Universität Stuttgart

Fachbereich Mathematik

k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis

Simon Müller, Jürgen Dippon

Preprint 2011/014

Universität Stuttgart

Fachbereich Mathematik

k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis

Simon Müller, Jürgen Dippon

Preprint 2011/014

Fachbereich Mathematik
Fakultät Mathematik und Physik
Universität Stuttgart
Pfaffenwaldring 57
D-70 569 Stuttgart

E-Mail: preprints@mathematik.uni-stuttgart.de
WWW: http://www.mathematik.uni-stuttgart.de/preprints

ISSN **1613-8309**

Abstract

This paper examines the k-nearest neighbours method in functional non-parametric regression for α-mixing data. We prove almost complete convergence and give the almost complete convergence rate the k-NN kernel estimate. The results are obtained on the one hand by using results on the classical functional kernel estimate, where a deterministic bandwidth sequence is used, and on the other hand by applying lemmas from Bradley and Burba et al. The latter one was already used for the k-nearest neighbours kernel estimate for independent data. Finally, we give an outline on how to avoid the drawback of susceptibility of the k-nearest neighbours kernel estimate to outliers.

1 Introduction

In this paper we examine the functional k-nearest neighbours, shortly k-NN, non-parametric regression estimate in case of α-mixing data. The classical functional non-parametric regression estimate (see e.g. by [11]) depends on a real-valued non-random bandwidth sequence h_n. On the contrary, the smoothing parameter of the k-NN regression estimate depends on the numbers of neighbours at the point of interest at which we want to make a prediction. In cases where data is sparse, the k-NN kernel estimate has a significant advantage over the classical kernel estimate. The k-NN kernel estimate is also automatically able to take into account the local structure of the data. This advantage, however, may turn into a disadvantage. If there is an outlier in the dataset, the local prediction may be bad. To avoid this, a robust non-parametric regression ansatz may be chosen (see Section 6 of this paper). Selecting the bandwidth depending on the data turns the bandwidth into a random variable. Hence we are no longer able to use the same techniques in the consistency proofs as in the case of a non-random bandwidth sequence.

The k-NN kernel estimate is a widely studied if the explanatory variable is an element of a finite-dimensional space, see Györfi et al. [12]. For functional explanatory variable and with real-valued response variable, two different approaches for the k-NN regression estimation exist. The first one, published by Laloë [13], examines a k-NN kernel estimate when the functional variable is an element of a separable Hilbert space. For that case Laloë establishes a weak consistency result. However, his ansatz is not completely functional. Laloë's strategy is to reduce the dimension of the input variable by using a projection onto a finite-dimensional subspace and then applying multivariate techniques on the projected data. The second result, from Burba et al. [5], is based on a pure functional approach instead. Burba et al. examine the problem on a semi-metric functional space. They proved almost complete convergence and almost complete convergence rates for independent data. Furthermore, Burba et al. extended a lemma that we will also use in our proofs. This lemma originates from Collomb [7]. We will cite it in Section 4 and make some additional comments on it. Additionally, the k-NN kernel estimate is examined for classification in infinite dimension by Cérou and Guyader [6] and there exists a convergence result for the k-NN regression estimate when the response is an element of a Hilbert space (see Lian [14]).

In the case of finite-dimensional explanatory variable, the k-NN kernel estimate for α-mixing random variables is treated by Tran [18] and Lu and Cheng [15]. Both results are based on Collomb's [7] results. We combined their idea with Burba et alii's [5] results to prove consistency and the rate.

This paper is organised as follows. In Section 2 we present the k-NN kernel estimate. Afterwards, we introduce the assumptions and the main result, the almost complete convergence and the almost complete convergence rate. In Section 4, some technical auxiliary results are deployed and in Section 5, we show the proofs of our main result. In the end, we outline some applications and discuss how to get a robust k-NN kernel estimate.

2 Method and Assumptions

Let $(X_i, Y_i)_{i=1}^n$ be a dependent sequence identically distributed as (X, Y), the latter being a random pair with values in the measurable space $(\mathrm{E} \times \mathbb{R}, \mathcal{E}_d \otimes \mathcal{B})$. Here (E, d) is a semi-metric space and

$\mathcal{E}_d$ is the σ–algebra generated by the topology of E that is defined by the semi-metric d, and $\mathcal{B}$ is the Borel σ–algebra. In order to characterise the model of dependence, we use the notion of α-mixing.

We examine the k-NN kernel estimate that is defined for $x \in$ E as

$$\hat{m}_{\text{k-NN}}(x) = \sum_{i=1}^{n} Y_i \frac{K\left(H_{n,k}^{-1} d(x, X_i)\right)}{\sum_{i=1}^{n} K\left(H_{n,k}^{-1} d(x, X_i)\right)}, \text{ if } \sum_{j=1}^{n} K\left(H_{n,k}^{-1} d(x, X_i)\right) \neq 0, \tag{1}$$

otherwise $\hat{m}_{\text{k-NN}}(x) = 0$. $K : \mathbb{R} \to \mathbb{R}^+$ is a kernel function and $H_{n,k}$ is the bandwidth that is defined as

$$H_{n,k} := d(x, X_{(k)}), \tag{2}$$

where the sequence $(X_{(i)}, Y_{(i)})_{i=1}^{n}$ is the re-indexed sequence $(X_i, Y_i)_{i=1}^{n}$ such that

$$d(x, X_{(1)}) \leq d(x, X_{(2)}) \leq \ldots \leq d(x, X_{(n)}).$$

From now on, when we refer to the bandwidth of the k-NN kernel estimate, we mean the number of neighbours k we are considering.

To prove the almost complete convergence of the k-NN kernel estimate, we need some results of the Nadaraya-Watson kernel estimate. Therefore, we present this kernel estimate here. Let $x \in$ E, then

$$\hat{m}(x) = \sum_{i=1}^{n} Y_i \frac{K\left(h_n^{-1} d(x, X_i)\right)}{\sum_{i=1}^{n} K\left(h_n^{-1} d(x, X_i)\right)}, \quad \text{if} \quad \sum_{j=1}^{n} K\left(h_n^{-1} d(x, X_i)\right) \neq 0, \tag{3}$$

otherwise $\hat{m}(x) = 0$. K is a kernel function and $h := h_n$ is a non-random bandwidth. Hereafter, the notion kernel estimate will refer to the Nadaraya-Watson kernel estimate.

We consider two types of non-parametric models. The first one is called *continuity-type*, which means that the regression function m is continuous and the second one, the *Hölder-type*, which means that the regression function m is Hölder continuous with constant $\beta > 0$.

Prior to the presentation of our main results, we outline the assumptions.

