{E}[U_N(\mathbf{h})]$ is minimized for $\mathbf{h}=\mathbf{b}$ and $\|\cdot\|_\Gamma$ appears as the natural semi-norm (or norm if Γ is invertible) of the problem.

Minimizing with respect to $\mathbf{h}$ the contrast $\frac{1}{N}\sum_{i=1}^{N}\int_0^{\tau}(Y_i(t)-\mathbf{h}(t)^T\mathbf{X}_i(t))^2dt$ is equivalent to minimizing :

$$U_N(\mathbf{h})=\frac{1}{N}\sum_{i=1}^{N}\int_0^{\tau}(\mathbf{h}(t)^T\mathbf{X}_i(t))^2dt-\frac{2}{N}\sum_{i=1}^{N}\int_0^{\tau}Y_i(t)\mathbf{h}(t)^T\mathbf{X}_i(t)dt$$

and taking the expectation :

$$\begin{aligned}\mathbb{E}[U_N(\mathbf{h})] &= \mathbb{E}\left\{\int_0^{\tau}(\mathbf{h}(t)^T\mathbf{X}_1(t))^2dt\right\}-2\mathbb{E}\left[\int_0^{\tau}Y_1(t)\mathbf{h}(t)^T\mathbf{X}_1(t)dt\right]\\ &= \int_0^{\tau}\mathbf{h}(t)^T\underbrace{\mathbb{E}[\mathbf{X}_1(t)\mathbf{X}_1(t)^T]}_{:=\Gamma(t)}\mathbf{h}(t)dt-2\int_0^{\tau}\mathbf{b}(t)^T\mathbb{E}[\mathbf{X}_1(t)\mathbf{X}_1(t)^T]\mathbf{h}(t)dt\\ &:= \|\mathbf{h}\|_{\Gamma}^2-2\langle\mathbf{b};\mathbf{h}\rangle_{\Gamma}=\|\mathbf{h}-\mathbf{b}\|_{\Gamma}^2-\|\mathbf{b}\|_{\Gamma}^2\end{aligned}$$

Thus, $\mathbb{E}[U_N(\mathbf{h})]$ is minimized for $\mathbf{h}=\mathbf{b}$ and $\|\cdot\|_{\Gamma}$ appears as the natural semi-norm (or norm if Γ is invertible) of the problem.

The estimator $\widehat{\mathbf{b}}_{\mathbf{m}}(t) = (\widehat{b}_1(t), \ldots, \widehat{b}_K(t))^T$ where $\widehat{b}_k(t) = \sum_{j=1}^{m_k} \widehat{\beta}_{k,j}\varphi_j(t)$, or equivalently the coefficients :

$\widehat{B}_{\mathbf{m}} = (\underbrace{\widehat{\beta}_{1,1}, \ldots, \widehat{\beta}_{1,m_1}}_{b_1(t)}, \underbrace{\widehat{\beta}_{2,1}, \ldots, \widehat{\beta}_{2,m_2}}_{b_2(t)}, \ldots, \underbrace{\widehat{\beta}_{K,1}, \ldots, \widehat{\beta}_{K,m_K}}_{b_K(t)})^T$ are obtained by algebraic resolution through a standard gradient calculation associated with the least squares contrast : $\vec{\nabla} U_N(\widehat{B}_{\mathbf{m}}) = \vec{0}$ which leads to

$$\underbrace{\begin{pmatrix} \widehat{\Psi}_{m_1,m_1} & \cdots & \widehat{\Psi}_{m_1,m_K} \\ \vdots & & \vdots \\ \widehat{\Psi}_{m_K,m_1} & \cdots & \widehat{\Psi}_{m_K,m_K} \end{pmatrix}}_{:=\widehat{\Psi}_{\mathbf{m}}} \begin{pmatrix} \widehat{\beta}_{1,1} \\ \cdots \\ \widehat{\beta}_{1,m_1} \\ \vdots \\ \widehat{\beta}_{K,1} \\ \cdots \\ \widehat{\beta}_{K,m_K} \end{pmatrix} = \underbrace{\begin{pmatrix} \widehat{V}_{1,m_1} \\ \vdots \\ \widehat{V}_{K,m_K} \end{pmatrix}}_{\widehat{V}_{\mathbf{m}}}$$

with, $\widehat{\Psi}_{m_j,m_k} = \left(\int_0^{\tau} \varphi_p(t)\varphi_q(t) \frac{1}{N} \sum_{i=1}^{N} X_{i,j}(t) X_{i,k}(t) dt\right)_{1 \le p \le m_j; 1 \le q \le m_k}$.

