Bootstrap confidence intervals in functional regression under dependence

Juan Manuel Vilar Fernández
Paula Raña, Germán Aneiros and Philippe Vieu

Departamento de Matemáticas, Universidade da Coruña

Galician Seminar of Nonparametric Statistical Inference
8-9 June 2016

Spanish Electricity Market

OMIE: 'Operador del Mercado Ibérico de Energía'

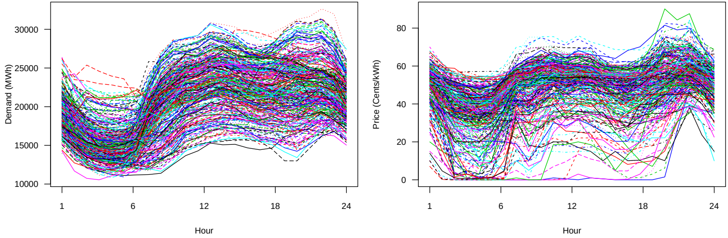
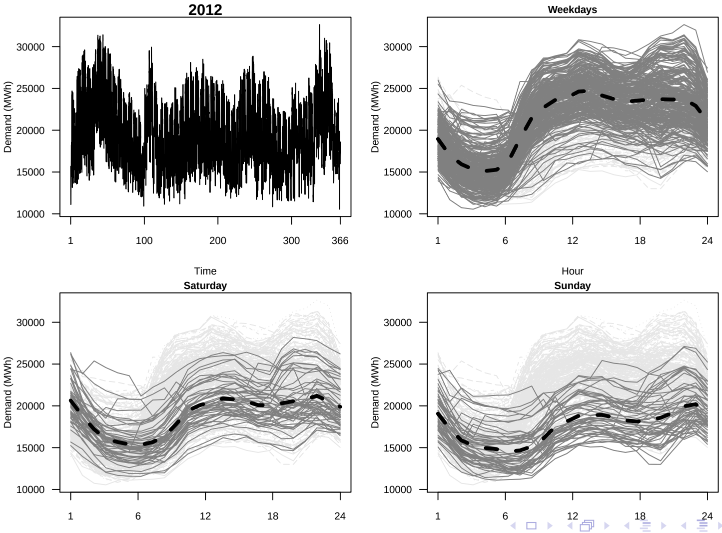
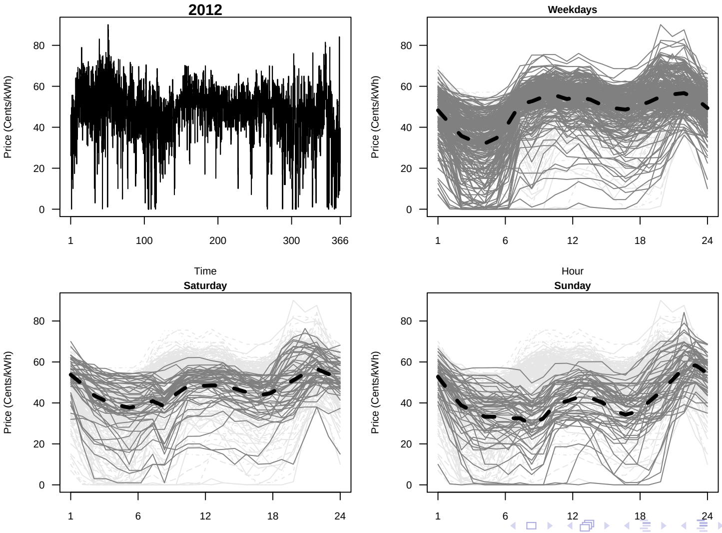


Figure : Electricity demand and price daily curves in 2012.

Electricity demand



Electricity price



Functional time series

Daily curves of electricity demand or price along 2012: $\{\chi_i\}_{i=1}^{365}$
Discretized curves: $\chi_i(t_j), j = 1, \ldots, 24$.

Functional time series: $\{\chi_i\}_{i=1}^{n}$
Real-valued continuous time stochastic process $\{\chi(t)\}_{t \in R}$
Seasonal process, with seasonal length τ,
observed on the interval $(a, b]$ with $b = a + n\tau$.

Functional time series, $\{\chi_i\}_{i=1}^{n}$, is defined in terms of $\{\chi(t)\}_{t \in R}$ as:

$$\chi_i(t) = \chi(a + (i-1)\tau + t) \text{ with } t \in [0, \tau).$$

Prediction with functional regression

Objective

Predict next–day electricity demand/price in Spain during 2012.

$$\{\chi_i\}_{i=1}^{N} \longrightarrow \chi_{N+1}$$

Functional Autoregressive models

- Functional nonparametric regression.
- Semi-functional partial linear regression.

Covariates

- Electricity demand: meteorological variables, temperature.
- Electricity price: demand, wind power production.

Funtional Nonparametric Regression

Functional explanatory variable and scalar response

Autoregressive model

$$G(\chi_{i+1}) = m(\chi_i) + \varepsilon_i \ (i = 1, \ldots, n)$$

General model

$Y_i = m(\chi_i) + \varepsilon_i \ , i = 1, \ldots, n$ where $\{(\chi_i, Y_i)\}$ is α-mixing

Functional kernel estimator

$$\widehat{m}_h(\chi) = \frac{\sum_{i=1}^n K(d(\chi_i, \chi)/h) Y_i}{\sum_{i=1}^n K(d(\chi_i, \chi)/h)} = \sum_{i=1}^n w_h(\chi_i, \chi) Y_i$$

Semi-Funtional Partial Linear Regression

Functional nonparametric explanatory variable, scalar linear-effect covariate and scalar response

Autoregressive model

$$G(\chi_{i+1}) = \boldsymbol{X}_i^T \beta + m(\chi_i) + \varepsilon_i,\ i = 1, \ldots, n$$

General model

$Y_i = \boldsymbol{X}_i^T \beta + m(\chi_i) + \varepsilon_i,\ i = 1, \ldots, n$, where $\{(\boldsymbol{X}_i, \chi_i, Y_i)\}$ is α-mixing

Estimators

Denote

$$\mathbf{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^T, \ \mathbf{Y} = (Y_1, \ldots, Y_n)^T, \ \mathbf{W}_h = (w_h(\chi_i, \chi_j))$$

and, for any $(n \times q)$ matrix $\mathbf{A}$ $(q \geq 1)$,

$$\widetilde{\mathbf{A}}_h = (\mathbf{I} - \mathbf{W}_h)\mathbf{A}.$$

Estimators

$$\widehat{\beta}_h = (\widetilde{\mathbf{X}}_h^T \widetilde{\mathbf{X}}_h)^{-1} \widetilde{\mathbf{X}}_h^T \widetilde{\mathbf{Y}}_h \qquad \widehat{m}_h(\chi) = \sum_{i=1}^{n} w_h(\chi_i, \chi)(Y_i - \boldsymbol{X}_i^T \widehat{\beta}_h)$$

Nadaraya-Watson type weights $w_h(\chi_i, \chi) = \dfrac{K(d(\chi_i, \chi)/h)}{\sum_{i=1}^n K(d(\chi_i, \chi)/h)}$, where $K(\cdot)$ is a real function (the kernel) and $h > 0$ is a smoothing parameter.

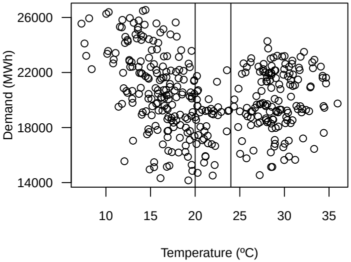
In practice: predict electricity demand

Functional nonparametric regression

- Functional explanatory variable: previous daily demand curves.
- Scalar response: electricity demand for the next day, fixed hour.

Semi-Functional Partial Linear Regression

- Scalar explanatory variable: daily temperature $\rightarrow$ linear effect.



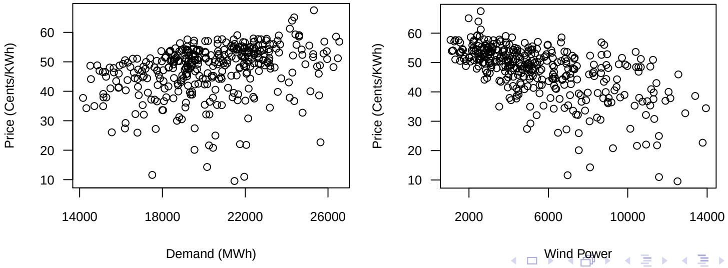
In practice: predict electricity price

Functional nonparametric regression

- Functional explanatory variable: previous daily price curves.
- Scalar response: electricity price for the next day, fixed hour.