(F) Let $x \in E$, then assume that the probability of observing the functional random variable X around x is strictly positive, that means

$$\forall \varepsilon > 0 : \; F_x(\varepsilon) := \mathrm{P}\left(d(x, X) \leq \varepsilon\right) > 0.$$

(K) We distinguish two types of kernel functions K:

(a) There exist constants $0 < C_1 < C_2 < \infty$ such that

$$\forall u \in \mathbb{R} : \; C_1 1_{[0,1]}(u) \leq K(u) \leq C_2 1_{[0,1]}(u).$$

(b) The kernel function K has its support in $[0, 1]$, K is differentiable in $[0, 1]$, $K(1) = 0$, and there exist two constants $-\infty < C_1 < C_2 < 0$ such that

$$\forall u \in [0, 1] : \; C_1 \leq K'(u) \leq C_2.$$

Furthermore, we need following technical assumption

$$\exists C > 0 \; \exists \varepsilon_0 > 0 \; \forall \varepsilon < \varepsilon_0 : \; \int_0^{\varepsilon} F_x(u) du > C \varepsilon F_x(\varepsilon),$$

for details see [11, p. 44]

(M) Assume that the conditional moments of Y are bounded,

$$\forall m \in \mathbb{N}: \ \mathrm{E}\left[|Y|^m | X = x\right] < \sigma_m(x) < \infty,$$

with $\sigma_m(\cdot)$ continuous at $x \in E$.

The sequence $((X_i, Y_i))$ is said to be α-mixing if

$$\alpha(n) := \sup_k \sup_{A \in \mathcal{A}_1^k} \sup_{B \in \mathcal{A}_{k+n}^\infty} \left|\mathrm{P}(A \cap B) - \mathrm{P}(A)\,\mathrm{P}(B)\right| \to 0$$

as $n \to \infty$, where $\mathcal{A}_l^m$ is the σ-algebra generated by $\{(X_i, Y_i), l \leq i \leq m\}$.

(A) Assume that the sequence (X_i, Y_i) is arithmetic α-mixing (or algebraic), namely we have for some $C > 0$ and rate $b > 1$: $\alpha(n) \leq Cn^{-b}$.

The terms of covariance, which are a measure of dependence, are here denoted by

$$s_{n,1}(x) := \sum_{i,j=1}^n |\mathrm{C}ov(\Delta_i(x), \Delta_j(x))| \text{ and } s_{n,2}(x) := \sum_{i,j=1}^n |\mathrm{C}ov(Y_i\Delta_i(x), Y_j\Delta_j(x))|,$$

where $\Delta_i(x) := K(h^{-1}d(x, X_i))/\mathrm{E}\left[K(h^{-1}d(x, X_1))\right]$. Note that we can split for example $s_{n,2}(x)$ as

$$s_{n,2}(x) = \underbrace{\sum_{i=1}^n \mathrm{V}ar\left[Y_i\Delta_i(x)\right]}_{I} + \underbrace{\sum_{\substack{i,j=1 \\ j \neq i}}^n \left|\mathrm{C}ov(Y_i\Delta_i(x), Y_j\Delta_j(x))\right|}_{II}. \tag{4}$$

Term II in (4) is a measure of the dependence of the random variables. We want to remark, if the X_i are α-mixing then also the $\Delta_i(x)$ are α-mixing, see e.g. Lemma 10.3 in [11, p. 155].

(D) Assume for the covariance term $s_n(x) := \max\{s_{n,1}(x), s_{n,2}(x)\}$ that there exists a $\theta > 2$ such that $s_n^{-(b+1)} = o\left(n^{-\theta}\right)$, where b is the rate of the mixing coefficient (see Condition (A)).

(B) Assume for the sequence of bandwidths $k := k_n$ that there exists a $\gamma \in (0, 1)$ such that $k \sim n^\gamma$.

Condition (B) is not to be more restrictive than in the independent case. However, for their consistency result Burba et al. [5] need the following two conditions, $k/n \to 0$ and $\log n/k \to 0$ as $n \to \infty$, so k must exceed logarithmic order. As Lian comments in [14], in most cases in the functional context the small ball probability is of exponential-type. Hence the convergence speed is logarithmic, no matter if the number of neighbours k increases logarithmically or polynomially. For example, if we have for the small ball probability

$$F_x(h) \sim \exp\left(-\frac{1}{h^\tau}\right), \text{ then } F^{-1}\left(\frac{k}{n}\right) \sim \left(\frac{1}{\log\left(\frac{k}{n}\right)}\right)^\tau,$$

where $F_x^{-1}(y) := \inf\{h | F_x(h) \geq y\}$ (see [14]). It can be easily seen that the order of k is less important for such small probabilities.

(D1) This condition is on the distribution of two distinct pairs (X_i, Y_i) and (X_j, Y_j). We assume that

$$\forall i \neq j: \ \mathrm{E}\left[Y_iY_j | X_iX_j\right] \leq C < \infty,$$

and the joint distribution functions $\mathrm{P}\left(X_i \in B(x,h),\ X_j \in B(x,h)\right)$ satisfy

$$\exists \varepsilon_1 \in (0,1]:\ 0 < G_x(h) = \mathcal{O}\left(F_x(h)^{1+\varepsilon_1}\right),$$

where

$$G_x(h) := \max_{i,j \in \{1,\ldots,n\}, i \neq j} \mathrm{P}\left(X_i \in B(x,h),\ X_j \in B(x,h)\right).$$

Condition (D1) is, as Ferraty and Vieu [11, p. 163] in Note 11.2 describe, not too restrictive. For example, if we choose $\mathrm{E} = \mathbb{R}^p$, then $\varepsilon_1 = 1$ as soon as each pair of random variables (X_i, X_j) has a bounded density $f_{i,j}$ with respect to the Lebesgue measure.

Next, we formulate a more general condition on the joint distribution function.

(D2) Define $\chi(x,h) := \max\left\{1, G_x(h)/(F_x(h))^2\right\}$ and $s := 1/(b+1)$ with b as the rate of the mixing coefficient. Then assume that

$$\frac{\log(n)\chi(x,h)^{1-s} n^{1+s}}{k^2} \to 0 \text{ as } n \to \infty.$$

3 Almost Complete Convergence and Almost Complete Convergence Rate

Before we present the consistency result of the k-NN estimate the almost complete convergence[1] result of the kernel regression estimate $\hat{m}(x)$ of Ferraty and Vieu [11] is presented. All results presented here assumes a sequence of α-mixing random variables $((X_i, Y_i))$ (see Condition (A)).

Theorem 3.1 (Ferraty and Vieu [11], p. 63)**.** *Assume that the regression function is of continuity-type, furthermore assume (F), (K), and (M). Additionally, suppose for the bandwidth that $h_n \to 0$ and $\log n/(nF_x(h_n)) \to 0$ as $n \to \infty$. Then we have for $x \in E$*

$$\lim_{n\to\infty} \hat{m}(x) = m(x) \quad \text{almost completely.}$$

The following theorem gives almost complete rates[2].

Theorem 3.2 (Ferraty and Vieu [11], p. 80)**.** *Assume the same conditions as in Theorem 3.1, and a Hölder-type model instead of a continuity-type model. Then we have for $x \in E$*

$$\hat{m}(x) - m(x) = \mathcal{O}\left(h^\beta\right) + \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n(x)\log n}}{n}\right).$$

Now we state the almost complete convergence result for the non-parametric k-NN kernel estimate, introduced in (1).

