Filomat 36:14 (2022), 4919–4933
https://doi.org/10.2298/FIL2214919B

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Local Linear Estimation of the Trimmed Regression for Censored Data

Ataouia Bakhtaoui[a], **Faiza Limam-Belarbi**[a]

[a]*Laboratory of Statistics and Stochastic Processes, Department of Probability and Statistics, Institution Djillali Liabes University, Sidi Bel Abbes 22000, Algeria*

Abstract. We introduce a non-parametric estimation of the trimmed regression by using the local linear method of a censored scalar response variable, given a functional covariate. The main result of this work is the establishment of almost complete convergence for the constructed estimator. A simulation study is carried out to compare the finite sample performance based on the mean square error between the classic local linear regression estimator and the trimmed local linear regression estimator. Moreover, a real data study is used to illustrate our methodology.

1. Introduction

The theory of nonparametric estimation has developed considerably over the past two decades because it is important to the field of research in the statistic.

In statistics, a trimmed estimator is an estimator derived from another estimator by excluding some of the extreme values. In particular, it is defined by minimizing a robust measure of the scatter of the residuals.

Linear regression is one of the most widely used statistical models to modeling the relationship between a scalar response and one or more explanatory variables (see [10]) on the importance of this approach.

It is well known that functional statistics have experienced very significant development in recent years. In the case of functional data analysis, the local linear method dates back to [2–3], [7–9], [13], and [19]. Whereas, in the case of a linear model and censored variables, see [4–6], and [18]. Recently, many topics concerning the analysis of functional and censored data have been developed. We refer to [1], [16–17], and [20].

As far as we know, the local linear estimation of the trimmed regression combining censored and functional data has not been studied in the statistical literature. Therefore, we construct in Section 2 of our paper a new non-parametric robust estimator of the regression function based on the idea of least trimmed squares (See [22]), (p.135,181,190) and [12]. Specifically, we use also the functional local linear procedure proposed by [3]. We fix the notations, assumptions, and state our main result in section 3. In Section

2020 *Mathematics Subject Classification*. Primary 62G35; Secondary 62R10 62G05, 62G08, 62N01
Keywords. Nonparametric robustness, Functional data analysis, Nonparametric regression, Nonparametric estimation, Robust estimation, Censored data.
Received: 18 August 2021; Revised: 03 December 2021; Accepted: 03 June 2022
Communicated by Biljana Popović
Email addresses: ataouia.30.01.1993@gmail.com (Ataouia Bakhtaoui), faiza_belarbi@yahoo.fr (Faiza Limam-Belarbi)

4, a numerical study and a real example show the performances of the proposed methodology for finite samples. We give the proof of the result in section 5. Finally, we give our conclusion, in the last section.

2. The trimmed local linear estimator

Let $(X_i, Y_i)_{i=1,...,n}$ be n copies of (r.v.), independent and identically distributed as (X, Y). The latter is valued in $\mathcal{F} \times \mathbb{R}$, where $(\mathcal{F}, d)$ is a semi-metric space (i.e. X is a functional random variable (f.r.v) and d a semi-metric).
In the censoring model, we observe the censored lifetimes C instead of observing the lifetimes Y (with common unknown absolutely continuous distribution function F with density f). Supposing that (C_i) is a sequence of independent and identically distributed censoring random variable with common unknown continuous survival function $\bar{G}(.)$. The continuity of G allows using the convergence results for [14] estimator of G.
We remark the pairs $(T_{(i)}, \delta_{[i]})$ where

$$T_{(i)} = \min(Y_i, C_i) \qquad \text{and} \qquad \delta_{[i]} = \mathbb{1}_{\{Y_i \leq C_i\}}.$$

Where $\mathbb{1}$ denotes the indicator of no censoring. $(X_i, Y_i)_{i=1,...,n}$ and $(C_i)_{i=1,...,n}$ are independent.
For $x \in \mathcal{F}$, The nonparametric trimmed regression, denoted by θ_x is solution with respect to t of following problem

$$\mathbb{E}\left[\frac{\delta\psi(T-t)}{\overline{G}(t)}|X = x\right] = 0 \ \ t \in \mathbb{R}$$

where
✓ $\psi(y) = \mathbb{1}_{|y| \leq q}$; With $\mathbb{1}$ is indicator function,
✓ $q = F^{-1}(1-\alpha/2)$ for $\alpha \in \left]0, \frac{1}{2}\right[$ with F is a cumulative distribution function which has a symmetric density; and

$$\begin{aligned}\overline{G}_n(t) &= 1 - G_n(t) \\ &= \begin{cases} \prod_{i=1}^{n}\left(1 - \frac{1-\delta_{[i]}}{n-i+1}\right)^{\mathbb{1}_{\{T_{(i)} \leq t\}}} & \text{if } t \leq T_{(n)} \\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

where $T_{(1)} \leq ... \leq T_{(n)}$ are the order statistics of $T_{(i)}$ and $\delta_{[i]}$ concomitant with $T_{(i)}$.
The basic idea of trimmed approach is to replace the set of n residuals by the subset which contains just the residuals which are between $-q$ and q. Precisely, it keeps the square function and eliminates a percentage of residuals.(See [22])for more discussions).
Our main purpose of this paper is to study the functional local linear estimate of the trimmed regression function by adopting the fast functional local modeling proposed by [3] for which the function θ_x is approximated by

$$\forall\, \mathcal{X} \quad \text{in neighborhood of } x \qquad \theta_{\mathcal{X}} = a + b\beta(x, \mathcal{X})$$

Where a and b are estimated by $\hat{a}$ and $\hat{b}$ are solution of

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \frac{\delta_{[i]}}{\overline{G}(t_i)} \varrho\left(T_{(i)} - a - b\beta(X_i, x)\right) K\left(h^{-1}\Delta(x, X_i)\right)$$

Where
✓ ϱ is the primitive of ψ
✓ $\beta(.,.)$ is a known function from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$ such that, $\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0$,
✓ K is a kernel function and $h = h_n$ (to simplify the notations) is a sequence of positive real numbers which goes to zero as n goes to infinity,
✓ d denotes the semi-metric and $d(.,.) = |\Delta(.,.)|$ is a function of $\mathcal{F}^2$. A natural estimator of θ_x denoted by $\hat{\theta}_x$.