and $\widehat{V}_{j,m_j} = \left(\frac{1}{N} \int_0^{\tau} \varphi_p(t) \sum_{i=1}^{N} X_{i,j}(t) Y_i(t) dt\right)_{1 \le p \le m_j}$

In order to ensure the existence of the estimator, conditions for the invertibility of $\widehat{\Psi}_{\mathbf{m}}$ must be imposed.

Recall that the semi-norm associated with the contrast is

$$\|\mathbf{h}\|_\Gamma^2 = \int_0^\tau \mathbf{h}(t)^T \Gamma(t) \mathbf{h}(t) dt \text{ with } \Gamma(t) = \mathbb{E}[\mathbf{X}_1(t)\mathbf{X}_1(t)^T]$$

and its empirical equivalent $\|\mathbf{h}\|_N^2 = \int_0^\tau \mathbf{h}(t)^T \Gamma_N(t) \mathbf{h}(t) dt$ with $\Gamma_N(t) := \left(\frac{1}{N} \sum_{i=1}^N X_{i,j}(t) X_{i,k}(t) \right)_{1 \leq j,k \leq K}$ and $\mathbb{E}[\Gamma_N(t)] = \Gamma(t)$.

Condition $(\mathcal{A}_S)$ (Identifiability)

(i) $\forall t \in [0, \tau]$, the matrix $\Gamma(t)$ is invertible and

$$\mathfrak{f}_\Gamma := \sup_{t \in [0,\tau]} \|\Gamma(t)^{-1}\|_{\mathrm{op}} < +\infty.$$

(ii) $\forall t \in [0, \tau]$, $\forall N \geq 1$, $\Gamma_N(t)$ is almost surely invertible.

$(\mathcal{A}_S)$ implies that $\| \cdot \|_\Gamma$ and $\| \cdot \|_N$ are norms.

Proposition

If the condition $(\mathcal{A}_S)$ is satisfied, then the matrix $\widehat{\Psi}_{\mathbf{m}}$ (resp. $\Psi_{\mathbf{m}}$) is symmetric positive definite.

Thus, the least squares estimator
$\widehat{\mathbf{b}}_{\mathbf{m}}(t) = (\widehat{b}_1(t), \ldots, \widehat{b}_K(t))^T$ where $\widehat{b}_k(t) = \sum_{j=1}^{m_k} \widehat{\beta}_{k,j}\varphi_j(t)$,

or equivalently the vector of coefficients :

$\widehat{B}_{\mathbf{m}} = (\widehat{\beta}_{1,1}, \ldots, \widehat{\beta}_{1,m_1}, \widehat{\beta}_{2,1}, \ldots, \widehat{\beta}_{2,m_2}, \ldots, \widehat{\beta}_{K,1}, \ldots, \widehat{\beta}_{K,m_K})^T$ are obtained by :

$$\widehat{B}_{\mathbf{m}} = \widehat{\Psi}_{\mathbf{m}}^{-1} \widehat{V}_{\mathbf{m}}$$

Outline

We further assume that :

Condition $(\mathcal{A}_{X,p})$

a) There exists an integer $p \geq 1$, such that $G_p^p := \max_{j=1,\dots,K} \sup_{t\in[0,\tau]} \mathbb{E}[|X_{1,j}|^p(t)] < +\infty$.

b) ... and some other technical conditions satisfied by gaussian processes or Poisson processes...

Condition $(\mathcal{A}_b)$

The functions $b_k : \mathbb{R}^+ \to \mathbb{R}$, $k = 1, \dots, K$, are measurable and bounded on $[0, \tau]$ (so they belong to $\mathbb{L}^2_\tau$).