Theorem 3.3. *In the case of a continuity-type model, we suppose condition (F) for the small ball probability, (K) for the kernel function, (B) for the bandwidth k. Either assume that Condition (D1) holds with $b > \max\{3/(2\gamma) - 1, (2-\gamma)/(\varepsilon_1(1-\gamma))\}$, where γ is the constant in Condition (B) and ε_1 the constant in Condition (D1). Or assume that Condition (D2) is enforced, with rate $b > \frac{3}{2\gamma} - 1$. Then we have for $x \in E$*

$$\lim_{n\to\infty} \hat{m}_{\text{k-NN}}(x) = m(x) \quad \text{almost completely.}$$

---

[1] We recall that $X_n \to X$ almost completely if and only if $\forall \varepsilon > 0:\ \sum_{n=1}^{\infty} \mathrm{P}\left(|X_n - X| > \varepsilon\right) < \infty$

[2] We recall that $X_n - X = \mathcal{O}_{a.co.}(u_n)$ if and only if $\exists \varepsilon_0 > 0:\ \sum_{n=1}^{\infty} \mathrm{P}\left(|X_n - X| > \varepsilon_0 u_n\right) < \infty$

Theorem 3.4. *In the case of a Hölder-type model, we suppose condition (F) for the small ball probability, (K) the kernel function, (B) the bandwidth k.*

If Condition (D1) holds with $b > \max\{3/(2\gamma) - 1, (2-\gamma)/(\varepsilon_1(1-\gamma))\}$, where γ is the constant in Condition (B) and ε_1 the constant in Condition (D1). Then we have $x \in E$

$$\hat{m}_{\text{k-NN}}(x) - m(x) = \mathcal{O}\left(\left(F_x^{-1}\left(\frac{k}{n}\right)\right)^{\beta}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{k}}\right). \tag{5}$$

If (D2) holds instead of (D1) with $b > 3/(2\gamma) - 1$, then we have for $x \in E$

$$\begin{aligned}\hat{m}_{\text{k-NN}}(x) - m(x) &= \mathcal{O}\left(\left(F_x^{-1}\left(\frac{k}{n}\right)\right)^{\beta}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{k}}\right)\\ &\quad + \mathcal{O}_{a.co.}\left(\sqrt{\frac{n^{1+s}\log n}{k^2}\chi\left(x, F_x^{-1}\left(\frac{k}{n}\right)\right)^{1-s}}\right),\end{aligned} \tag{6}$$

where $\chi(x,h) := \max\left\{1, G_x(h)/(F_x(h))^2\right\}$.

The covariance term $s_n(x)$ disappears in (5). The Condition (D1) and the condition on the rate b implies that term II in (4) decays faster than term I. We get $s_n(x) = \mathcal{O}\left(n(F_x(h))^{-1}\right)$, see Lemma 11.5 in [11, p. 166]. If Condition (D2) instead of (D1) is assumed we get three terms for the rate (see (6)). The first one in (6) has its origin in the regularity of the regression function, the second one stems from term I in (4) and the third one represents the dependence of the random variables (compare term II in (4)).

4 Technical Tools

Because of the randomness of the smoothing parameter $H_{n,k}$, it is not possible to use the same tools for proving the consistency as in the kernel estimation. The necessary tools are presented in this section. The following two lemmas of Burba et al. [5] are generalisations of a result firstly presented by Collomb [7]. In our opinion, Burba et alii's [5] Lemmas 4.1 and 4.2 are valid for dependent random variables as in the original lemma from Collomb [7]. We checked the proof from Burba et al. against Collomb's proof; we did not find any reason why Burba et al. [5] assume independence. On reflection, this assumption appears unnecessary.

Let $(A_i, B_i)_{i=1}^n$ be a sequence of random variables with values in $(\Omega \times \mathbb{R}, \mathcal{A} \otimes \mathcal{B})$, not necessarily identically distributed or independent. Let $k : \mathbb{R} \times \Omega \to \mathbb{R}^+$ be a measurable function with the property

$$z \leq z' \Rightarrow \forall \omega \in \Omega : \; k(z,\omega) \leq k(z',\omega).$$

Let H be a real-valued random variable. Then define

$$\forall n \in \mathbb{N} : \; c_n(H) = \frac{\sum_{i=1}^n B_i k(H, A_i)}{\sum_{i=1}^n k(H, A_i)}. \tag{7}$$

Lemma 4.1 (Burba et al. [5])**.** *Let (D_n) be a sequence of real random variables and (u_n) be a decreasing sequence of positive numbers.*

- *If $l = \lim u_n \neq 0$ and if, for all increasing sequences $\beta_n \in (0,1)$, there exist two sequences of real random variables $(D_n^-(\beta_n))$ and $(D_n^+(\beta_n))$ (depending on the sequence (β_n)) such that*

  (L1) $\forall n \in \mathbb{N}$ $D_n^- \leq D_n^+$ and $1_{[D_n^- \leq D_n \leq D_n^+]} \to 1$ almost completely,

(L2) $\frac{\sum_{i=1}^{n} k(D_n^-,A_i)}{\sum_{i=1}^{n} k(D_n^+,A_i)} - \beta_n = \mathcal{O}_{a.co.}(u_n)$,

(L3) *Assume there exists a real positive number* c *such that*
$c_n(D_n^-) - c = \mathcal{O}_{a.co.}(u_n)$ *and* $c_n(D_n^+) - c = \mathcal{O}_{a.co.}(u_n)$.

Then $c_n(D_n) - c = \mathcal{O}_{a.co.}(u_n)$.

- *If* $l = 0$ *and if (L1), (L2), and (L3) hold for any increasing sequence* $\beta_n \in (0,1)$ *with limit* 1, *the same conclusion holds.*

Lemma 4.2 (Burba et al. [5]). *Let* (D_n) *be a sequence of real random variables and* (v_n) *a decreasing positive sequence.*

- *If* $l' = \lim v_n \neq 0$ *and if, for all increasing sequences* $\beta_n \in (0,1)$, *there exist two sequences of real random variables* $(D_n^-(\beta_n))$ *and* $(D_n^+(\beta_n))$ *such that*

(L1') $D_n^- \leq D_n^+ \; \forall n \in \mathbb{N}$ *and* $1_{[D_n^- \leq D_n \leq D_n^+]} \to 1$ *almost completely,*

(L2') $\frac{\sum_{i=1}^{n} k(D_n^-,A_i)}{\sum_{i=1}^{n} k(D_n^+,A_i)} - \beta_n = o_{a.co.}(v_n)$,

(L3') *Assume there exists a real positive number* c *such that*
$c_n(D_n^-) - c = o_{a.co.}(v_n)$ *and* $c_n(D_n^+) - c = o_{a.co.}(v_n)$.

Then $c_n(D_n) - c = o_{a.co.}(v_n)$,

- *If* $l' = 0$ *and if (L1'), (L2'), and (L3') are checked for any increasing sequence* $\beta_n \in (0,1)$ *with limit* 1, *the same result holds.*

Burba et al. [5] use in their consistency proof of the k-NN kernel estimate for independent data a Chernoff-type exponential inequality to check Conditions (L1) or (L1'). In the case of α-mixing random variables however, we cannot use that exponential inequality. Instead we use the following lemma of Bradley [3] and Lemma 4.4.