3. Assumptions and main result

In what follows, when no confusion is possible, we will denote by C_1 and C_2 some strictly positive generic constants. Moreover, x denotes a fixed point in $\mathcal{F}$, $\mathcal{N}_x$ denotes a fixed neighborhood of x. In the remainder of this paper, we set:

$$\begin{cases} K_i = K(h^{-1}\Delta(x, X_i)) \text{ and } \beta_i = \beta(X_i, x) \text{ for i=1,.....,n} \\ \phi_x(u) = \mathbb{P}(|\Delta(x, X)| \leq u) \text{ where } B(x, u) = \{x' \in \mathcal{F} / |\Delta(x', x)| \leq u\} \end{cases}$$

(H1) $\forall u > 0 \quad \phi_x(u) = \phi_x(-u, u) > 0$ and there exists a function $\Phi_x(.)$ such that:

$$\forall t \in (-1, 1), \lim_{h \longrightarrow 0} \frac{\phi_x(th, h)}{\phi_x(h)} = \Phi_x(t)$$

(H2) The function τ is such that:

$$\tau_\lambda(x, .) := \mathbb{E}\left[\left|\frac{\delta\psi^\lambda(Y_i - .)}{\overline{G}^\lambda(t)}\right| |X = x\right] \text{ is of class } C^1 \text{ on } [\theta_x - \Delta, \theta_x + \Delta], \Delta > 0 \text{ and } \lambda \in \{1, 2\}.$$

$$\begin{cases} \forall (t_1, t_2) \in [\theta_x - \Delta, \theta_x + \Delta] \times [\theta_x - \Delta, \theta_x + \Delta], \forall (x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x \text{ and for } (b_1, b_2) > 0 \\ (i) \qquad |\tau_\lambda(x_1, t_1) - \tau_\lambda(x_2, t_2)| \leq C_1 d^{b_1}(x_1, x_2) + |t_1 - t_2|^{b_2}. \\ (ii) \qquad |\nu(x_1, t_1) - \nu(x_2, t_2)| \leq C_2 d^{b_1}(x_1, x_2) + |t_1 - t_2|^{b_2}, \text{where } \nu(x, .) = \frac{d}{dt}\tau_1(x, .). \end{cases}$$

(H3) The functional operator β satisfies the following three conditions:

- $\forall \mathcal{X} \in \mathcal{F}, C_1|\Delta(x, \mathcal{X})| \leq |\beta(\mathcal{X}, x)| \leq C_2|\Delta(x, \mathcal{X})|$;
- $\sup_{r \in B(x,u)} |\beta(r, x)| - |\Delta(x, r)| = o(u)$

(H4) K is a differentiable function, knowing that its support is $[-1, 1]$ such that

$$D = \begin{pmatrix} K(1) - \int_{-1}^{1} tK'(t)\Phi_x(t)dt & K(1) - \int_{-1}^{1} (tK(t))'\Phi_x(t)dt \\ K(1) - \int_{-1}^{1} (tK(t))'\Phi_x(t)dt & K(1) - \int_{-1}^{1} (t^2K(t))'\Phi_x(t)dt \end{pmatrix}$$

is a positive definite matrix.
(H5) h is a positive sequel such as $\lim_{n \longrightarrow \infty} h = 0$ and $\lim_{n \longrightarrow \infty} \frac{\log n}{n\phi_x(h)} = 0$.
Our main result is the following theorem

Theorem 3.1. *Under hypotheses (H1)-(H5) and if* $\nu(x, \theta_x)>0$ *we have*

$$|\hat{\theta}_x - \theta_x| = O(h^{\min(b_1, b_2)}) + O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right) \qquad a.co$$

4. Numerical studies

4.1. Simulation study on the finite samples

The main objective of this subsection is to evaluate the performance of the trimmed approach by comparing between the local linear method and the classical kernel method (constant local), on finite samples. We generate our functional observations $X_i, \{i = 1, ..., n\}$ by using the following process:

$$X_i(t) = \sin(4(b_i - t)\pi) + a_i t^2,$$

with b_i is distributed as $\mathcal{N}(0, 1)$, while the n random variables a_i 's are generated according to a $\mathcal{N}(4, 3)$ distribution. All the curves X_i are discretized on the same grid which is composed of 150 equidistant values in $[0, 1]$ and they are represented in Figure 1.

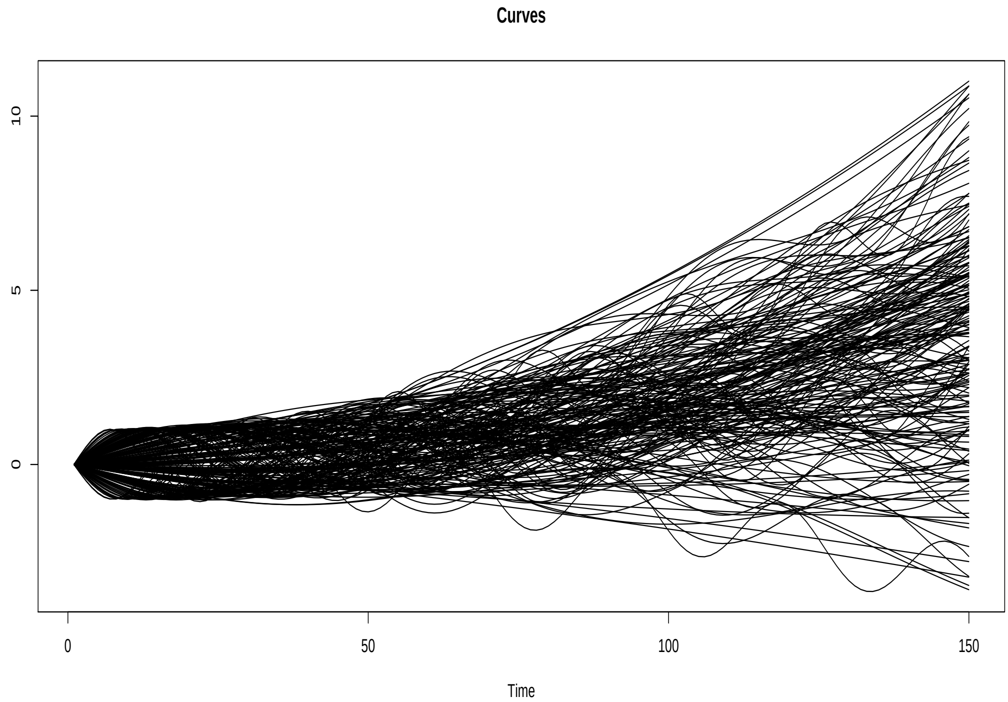

Figure 1: The curves $X_{i=1,...,150}(t),\ \ t \in [0, 1]$.

Then, we define the scalar response variables Y_i by

$$Y_i = r(X_i) + \varepsilon_i,\ \ i = 1, ..., n$$

where the $\varepsilon_i \rightsquigarrow \exp(0.5)$, and the operator r is defined by

$$r(X_i) = \exp\left\{-\int_0^1 \frac{dt}{1 + X_i^2(t)}\right\} \text{ for } i = 1, \dots, n.$$

Recall now that our estimator $\widehat{\theta}_x$ which is defined by the solution of the following minimization problem:

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \frac{\delta_{[i]}}{\overline{G}(t_i)} \varrho\left(T_{(i)} - a - b\beta(X_i, x)\right) K\left(h^{-1}\Delta(x, X_i)\right)$$

where ϱ is the primitive function of ψ,

We obtain the definition of the constant local regression trimmed (respectively the local linear regression trimmed) if we work with a function $\psi(y) = y\mathbb{1}_{|y|\leq q}$ where $\mathbb{1}$ is an indicator function and $b = 0$ (respectively $\psi(y) = y\mathbb{1}_{|y|\leq q}$, and $b \neq 0$)