Semi-Functional Partial Linear Regression

- Scalar explanatory variable: daily demand, wind power.



Naive bootstrap

From a general functional nonparametric regression model, $Y_i = m(\chi_i) + \varepsilon_i$, built from the sample $\mathcal{S} = \{(\chi_i, Y_i)\}_{i=1}^n$:

Homoscedastic model → Naive bootstrap

1. Construct the residuals $\widehat{\varepsilon}_{i,b} = Y_i - \widehat{m}_b(\chi_i)$, $i = 1, \ldots, n$.
2. Draw n i.i.d random variables $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from the empirical distribution function of $(\widehat{\varepsilon}_{1,b} - \overline{\widehat{\varepsilon}}_b, \ldots, \widehat{\varepsilon}_{n,b} - \overline{\widehat{\varepsilon}}_b)$, where $\overline{\widehat{\varepsilon}}_b = n^{-1} \sum_{i=1}^n \widehat{\varepsilon}_{i,b}$.
3. Obtain $Y_i^* = \widehat{m}_b(\chi_i) + \varepsilon_i^*$, $i = 1, \ldots, n \;\Rightarrow \mathcal{S}^* = \{(\chi_i, Y_i^*)\}_{i=1}^n$
4. Define $\widehat{m}_{hb}^*(\chi) = \dfrac{\sum_{i=1}^n K(d(\chi_i, \chi)/h) Y_i^*}{\sum_{i=1}^n K(d(\chi_i, \chi)/h)}$.

Wild bootstrap

Heteroscedastic model $\rightarrow$ Wild bootstrap

1. Construct the residuals $\widehat{\varepsilon}_{i,b} = Y_i - \widehat{m}_b(\chi_i)$, $i = 1, \ldots, n$.
2. Define $\varepsilon_i^* = \widehat{\varepsilon}_{i,b} V_i$, $i = 1, \ldots, n$, where $V_1, \ldots, V_n$ are i.i.d. random variables that are independent of the data $\mathcal{S}$ and that satisfy $E(V_1) = 0$ and $E(V_1^2) = 1$.
3. Obtain $Y_i^* = \widehat{m}_b(\chi_i) + \varepsilon_i^*$, $i = 1, \ldots, n \ \Rightarrow \mathcal{S}^* = \{(\chi_i, Y_i^*)\}_{i=1}^n$
4. Define $\widehat{m}_{hb}^*(\chi) = \dfrac{\sum_{i=1}^n K(d(\chi_i, \chi)/h) Y_i^*}{\sum_{i=1}^n K(d(\chi_i, \chi)/h)}$.

Notation

For a given fixed element χ_0 of the space $\mathcal{H}$, we denote:

$$\begin{aligned} B(\chi_0, l) &= \{\chi_1 \in \mathcal{H} \text{ such that } d(\chi_1, \chi_0) \leq l\}, \\ F_{\chi_0}(l) &= P(\boldsymbol{\chi} \in B(\chi_0, l)) \text{ for } l > 0, \\ \varphi_{\chi_0}(s) &= E(m(\boldsymbol{\chi}) - m(\chi_0) | d(\boldsymbol{\chi}, \chi_0) = s) \\ \tau_{h\chi_0}(s) &= F_{\chi_0}(hs)/F_{\chi_0}(h) \text{ for } s \in (0, 1] \end{aligned}$$

and

$$\tau_{0\chi_0}(s) = \lim_{h \downarrow 0} \tau_{h\chi_0}(s).$$

Notation

$$\begin{aligned} M_{0\chi_0} &= K(1) - \int_0^1 (sK(s))' \tau_{0\chi_0}(s) ds, \\ M_{1\chi_0} &= K(1) - \int_0^1 K'(s) \tau_{0\chi_0}(s) ds, \\ M_{2\chi_0} &= K^2(1) - \int_0^1 (K^2(s))' \tau_{0\chi_0}(s) ds \end{aligned}$$

and

$$\Theta(s) = \max\{\max_{i \neq j} P(d(\boldsymbol{\chi}_i, \chi_0) \leq s, d(\boldsymbol{\chi}_j, \chi_0) \leq s), F^2_{\chi_0}(s)\}.$$

Assumptions for the convergence of $\widehat{m}_h(\chi)$

Distribution

$m(\cdot)$ and $\sigma_\varepsilon^2(\cdot)$ are continuous on a neighbourhood of χ_0; $\sigma_\varepsilon^2(\chi_0) > 0$.

$$F_{\chi_0}(0) = 0 \text{ and } \varphi_{\chi_0}(0) = 0 \text{ and } \varphi'_{\chi_0}(0) \text{ exists.}$$

$$\forall s \in [0,1], \quad \lim_{n\to\infty} \tau_{h\chi_0}(s) = \tau_{0\chi_0}(s) \text{ with } \tau_{0\chi_0}(s) \neq 1_{[0,1]}(s).$$

Moments

$$\exists p > 2, \ \exists M > 0 \text{ such that } \mathbb{E}(|\varepsilon|^p|\boldsymbol{\chi}) \leq M \ a.s.$$

$$\max\{\mathbb{E}(|Y_i Y_j|^p|\boldsymbol{\chi}_i, \boldsymbol{\chi}_j), \mathbb{E}(|Y_i|^p|\boldsymbol{\chi}_i, \boldsymbol{\chi}_j)\} \leq M \ a.s. \ \forall i,j \in \mathbb{Z}.$$

Small ball probabilities

$$h(nF_{\chi_0}(h))^{1/2} = \mathcal{O}(1) \text{ and } \lim_{n\to\infty} nF_{\chi_0}(h) = \infty \ .$$

Assumptions for the convergence of $\widehat{m}_h(\chi_0)$

Kernel

$K(\cdot)$ is supported on $[0,1]$, has a continuous derivative on $[0,1)$, $K'(s) \leq 0$ for $s \in [0,1)$ and $K(1) > 0$.

Dependence structure

$\{(\chi_i, Y_i)\}_{i=1}^n$ comes from a α-mixing process with α-mixing coefficients $\alpha(n) \leq Cn^{-a}$, where a is given by:

$$\exists v > 0 \text{ such that } \Theta(h) = \mathcal{O}(F_{\chi_0}(h)^{1+v}) \text{ with } a > \frac{(1+v)p-2}{v(p-2)}$$

$$\exists \gamma > 0/ \quad nF_{\chi_0}(h)^{1+\gamma} \to \infty \text{ and } a > \max\left\{\frac{4}{\gamma}, \frac{p}{p-2} + \frac{2(p-1)}{\gamma(p-2)}\right\}$$

[0] Delsol (2009)

Assumptions for the convergence of $\widehat{m}^*_{hb}(\chi_0)$

Moments

Function $\mathbb{E}(|Y||\boldsymbol{\chi} = \cdot)$ is continuous on a neighbourhood of χ_0, and $\sup_{d(\chi_1,\chi_0)<\delta} \mathbb{E}(|Y|^q|\boldsymbol{\chi} = \chi_1) < \infty$ for some $\delta > 0; \forall q \geq 1$.

Distribution

$\forall(\chi_1, s)$ in neighbourhood of $(\chi_0, 0)$, $\varphi_{\chi_1}(0) = 0$, $\exists\varphi'_{\chi_1}(s), \varphi'_{\chi_1}(0) \neq 0$ and $\varphi'_{\chi_1}(s)$ uniformly Lipschitz continuous, order $0 < \alpha \leq 1$ in (χ_1, s).

$\forall\chi_1 \in \mathcal{H}$, $F_{\chi_1}(0) = 0$ and $F_{\chi_1}(t)/F_{\chi_0}(t)$ Lipschitz continuous, order α in χ_1, uniformly in t in neighbourhood of 0.

---

[0] Ferraty, van Keilegom and Vieu (2010)

Assumptions for the convergence of $\widehat{m}^*_{hb}(\chi_0)$

Distribution

$\forall \chi_1 \in \mathcal{H}$ and $\forall s \in [0,1]$, $\tau_{0\chi_1}(s)$ exists, $\sup_{\chi_1 \in H, s\in[0,1]} |\tau_{h\chi_1}(s) - \tau_{0\chi_1}(s)| = o(1)$, $M_{0\chi_0} > 0$, $M_{2\chi_0} > 0$, $\inf_{d(\chi_1,\chi_0)<\varepsilon} M_{1\chi_0} > 0$ for some $\varepsilon > 0$, and $M_{k\chi_1}$ is Lipschitz continuous of order α for $k = 0, 1, 2$.