Lemma 4.3 (Bradley [3] p. 20). *Let* (X,Y) *be a* $\mathbb{R}^r \times \mathbb{R}$ *valued random vector, such that* $Y \in L_p(P)$ *for some* $p \in [1,\infty]$. *Let* d *be a real number such that* $\|Y+d\|_p > 0$ *and* $\varepsilon \in (0, \|Y+d\|_p]$. *Then there exists a random variable* Z *such that*

- $P_Z = P_Y$ *and* Z *is independent of* X,
- $\mathrm{P}(|Z-Y| > \varepsilon) \leq 11 \left(\frac{\|Y+d\|_p}{\varepsilon}\right)^{\frac{p}{2p+1}} [\alpha(\sigma(X),\sigma(Y))]^{\frac{p}{2p+1}}$, *where* $\sigma(X)$ *is the* σ*-Algebra generated by* X.

The following lemma is needed in our proofs for technical reasons.

Lemma 4.4. *Let* (X_i) *be an arithmetically* α*-mixing sequence in the semi-metric space* (E,d), $\alpha(n) \leq cn^{-b}$, *with* $b, c > 0$. *Define* $\Delta_i(x) := 1_{B(x,h)}(X_i)$. *Then we have*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} |\mathrm{Cov}\,(\Delta_i(x),\Delta_j(x))| = \mathcal{O}\,(nF_x(h)) + \mathcal{O}\left(\chi(x,h)^{1-s} n^{1+s}\right),$$

where $\chi(x,h) := \max\{G_x(h), F_x(h)^2\}$ *and* $s = 1/(b+1)$.

Proof. The proof of this lemma is identical to that of Lemma 3.2 of [10], except for the choice of the parameter s.

□

5 Proofs

Proof. Theorem 3.3
To prove this theorem we apply Lemma 4.2. The main difference to the proof of the independent case in [5] concerns verification of (L1'). To verify (L2') and (L3') we need only small modifications.

Let $v_n = 1$, $c_n(H_{n,k}) = \hat{m}_{\text{k-NN}}(x)$ and $c = m(x)$. Choose $\beta \in (0,1)$ arbitrarily, D_n^+ and D_n^- such that $F_x(D_n^+) = \frac{1}{\sqrt{\beta}}\frac{k}{n}$ and $F_x(D_n^-) = \sqrt{\beta}\frac{k}{n}$. Define $h^+ := D_n^+ = F^{-1}\left(\sqrt{\beta}\frac{k}{n}\right)$ and $h^- := D_n^- = F^{-1}\left(\frac{1}{\sqrt{\beta}}\frac{k}{n}\right)$.

To apply Theorem 3.1, we have to show that the covariance term s_n fulfills following condition: there exists a $\theta > 2$ such that

$$s_n^{-(b+1)} = o\left(n^{-\theta}\right), \tag{8}$$

where b is the rate of the mixing coefficient. If (D1) and the condition on the rate b of the mixing coefficient holds, we have by Lemma 11.5 in [11, p. 166]

$$s_n(x) = \mathcal{O}\left(\frac{n}{F_x(h^+)}\right) = \mathcal{O}\left(\frac{n^2}{k}\right).$$

The same is true for the bandwidth h^-. It can be easily seen that there exists an $\theta > 2$ such that (8) holds. In the case of (D2), we have

$$s_n(x) = \mathcal{O}\left(\frac{n^2}{k}\right) + \mathcal{O}\left(\chi(x,h^+)^{1-s}n^{1+s}\right).$$

Since $\chi(x,h^+)^{1-s}n^{1+s} > 0$ for all n, it turns out that (8) holds under Condition (D2) as well.

Consequently, we are able to apply Theorem 3.1 to quarantee

$$c_n(D_n^+) \to c \text{ almost completely, and } c_n(D_n^-) \to c \text{ almost completely.}$$

Thus Condition (L3') is verified.

In [11, p. 162] Ferraty and Vieu proved under the conditions of Theorem 3.1 that

$$\frac{1}{nF_x(h)}\sum_{i=1}^{n} K(h^{-1}d(x,X_i)) \to 1 \text{ almost completely.} \tag{9}$$

By (9) we have

$$\frac{1}{nF_x(h^+)}\sum_{i=1}^{n} K({h^+}^{-1}d(x,X_i)) \to 1 \text{ almost completely and}$$
$$\frac{1}{nF_x(h^-)}\sum_{i=1}^{n} K({h^-}^{-1}d(x,X_i)) \to 1 \text{ almost completely.}$$

We get

$$\frac{\sum_{i=1}^{n} K({h^+}^{-1}d(x,X_i))}{\sum_{i=1}^{n} K({h^-}^{-1}d(x,X_i))} \to \beta.$$

Condition (L2') is proved.

Finally, we check (L1'),

$$\forall\varepsilon > 0: \sum_{n=1}^{\infty} \mathrm{P}\left(|1_{\{D_n^- \le H_{n,k} \le D_n^+\}} - 1| > \varepsilon\right) < \infty.$$

Let $\varepsilon > 0$ be fixed. We know that

$$\mathrm{P}\left(|1_{\{D_n^- \le H_{n,k} \le D_n^+\}} - 1| > \varepsilon\right) \le \mathrm{P}\left(H_{n,k} < D_n^-\right) + \mathrm{P}\left(H_{n,k} > D_n^+\right). \tag{10}$$

For the two terms in (10) we obtain

$$\begin{aligned}\mathrm{P}\left(H_{n,k} < D_n^-\right) &\le \mathrm{P}\left(\sum_{i=1}^n 1_{B(x,D_n^-)}(X_i) > k\right)\\ &\le \mathrm{P}\left(\sum_{i=1}^n \left(1_{B(x,D_n^-)}(X_i) - F_x(D_n^-)\right) > k - nF_x(D_n^-)\right)\\ &=: \mathrm{P}_{1n}\end{aligned} \tag{11}$$

and

$$\begin{aligned}\mathrm{P}\left(H_{n,k} > D_n^+\right) &\le \mathrm{P}\left(\sum_{i=1}^n 1_{B(x,D_n^+)}(X_i) < k\right)\\ &\le \mathrm{P}\left(\sum_{i=1}^n \left(1_{B(x,D_n^+)}(X_i) - F_x(D_n^+)\right) < k - nF_x(D_n^+)\right)\\ &=: \mathrm{P}_{2n}\end{aligned} \tag{12}$$

In the second step of (11) and (12), we centred the random variables.

At this step, Burba et al. [5] use the independence of the random variables. The plan here is to split the data into a block scheme as is done by Modha and Masry [16], Oliveira [17], Tran [18] or Lu and Cheng [15]. Afterwards we are applying Lemma 4.3. Divide the set $\{1, \dots, n\}$ into blocks of length $2l_n$, set $m_n = [n/2l_n]$, where $[\cdot]$ is the Gaussian bracket[3] and $f_n = n - 2l_nm_n < 2l_n$. The sequences are chosen such that $m_n \to \infty$ and $f_n \to \infty$. l_n is specified later on in the proof, see (16). By this choice we have $n = 2l_nm_n + f_n$.