Our purpose is to compare the Mean Square Error (MSE) of the estimator of the Classic Local Linear (CLL) and the Trimmed Local Linear regression estimator (TLL) with the censored data set. To this end, we choose $\alpha = 0.75$ and F is the distribution function of ε_i. We select the optimal bandwidth h, for the two regressions models, by the cross validation method on the k nearest neighbors in a local way and we use the quadratic kernel which is defined by

$$K(x) = \frac{3}{4}\left(1 - x^2\right)\mathbb{1}_{[-1,1]}.$$

On the other hand, the choice of the locating functions has an important role in the local linear estimate. It's clear that, we can take $\beta(.,.) = \Delta(.,.)$ in theory but the operators β and Δ do not play the same role. Furthermore, the choices of two locating functions δ and β depend on the shape of data. Here the curves X_i 's are sufficiently smooth, then we take the following types

$$\delta\left(x, x'\right) = \left(\int_0^1 \left(x^{(i)}(t) - x'^{(i)}(t)\right)^2 dt\right)^{1/2} \text{ and } \beta\left(x, x'\right) = \int_0^1 \theta(t)\left(x^{(j)}(t) - x^{(j)}(t)\right) dt$$

where $x^{(i)}$ denotes the i^{th} derivative of the curve x and θ is the eigenfunction of the empirical covariance operator $\frac{1}{n}\sum_{i=1}^{n}\left(X_i^{(j)} - \overline{X^{(j)}}\right)^t \left(\left(X_i^{(j)} - \overline{X^{(j)}}\right)\right)$ associated to the q greatest eigenvalues.

In this simulation study, we have worked with several values of i, q and j, but, for the sake of shortness, we present only the results of the case where $i = 2$, $j = 1$ and $q = 3$.

In this simulation, to illustrate the performance of our estimator, we proceed as follows.
- Step 1. We generate $n = 150$ independent replications of $(X_i, T_i, \Delta_i)_{i=1,\ldots,n}$
- Step 2. We divide our observations into two subsets:

$$\begin{aligned} &-(X_i, T_i, \Delta_i)_{i=1,\ldots,100}, \text{ training sample.} \\ &-(X_i, T_i, \Delta_i)_{j=101,\ldots,150}, \text{ test sample.} \end{aligned}$$

- Step 3. We calculate the two estimators by using the learning sample and we find the Classic Local Linear (CLL) and the Trimmed Local Linear (TLL) estimators.
- Step 4. We present our results by plotting the boxplot of the prediction error are represented in (Figure 2.) and we compute the empirical mean square error with :

$$\frac{1}{50}\sum_{i=101}^{150}\left(Y_i - \widehat{\theta}_{X_i}\right)^2$$

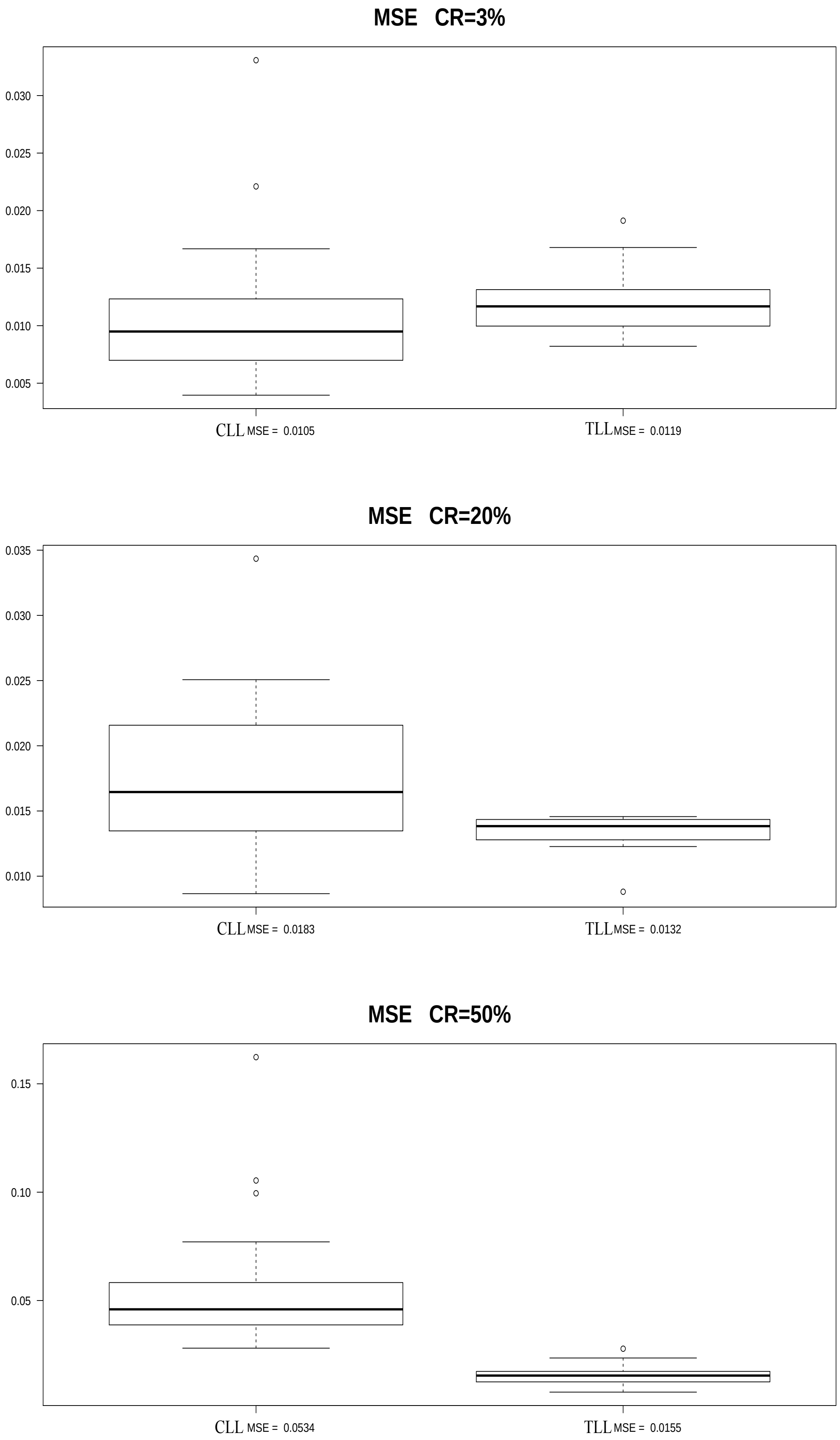

Figure 2: Comparison results betwenn CLL estimator and the TLL estimator.

It is clear to our estimator has a good performance and seems to outperform the classic local linear estimator even for censored data. This is confirmed by the mean squared error MSE(CLL) and MSE(TLL) This conclusion shows the good performance of our approach.

Table 4.1: MSE for the CLL estimator and the TLL estimator according to numbers of introduced artificial outliers and Censorship Rate(CR= 20%).

CR= 20 %		
number of artificial outliers	CLL MSE	TLL MSE
0	0.018250	0.013165
5	2.552174	0.017595
10	5.606158	0.026044
30	7.062188	0.047918

We observe that in the presence of outliers (0, 5, 10, 20) with Censorship Rate (CR = 20%), the trimmed local linear regression gives better results than the classic local linear method, in sense that, even if the MSE value of both methods increases relative to the number of the perturbed points, but it remains very low for the robust Local linear regression.