$\forall n\, \exists r_n \geq 1$, $l_n > 0$ and curves $\chi_{1n}, \ldots, \chi_{r_n n}$ such that $B(\chi_0, h) \subset \cup_{k=1}^{r_n} B(\chi_{kn}, l_n)$, $r_n = \mathcal{O}(n^{b/h})$ and $l_n = o(b(nF_{\chi_0}(h))^{-1/2})$, $\inf_{d(\chi_1,\chi_0)<\varepsilon} M_{1\chi_0} > 0$ for some $\varepsilon > 0$, $M_{k\chi_1}$ is Lipschitz continuous of order α for $k = 0, 1, 2$.

Small ball probabilities

$$\max\{b, h/b, b^{1+\alpha}(nF_{\chi_0}(h))^{1/2}, (F_{\chi_0}(h)/F_{\chi_0}(b))\log n, n^{1/p}F_{\chi_0}(h)^{1/2}\log n\} = o(1)$$

$$\max\{bh^{\alpha-1}, F_{\chi_0}(b)^{-1}h/b\} = \mathcal{O}(1) \text{ and } \lim_{n\to\infty} F_{\chi_0}(b+h)/F_{\chi_0}(b) = 1.$$

[0] Ferraty, van Keilegom and Vieu (2010)

Validity of the bootstrap

Theorem

Under previous assumptions, for the wild bootstrap procedure, we have that

$$\sup_{y\in\mathbb{R}} \left| P^{\mathcal{S}}\left(\sqrt{nF_\chi(h)}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) \leq y\right) - P\left(\sqrt{nF_\chi(h)}(\widehat{m}_h(\chi) - m(\chi)) \leq y\right) \right| \to 0 \ a.s.$$

In addition, if the model is homoscedastic (i.e. $\sigma^2_\varepsilon(\cdot) = \sigma^2_\varepsilon$), then the same result holds for the naive bootstrap.

Background

Result for independent data

Ferraty, Van Keilegom and Vieu (2010) On the Validity of the Bootstrap in Non-Parametric Functional Regression.

Asymptotic distribution of $\widehat{m}_h(\chi) - m(\chi)$ for independent data

Ferraty, Mas and Vieu (2007) Nonparametric Regression on Functional data: Inference and Practical Aspects.

Asymptotic distribution of $\widehat{m}_h(\chi) - m(\chi)$ for dependent data

Delsol (2009) Advances on asymptotic normality in non-parametric functional time series analysis.

Proof outline

$$P^{\mathcal{S}}\left(\sqrt{nF_\chi(h)}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) \leq y\right) - P\left(\sqrt{nF_\chi(h)}(\widehat{m}_h(\chi) - m(\chi)) \leq y\right) = T_1(y) + T_2(y) + T_3(y)$$

where

$$T_1(y) = P^{\mathcal{S}}\left(\sqrt{nF_\chi(h)}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) \leq y\right) - \Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right) - \widehat{m}_b(\chi)\right)}{\sqrt{nF_\chi(h)Var^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right)}}\right)$$

[0] $P^{\mathcal{S}}$: probability conditionally on $\mathcal{S} = \{(\chi_i, Y_i)\}_{i=1}^n$

Proof outline

$$T_2(y) = \Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right) - \widehat{m}_b(\chi)\right)}{\sqrt{nF_\chi(h)\,Var^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right)}}\right) -$$

$$\Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}\left(\widehat{m}_h(\chi)\right) - m(\chi)\right)}{\sqrt{nF_\chi(h)\,Var\left(\widehat{m}_h(\chi)\right)}}\right)$$

and

$$T_3(y) = \Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}\left(\widehat{m}_h(\chi)\right) - m(\chi)\right)}{\sqrt{nF_\chi(h)\,Var\left(\widehat{m}_h(\chi)\right)}}\right) -$$

$$P\left(\sqrt{nF_\chi(h)}(\widehat{m}_h(\chi) - m(\chi)) \le y\right)$$

Proof outline

$$T_3(y) = \Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}\left(\widehat{m}_h(\chi)\right) - m(\chi)\right)}{\sqrt{nF_\chi(h) Var\left(\widehat{m}_h(\chi)\right)}}\right) - P\left(\sqrt{nF_\chi(h)}(\widehat{m}_h(\chi) - m(\chi)) \leq y\right)$$

Delsol (2009)

$$\frac{\widehat{m}_h(\chi) - \mathbb{E}\left(\widehat{m}_h(\chi)\right)}{\sqrt{Var\left(\widehat{m}_h(\chi)\right)}} \xrightarrow{d} N(0,1), \text{a.s.}$$

$T_3(y) \longrightarrow 0$ a.s. for any fixed value of y.

Proof outline

$$T_1(y) = P^{\mathcal{S}}\left(\sqrt{nF_\chi(h)}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) \leq y\right) -$$

$$\Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right) - \widehat{m}_b(\chi)\right)}{\sqrt{nF_\chi(h)\,Var^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right)}}\right)$$

Lemma: adapted from Ferraty, van Keilegom and Vieu (2010)

$$\frac{\widehat{m}^*_{hb}(\chi) - \mathbb{E}^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right)}{\sqrt{Var^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right)}} \xrightarrow{d} N(0,1), \quad a.s.(P^{\mathcal{S}})$$

$T_1(y) \longrightarrow 0$ a.s. for any fixed value of y.

Proof outline

$$T_2(y) = \Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right) - \widehat{m}_b(\chi)\right)}{\sqrt{nF_\chi(h)Var^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right)}}\right) - \Phi\left(\frac{y - \sqrt{nF_\chi(h)}\left(\mathbb{E}\left(\widehat{m}_h(\chi)\right) - m(\chi)\right)}{\sqrt{nF_\chi(h)Var\left(\widehat{m}_h(\chi)\right)}}\right)$$

Lemma: adapted from Ferraty, van Keilegom and Vieu (2010)

$$\left|\sqrt{nF_\chi(h)}\left(\mathbb{E}\left(\widehat{m}_h(\chi)\right) - m(\chi) - \mathbb{E}^{\mathcal{S}}\left(\widehat{m}^*_{hb}(\chi)\right) + \widehat{m}_b(\chi)\right)\right| \to 0 \text{ a.s.}$$

$$\sup_{y\in\mathbb{R}} |T_2(y)| \to 0 \text{ a.s.},$$

Simulation procedure: building confidence intervals

Given a curve χ and the FNP regression model

$$Y_i = m(\boldsymbol{\chi}_i) + \varepsilon_i \ (i = 1, \ldots, n),$$

where the process $\{(\boldsymbol{\chi}_i, Y_i)\}$ is α-mixing and identically distributed as $(\boldsymbol{\chi}, Y)$, and χ is observed from $\boldsymbol{\chi}$, the true, bootstrap and asymptotic $(1-\alpha)$-confidence intervals for $m(\chi)$ were constructed:

$$I^{true}_{\chi,1-\alpha} = (\widehat{m}_h(\chi) + q^{true}_{\alpha/2}(\chi), \widehat{m}_h(\chi) + q^{true}_{1-\alpha/2}(\chi))$$
$$I^{*}_{\chi,1-\alpha} = (\widehat{m}_h(\chi) + q^{*}_{\alpha/2}(\chi), \widehat{m}_h(\chi) + q^{*}_{1-\alpha/2}(\chi))$$
$$I^{asymp}_{\chi,1-\alpha} = (\widehat{m}_h(\chi) + q^{asymp}_{\alpha/2}(\chi), \widehat{m}_h(\chi) + q^{asymp}_{1-\alpha/2}(\chi))$$

Theoretical quantiles

Theoretical quantiles ($q_p^{true}(\chi)$)

1. Generate n_{MC} samples $\{(\chi_i^s, Y_i^s), i = 1, \dots, n\}_{s=1}^{n_{MC}}$ from FNP Model.
2. Carry out $n_{MC} = 2000$ estimates $\left\{\widehat{m}_h^s(\chi)\right\}_{s=1}^{n_{MC}}$, where $\widehat{m}_h^s(\cdot)$ is the functional kernel estimator derived from the s^{th} sample $\{(\chi_i^s, Y_i^s)\}_{i=1}^{n}$.
3. Compute the set of approximation errors $ERRORS.MC = \left\{\widehat{m}_h^s(\chi) - m(\chi)\right\}_{s=1}^{n_{MC}}$.
4. Compute the theoretical quantile, $q_p^{true}(\chi)$, from the quantile of order p of $ERROR.MC$.