Condition $(\mathcal{A}_\varphi)$

There exists $\omega > 0$, and $\exists c_\varphi > 0$ such that $L(S_m) := \sup_{t\in[0,\tau]} \sum_{j=1}^m \varphi_j^2(t) \leq c_\varphi\, m^\omega$, for all $m \geq 1$.

We study a truncated version of the estimator :

$$\widetilde{\mathbf{b}}_{\mathbf{m}} = \widehat{\mathbf{b}}_{\mathbf{m}} \mathbf{1}_{\Lambda_N},$$

where $\Lambda_N = \{\forall t \in [0,\tau], \ \ \|\Gamma_N(t)^{-1}\|_{\mathrm{op}} \leq \mathfrak{c}_1 N^{\mathfrak{c}_2}\}$, and $\mathfrak{c}_1, \mathfrak{c}_2 > 0$.

Théorème

Suppose that $(\mathcal{A}_S)$, $(\mathcal{A}_{X,p})$, $(\mathcal{A}_b)$ and $(\mathcal{A}_\varphi)$ are satisfied, and let $\mathbf{m} = (m_1, \cdots m_K)$ such that $|\mathbf{m}| := m_1 + \cdots + m_K \leq N$,
Moreover, suppose that

$$\sup_{t\in[0,\tau]} \mathbb{E}[\varepsilon^4(t)] < +\infty \text{ and } \sup_{j=1,\dots,K} \sup_{t\in[0,\tau]} \mathbb{E}[X_{1,j}^4(t)\sigma^4(t,\mathbf{X}_1(t))] < +\infty,$$

and $\sup_{t\in[0,\tau],x\in\mathbb{R}^K} \sigma^2(t,x) := \|\sigma\|_\infty^2 < +\infty$.
The estimator $\widetilde{\mathbf{b}}_{\mathbf{m}}$ of $\mathbf{b}(t) = (b_1(t), \dots, b_K(t))^T$ satisfies, for $p \geq (2\mathfrak{c}_2 + 3) \vee 4$ and a constant $c > 0$,

$$\mathbb{E}[\|\widetilde{\mathbf{b}}_{\mathbf{m}} - \mathbf{b}\|_N^2] \leq \inf_{\mathbf{h}\in S_{\mathbf{m}}} \|\mathbf{h} - \mathbf{b}\|_\Gamma^2 + 2\tau\|\sigma\|_\infty^2 \frac{|\mathbf{m}|}{N} + \frac{c}{N},$$

Convergence Rate

With a regularity hypothesis on the functions $b_1, \dots, b_K$, we control the approximation error (bias).

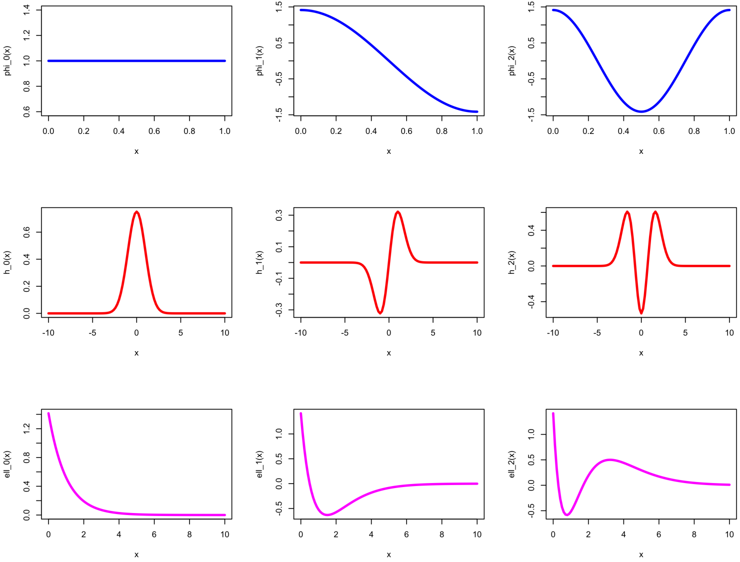
Proposition

Suppose that for all $j \in \{1, \dots, K\}$, the function b_j belongs to a function space such that $\inf_{h \in S_{m_j}} \|b_j - h\|^2 \le L_j m_j^{-2s_j}$. If we choose each $m_j = O(N^{1/(2s_j+1)})$ and with $s_{(1)} = \min_{j=1,\dots,K} s_j$, there exists a constant $C > 0$ such that,

$$\mathbb{E}[\|\widetilde{\mathbf{b}}_{\mathbf{m}} - \mathbf{b}\|_\Gamma^2] \le C\, N^{-2s_{(1)}/(2s_{(1)}+1)}.$$