4.2. *Real data application*

We apply the theoretical results obtained in the previous subsection to real data. More specifically, we examine the performance of the trimmed regression estimator than the local linear method. For this purpose application, we consider the spectroscopic dataset, are available from http://www.models.kvl.dk/NIRsoil. The data concern spectra of 108 soil samples measured by Near Infrared Reflectance (NIR), in the range $0-1050$ nanometre (nm) with a 2 nm resolution (see[21]). The aim is to analyse relationships between the NIR data (X-variables), and the chemical and microbiological data (Y -variables). Hence, $X_i(t)$ is the reflectance of the i^{th} sample of soil at wavelength t, where $t \in \{0, ..., 1050\}$. Let Y_1 and Y_2 be two response variables which correspond to soil organic matter and ergosterol concentration, respectively (see Figures 5 and 6). The functional covariates in Figure 5 shows the 108 NIR reflectance spectra.

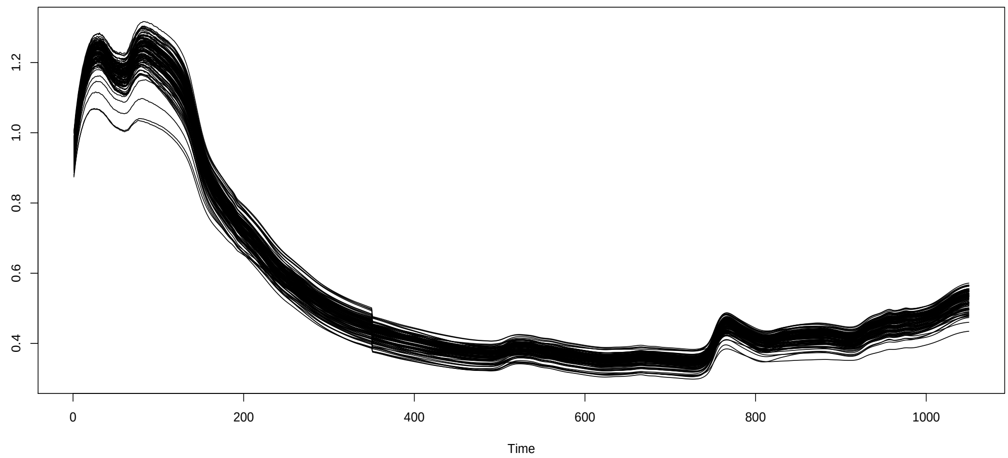

Figure 3: Curves of 108 NIR spectra.

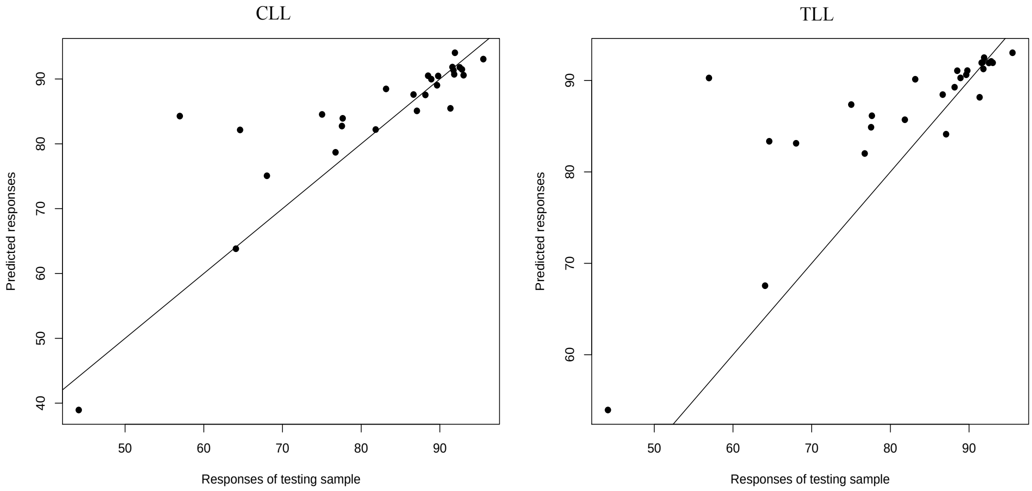

Figure 4: The distribution of 108 values of Y_1.

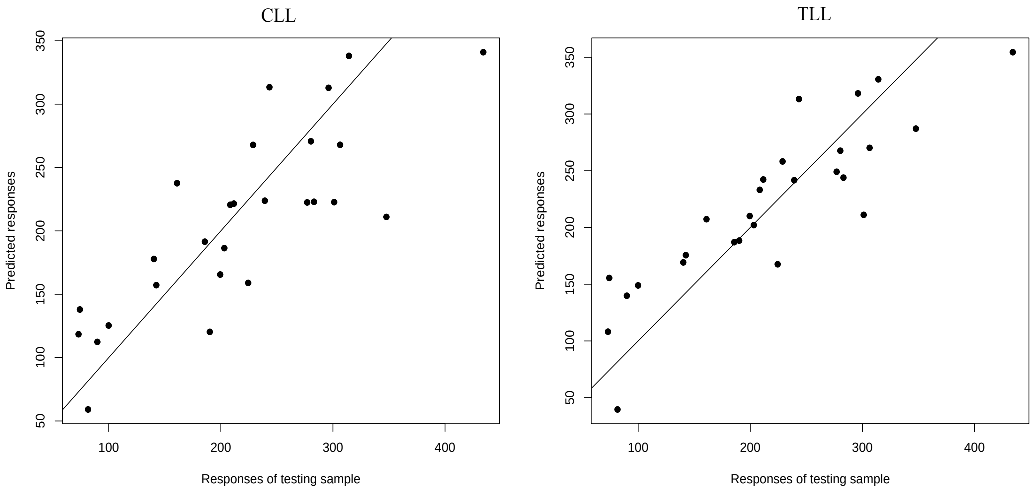

Figure 5: The distribution of 108 values of Y_2.

Applied to NIR data the MAD-Median method identifies 21 outliers for Y_1 and 1 outlier for Y_2. Recall that we are interested to build two models: $Y_1 = r_1(X) + \varepsilon_1$ and $Y_2 = r_2(X) + \varepsilon_2$, where $r_1(x) = \mathbb{E}(Y_1|X = x)$ and $r_2(x) = \mathbb{E}(Y_2|X = x)$. Furthermore, the dataset was randomly split into a learning sample (72 curves) used to build the estimators, and a testing sample (36 curves) which allows computing the MSE. We note that the result of our simulation study is evaluated over 400 independent replications and its sensitivity to grid sizes or to size of test sample and training sample is not very substantial. Because of the smoothness of the NIR curves, we use the semi-metric based on the second order derivatives, where the curves are replaced by their B-spline expansion. Here, the best results in terms of prediction are obtained for a number of interior knots needed for defining the B-spline basis, equal to 40 . Therefore, we chosen the smoothing parameter h via a local cross-validation method on the number of nearest-neighbors. It can be seen that, in the presence of outliers, the TLL regression estimator performs better than the CLL method. This is confirmed by the MSE obtained respectively in the two cases of study.