Firstly, we examine term P_{1n}. Let

$$U_n(j) := \sum_{i=(j-1)l_n+1}^{jl_n} \left(1_{B(x,D_n^-)}(X_i) - F_x(D_n^-)\right),$$

and define

$$\begin{aligned}B_{n1} &:= \sum_{j=1}^{m_n} U_n(2j-1), \quad B_{n2} := \sum_{j=1}^{m_n} U_n(2j), \text{ and}\\ R_n &:= \sum_{i=2l_nm_n+1}^{n} \left(1_{B(x,D_n^-)}(X_i) - F_x(D_n^-)\right).\end{aligned}$$

We get

$$\begin{aligned}\mathrm{P}_{1n} &\le \mathrm{P}\left(B_{n1} > \frac{k - nF_x(D_n^-)}{3}\right) + \mathrm{P}\left(B_{n2} > \frac{k - nF_x(D_n^-)}{3}\right)\\ &\quad + \mathrm{P}\left(R_n > \frac{k - nF_x(D_n^-)}{3}\right)\\ &=: \mathrm{P}_{1n}^{(1)} + \mathrm{P}_{1n}^{(2)} + \mathrm{P}_{1n}^{(3)}\end{aligned} \tag{13}$$

Let us consider $\mathrm{P}_{1n}^{(1)}$. Lemma 4.3 with $d := l_nm_n$ leads to $0 < l_nm_n \le \|U_n(2j-1) + d_n\|_\infty \le 2l_n + l_nm_n$. Because of $m_nl_n = \mathcal{O}(n)$ and $k/n \to 0$, we have

$$\varepsilon := \frac{k - nF_x(D_n^-)}{6m_n} = \frac{k(1-\sqrt{\beta})}{6m_n} \in (0, \|U_n(2j-1) + d_n\|_\infty].$$

[3] $[x] = \max\{y \in \mathbb{Z} |\ z \le x\}$, $x \in \mathbb{R}$

This choice of ε is motivated by (15) below. By Lemma 4.3 we can construct $(\tilde{U}_n(2j-1))_{j=1}^{m_n}$ such that

- the random variables $(\tilde{U}_n(2j-1))_{j=1}^{m_n}$ are independent,
- $\tilde{U}_n(2j-1)$ has the same distribution as $U_n(2j-1)$ for $j = 1, \ldots, m_n$,
- and

$$\begin{aligned} \mathrm{P}\left(|\tilde{U}_n(2j-1) - U_n(2j-1)| > \varepsilon\right) \leq & 11 \left(\frac{\|U_n(2j-1) + d\|_\infty}{\varepsilon}\right)^{\frac{1}{2}} \cdot \\ & \cdot \sup |\mathrm{P}(AB) - \mathrm{P}(A)\mathrm{P}(B)|, \end{aligned}$$

where the supremum is taken over all sets $A, B \in \sigma(U_n(1), U_n(3), \ldots, U_n(2m_n-1))$.

This leads to

$$\begin{aligned} \mathrm{P}_{1n}^{(1)} &= \mathrm{P}\left(\sum_{j=1}^{m_n}\left[\tilde{U}_n(2j-1) + (U_n(2j-1) - \tilde{U}_n(2j-1))\right] > \frac{k - nF_x(D_n^-)}{3}\right) \\ &\leq \mathrm{P}\left(\sum_{j=1}^{m_n}\tilde{U}_n(2j-1) > \frac{k - nF_x(D_n^-)}{6}\right) \\ &\quad + \mathrm{P}\left(\sum_{j=1}^{m_n}(U_n(2j-1) - \tilde{U}_n(2j-1)) > \frac{k - nF_x(D_n^-)}{6}\right) \\ &=: \mathrm{P}_{1n}^{(11)} + \mathrm{P}_{1n}^{(12)}. \end{aligned} \tag{14}$$

Applying Lemma 4.3 on $\mathrm{P}_{1n}^{(12)}$,

$$\begin{aligned} \mathrm{P}_{1n}^{(12)} &\leq \sum_{j=1}^{m_n}\mathrm{P}\left((U_n(2j-1) - \tilde{U}_n(2j-1)) > \frac{k - nF_x(D_n^-)}{6m_n}\right) \\ &\leq m_n\left(\frac{6m_n l_n(m_n+1)}{k(1-\sqrt{\beta})}\right)^{\frac{1}{2}}\alpha(l_n) \\ &\leq C\frac{n^2}{l_n^{\frac{3}{2}}k}\alpha(l_n). \end{aligned} \tag{15}$$

We choose the sequence l_n such that

$$l_n^a = \frac{n^2}{2^a r^a k}, \tag{16}$$

where r is a positive constant specified below and $a > 2/\gamma - 1$. By the condition on the mixing coefficient b and some calculations

$$\frac{n^2}{l_n^{3/2}k}\alpha(l_n) = Cn^{(2-\gamma)(a-3/2-b)/a}$$

Consequently, by the assumptions we arrive at

$$\sum_{n=1}^{\infty}\mathrm{P}_{1n}^{(12)} < \infty. \tag{17}$$

Apply now Markov's inequality on term $\mathrm{P}_{1n}^{(11)}$ for some $t > 0$,

$$\begin{aligned}
&\mathrm{P}\left(\sum_{j=1}^{m_n} \tilde{U}_n(2j-1) > \frac{k - nF_x(D_n^-)}{6}\right) \\
&\leq \exp\left(-t\frac{k - nF_x(D_n^-)}{6}\right)\mathrm{E}\left[\exp\left(t\sum_{j=1}^{m_n} \tilde{U}_n(2j-1)\right)\right]. \qquad (18)
\end{aligned}$$

Due to the independence of the random variables $(\tilde{U}_n(2j-1))_{j=1}^{m_n}$, we have

$$\mathrm{E}\left[\exp\left(t\sum_{j=1}^{m_n} \tilde{U}_n(2j-1)\right)\right] = \prod_{j=1}^{m_n} \mathrm{E}\left[\exp\left(t\tilde{U}_n(2j-1)\right)\right]. \qquad (19)$$

Choose $t := r\log n/k$, then we obtain together with l_n as defined in (16)

$$t|\tilde{U}_n(2j-1)| \leq \frac{2rl_n \log n}{k} = \log n\left(\frac{n^2}{k^{a+1}}\right)^{\frac{1}{a}}.$$

In this step, we need the number of neighbours to be a power in n, i.e. $k \sim n^\gamma$. By the choice of $a > 2/\gamma - 1$, we have for large n that $t|\tilde{U}_n(2j-1)| \leq 1$. For large n we have

$$\exp\left(t\tilde{U}_n(2j-1)\right) \leq 1 + t\tilde{U}_n(2j-1) + t^2\tilde{U}_n(2j-1)^2.$$