Example : Sobolev space associated with the (orthonormal) basis $(\varphi_j)_j$.
$W(s, L) = \{f = \sum_{j \ge 0} \theta_j \varphi_j,$ such that $\forall J \ge 1, \sum_{j \ge J} \theta_j^2 \le L^2 J^{-2s}\}$.
If $b_j \in W(s_j, L_j)$, $j \in \{1, \dots, K\}$, we can establish that $\inf_{h \in S_{m_j}} \|b_j - h\|^2 \le L_j^2 m_j^{-2s_j}$.

Outline



First bases functions Trigo [T], Hermite [H] et Laguerre [L].

Outline

Model Selection

We need to choose the dimension $\mathbf{m} = (m_1, \cdots, m_K)$ from the family of dimensions :

$$\mathcal{M}_N = \left\{\mathbf{m} \in \{1, \ldots, N\}^K, \quad \forall i \in \{1, \ldots, K\},\ m_i \leq N\right\}.$$

And we select

$$\widehat{\mathbf{m}} \in \arg\min_{\mathbf{m} \in \mathcal{M}_N} \left[U_N(\widehat{\mathbf{b}}_{\mathbf{m}}) + \mathrm{pen}(\mathbf{m})\right], \quad \mathrm{pen}(\mathbf{m}) = \kappa \|\sigma\|_\infty^2 \frac{|\mathbf{m}|}{N},$$

i.e., penalizing the least squares criterion with a term $\mathrm{pen}(\mathbf{m})$ of the order of the variance.

In practice, we still need to estimate $\|\sigma\|_\infty^2$, which is unknown, by evaluating the supremum of the residual variance.

Model Selection

We need to choose the dimension $\mathbf{m} = (m_1, \cdots, m_K)$ from the family of dimensions :

$$\mathcal{M}_N = \left\{\mathbf{m} \in \{1, \ldots, N\}^K, \quad \forall i \in \{1, \ldots, K\},\ m_i \leq N\right\}.$$

And we select

$$\widehat{\mathbf{m}} \in \arg\min_{\mathbf{m} \in \mathcal{M}_N} \left[U_N(\widehat{\mathbf{b}}_{\mathbf{m}}) + \mathrm{pen}(\mathbf{m})\right], \quad \mathrm{pen}(\mathbf{m}) = \kappa \|\sigma\|_\infty^2 \frac{|\mathbf{m}|}{N},$$

i.e., penalizing the least squares criterion with a term $\mathrm{pen}(\mathbf{m})$ of the order of the variance.
In practice, we still need to estimate $\|\sigma\|_\infty^2$, which is unknown, by evaluating the supremum of the residual variance.

Outline

Illustration 1 : $K = 3$ covariates

Simulation of $N = 200$ trajectories of :

$$Y_i(t) = b_1(t)X_{1,i}(t) + b_2(t)X_{2,i}(t) + b_3(t)X_{3,i}(t) + \varepsilon_i(t)$$

with

$$b_1(t) = \cos(2.8\pi t),\ b_2(t) = 0.25\exp(-t/3) - 2\exp(-2t),\ b_3(t) = 2t^2.$$

- $X_1(t) = 1$,
- $X_2(t)$ a Poisson process with parameter $\lambda = 0.5$,
- $X_{3,i}(t) = \mu_X(t) + \sum_{j=1}^{10} \rho_j \xi_{i,j} \phi_j(t)$ where $\mu_X(t) = t + \sin(t)$, and for $j \geq 1$, $\phi_j(t) = \sqrt{2}\sin((j - 1/2)\pi t)$, $\rho_j = 1/((j - 1/2)\pi)$, and the $\xi_{i,j}$ are i.i.d. $\mathcal{N}(0, 1)$.
- The noise $\varepsilon(t) = c_\varepsilon \sum_{j=11}^{20} \rho_j \xi_{i,j} \phi_j(t)$ with $c_\varepsilon = 7$ chosen to obtain a signal-to-noise ratio of approximately 2.