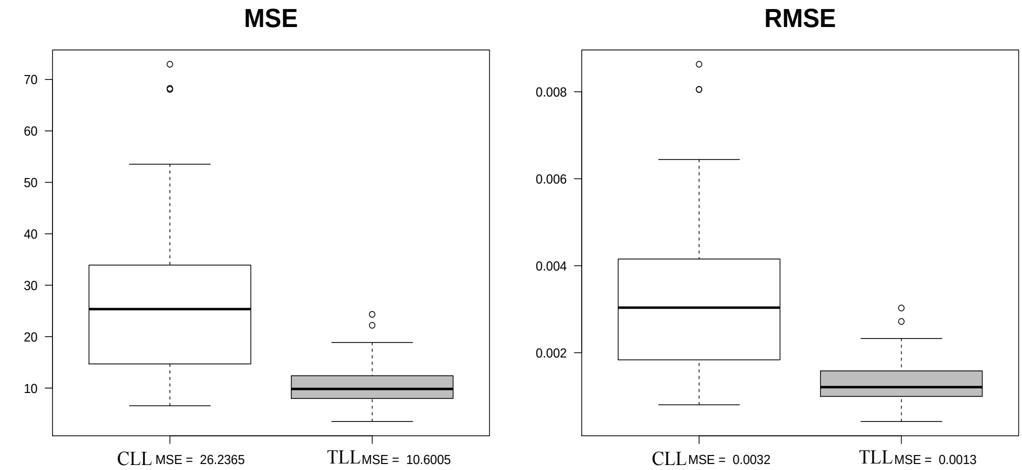

Figure 6: Box plots of the MSE Y_1.

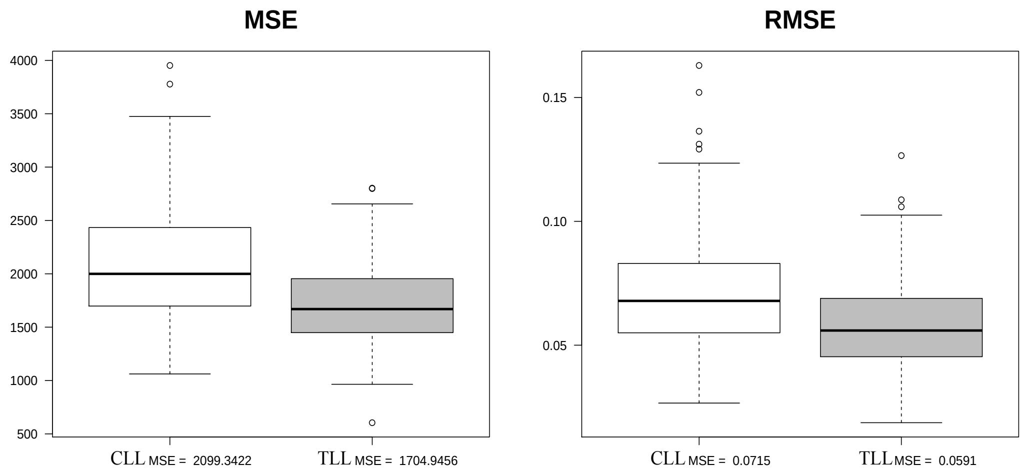

Figure 7: Box plots of the MSE for Y_2.

5. Appendix

First, we indicate the following lemmas that are necessary to establish our asymptotic result

Lemma 5.1. *Let V_n a sequence of vector function that satisfies*
❶ $-\Delta' V_n(\lambda\Delta) \geq -\Delta' V_n(\Delta), \forall\lambda \geq 1$.
❷ $\sup_{\|\Delta\|\leq M} \|V_n(\Delta) + \lambda_0 D\Delta - V_n(\Delta_0)\| = o_p(1)$ *for* $\frac{A}{\lambda_0\lambda_1(D)} < M < \infty$.
for $\|V_n(\Delta_n)\| = o_p(1)$

$$\|\Delta_n\| \leq M, \textit{in probability}$$

for $\|V_n(\Delta_n)\| = o_{a.co}(1)$ *and if* $\|V_n(\Delta_0)\| = o_{a.co}(1)$ *and* $\sup_{\|\Delta\|\leq M} \|V_n(\Delta) + \lambda_0 D\Delta - V_n(\Delta_0)\| = o_p(1)$*, we obtain*

$$\|\Delta_n\| \leq M, \textit{in almost completely}$$

Where:

- *D is a positive matrix,*
- *Δ_n is a vector,*
- *$V_n(\Delta_0)$ vectorial sequence such that $\mathbb{P}(\|V_n(\Delta_0)\| \geq A) \longrightarrow 0$,*
- *$\lambda_1(D)$ represents the minimum eigenvalue of D and $\lambda_0 > 0$.*

Proof. The demonstration of the two results is similar. It is based on the same arguments as in [15]. For the sake of brevity, we prove the second case which is more general.
Indeed, for $\eta > 0$, we have

$$\begin{aligned}\mathbb{P}(\|\Delta_n\| \geq M) &= \mathbb{P}(\|\Delta_n\| \geq M, \|V_n(\Delta_n)\| < \eta) + \mathbb{P}(\|V_n(\Delta_n)\| \geq \eta) \\ &\leq \underbrace{\mathbb{P}\left(\inf_{\|\Delta\| \leq M} \|V_n(\Delta_n)\| < \eta\right)}_{I} + \underbrace{\mathbb{P}(\|V_n(\Delta_n)\| \geq \eta)}_{II} \\ &< \infty\end{aligned}$$

The demonstration of II is simple. Since $\|V_n(\Delta_n)\| = o_{a.co}(1)$, we have

$$\sum_n \mathbb{P}(\|V_n(\Delta_n)\| \geq \eta) < \infty.$$

So, all that remains to show

$$I = \sum_n \mathbb{P}\left(\inf_{\|\Delta\| \leq M} \|V_n(\Delta_n)\| < \eta\right) < \infty$$

Hence by the Schwarz inequality

$$\Delta_1' V_n(\Delta) \leq \|\Delta_1\|.\|V_n(\Delta)\|$$

We have;

$$V_n(\Delta) \geq -\frac{\Delta_1' V_n(\Delta)}{\|\Delta_1\|} \tag{1}$$

For $\|\Delta\| \geq M$, let $\|\Delta_1\| = M$ and $\Delta = \lambda\Delta_1$, $\lambda \geq 1$. we obtain

$$-\frac{\Delta_1' V_n(\Delta)}{\|\Delta_1\|} = -\frac{\Delta_1' V_n(\lambda\Delta_1)}{M}$$

and for condition ❶ of **Lemma 5.1.** we have:

$$-\frac{\Delta_1' V_n(\Delta)}{\|\Delta_1\|} = -\frac{\Delta_1' V_n(\lambda\Delta_1)}{M} \geq -\frac{\Delta_1' V_n(\Delta_1)}{M} \tag{2}$$

So for (1) and (2) we obtain

$$V_n(\Delta) \geq -\frac{\Delta_1' V_n(\Delta_1)}{M}$$

Thus

$$\inf_{\|\Delta\| \geq M} V_n(\Delta) \geq \inf_{\|\Delta_1\| = M} -\frac{\Delta_1' V_n(\Delta_1)}{M}$$