Bootstrap quantiles

Bootstrap quantiles ($q_p^*(\chi)$)

1. Generate the sample $\mathcal{S} = \{(\chi_1, Y_1), \ldots, (\chi_n, Y_n)\}$ from FNP Model.
2. Compute $\widehat{m}_b(\chi)$ over the dataset $\mathcal{S}$.
3. Repeat $B = 500$ times the bootstrap algorithm over $\mathcal{S}$ by using i.i.d. random variables V_i drawn from the two Dirac distributions $0.1(5+\sqrt{5})\delta_{(1-\sqrt{5})/2} + 0.1(5-\sqrt{5})\delta_{(1+\sqrt{5})/2}$, giving the B estimates $\left\{\widehat{m}_{hb}^{*,r}(\chi)\right\}_{r=1}^{B}$.
4. Compute set of bootstrap errors $ERRORS.BOOT \left\{\widehat{m}_{hb}^{*,r}(\chi) - \widehat{m}_b(\chi)\right\}_{r=1}^{B}$.
5. Compute the bootstrap quantile, $q_p^*(\chi)$, from the quantile of order p of $ERRORS.BOOT$.

Asymptotic quantiles

Asymptotic quantiles ($q_p^{asymp}(\chi)$)

1. Generate the sample $\mathcal{S} = \{(\chi_1, Y_1), \ldots, (\chi_n, Y_n)\}$ from FNP Model.
2. Use the sample $\mathcal{S}$ to estimate the constants $F_\chi(h)$, $M_{1\chi}$, $M_{2\chi}$ and σ_ε as suggested in Delsol (2009).
3. Compute the asymptotic quantile, $q_p^{asymp}(\chi)$, from the quantile of order p of the corresponding normal distribution.

Simulation procedure

- $\widehat{m}_h(\chi)$ in each of the three intervals was obtained from $\mathcal{S}$
- Test sample $\mathcal{C} = \{\chi_1, \ldots, \chi_{n_\mathcal{C}}\}$, consisting in $n_\mathcal{C} = 100$ independent curves
- Empirical coverages: repeat the procedure $M = 500$ times and computing the proportion of times that each interval contains the value $m(\chi)$

Model 1: Smooth curves

FNP regression model

$$Y_i = m(\chi_i) + \varepsilon_i$$

Functional covariate

$$\chi_i(t_j) = \cos(a_i + \pi(2t_j - 1))$$

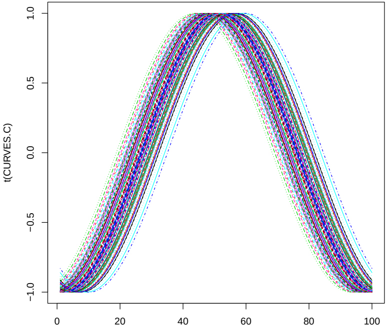
Regression operator

$$m(\chi) = \frac{1}{2\pi} \int_{1/2}^{3/4} (\chi'(t))^2 dt$$

- $\{a_i\} \sim$ AR(1) gaussian process with correlation coefficient $\rho_a = 0.7$ and variance $\sigma_a^2 = 0.05$
- $0 = t_1 < \cdots < t_{100} = 1$
- $\{\varepsilon_i\} \sim N(0, \sigma^2)$, $\sigma^2 = 0.1\, Var(m(\chi_1), \ldots, m(\chi_n))$
- semi-metric $d_1^{deriv}(\cdot, \cdot)$

$$d_1^{deriv}(\chi_i, \chi_j) = \sqrt{\int_0^1 (\chi_i'(t) - \chi_j'(t))^2 dt},$$

Simulated data: Model 1



Model 2: Rough curves

FNP regression model

$$Y_i = m(\chi_i) + \varepsilon_i$$

Functional covariate

$$\chi_i(t_j) = b_{2i}\cos(b_{1i}t_j) + \sum_{k=1}^{j} B_{ik}/b$$

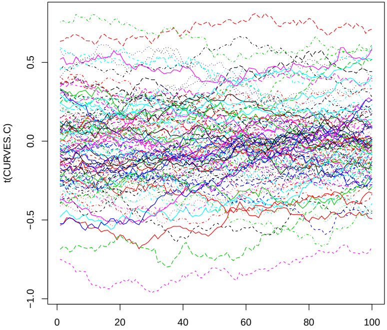
Regression operator

$$m(\chi) = \int_0^{\pi} (\chi(t))^2 dt$$

- $\{b_{1i}\} \sim$MA(1), $\{b_{2i}\} \sim$ AR(1) with $\theta_{b_1} = -0.5$ and $\rho_{b_2} = 0.9$ and $\sigma^2_{b_1} = \sigma^2_{b_2} = 0.1$
- $b = 5,\ B_{ik} \sim N(0, 0.1)$
- $0 = t_1 < \cdots < t_{100} = \pi$
- $\{\varepsilon_i\} \sim N(0, \sigma^2)$, $\sigma^2 = 0.1\,Var(m(\chi_1), \ldots, m(\chi_n))$
- semi-metric $d_4^{proj}(\cdot,\cdot)$

$$d_4^{proj}(\chi_i, \chi_j) = \sqrt{\sum_{k=1}^{4}\left(\int_0^{\pi}(\chi_i(t) - \chi_j(t))v_k(t)dt\right)^2}.$$

Simulated data: Model 2



Average over $\mathcal{C}$ of the empirical coverage of the true, bootstrap and asymptotic confidence intervals.

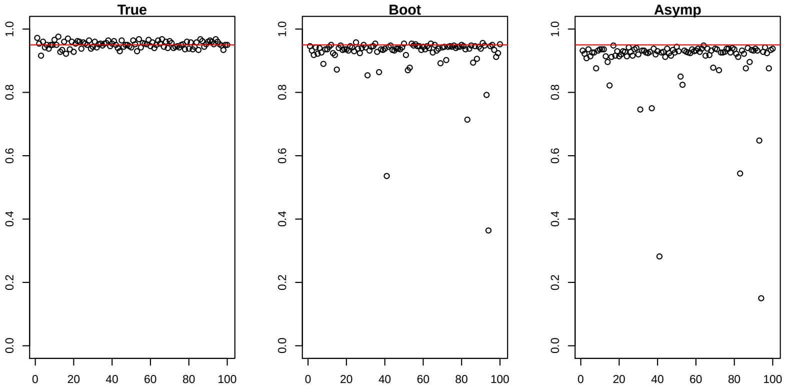
Model 1: smooth curves

$1-\alpha$	0.95		0.90	
n	100	250	100	250
I^{true}	0.95 (0.12)	0.95 (0.01)	0.90 (0.02)	0.90 (0.02)
I^{*}	0.89 (0.12)	0.92 (0.08)	0.85 (0.12)	0.88 (0.08)
I^{asymp}	0.85 (0.14)	0.90 (0.11)	0.79 (0.14)	0.84 (0.12)

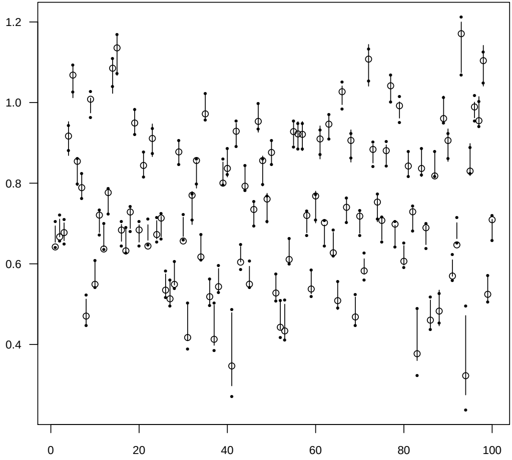
Model 2: rough curves

$1-\alpha$	0.95		0.90	
n	100	250	100	250
I^{true}	0.95 (0.01)	0.95 (0.01)	0.90 (0.02)	0.90 (0.02)
I^{*}	0.80 (0.18)	0.89 (0.07)	0.77 (0.18)	0.86 (0.07)
I^{asymp}	0.76 (0.17)	0.82 (0.06)	0.69 (0.16)	0.75 (0.06)

Model 1: CI coverage.