The random variable $\tilde{U}_n(2j-1)$ has the same distribution as the centred random variable $U_n(2j-1)$. Hence we know that the expectation of the linear term is zero, $\mathrm{E}\left[\tilde{U}_n(2j-1)\right] = 0$. With this and $1 + x \leq \exp(x)$ we get

$$\mathrm{E}\left[\exp\left(t\tilde{U}_n(2j-1)\right)\right] \leq 1 + \mathrm{E}\left[t^2\tilde{U}_n(2j-1)^2\right] \leq \exp\left(t^2\mathrm{E}\left[\tilde{U}_n(2j-1)^2\right]\right). \qquad (20)$$

Furthermore, because $\tilde{U}_n(2j-1)$ and $U_n(2j-1)$ have the same distribution function and by some calculations, it follows that

$$\sum_{j=1}^{m_n} \mathrm{E}\left[\tilde{U}_n(2j-1)^2\right] \leq \sum_{i,j=1}^{n} \left|\mathrm{Cov}(1_{B(x,D_n^-)}(X_i), 1_{B(x,D_n^-)}(X_j))\right|.$$

Since $F_x(D_n^-) = \sqrt{\beta}\frac{k}{n}$ and $k \sim n^\gamma$, we know that $F_x(D_n^-) = \mathcal{O}\left(n^{\gamma-1}\right)$. We apply Lemma 4.4 and get in the case of (D2)

$$\sum_{j=1}^{m_n} \mathrm{E}\left[\tilde{U}_n(2j-1)^2\right] \leq C_1\sqrt{\beta}k + C_2\chi(D_n^-)^{1-s}n^{1+s}, \qquad (21)$$

and in the case of (D1)

$$\sum_{j=1}^{m_n} \mathrm{E}\left[\tilde{U}_n(2j-1)^2\right] \leq C_1\sqrt{\beta}k.$$

Below, we present the arguments if Condition (D2) holds, because in the case of (D1) the rationale follows the same line. By (19), (20), (21), and $t := r\log n/k$, we have for the second term in (18)

$$\begin{aligned}
\mathrm{E}\left[\exp\left(t\sum_{j=1}^{m_n} \tilde{U}_n(2j-1)\right)\right] &\leq \exp\left(C_1\sqrt{\beta}r^2\frac{(\log n)^2}{k}\right)\cdot \\
&\quad \cdot \exp\left(C_2\sqrt{\beta}r^2\frac{(\log n)^2\chi(D_n^-)^{1-s}n^{1+s}}{k^2}\right). \qquad (22)
\end{aligned}$$

By $k \sim n^{\gamma}$, we know that the first term in (22) satisfies

$$\exp\left(C_1\sqrt{\beta}r^2\frac{(\log n)^2}{k}\right) \to 1 \text{ as } n \to \infty.$$

If (D2) holds, we have for the second term in (22)

$$\exp\left(C_2\sqrt{\beta}\mu^2\frac{(\log n)^2\chi(D_n^-)^{1-s}n^{1+s}}{k^2}\right) \to 1 \text{ as } n \to \infty.$$

Since $F_x(D_n^-) = \sqrt{\beta}\frac{k}{n}$, $t = r\log n/k$, and by choosing $r > 6/(1-\sqrt{\beta})$, we find for the first term in (18)

$$\begin{aligned}\exp\left(-t\frac{k-nF_x(D_n^-)}{6}\right) &= \exp\left(-\frac{r(1-\sqrt{\beta})}{6}\log(n)\right)\\ &= n^{-\frac{r(1-\sqrt{\beta})}{6}}\end{aligned}$$

By this,

$$\sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(11)} < \infty \tag{23}$$

Now, combine relations (17) and (23) to obtain

$$\sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(1)} \le \sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(11)} + \sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(12)} < \infty.$$

By similar arguments as for $\mathrm{P}_{1n}^{(1)}$ we receive

$$\sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(2)} < \infty.$$

Finally, we examine

$$\mathrm{P}_{1n}^{(3)} = \mathrm{P}\left(R_n > \frac{k-nF_x(D_n^-)}{3}\right).$$

We know that $|R_n| \le 4l_n$ and $k - nF_x(D_n^-)/3 = \mathcal{O}(k)$. Together with the choice of l_n in (16) and the condition on the parameter $a > 2/\gamma - 1$ we have $k > l_n$ for large n. This implies

$$\sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(3)} < \infty.$$

Finally, we get

$$\sum_{n=1}^{\infty} \mathrm{P}_{1n} \le \sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(1)} + \sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(2)} + \sum_{n=1}^{\infty} \mathrm{P}_{1n}^{(3)} < \infty.$$

Analysis of P_{2n} is similar to that of P_{1n}. This finishes the proof of Condition (L1'), which states that

$$1_{[D_n^- \le D_n \le D_n^+]} \to 1 \text{ almost completely.}$$

Now, we are in the position to apply Lemma 4.2 to obtain the desired result,

$$\lim_{n\to\infty} \hat{m}_{\text{k-NN}}(x) = m(x) \text{ almost completely.}$$

$\square$

Proof. **Theorem 3.4**
To prove this theorem we use Lemma 4.1 from Burba et al. [5]. The conditions of Lemma 4.1 are proven in a similar manner as in the proof of Theorem 3.4. Condition (L1) is the same as (L1') of Lemma 4.2. So the proof can be omitted here. Conditions (L2) and (L3) are checked in a similar way as in the proof of Theorem 3.3. In [11, p. 162] Ferraty and Vieu prove under the conditions of Theorem 3.2 that

$$\frac{1}{n}\sum_{i=1}^{n} K(h^{-1}d(x,X_i)) = \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n(x)\log n}}{n}\right). \tag{24}$$

Choose β_n as an increasing sequence in $(0,1)$ with limit 1. Furthermore, choose D_n^+ and D_n^- such that

$$F_x(D_n^+) = \frac{1}{\sqrt{\beta_n}}\frac{k}{n} \quad \text{and} \quad F_x(D_n^-) = \sqrt{\beta_n}\frac{k}{n}.$$

If (D1) holds, then

$$s_n(x) = \mathcal{O}\left(\frac{n}{F_x(h^+)}\right) = \mathcal{O}\left(\frac{n^2}{k}\right). \tag{25}$$

The same is true for the bandwidth h^-. In the case of (D2), we have for both bandwidth sequences h^- and h^+

$$s_n(x) = \mathcal{O}\left(\frac{n^2}{k}\right) + \mathcal{O}\left(\chi(x,h)^{1-s}n^{1+s}\right). \tag{26}$$

Now we are able to apply Theorem 3.2 with

$$h^+ = D_n^+ = F^{-1}\left(\sqrt{\beta_n}\frac{k}{n}\right) \text{ and } h^- = D_n^- \;=\; F^{-1}\left(\frac{1}{\sqrt{\beta_n}}\frac{k}{n}\right)$$

to get

$$\begin{aligned} c_n(D_n^+) &= \mathcal{O}\left(\left(F_x^{-1}\left(\frac{k}{n}\right)\right)^{\beta}\right) + \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n(x)\log n}}{n}\right) \text{ and} \\ c_n(D_n^-) &= \mathcal{O}\left(\left(F_x^{-1}\left(\frac{k}{n}\right)\right)^{\beta}\right) + \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n(x)\log n}}{n}\right). \end{aligned}$$