Which implies

$$\begin{aligned}
\mathbb{P}\left(\inf_{\|\Delta\|\geq M} V_n(\Delta) < \eta\right) &\leq \mathbb{P}\left(\inf_{\|\Delta_1\|=M}\left[-\frac{\Delta_1' V_n(\Delta_1)}{M}\right] < \eta\right)\\
&= \mathbb{P}\left(\inf_{\|\Delta_1\|=M}\left[-\Delta_1' V_n(\Delta_1)\right] < \eta M\right)\\
&\leq \mathbb{P}\left(\inf_{\|\Delta_1\|=M}\left[-\Delta_1' V_n(\Delta_1)\right] < \eta M, \inf_{\|\Delta_1\|=M}[-\Delta_1'(-\lambda_0 D\Delta_1 + V_n(\Delta_0))] \geq 2\eta M\right)\\
&+ \mathbb{P}\left(\inf_{\|\Delta_1\|=M}[-\Delta_1'(-\lambda_0 D\Delta_1 + V_n(\Delta_0))] \geq 2\eta M\right)\\
&= P_1 + P_2.
\end{aligned}$$

$$\begin{aligned}
P_1 &= \mathbb{P}\left(\inf_{\|\Delta_1\|=M}\left[-\Delta_1' V_n(\Delta_1)\right] < \eta M, \inf_{\|\Delta_1\|=M}[-\Delta_1'(-\lambda_0 D\Delta_1 + V_n(\Delta_0))] \geq 2\eta M\right)\\
&= \mathbb{P}\left(\inf_{\|\Delta_1\|=M} V_n(\Delta_1) < \eta, \inf_{\|\Delta_1\|=M}(\lambda_0 D\Delta_1 - V_n(\Delta_0)) \geq 2\eta\right) \qquad \text{Car } \|\Delta_1\| = M\\
&\leq \mathbb{P}\left(\sup_{\|\Delta_1\|=M} \|V_n(\Delta_1) + \lambda_0 D\Delta_1 - V_n(\Delta_0)\| \geq \eta\right)\\
&< \infty \qquad \text{According to condition ❷ of } \textbf{Lemma 5.1.}
\end{aligned}$$

$$\begin{aligned}
P_2 &= \mathbb{P}\left(\inf_{\|\Delta_1\|=M}[-\Delta_1'(-\lambda_0 D\Delta_1 + V_n(\Delta_0))] \geq 2\eta M\right)\\
&= \mathbb{P}\left(\inf_{\|\Delta_1\|=M}(\Delta_1'\lambda_0 D\Delta_1 - \Delta_1' V_n(\Delta_0)) \geq 2\eta M\right)\\
&\leq \mathbb{P}\left(\inf_{\|\Delta_1\|=M}(\|\Delta_1'\|^2\lambda_0\lambda_1(D) - \|\Delta_1'\| V_n(\Delta_0)) \geq 2\eta M\right)\\
&\leq \mathbb{P}\left(\inf_{\|\Delta_1\|=M}(\|\Delta_1'\|\lambda_0\lambda_1(D) - V_n(\Delta_0)) \geq 2\eta\right)\\
&\leq \mathbb{P}\left(\sup_{\|\Delta_1\|=M}(\|\Delta_1'\|\lambda_0\lambda_1(D) - V_n(\Delta_0)) \geq 2\eta\right)\\
&\leq \mathbb{P}\left(\sup_{\|\Delta_1\|=M} V_n(\Delta_0)) \leq \|\Delta_1'\|\lambda_0\lambda_1(D) - 2\eta\right)\\
&< \infty \qquad \text{condition ❷ of } \textbf{Lemma 5.1.}
\end{aligned}$$

Finally, from P_1 and P_2, we conclude

$$\sum_n \mathbb{P}(\inf_{\Delta_n\geq M} \|V_n(\Delta)\| < \eta) < \infty.$$

□

Now, for the proof of the theorem we define, the vectorial sequences

$$V_n(\Delta) = \begin{cases} \frac{1}{n\phi_x(h)} \sum_{i=1}^n \varphi_i(\Delta) K_i \\ \frac{1}{nh\phi_x(h)} \sum_{i=1}^n \varphi_i(\Delta)\beta_i K_i \end{cases}$$

Where

$$\varphi_i(\Delta) = \frac{\delta_i\psi_i\big(T_i-(c+a)-(h_n^{-1}d+b)\beta_i\big)}{\overline{G}(t_i)} = \frac{\delta_i\big(T_i-(c+a)-\big(h_n^{-1}d+b\big)\beta_i\big)}{\overline{G}(t_i)}\mathbb{1}_{\left|\frac{\delta_i\left(T_i-(c+a)-\left(h_n^{-1}d+b\right)\beta_i\right)}{\overline{G}(t_i)}\right|\le q}$$

With:

$$\Delta = \begin{pmatrix} c \\ d \end{pmatrix},\ \Delta_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Delta_n = \begin{pmatrix} \hat{a}-a \\ h(\hat{b}-b) \end{pmatrix} \Leftrightarrow \Delta'_n = \sqrt{n\phi_x(h)}\Delta_n \text{ and } \lambda = \sqrt{n\phi_x(h)}$$

Now, the proof of the result of **Theorem 3.1** is based on the application of the second part of **Lemma 5.1.** to $(V_n, V_n(\Delta_0), \Delta_n)$.

Thus, **Theorem 3.1** is consequence of the following Lemmas:

Lemma 5.2. *Under assumptions (H1)-(H5), we have*

$$\|V_n(\Delta_0)\| = O(h^{\min(b_1,b_2)}) + O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), a.co$$

Proof. for $\ell = 0, 1$ we obtain

$$\begin{aligned}
V_n^{\ell+1}(\Delta_0) &= V_n(\Delta_0) - \mathbb{E}[V_n(\Delta_0)] \\
&= \frac{1}{n\phi_x(h)}\sum_{i=1}^{n}\varphi_i(\Delta_0)h^{-\ell}\beta_i^{\ell}K_i - \mathbb{E}\left[\frac{1}{n\phi_x(h)}\sum_{i=1}^{n}\varphi_i(\Delta_0)h^{-\ell}\beta_i^{\ell}K_i\right] \\
&= \frac{1}{n\phi_x(h)}\sum_{i=1}^{n}\left(\varphi_i(\Delta_0)h^{-\ell}\beta_i^{\ell}K_i - \mathbb{E}\left[\varphi_i(\Delta_0)h^{-\ell}\beta_i^{\ell}K_i\right]\right) \\
&= \frac{1}{n\phi_x(h)}\sum_{i=1}^{n} Z_i^{\ell+1} && (3) \\
&= O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right) && (4)
\end{aligned}$$

From (3) to (4) we apply the exponential inequality for unbounded variables, we get from **Corollary A.8.** in [11] when $a^2 = O\left(\frac{1}{\phi_x(h)}\right)$.