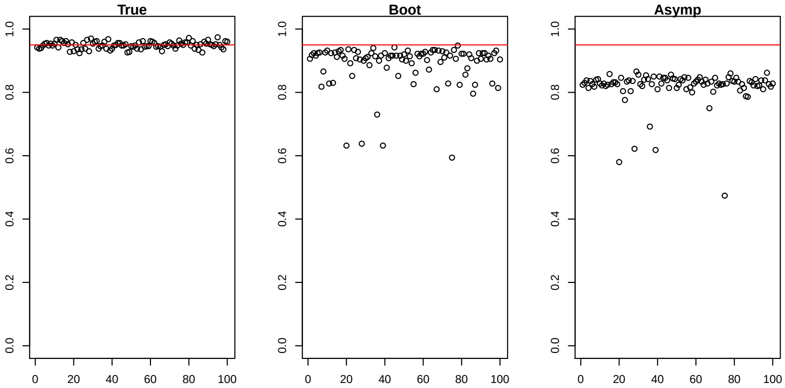


Model 1: Confidence interval for each χ in $\mathcal{C}$.

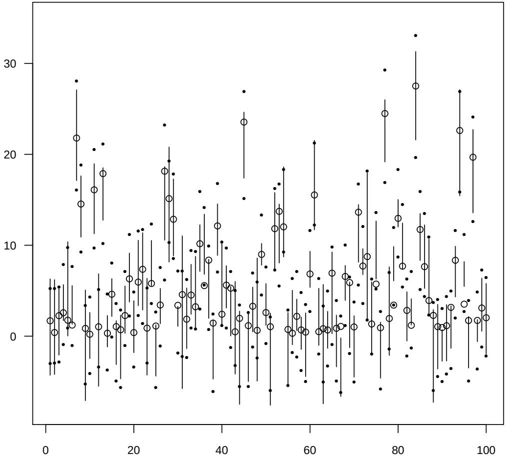


[0] Segment: bootstrap CI, points: true CI.

Model 2: CI coverage.



Model 2: Confidence interval for each χ in $\mathcal{C}$.



[0] Segment: bootstrap CI, points: true CI.

1. Introduction
2. Prediction with functional regression
3. Confidence intervals in FNP
   - Bootstrap
   - Asymptotic theory
   - Simulation study
   - Applications
4. Confidence intervals in SFPLR
   - Bootstrap
   - Asymptotic theory
   - Applications

Electricity demand

Dataset: workdays of the second quarter of the year 2012.

Predict one day (24 hours)

$$\chi_{i+1}(t) = m_t(\chi_i) + \varepsilon_{i,t} \ (t = 1, \ldots, 24, \ i = 1, \ldots, n);$$

Predict one hour for 21 days

$$\chi_{i+1,d}(9) = m_d(\chi_{i,d}) + \varepsilon_{i,d} \ (d = 1, \ldots, 21, \ i = 1, \ldots, n);$$

Confidence intervals for electricity demand

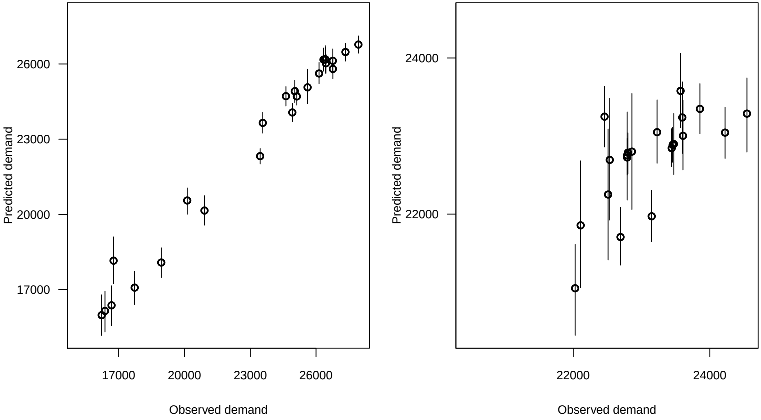


Figure : Left: Bootstrap CI for the 24 hours of Friday, June 29, 2012. Right: Bootstrap CI the workdays in June, 2012 (fixed hour: 09:00 a.m.).

Confidence intervals for electricity price

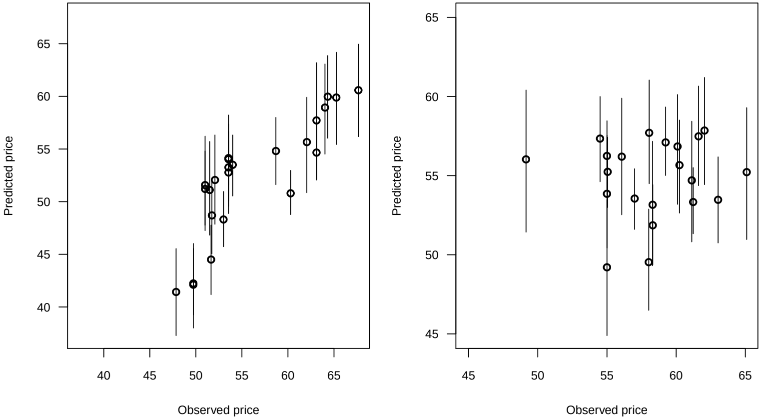


Figure : Left: Bootstrap CI for the 24 hours of Friday, June 29, 2012. Right: Bootstrap CI the workdays in June, 2012 (fixed hour: 09:00 a.m.).

Semi-Funtional Partial Linear Regression

Functional nonparametric explanatory variable, scalar linear-effect covariate and scalar response

Autoregressive model

$$G(\chi_{i+1}) = \boldsymbol{X}_i^T\beta + m(\chi_i) + \varepsilon_i,\ i = 1, \ldots, n$$

General model

$Y_i = \boldsymbol{X}_i^T\beta + m(\chi_i) + \varepsilon_i,\ i = 1, \ldots, n$, where $\{(\boldsymbol{X}_i, \chi_i, Y_i)\}$ is α-mixing

Estimators

$$\widehat{\beta}_h = (\widetilde{\mathbf{X}}_h^T\widetilde{\mathbf{X}}_h)^{-1}\widetilde{\mathbf{X}}_h^T\widetilde{\mathbf{Y}}_h \qquad \widehat{m}_h(\chi) = \sum_{i=1}^n w_h(\chi_i, \chi)(Y_i - \boldsymbol{X}_i^T\widehat{\beta}_h)$$

Naive bootstrap

Homoscedastic model

1. Construct the residuals $\widehat{\varepsilon}_{i,b} = Y_i - \boldsymbol{X}_i^T \widehat{\beta}_b - \widehat{m}_b(\chi_i)$, $i = 1, \ldots, n$.
2. Draw n i.i.d. random variables $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from the empirical distribution function of $(\widehat{\varepsilon}_{1,b} - \overline{\widehat{\varepsilon}}_b, \ldots, \widehat{\varepsilon}_{n,b} - \overline{\widehat{\varepsilon}}_b)$, where $\overline{\widehat{\varepsilon}}_b = n^{-1} \sum_{i=1}^n \widehat{\varepsilon}_{i,b}$.
3. Obtain $Y_i^* = \boldsymbol{X}_i^T \widehat{\beta}_b + \widehat{m}_b(\chi_i) + \varepsilon_i^*$, $i = 1, \ldots, n$.
4. Define
$$\widehat{\beta}_b^* = (\widetilde{\mathbf{X}}_b^T \widetilde{\mathbf{X}}_b)^{-1} \widetilde{\mathbf{X}}_b^T \widetilde{\mathbf{Y}^*}_b$$
and
$$\widehat{m}_{hb}^*(\chi) = \sum_{i=1}^n w_h(\chi_i, \chi)(Y_i^* - \boldsymbol{X}_i^T \widehat{\beta}_b^*),$$