That verifies Condition (L3'). Now, by (24) and the same choice of h^+ and h^- as above, we have

$$\begin{aligned} \frac{1}{nF_x(h^+)}\sum_{i=1}^{n} K({h^+}^{-1}d(x,X_i)) &= \sqrt{\beta_n}\frac{k}{n} + \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n(x)\log n}}{n}\right) \text{ and} \\ \frac{1}{nF_x(h^-)}\sum_{i=1}^{n} K({h^-}^{-1}d(x,X_i)) &= \sqrt{\beta_n}\frac{k}{n} + \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n(x)\log n}}{n}\right). \end{aligned}$$

By this, we obtain

$$\frac{\sum_{i=1}^{n} K({h^+}^{-1}d(x,X_i))}{\sum_{i=1}^{n} K({h^-}^{-1}d(x,X_i))} - \beta_n = \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n(x)\log n}}{n}\right).$$

To check Condition (L2') we estimate $s_n(x)$ by bounds obtained either by Condition (D1) and $b > (2-\gamma)/(\varepsilon_1(1-\gamma))$ or by (D2), see (25) or (26). This completes this proof. □

6 Applications and Related Results

Applications

In the context of functional data analysis the k-NN kernel estimate was first introduced in the monograph of Ferraty and Vieu [11]. There the authors give numerical examples for the k-NN kernel estimate. They tested it on different data sets, such as electrical consumption in the U.S. [11, p. 200]. In [9], Ferraty et al. examined a data set describing the El Niño phenomenon. Other interesting examples can be found in the R-package *fds (functional data sets)* or Bosq [4, pp. 247]. For both data sets the assumption of α-mixing is plausible. If we have for example a look on the electrical consumption data set, it makes sense that the electrical consumption of the year which we want to predict is more dependent on the near past than on years afterwards.

Related Results

Here we want to outline how to make a robust k-NN kernel estimate. As already mentioned in the introduction, the k-NN estimate is prone to outliers. This disadvantage can be treated by robust regression estimation. For functional data analysis Azzedine et al. [2] introduce a robust non-parametric regression estimate for independent data. Attouch et al. [1] prove the asymptotic normality for the non-parametric regression estimate for α-mixing data. Crambes et al. [8] present results dealing with the L_p error for independent and α-mixing data.

In robust estimation the non-parametric model θ_x can be defined as the root t of the following equation

$$\Psi(x,t) := \mathrm{E}\left[\psi_x(Y,t)|X=x\right] = 0. \tag{27}$$

The model θ_x is called the $\psi_x(Y,t)$-regression and is a generalisation of the classical regression function. If we choose for example $\psi_x(Y,t) = Y - t$ then we have $\theta_x = m(x)$.

In the case of α-mixing data Almost complete convergence and almost complete convergence rate are not yet proven for robust kernel estimate. These results can be easily obtained by arguments similar to those of Azzedine et al. [2] and those for the classical regression function estimation. By such a result and similar arguments as in this section: we get almost complete convergence and related rates for a robust k-NN non-parametric estimate.

Attouch et al. [1], Azzedine et al. [2], or Crambes et al. [8] suggest in their application the L_1-L_2 function $\psi(t) := t/\sqrt{(1+t^2)/2}$ and $\psi_x(t) := \psi\left(t/M(x)\right)$, where $M(x) := \mathrm{med}|Y - \mathrm{med}(Y|X=x)|$ with $\mathrm{med}(Y|X=x)$ as the conditional median of Y given $X=x$. We get the consistency for the kernel estimate of conditional distribution function directly by choosing in (7) for $B_i = 1_{(-\infty,y]}(Y_i)$ with Y_i as a real valued random variable distributed as Y, and by this a consistent kernel estimate of $\mathrm{med}(Y|X=x)$.

Alternatively, if one has consistency results for a robust k-NN kernel, we can choose $\psi_x(t) = 1_{[t\geq 0]} - 1/2$, to get immediately the consistency for the kernel estimate of the conditional distribution function.

References

[1] Attouch, M., Laksaci, A., and Saïd, E. O. Asymptotic normality of a robust estimator of the regression function for functional time series data. *Journal of the Korean Statistical Society 39*, 4 (2010), 489–500.

[2] Azzedine, N., Laksaci, A., and Ould-Saïd, E. On robust nonparametric regression estimation for a functional regressor. *Statistics & Probability Letters 78*, 18 (2008), 3216–3221.

[3] Bosq, D. *Nonparametric statistics for stochastic processes: estimation and prediction.* Lecture notes in statistics. Springer, 1998.

[4] Bosq, D. *Linear Processes in Function Spaces.* Springer, Berlin, 2000.

[5] Burba, F., Ferraty, F., and Vieu, P. k-Nearest Neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics 21*, 4 (2009), 453–469.

[6] Cérou, F., and Guyader, A. Nearest neighbor classification in infinite dimension. *ESAIM: P&S 10* (2005), 340–255.

[7] Collomb, G. Estimation de la regression par la methode des k points les plus proches avec noyau : quelques propriétés de convergence ponctuelle. In *Statistique non Paramétrique Asymptotique*, J.-P. Raoult, Ed., vol. 821 of *Lecture Notes in Mathematics.* Springer Berlin / Heidelberg, 1980, pp. 159–175.

[8] Crambes, C., Delsol, L., and Laksaci, A. Robust nonparametric estimation for functional data. *Journal of Nonparametric Statistics 20*, 7 (Oct. 2008), 573–598.

[9] Ferraty, F., Rabhi, A., and Vieu, P. Conditional Quantiles for Dependent Functional Data with Application to the Climatic El Niño Phenomenon. *Sankhya: The Indian Journal of Statistics 67* (2005), 378–398.

[10] Ferraty, F., and Vieu, P. Non-parametric models for functional data, with application in regression, time-series prediction and curve discrimination. *Journal of Nonparametric Statistics 16*, 1 & 2 (2004), 111–125.

[11] Ferraty, F., and Vieu, P. *Nonparametric Functional Data Analysis.* Springer, New York, 2006.

[12] Györfi, L., Kohler, M., Krzyzak, A., and Walk, H. *A Distribution-Free Theory of Nonparametric Regression.* Springer, New-york, 2002.

[13] Laloë, T. A k-nearest neighbor approach for functional regression. *Statistics & Probability Letters 78*, 10 (2008), 1189–1193.

[14] Lian, H. Convergence of functional k-nearest neighbor regression estimate with functional responses. *Electronic Journal of Statistics 5* (2011), 31–40.

[15] Lu, Z., and Cheng, P. Strong consistency of nearest neighbor kernel regression estimation for stationary dependent samples. *Science in China Series A: Mathematics 41*, 9 (1998), 918–926.

[16] Modha, D. S., and Masry, E. Minimum Complexity Regression Estimation with Weakly Dependent Observations. *IEEE Transactions on Information Theory 42*, 6 (1996), 2133–2145.

[17] Oliveira, P. E. Nonparametric density and regression estimation functional data. Tech. rep., Departamento de Matemática, Universidade de Coimbra, 2005.