By using the bounded of φ_i and K we have

$$|Z_i^1| \le C_1 \quad , \quad \mathbb{E}[Z_i^1]^2 \le C_1\phi_x(h)$$

and

$$|Z_i^2| \le C_2 h \quad , \quad \mathbb{E}[Z_i^2]^2 \le C_2 h^2\phi_x(h)$$

On the other hand, by (H2), we have

$$\begin{aligned}
\mathbb{E}[V_n^{\ell+1}(\Delta_0)] &= \mathbb{E}[V_n(\Delta_0) - \mathbb{E}[V_n(\Delta_0)]] \\
&= \mathbb{E}\left[\frac{1}{n\phi_x(h)}\sum_{i=1}^{n}(\varphi_i(\Delta_0)h^{-\ell}\beta_i^{\ell}K_i - \mathbb{E}\left[\varphi_i(\Delta_0)h^{-\ell}\beta_i^{\ell}K_i)\right]\right] \\
&= \frac{1}{\phi_x(h)}\mathbb{E}\left[\varphi_1(\Delta_0)h^{-\ell}\beta_1^{\ell}K_1\right] \\
&\le \frac{1}{\phi_x(h)}\mathbb{E}|\tau_1(x,\theta_x)h^{-\ell}\beta_1^{\ell}K_1 - \tau_1(X, a+b\beta_1)h^{-\ell}\beta_1^{\ell}K_1| \\
&= O(h^{b_1}) + o(h^{b_2})
\end{aligned}$$

□

Lemma 5.3. *Under assumptions (H1)-(H5), for all $\lambda_0 = \nu(x, \theta_x)$, we have*

$$\sup_{\|\Delta\|\leq M} \|V_n(\Delta) + \nu(x, \theta_x)D\Delta - V_n(\Delta_0)\| = O(h^{\min(b_1,b_2)}) + O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), a.co$$

Proof.

$$V_n(\Delta) + \nu(x, \theta_x)D\Delta - V_n(\Delta_0) = V_n(\Delta) + \nu(x, \theta_x)D\Delta - V_n(\Delta_0) - \mathbb{E}[V_n(\Delta) - V_n(\Delta_0)] + \mathbb{E}[V_n(\Delta) - V_n(\Delta_0)]$$

The proof of **Lemma 5.3** must be made on two parts

a) $\sup_{\|\Delta\|\leq M} \|\mathbb{E}[V_n(\Delta) - V_n(\Delta_0)] + \nu(x, \theta_x)D\Delta\| = O(h^{\min(b_1,b_2)})$;

b) $\sup_{\|\Delta\|\leq M} \|V_n(\Delta) - V_n(\Delta_0) - \mathbb{E}[V_n(\Delta) - V_n(\Delta_0)]\| = O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$.

for the result a), we use (H2)

$$\begin{aligned}
\mathbb{E}[V_n(\Delta) - V_n(\Delta_0)] &= \mathbb{E}\left[\frac{1}{n\phi_x(h)}\sum_{i=1}^{n}(\varphi_i(\Delta) - \varphi_i(\Delta_0))h^{-\ell}\beta_i^{\ell}K_i\right] \\
&= \frac{1}{n\phi_x(h)}\sum_{i=1}^{n}\mathbb{E}\left[(\varphi_i(\Delta) - \varphi_i(\Delta_0))h^{-\ell}\beta_i^{\ell}K_i\right] \\
&= \frac{1}{\phi_x(h)}\mathbb{E}\left[(\varphi_i(\Delta) - \varphi_i(\Delta_0))h^{-\ell}\beta_i^{\ell}K_i\right] \\
&= \frac{1}{\phi_x(h)}\mathbb{E}\left[(\tau_1(x, (c+a) + (h^{-1}d + b)\beta_1) - \tau_1(x, a + b\beta_1))h^{-\ell}\beta_1^{\ell}K_1\right] + O(h^{\min(b_1,b_2)}) \\
&= \frac{1}{\phi_x(h)}\mathbb{E}\left[\nu(x, a + b\beta_1)(1, h^{-1}\beta_1)\Delta h^{-\ell}\beta_1^{\ell}K_1\right] + O(h^{\min(b_1,b_2)}) + o(\|\Delta\|) \\
&= \nu(x, \theta_x)\frac{1}{\phi_x(h)}\mathbb{E}\left[(1, h^{-1}\beta_1)h^{-\ell}\beta_1^{\ell}K_1\right]\Delta + O(h^{\min(b_1,b_2)}) + o(\|\Delta\|) \\
&= \nu(x, \theta_x)\frac{1}{\phi_x(h)}\mathbb{E}\left[(h^{-\ell}\beta_1^{\ell}K_1, h^{-(\ell+1)}\beta_1^{\ell+1}K_1)\right]\Delta + O(h^{\min(b_1,b_2)}) + o(\|\Delta\|) \\
&= \nu(x, \theta_x)\frac{1}{\phi_x(h)}(h^{-\ell}\mathbb{E}[\beta_1^{\ell}K_1], h^{-(\ell+1)}\mathbb{E}[\beta_1^{\ell+1}K_1])\Delta + O(h^{\min(b_1,b_2)}) + o(\|\Delta\|)
\end{aligned}$$

Therefore,

$$\mathbb{E}[V_n(\Delta) - V_n(\Delta_0)] = \nu(x, \theta_x)\frac{1}{\phi_x(h)}\begin{pmatrix}\mathbb{E}[K_i] & h^{-1}\mathbb{E}[\beta_i K_i)] \\ h^{-1}\mathbb{E}[\beta_i K_i] & h^{-2}\mathbb{E}[\beta_i^2 K_i])\end{pmatrix}\Delta + O(h^{\min(b_1,b_2)}) + o(\|\Delta\|)$$

Using the same ideas of [9], under the second part of (H3)

$$h^{-a}\mathbb{E}[\beta^a K_i^b] = \phi_x(h)\left(K^b(1) - \int_{-1}^{1}(r^a K^b(r))'\Phi_x(r)dr\right)$$

Consequently

$$\sup_{\|\Delta\|\leq M} \|\mathbb{E}[V_n(\Delta) - V_n(\Delta_0)] + \nu(x, \theta_x)D\Delta + o(\|\Delta\|)\| = O(h^{\min(b_1,b_2)});$$

then, all it remains to show the result b), we use the compactness of the ball $B(0, M)$ in $\mathbb{R}^2$ and we write

$$B(0, M) \subset \bigcup_{j=1}^{d_n} B(\Delta_j, l_n) \quad \Delta_j = \begin{pmatrix} c_j \\ d_j \end{pmatrix}, \text{ and } l_n = d_n = \frac{1}{\sqrt{n}}$$

Let $j(\Delta) = \arg\min_j |\Delta - \Delta_j|$, in which

$$\begin{aligned}
\sup_{\|\Delta\|\leq M} \|V_n(\Delta) - V_n(\Delta_0) - \mathbb{E}[V_n(\Delta) - V_n(\Delta_0)] &\leq \underbrace{\sup_{\|\Delta\|\leq M} \|V_n(\Delta) - V_n(\Delta_j)\|}_{T_1} \\
&+ \underbrace{\sup_{\|\Delta\|\leq M} \|V_n(\Delta_j) - V_n(\Delta_0) - \mathbb{E}[V_n(\Delta_j) - V_n(\Delta_0)]\|}_{T_2} \\
&+ \underbrace{\sup_{\|\Delta\|\leq M} \|\mathbb{E}[V_n(\Delta) - V_n(\Delta_j)]\|}_{T_3}.
\end{aligned}$$