Wild bootstrap

Heteroscedastic model

1. Construct the residuals $\widehat{\varepsilon}_{i,b} = Y_i - \boldsymbol{X}_i^T \widehat{\beta}_b - \widehat{m}_b(\chi_i)$, $i = 1, \ldots, n$.
2. Define $\varepsilon_i^* = \widehat{\varepsilon}_{i,b} V_i$, $i = 1, \ldots, n$, where $V_1, \ldots, V_n$ are i.i.d. random variables that are independent of the data $\mathcal{S}$ and that satisfy $E(V_1) = 0$ and $E(V_1^2) = 1$.
3. Obtain $Y_i^* = \boldsymbol{X}_i^T \widehat{\beta}_b + \widehat{m}_b(\chi_i) + \varepsilon_i^*$, $i = 1, \ldots, n$.
4. Define
$$\widehat{\beta}_b^* = (\widetilde{\mathbf{X}}_b^T \widetilde{\mathbf{X}}_b)^{-1} \widetilde{\mathbf{X}}_b^T \widetilde{\mathbf{Y}^*}_b$$
and
$$\widehat{m}_{hb}^*(\chi) = \sum_{i=1}^{n} w_h(\chi_i, \chi)(Y_i^* - \boldsymbol{X}_i^T \widehat{\beta}_b^*),$$

Assumptions for the linear part of the SFPLR model

Semi-metric space

χ is valued in come given compact subset $\mathcal{C}$ of $\mathcal{H}$ such that
$\mathcal{C} \subset \bigcup_{k=1}^{\tau_n} \mathcal{B}(z_k, l_n)$, where $\tau_n l_n^{\gamma} = C$, $\tau_n \to \infty$ and $l_n \to 0$ as $n \to \infty$.

Kernel

K has support $[0, 1]$, Lipschitz continuous on $[0, \infty)$.
$\exists k / \forall u \in [0, 1], -K'(u) > k > 0$.

Smoothness

Denote $g_j(\chi) = E(X_{ij} | \chi_i = \chi)$, $1 \leq i \leq n, 1 \leq j \leq p$.
All the operators to be estimated are smooth, ie, for some $c < \infty$ and $\alpha > 0$, $\forall (u, v) \in \mathcal{C} \times \mathcal{C}, \forall f \in m, g_1, \ldots, g_p$: $|f(u) - f(v)| \leq c d(u, v)^{\alpha}$.

Assumptions for the linear part of the SFPLR model

Distributions

For the probability distribution of the infinite-dimensional process χ, it is assumed that exists F, positive valued function on $(0,\infty)$ and positive constants $\alpha_0, \alpha_1, \alpha_2$ such that, $\forall t \in \mathcal{C}, h > 0$:

$$\int_0^1 F(hs)ds > \alpha_0 F(h) \text{ and } \alpha_1 F(h) \leq P(\chi \in \mathcal{B}(t,h)) \leq \alpha_2 F(h).$$

The joint probability distribution of $(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j)$ is assumed that exists a function $\psi(h) = cF(h)^{1+\varepsilon}$ $(c > 0, 0 \leq \varepsilon \leq 1)$ and positive constants α_3, α_4 such that $\forall t \in \mathcal{C}, h > 0$:

$$0 < \alpha_3 \psi(h) \leq \sup_{i \neq j} P[(\boldsymbol{\chi}_j, \boldsymbol{\chi}_j) \in \mathcal{B}(t,h) \times \mathcal{B}(t,h)] \leq \alpha_4 \psi(h).$$

Assumptions for the linear part of the SFPLR model

Dependence structure

$\{(\boldsymbol{X}_i, \chi_i, Y_i)\}_{i=1}^n$ come from some stationary strong mixing process, with mixing coefficients $\{\alpha(n)\}$ that verify

$$\alpha(n) \leq cn^{-a}, a > 4.5.$$

while

$$\eta_i \text{ is independent of } \varepsilon_i, (i = 1, \ldots, n),$$

where $\eta_i = (\eta_{i1}, \ldots, \eta_{ip})^T$,
$\eta_{ij} = X_{ij} - E(X_{ij}|\chi_i) = X_{ij} - g_j(\chi), j = 1, \ldots, p.$

Assumptions for the linear part of the SFPLR model

Moments

Denote $V_\varepsilon = E(\varepsilon\varepsilon^T), \varepsilon^T = (\varepsilon_1, \ldots, \varepsilon_n), \eta^T = (\eta_1, \ldots, \eta_n)$.

$$E|Y_1|^r + E|X_{11}|^r + \ldots + E|X_{1p}|^r < \infty \text{ for some } r > 4.$$

$$sup_{i,j} E(|Y_i Y_j||(\chi_i, \chi_j)) < \infty$$

$$\max_{1\leq j\leq p} \sup_{i_1,i_2} E(|X_{i_1 j} X_{i_2 j}||(\chi_{i_1,j}\chi_{i_2,j})) < \infty$$

$$B = E(\eta_1\eta_1^T), C = \lim_{n\to\infty} n^{-1} E(\eta^T V_\varepsilon \eta).$$

B and C are positive definite matrix.

Assumptions for the linear part of the SFPLR model

Moments

$$s_n^{\frac{r(a+1)}{2(a+r)}} = o(n^\theta) \text{ for some } \theta > 2,$$

where $s_n = \sup_{\chi \in \mathcal{C}} (s_{n,1}(\chi) + s_{n,2}(\chi) + s_{n,3}(\chi))$, with

$$s_{n,1}(\chi) = \sum_{i=1}^{n} \sum_{j=1}^{n} |Cov(\Delta_i(\chi), \Delta_j(\chi)| \text{ with } \Delta_i(\chi) = K(\frac{d(\chi_i, \chi)}{h})$$

$$s_{n,2}(\chi) = \sum_{i=1}^{n} \sum_{j=1}^{n} |Cov(\Gamma_i(\chi), \Gamma_j(\chi)| \text{ with } \Gamma_i(\chi) = Y_i K(\frac{d(\chi_i, \chi)}{h})$$

$$s_{n,3}(\chi) = \max_{1 \le k \le p} \sum_{i=1}^{n} \sum_{j=1}^{n} |Cov(\Gamma_{ik}(\chi), \Gamma_{jk}(\chi)| \text{ with } \Gamma_{ik}(\chi) = X_{ik} K(\frac{d(\chi_i, \chi)}{h})$$

Assumptions for the linear part of the SFPLR model

Small ball probabilities

In order to manage the convergence rates found in the development of the Theorem, it is necessary to consider the following assumptions:

$$nh^{4\alpha} \to 0,\ F(h)^{-1} n^{-1/4+1/r} \log n \to 0,\ nF(h)^{\frac{\varepsilon a(r-2)}{r}-1} = \mathcal{O}(1)$$

$$F(h)^{-2} \left(n^{1-\frac{\theta(a+r)}{r(a+1)}}\right)^{-2} \log n = \mathcal{O}(1) \text{ as } n \to \infty$$

where $\alpha > 0, 0 \leq \varepsilon \leq 1, a > 4.5, r > 4$ and $\theta > 2$.

Validity of the bootstrap for the linear part

Theorem (Naive)

Under previous assumptions, if the model is homoscedastic and $\mathbf{a} \in \mathbb{R}^p$, *for the naive bootstrap we have:*

$$\sup_{y\in\mathbb{R}} \left| P^{\mathcal{S}}\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b^* - \widehat{\beta}_b) \leq y\right) - P\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b - \beta) \leq y\right)\right| \to_P 0$$

Theorem (Wild)

Under previous assumptions if, in addition $|\varepsilon_i| < \infty, i = 1, \ldots, n$, $F(h)^{-1} n^{-1/4+1/r} logn(loglogn)^{1/4} \to 0$, $\mathbb{E}|\eta\eta^T| < \infty$, $\mathbb{E}|\eta|^3 < \infty$ *and* $\mathbf{a} \in \mathbb{R}^p$, *for the wild bootstrap procedure we have that*

$$\sup_{y\in\mathbb{R}} \left| P^{\mathcal{S}}\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b^* - \widehat{\beta}_b) \leq y\right) - P\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b - \beta) \leq y\right)\right| \to_P 0$$

Validity of the bootstrap for the nonparametric part

Theorem (Naive and Wild bootstrap)

Under previous assumptions, if $||\boldsymbol{X}_i||_\infty \leq C < \infty$*, we have:*

$$\sup_{y\in\mathbb{R}} \left| P^{\mathcal{S}}\left(\sqrt{nF(h)}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) \leq y\right) - P\left(\sqrt{nF(h)}(\widehat{m}_h(\chi) - m(\chi)) \leq y\right)\right| \to_P 0$$

Electricity demand

Dataset: workdays of the second quarter of the year 2012.