$b_1(t) = \cos(2.8\pi t),\ b_2(t) = 0.25\exp(-t/3) - 2\exp(-2t),\ b_3(t) = 2t^2.$

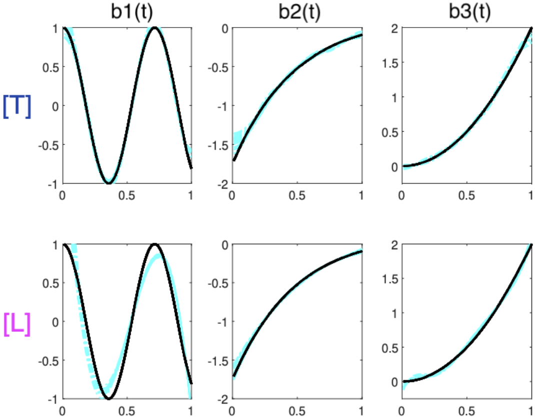


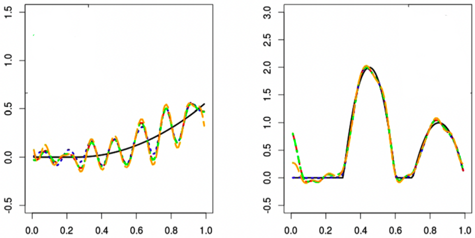
Figure – Estimation of $b_1(t)$, $b_2(t)$, and $b_3(t)$, with the [T] basis (first row) and the [L] basis (second row) for 25 repetitions ($N = 200$). Average dimensions selected : (5.6, 3.6, 5.5) for [T] and (6.0, 2.0, 4.0) for [L].

Illustration 2 : from Manrique et al. (2018)

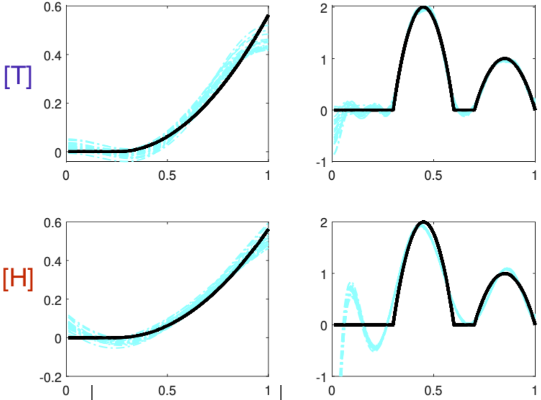
We compare our method to that of Manrique, Crambes & Hilgert (2018) :

$$Y_i(t) = b_1(t) + b_2(t)X_{i,1}(t) + \sigma(t)\varepsilon_i(t), \quad i = 1 \cdots, N.$$

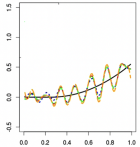
Figure from Manrique et al. (2018).



- $K = 1$ and the explanatory functional covariate $X_{i,1}$ is centered.
- "Ridge" regularization of the functional linear regression estimator.
- Convergence results under fairly general conditions (in particular, non-compact support).
- No guarantees of optimality and difficult to generalize for $K > 1$.



	Manrique *et al.*	Basis [T]	Basis [H]
WASE	0.0150 (0.0097)	0.0032 (0.0015)	0.0642(0.0088)



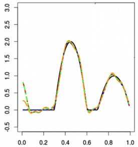


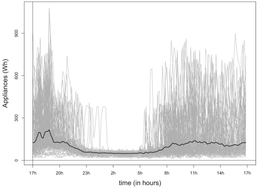
Illustration 3 : Real data

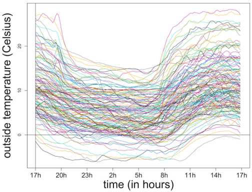
Electric consumption of devices in an eco-friendly house in Stambrugge, Belgium, from Candanedo et al. (2017)

Illustration 3 : Real data

19,728 measurements every 10 minutes over 137 days (144 measurements per day) of the electricity consumption of devices $(Y_i(t)), i = 1, \cdots, 137$, $t \in [0, 24]$. $t = 0$ at 5 p.m, $t = 144$ at 4 :50 p.m. (+1 day)