Concerning T_1 we use the boundness of the kernel K

$$\sup_{\|\Delta\|\leq M} \|V_n(\Delta) - V_n(\Delta_j)\| \leq \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} \tilde{\upsilon}_i$$

The function $\varphi_i(.)$ is locally Lipschitzian on $[-q;q]$, so we can write

$$\tilde{\upsilon}_i = \sup_{\|\Delta\|\leq M} |\varphi(\Delta) - \varphi(\Delta_j)| \left\| \begin{pmatrix} 1 \\ h_n^{-1}\beta_i \end{pmatrix} \right\| K_i \leq Cl_n$$

Then

$$\sup_{\|\Delta\|\leq M} \|V_n(\Delta) - V_n(\Delta_j)\| \leq \frac{Cl_n}{\phi_x(h)} = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

Finally

$$\sup_{\|\Delta\|\leq M} \|V_n(\Delta) - V_n(\Delta_j)\| = O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

For the quantity T_3, analogously to first term and we use the fact that $\mathbb{E}[K_i] = O(\phi_x(h))$(see ([10])) we obtain

$$\sup_{\|\Delta\|\leq M} \|\mathbb{E}[V_n(\Delta) - V_n(\Delta_j)]\| \leq \frac{Cl_n}{\phi_x(h)} = o_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

Now, for T_2 we can put

$$W_n^{\ell+1}(\Delta_j) = V_n(\Delta_j) - V_n(\Delta_0) - \mathbb{E}[V_n(\Delta_j) - V_n(\Delta_0)]$$

We obtain

$$\begin{aligned} W_n^{\ell+1}(\Delta_j) &= \frac{1}{nh^{\ell}\phi_x(h)} \sum_{i=1}^{n} ((\varphi_i(\Delta_j) - \varphi_i(\Delta_0))\beta_i^{\ell} K_i - \mathbb{E}[(\varphi_i(\Delta_j) - \varphi_i(\Delta_0))\beta_i^{\ell} K_i]) \\ &= \frac{1}{nh^{\ell}\phi_x(h)} \sum_{i=1}^{n} \Upsilon_i^{\ell+1} \end{aligned}$$

by to the boundedness of K and $\varphi_i(.)$ we obtain

$$|\Upsilon_i^{\ell+1}| \leq C_{\ell+1} \quad \text{and} \quad \mathbb{E}[\Upsilon_i^{\ell+1}]^2 \leq C_{\ell+1} h^{\ell} \phi_x(h)$$

Now, We apply the exponential inequality for unbounded variables. we obtain

$$\sup_{\|\Delta\|\leq M} \|V_n(\Delta_j) - V_n(\Delta_0) - \mathbb{E}[V_n(\Delta_j) - V_n(\Delta_0)]\| = O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

□

6. Conclusion

In this work, we have positioned our contribution in the extensive literature of nonparametric analysis of censored functional data. by giving almost complete convergence. The results obtained, in the simulation part, confirm that the trimmed local linear estimator has an effective and good performance statistically.

Acknowledgments

The authors are very pleased to thank the editor and the reviewers for their helpful suggestions and comments.

References

[1] L.Ait Hennania, M.Lemdania, E.Ould Said, Robust regression analysis for a censored response and functional regressors, Journal of Nonparametric Statistics, https://doi.org/10.1080/10485252.2018.1546386.
[2] A. Baìllo, A.Grané, Local linear regression for functional predictor and scalar response, Multivariate Analysis 100 (2009) 102–111.
[3] J.Barrientos-Marin, F. Ferraty, P.Vieu, Locally modeled regression and functional data, J. Nonparametr. Stat 22 (2010) 617–632.
[4] I.Basak, Robust regression with censored data, Naval Res Logist 39 (1992) 323–344
[5] T.Bednarsky, On robust causality nonresponse testing in duration studies under the Cox model, J. Nonparametr. Stat 22 (2010) 617–632.
[6] R.Beran, Nonparametric regression with randomly censored survival data, Technical Report. University of California, Berkeley 1981
[7] A.Berlinet, A.Elmine, A.Mas, Local linear estimate of nonparametric robust regression in function date, Ann. Inst. Statist. Math 63 (2011) 1045–1075.
[8] A.Chouaf, Modelization local linear regression for functional random variables, International Journal of Statistics & Economics 16 (2015) 54–68.
[9] J.Demongeot, A.Laksaci, F.Madani, M.Rachdi, Modelization local linear regression for functional random variables, Statistics 47 (2013) 26–44.
[10] J.Fan, I.Gijbels, Local polynomial modelling and its applications, Vol. 1 of Monographs on Statistics and Applied Probability 66, Chapman & Hall, 2-6 Boundary Row, London SE1 8HN, UK 1996
[11] F.Ferraty, P.Vieu, Nonparametric functional data analysis, Vol. 1 of Springer Series in Statistics, Springer Science +Business Media, New York Inc., 233 Spring Street, New York,NY10013, USA, 2006
[12] F.Hampel, E.Ronchetti, P.Rousseeuw, W.Stahel, Robust Statistics : The Approach Based on Influence Functions, Vol. 1 of Wiley Series in Probability and Statistics, John Wiley & Sons, New York, 1986
[13] Z.Kaid, A.Laksaci, Functional quantile regression :local linear modelisation, Functional Statistics and Related Fields 20 (2017) 155–156.
[14] E.L.Kaplan and Paul Meier, Nonparametric Estimation from Incomplete Observations, Journal of the American Statistical Association 53(282) (1958) 457–481
[15] R. Koenker, Q. Zhao, Conditional quantile estimation and inference for ARCH models, Econometric theory 1996 793–813
[16] S.Leulmi, Local linear estimation of the conditional quantile for censored data and functional regressors, Communications in Statistics-Theory and Methods, https://doi.org/10.1080/03610926.2019.1692033.
[17] S.Leulmi, Nonparametric local linear regression estimation for censored data and functional regressors, J. Korean Stat. Soc. 51, 25–46 (2022), https://doi.org/10.1007/s42952-020-00080-7
[18] J.Li, M.Zheng, Robust estimation of multivariate regression model, Stat Papers 50 (2009) 81–100
[19] I.Massim, B.Mechab, Local linear estimation of the conditional hazard function, Journal of Statistics & Economics 17 (2016) 1–11.
[20] B.Mechab, N.Hamidi, S.Benaissa, Nonparametric estimation of the relative error in functional regression and censored data, Chilean Journal of Statistics, 10(2) (2019) 177–195.
[21] R.Rinnan, A.A.Rinnan, Application of near infrared reflectance (NIR) and fluorescence spectroscopy to analysis of microbiological and chemical properties of arctic soil. Soil Biology and Biochemistry, 39 (2007) 1664–1673.
[22] P.J.Rousseeuw, A.M.Leroy, Robust Regression and Outlier Detection Vol. 1 of Wiley series in probability and mathematical statistics. Applied probability and statistics, John Wiley & Sons, New York, 1987.