Predict one day (24 hours)

$$\chi_{i+1}(t) = \boldsymbol{X}_i^T\beta + m_t(\chi_i) + \varepsilon_{i,t} \; (t = 1, \ldots, 24, \; i = 1, \ldots, n);$$

Temperature covariates: $\boldsymbol{X}_i = (X_{i1}, X_{i2})^T = (HDD_i, CDD_i)^T$

Model	length: mean (sd)
FNP	1045.92 (353.44)
SFPLR	969.92 (250.00)

Confidence intervals for electricity demand

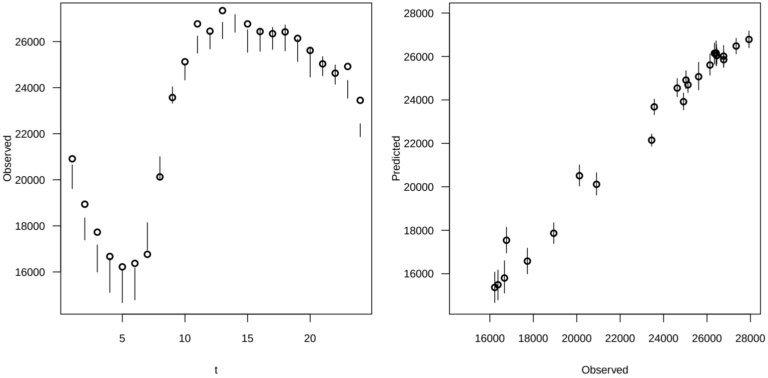


Figure : Bootstrap CI for the 24 hours of Friday, June 29, 2012.

Electricity price

Dataset: workdays of the second quarter of the year 2012.

Predict one day (24 hours)

$$\chi_{i+1}(t) = \mathbf{X}_i^T \beta + m_t(\chi_i) + \varepsilon_{i,t} \;\; (t = 1, \ldots, 24, \;\; i = 1, \ldots, n);$$

Covariates: $\mathbf{X}_i = (X_{i1}, X_{i2})^T = (Demand_i, Wind_i)^T$

Model	length: mean (sd)
FNP	7.44 (1.63)
SFPLR (Demand)	6.50 (1.55)
SFPLR (Demand+Wind)	8.40 (1.21)

Confidence intervals for electricity price

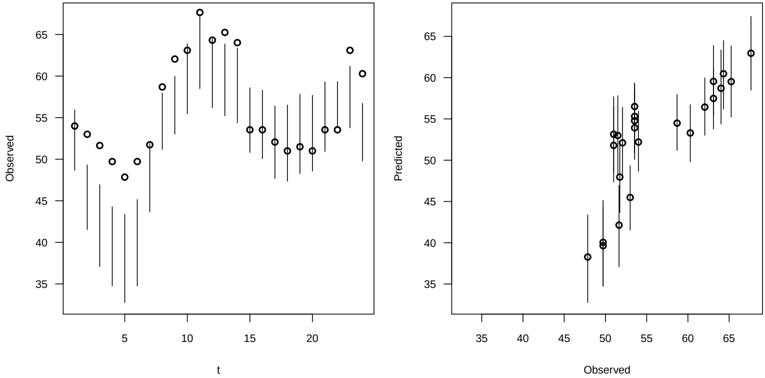


Figure : Bootstrap CI for the 24 hours of Friday, June 29, 2012.

Electricity price

Predict one hour for 21 days

$$\chi_{i+1,d}(20) = \boldsymbol{X}_i^T \beta + m_d(\chi_{i,d}) + \varepsilon_{i,d} \ (d = 1, \ldots, 21, \ i = 1, \ldots, n);$$

Covariates: $\boldsymbol{X}_i = (X_{i1}, X_{i2})^T = (Demand_i, Wind_i)^T$

Model	length: mean (sd)
FNP	6.21 (1.54)
SFPLR (Demand)	6.23 (1.57)
SFPLR (Demand+Wind)	8.34 (2.64)

Confidence intervals for electricity price

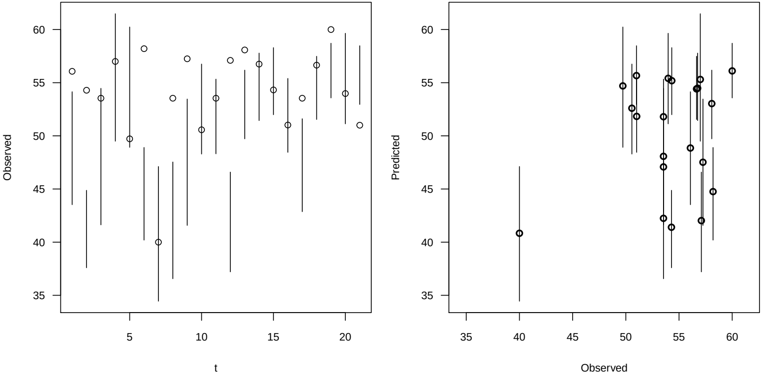


Figure : Bootstrap CI the workdays in June, 2012 (fixed hour: 20:00 a.m.).

References

- Ferraty, F., Van Keilegom, I. and Vieu, P. (2010), On the Validity of the Bootstrap in Non-Parametric Functional Regression, *Scandinavian Journal of Statistics*, 37, 286–306.
- Delsol, L. (2009) Advances on asymptotic normality in non-parametric functional time series analysis, *Statistics*, 43:1, 13–33.
- Raña, P., Aneiros, G., Vilar, J. and Vieu, P. Bootstrap confidence intervals in functional nonparametric regression under dependence. *(Submitted)*.
- Aneiros, G. and Vieu, P. (2008) Nonparametric time series prediction: A semi-functional partial linear modeling, *Journal of Multivariate Analysis*, 99, 834–857.
- Raña, P., Aneiros, G., Vilar, J. and Vieu, P. Bootstrap confidence intervals in semi-functional partial linear regression. *(Preprint)*.

Thanks for your attention!

Proofs outline: linear part

$$P^{\mathcal{S}}\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b^* - \widehat{\beta}_b) \leq y\right) - P\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b - \beta) \leq y\right) = T_1(y) + T_2(y)$$

where $\mathbf{a}$ is a constant vector in $\mathbb{R}^p$,

$$T_1(y) = P^{\mathcal{S}}\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b^* - \widehat{\beta}_b) \leq y\right) - \Phi\left(\frac{y}{\sqrt{\mathbf{a}^T\mathbf{A}\mathbf{a}}}\right)$$

$$T_2(y) = \Phi\left(\frac{y}{\sqrt{\mathbf{a}^T\mathbf{A}\mathbf{a}}}\right) - P\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b - \beta) \leq y\right).$$

Proofs outline: linear part

$$T_2(y) = \Phi\left(\frac{y}{\sqrt{\mathbf{a}^T\mathbf{A}\mathbf{a}}}\right) - P\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b - \beta) \leq y\right).$$

Theorem 1, Aneiros and Vieu (2008)

$$\sqrt{n}(\widehat{\beta}_h - \beta) \longrightarrow^D N(0, \mathbf{A}) \text{where } \mathbf{A} = \mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1}.$$

$T_2(y) \longrightarrow 0$ for any fixed value of y.

Proofs outline: linear part

$$T_1(y) = P^{\mathcal{S}}\left(\sqrt{n}\mathbf{a}^T(\widehat{\beta}_b^* - \widehat{\beta}_b) \leq y\right) - \Phi\left(\frac{y}{\sqrt{\mathbf{a}^T\mathbf{A}\mathbf{a}}}\right)$$

Lemma

$\sqrt{n}(\widehat{\beta}_b^* - \widehat{\beta}_b) \xrightarrow{d}_P N(0, \mathbf{A})$,conditionally on the sample $\mathcal{S}$.

$T_1(y) \longrightarrow 0$ for any fixed value of y.