$X_1(t) = 100$ $X_2(t)$: outdoor temperature $X_3(t)$: lighting





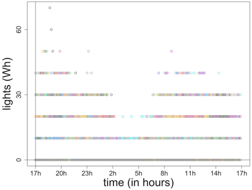


Illustration 3 : Real data

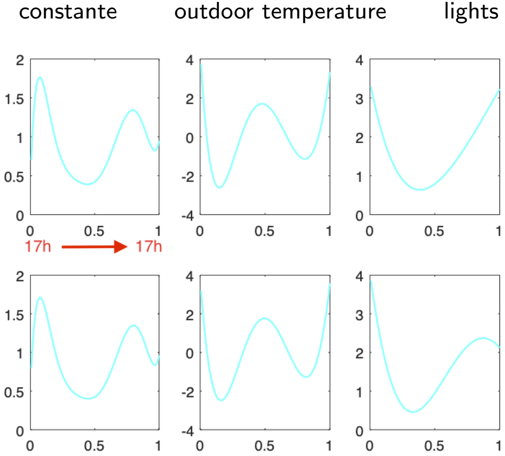


Figure – Electricity data. Estimation de $b_1(t)$ (left), $b_2(t)$ (middle) et $b_3(t)$ (right), base [L] (first line) et base [H] (second line). Selected dimensions are (8,5,5) and (8,5,4).

Conclusion

- We consider various continuous or non-continuous explanatory processes.
- Our method allows the simultaneous estimation of K functions : the method is computationally simple and fast.
- The model selection procedure allows for anisotropy.
- The convergence rates are proved to be optimal.
- Thanks to the additivity of the model, there is no curse of dimensionality. The convergence rate of $\widehat{b}_{\mathbf{m}}$ is associated with the lowest regularity of the functions to estimate b_k, $k = 1, \ldots, K$.

References

1 Brunel, É., Comte, F., Duval, C. (2024) Nonparametric estimation for additive concurrent regression models, Bernoulli, to appear.

2 Brunel, É., Mas, A., and Roche, A. (2016). Non-asymptotic adaptive prediction in functional linear models. Journal of Multivariate Analysis, 143 :208–232.

3 Cai, T. T., Hall, P., et al. (2006). Prediction in functional linear regression. The Annals of Statistics, 34(5) :2159–2179.

4 Cai, T. T. and Yuan, M. (2012). Minimax and adaptive prediction for functional linear regression. Journal of the American Statistical Association, 107(499) :1201–1216.

5 Candanedo, L. M., Feldheim, V., and Deramaix, D. (2017). Data driven prediction models of energy use of appliances in a low-energy house. Energy and Buildings, 140.

6 Cardot, H., Mas, A., and Sarda, P. (2007). Clt in functional linear regression models. Probability Theory and Related Fields, 138(3-4) :325–361.

7 Comte, F. , Genon-Catalot, V. (2024). Nonparametric estimation for i.i.d. Stochastic Differential Equations with space-time dependent coefficients. *hal-04139052*. To appear in *SIAM/ASA Journal of Uncertainty Quantification*.

8 Chagny, G., Meynaoui, A., Roche, A. (2022) Adaptive nonparametric estimation in the functional linear model with functional output., hal-03579232

9 Crambes, C. and Mas, A. (2013). Asymptotics of prediction in functional linear regression with functional outputs. Bernoulli, 19(5B) :2627–2651.

10 Ferraty, F., Van Keilegom, I., Vieu,P. (2012). Regression when both response and predictor are functions. J. Multivariate Analysis, 109, 10-28.

11 Hilgert, N., Mas, A., Verzelen, N., et al. (2013). Minimax adaptive tests for the functional linear model. Annals of Statistics, 41(2) :838–869.

12 Li, Y. and Hsing, T. (2007). On rates of convergence in functional linear regression. J. Multivariate Anal., 98(9) :1782–1804.

13 Manrique, T., Crambes, C. , Hilgert, N. (2018) Ridge regression for the functional concurrent model. Electron. J. Stat. 12, no.1, 985-1018.