[18] Tran, L. T. Nonparametric Estimation for Time Series by Local Average Estimators. *The Annals of Statistics 42*, 2 (1993), 1040–1057.

Simon Müller
Pfaffenwaldring 57
70569 Stuttgart
Germany
E-Mail: Simon.Mueller@mathematik.uni-stuttgart.de
WWW: http://www.isa.uni-stuttgart.de/LstStoch/Mueller/

Jürgen Dippon
Pfaffenwaldring 57
70569 Stuttgart
Germany
E-Mail: dippon@mathematik.uni-stuttgart.de
WWW: http://www.isa.uni-stuttgart.de/LstStoch/Dippon/

Erschienene Preprints ab Nummer 2007/001

Komplette Liste: `http://www.mathematik.uni-stuttgart.de/preprints`

2011/014 *Müller, S.; Dippon, J.:* k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis

2011/013 *Knarr, N.; Stroppel, M.:* Unitals over composition algebras

2011/012 *Knarr, N.; Stroppel, M.:* Baer involutions and polarities in Moufang planes of characteristic two

2011/011 *Knarr, N.; Stroppel, M.:* Polarities and planar collineations of Moufang planes

2011/010 *Jentsch, T.; Moroianu, A.; Semmelmann, U.:* Extrinsic hyperspheres in manifolds with special holonomy

2011/009 *Wirth, J.:* Asymptotic Behaviour of Solutions to Hyperbolic Partial Differential Equations

2011/008 *Stroppel, M.:* Orthogonal polar spaces and unitals

2011/007 *Nagl, M.:* Charakterisierung der Symmetrischen Gruppen durch ihre komplexe Gruppenalgebra

2011/006 *Solanes, G.; Teufel, E.:* Horo-tightness and total (absolute) curvatures in hyperbolic spaces

2011/005 *Ginoux, N.; Semmelmann, U.:* Imaginary Kählerian Killing spinors I

2011/004 *Scherer, C.W.; Köse, I.E.:* Control Synthesis using Dynamic D-Scales: Part II — Gain-Scheduled Control

2011/003 *Scherer, C.W.; Köse, I.E.:* Control Synthesis using Dynamic D-Scales: Part I — Robust Control

2011/002 *Alexandrov, B.; Semmelmann, U.:* Deformations of nearly parallel G_2-structures

2011/001 *Geisinger, L.; Weidl, T.:* Sharp spectral estimates in domains of infinite volume

2010/018 *Kimmerle, W.; Konovalov, A.:* On integral-like units of modular group rings

2010/017 *Gauduchon, P.; Moroianu, A.; Semmelmann, U.:* Almost complex structures on quaternion-Kähler manifolds and inner symmetric spaces

2010/016 *Moroianu, A.; Semmelmann,U.:* Clifford structures on Riemannian manifolds

2010/015 *Grafarend, E.W.; Kühnel, W.:* A minimal atlas for the rotation group $SO(3)$

2010/014 *Weidl, T.:* Semiclassical Spectral Bounds and Beyond

2010/013 *Stroppel, M.:* Early explicit examples of non-desarguesian plane geometries

2010/012 *Effenberger, F.:* Stacked polytopes and tight triangulations of manifolds

2010/011 *Györfi, L.; Walk, H.:* Empirical portfolio selection strategies with proportional transaction costs

2010/010 *Kohler, M.; Krzyżak, A.; Walk, H.:* Estimation of the essential supremum of a regression function

2010/009 *Geisinger, L.; Laptev, A.; Weidl, T.:* Geometrical Versions of improved Berezin-Li-Yau Inequalities

2010/008 *Poppitz, S.; Stroppel, M.:* Polarities of Schellhammer Planes

2010/007 *Grundhöfer, T.; Krinn, B.; Stroppel, M.:* Non-existence of isomorphisms between certain unitals

2010/006 *Höllig, K.; Hörner, J.; Hoffacker, A.:* Finite Element Analysis with B-Splines: Weighted and Isogeometric Methods

2010/005 *Kaltenbacher, B.; Walk, H.:* On convergence of local averaging regression function estimates for the regularization of inverse problems

2010/004 *Kühnel, W.; Solanes, G.:* Tight surfaces with boundary

2010/003 *Kohler, M; Walk, H.:* On optimal exercising of American options in discrete time for stationary and ergodic data

2010/002 *Gulde, M.; Stroppel, M.:* Stabilizers of Subspaces under Similitudes of the Klein Quadric, and Automorphisms of Heisenberg Algebras

2010/001 *Leitner, F.:* Examples of almost Einstein structures on products and in cohomogeneity one

2009/008 *Griesemer, M.; Zenk, H.:* On the atomic photoeffect in non-relativistic QED

2009/007 *Griesemer, M.; Moeller, J.S.:* Bounds on the minimal energy of translation invariant n-polaron systems

2009/006 *Demirel, S.; Harrell II, E.M.:* On semiclassical and universal inequalities for eigenvalues of quantum graphs

2009/005 *Bächle, A, Kimmerle, W.:* Torsion subgroups in integral group rings of finite groups

2009/004 *Geisinger, L.; Weidl, T.:* Universal bounds for traces of the Dirichlet Laplace operator

2009/003 *Walk, H.:* Strong laws of large numbers and nonparametric estimation

2009/002 *Leitner, F.:* The collapsing sphere product of Poincaré-Einstein spaces

2009/001 *Brehm, U.; Kühnel, W.:* Lattice triangulations of E^3 and of the 3-torus

2008/006 *Kohler, M.; Krzyżak, A.; Walk, H.:* Upper bounds for Bermudan options on Markovian data using nonparametric regression and a reduced number of nested Monte Carlo steps

2008/005 *Kaltenbacher, B.; Schöpfer, F.; Schuster, T.:* Iterative methods for nonlinear ill-posed problems in Banach spaces: convergence and applications to parameter identification problems

2008/004 *Leitner, F.:* Conformally closed Poincaré-Einstein metrics with intersecting scale singularities

2008/003 *Effenberger, F.; Kühnel, W.:* Hamiltonian submanifolds of regular polytope

2008/002 *Hertweck, M.; Hofert, C.R.; Kimmerle, W.:* Finite groups of units and their composition factors in the integral group rings of the groups $PSL(2,q)$

2008/001 *Kovarik, H.; Vugalter, S.; Weidl, T.:* Two dimensional Berezin-Li-Yau inequalities with a correction term

2007/006 *Weidl, T.:* Improved Berezin-Li-Yau inequalities with a remainder term

2007/005 *Frank, R.L.; Loss, M.; Weidl, T.:* Polya's conjecture in the presence of a constant magnetic field

2007/004 *Ekholm, T.; Frank, R.L.; Kovarik, H.:* Eigenvalue estimates for Schrödinger operators on metric trees

2007/003 *Lesky, P.H.; Racke, R.:* Elastic and electro-magnetic waves in infinite waveguides

2007/002 *Teufel, E.:* Spherical transforms and Radon transforms in Moebius geometry

2007/001 *Meister, A.:* Deconvolution from Fourier-oscillating error densities under decay and smoothness restrictions