Proofs outline: linear part

Proof of the Lemma:
For a given function $g(\cdot) = m(\cdot)$ or $g(\cdot) = \widehat{m}_b(\cdot)$, we denote

$$\widetilde{g}_b(\chi) = g(\chi) - \sum_{i=1}^{n} w_b(\boldsymbol{\chi}_i, \chi) g(\boldsymbol{\chi}_i).$$

Then, one can write

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}_b^* - \widehat{\boldsymbol{\beta}}_b) = (n^{-1} \widetilde{\mathbf{X}}_b^T \widetilde{\mathbf{X}}_b)^{-1} n^{-1/2} (S_{n1}^* - S_{n2}^* + S_{n3}^*).$$

Asymptotic normality is obtained by:

$$S_{n1}^* - S_{n2}^* + S_{n3}^* = \sum_{i=1}^{n} \eta_i \varepsilon_i^* + o_P(n^{1/2}) \; (P^{\mathcal{S}}),$$

and

$$n^{-1/2} \sum_{i=1}^{n} \eta_i \varepsilon_i^* \xrightarrow{D} N(0, \mathbf{C}) \text{ , in } P^{\mathcal{S}},$$

Proofs outline: nonparametric part

$$\sup_{y\in\mathbb{R}} \left| P^{\mathcal{S}} \left(\sqrt{nF(h)}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) \leq y \right) - P \left(\sqrt{nF(h)}(\widehat{m}_h(\chi) - m(\chi)) \leq y \right) \right| \to_P 0$$

$$(nF(h))^{1/2}(\widehat{m}_h(\chi) - m(\chi)) \longrightarrow N(0, \sigma^2(\chi))$$

$$(nF(h))^{1/2}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) \longrightarrow N(0, \sigma^2(\chi))$$

Proofs outline: nonparametric part

$$(nF(h))^{1/2}(\widehat{m}_h(\chi) - m(\chi)) =$$

$$(nF(h))^{1/2}(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)(Y_i - \boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_h) - m(\chi))) =$$

$$(nF(h))^{1/2}(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)(\boldsymbol{X}_i^T \boldsymbol{\beta} + m(\boldsymbol{\chi}_i) + \varepsilon_i - \boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_h) - m(\chi))) =$$

$$(nF(h))^{1/2}(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)(m(\boldsymbol{\chi}_i) + \varepsilon_i) - m(\chi)) -$$

$$-(nF(h))^{1/2} \sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi) \boldsymbol{X}_i^T (\widehat{\boldsymbol{\beta}}_h - \boldsymbol{\beta}) =$$

$$S_1(\chi) - S_2(\chi)$$

Proofs outline: nonparametric part

$$
\begin{aligned}
S_1(\chi) &= (nF(h))^{1/2}\Big(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i,\chi)(m(\boldsymbol{\chi}_i)+\varepsilon_i) - m(\chi)\Big) = \\
&= (nF(h))^{1/2}(\widehat{m}_h^{NP}(\chi) - m^{NP}(\chi))
\end{aligned}
$$

Delsol (2009)

$$
(nF(h))^{1/2}(\widehat{m}_h^{NP}(\chi) - m^{NP}(\chi)) \longrightarrow^{D} N(0,\sigma^2(\chi))
$$

$$
S_1(\chi) \longrightarrow^{D} N(0,\sigma^2(\chi))
$$

Proofs outline: nonparametric part

$$S_2(\chi) = (nF(h))^{1/2} \sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi) \mathbf{X}_i^T (\widehat{\beta}_h - \beta)$$

Theorem 1, Aneiros and Vieu (2008)

$$\sqrt{n}(\widehat{\beta}_h - \beta) \longrightarrow^D N(0, \mathbf{A}) \text{where } \mathbf{A} = \mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1}.$$

Lemma

$\max |w_h(\boldsymbol{\chi}_i, \chi)| = \mathcal{O}((nF(h))^{-1})$

$$S_2(\chi) = o_P(1)$$

Proofs outline: nonparametric part

$$
\begin{aligned}
&(nF(h))^{1/2}(\widehat{m}^*_{hb}(\chi) - \widehat{m}_b(\chi)) = \\
&(nF(h))^{1/2}\Big(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)(Y_i^* - \boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_b^*) - \widehat{m}_b(\chi)\Big) = \\
&(nF(h))^{1/2}\Big(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)(\boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_b + \widehat{m}_b(\boldsymbol{\chi}_i) + \varepsilon_i^* - \boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_b^*) - \widehat{m}_b(\chi)\Big) \\
&(nF(h))^{1/2}\Big(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)(\widehat{m}_b(\boldsymbol{\chi}_i) + \varepsilon_i^*) - \widehat{m}_b(\chi)\Big) - \\
&-(nF(h))^{1/2} \sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)\boldsymbol{X}_i^T(\widehat{\boldsymbol{\beta}}_b^* - \widehat{\boldsymbol{\beta}}_b) = \\
&S_1^*(\chi) - S_2^*(\chi)
\end{aligned}
$$

Proofs outline: nonparametric part

$$\begin{aligned} S_1^*(\chi) = & \quad (nF(h))^{1/2}\Big(\sum_{i=1}^{n} w_h(\boldsymbol{\chi}_i, \chi)(\widehat{m}_b(\boldsymbol{\chi}_i) + \varepsilon_i^*) - \widehat{m}_b(\chi)\Big) \\ = & \quad S_{1,1}^*(\chi) + S_{1,2}^*(\chi) \end{aligned}$$

$S_{1,1}^*(\chi)$ contains the nonparametric part of the expression.
$S_{1,2}^*(\chi)$ contains the linear part of the expression.

Proofs outline: nonparametric part

$$S^*_{1,1}(\chi) = \qquad (nF(h))^{1/2}(\widehat{m}^{*NP}_{hb}(\chi) - \widehat{m}^{NP}_{b}(\chi)) \longrightarrow^{D} N(0, \sigma^2(\chi))$$

Raña, Aneiros, Vilar and Vieu

$$\sup_{y\in\mathbb{R}} \left| P^{\mathcal{S}}\left(\sqrt{nF_\chi(h)}(\widehat{m}^{*NP}_{hb}(\chi) - \widehat{m}^{NP}_{b}(\chi)) \leq y\right) - P\left(\sqrt{nF_\chi(h)}(\widehat{m}^{NP}_{h}(\chi) - m^{NP}(\chi)) \leq y\right) \right| \to 0 \text{ a.s.}$$

Delsol (2009)

$$(nF(h))^{1/2}(\widehat{m}^{NP}_{h}(\chi) - m^{NP}(\chi)) \longrightarrow^{D} N(0, \sigma^2(\chi))$$

Proofs outline: nonparametric part

$$S^*_{1,2}(\chi) = (nF(h))^{1/2}(\sum_{i=1}^{n} w_h(\chi_i,\chi)[\sum_{j=1}^{n} w_b(\chi_j,\chi_i)\mathbf{X}_j^T(\beta-\widehat{\beta}_b)+$$

$$+\mathbf{X}_j^T(\beta-\widehat{\beta}_b)-\sum_{l=1}^{n} w_b(\chi_l,\chi_j)\mathbf{X}_l^T(\beta-\widehat{\beta}_b)-\frac{1}{n}\sum_{k=1}^{n}(\mathbf{X}_k^T(\beta-\widehat{\beta}_b) -$$

$$\sum_{l=1}^{n} w_b(\chi_l,\chi_k)\mathbf{X}_l^T(\beta-\widehat{\beta}_b))] - \sum_{i=1}^{n} w_b(\chi_i,\chi)\mathbf{X}_i^T(\beta-\widehat{\beta}_b))$$

Aneiros and Vieu (2008)

$$\sqrt{n}(\widehat{\beta}_h-\beta) \longrightarrow^D N(0,\mathbf{A})$$

Assumption

$$||\mathbf{X}_i|| \leq C < \infty$$

$$S^*_{1,2}(\chi) = o_P(1)(P^{\mathcal{S}}).$$

Proofs outline: nonparametric part

$$S_2^*(\chi) = (nF(h))^{1/2} \sum_{i=1}^{n} w_h(\chi_i, \chi) \mathbf{X}_i^T (\widehat{\beta}_b^* - \widehat{\beta}_b) = o_P(1)(P^{\mathcal{S}})$$

Raña, Aneiros, Vilar and Vieu

$$\sup_{y \in \mathbb{R}} \left| P^{\mathcal{S}} \left(\sqrt{n} \mathbf{a}^T (\widehat{\beta}_b^* - \widehat{\beta}_b) \leq y \right) - P \left(\sqrt{n} \mathbf{a}^T (\widehat{\beta}_b - \beta) \leq y \right) \right| \to_P 0$$

Aneiros and Vieu (2008)

$$\sqrt{n}(\widehat{\beta}_h - \beta) \longrightarrow^D N(0, \mathbf{A})$$

Assumption

$$||\mathbf{X}_i|| \leq C < \infty$$