ESAIM: PS 18 (2014) 130-144
DOI: 10.1051/ps/2013030

ESAIM: Probability and Statistics
www.esaim-ps.org

DENSITY SMOOTHNESS ESTIMATION PROBLEM USING A WAVELET APPROACH

Karol Dziedziul[1] and Bogdan Ćmiel[2]

Abstract. In this paper we consider a smoothness parameter estimation problem for a density function. The smoothness parameter of a function is defined in terms of Besov spaces. This paper is an extension of recent results (K. Dziedziul, M. Kucharska, B. Wolnik, *Estimation of the smoothness parameter*). The construction of the estimator is based on wavelets coefficients. Although we believe that the effective estimation of the smoothness parameter is impossible in general case, we can show that it becomes possible for some classes of the density functions.

Mathematics Subject Classification. 62G05, 62G07.

Received July 17, 2012. Revised October 29, 2012.

1. Introduction

Let us define a smoothness parameter of a function $f \in L^p(\mathbb{R})$, $1 \leq p < \infty$ or $f \in C_b(\mathbb{R})$ in case $p = \infty$ ($C_b(\mathbb{R})$ is the space of all continues and bounded functions) as

$$s_p^* = s_p^*(f) = \sup\{s : f \in B_{p\infty}^s\}. \tag{1.1}$$

We use usual convention $\sup \emptyset = 0$ and $\sup \mathbb{R} = \infty$. A definition of Besov spaces $B_{p\infty}^s$ is given below. The smoothness parameter for Hölder–Zygmund space, *i.e.* $p = \infty$ was introduced in [12,19]. The case for $p = 2$ was introduced in [11].

The smoothness and regularity of the density function or regression function is essential in analyzing the rate of convergence of the estimators of that functions. We usually assume that the considered functions are differentiable (see [14, 23]) or we define the smoothness in terms of Sobolev or Besov spaces (see [4, 9, 10]), which allow us to replace a discrete smoothness scale by continuous one. The adaptive methods presented in articles [6, 7, 12, 16] make possible to achieve the optimal rate of convergence (or almost optimal) of a density estimator when the real smoothness is unknown. There are also non-adaptive methods of estimation which require the knowledge of the number of derivatives or the exact value of the parameters of Sobolev or Besov space. In practice, there is a fundamental problem with verifying the smoothness assumption. There are some methods developed in [18], (see also a discussion in [16]), where Fourier coefficients are used to determine

Keywords and phrases. Estimation, wavelets, Besov spaces, smoothness parameter.

[1] Faculty of Applied Mathematics, Gdańsk University of Technology, ul. G. Narutowicza 11/12, 80-233 Gdańsk, Poland. kdz@mifgate.pg.gda.pl

[2] Faculty of Applied Mathematics, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Cracov, Poland. cmielbog@gmail.com

Article published by EDP Sciences

© EDP Sciences, SMAI 2013

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

whether a function belongs to an anisotropic Sobolev space. In some classical papers [1,17] the smoothness tests are also studied. In [1] the Sobolev spaces are used which corresponds to the Besov spaces B^s_{22} so it is in our area of interest. The paper [13] gives a new light on our effort. This paper confirms the need of estimating the family of smoothness parameters.

The effective estimation of the smoothness parameter over some wide function classes (like Besov spaces) seems to be impossible. This problem was previously analyzed by Low [20] and Cai and Low [5]. This of course does not mean that it is impossible in any case. We show in this paper that there exists some classes of the density functions, where we can effectively estimate the smoothness parameter of those densities.

On the other hand we have now results of Bull [2,3], but the relation between our result and his results are like between a parametric estimation and a confidance interval estimation. In this paper we use the methods developed in [11]. We concentrate on an estimation of the smoothness parameter. We define a class of the density functions for which the smoothness estimator is strongly consistent. In consequence we can analyze an asymptotic behaviour of such estimator and use it in the testing hypothesis problem. This will be done in next paper. It is also worth to see that our methods allowed to take care of the "Change-point" of the function smoothness from [17] *i.e.* to examine the defect of the function smoothness. Although we define the smoothness in a global way, we can study the problem locally by using wavelet approach (for example, we can check the smoothness in some neighbourhood of the edge of the support).

The paper is organized as follows. We introduce a wavelet analysis of the Besov spaces and a smoothness parameter in Section 2. In Section 3 we consider a case of piecewise-smooth density function. In Section 4 we show that our estimator is "pseudo-consistent" in general case and strongly consistent in the case, where we assume that the density function is some piecewise-smooth function. In Section 5 we demonstrate the behaviour of our estimator in a numerical experiment.

2. The smoothness parameter

Let be given r-regular multiresolution analysis (for the definition see [8,21], Defs. 1, 2 in Sect. 2.2). By [21], Theorem 3 in Section 3.8 we know that for each integer $r \geq 1$ there exists r-regular multiresolution analysis such that an associated scaling function (a father wavelet) ϕ and a wavelet (mother wavelet) ψ have compact support. In this paper we assume that ϕ and ψ have compact support. We will use following notation:

$$\phi_k(x) = \phi(x-k) \quad \psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),\ x \in \mathbb{R}.$$

We assume that $\operatorname{supp}\psi = [0, S(r)]$.

Note that by [8], Corollary 5.5.2 if we have r-regular multiresolution analysis then

$$\int_{-\infty}^{\infty} x^p \psi(x)\mathrm{d}x = 0 \quad \text{for } p = 0, 1, \ldots, r. \tag{2.1}$$

The wavelet coefficients of some function f are denoted by

$$\alpha_k := \alpha_k(f) = \int_{\mathbb{R}} \phi_k(x) f(x)\mathrm{d}x, \quad \beta_{jk} = \beta_{jk}(f) := \int_{\mathbb{R}} \psi_{jk}(x) f(x)\mathrm{d}x.$$

Let us recall some facts of the wavelets theory and Besov spaces B^s_{pq}, (see [15]). The full characterization of Besov spaces by spline basis coefficients was made in [22]. We will present the Besov space $B^s_{p,\infty}$ characterization in terms of wavelet coefficients.

A function $f \in L^p(\mathbb{R})$ with the following expansion $f = \sum_{k\in\mathbb{Z}} \alpha_k\phi_k + \sum_{j=1}^{\infty}\sum_{k\in\mathbb{Z}} \beta_{jk}\psi_{jk}$ belongs to the Besov space $B^s_{p\infty}$, for $s < r$ and $1 \leq p < \infty$, if and only if

$$\sup_{j\geq 0}(2^{j(s-1/p+1/2)}||\beta_{j\cdot}||_p) < \infty \quad \text{and} \quad ||\alpha||_p < \infty,$$

where $\|\cdot\|_p$ denote l^p norm of the sequences $\beta_{j\cdot} = \{\beta_{jk}\}_{k\in\mathbb{Z}},\ j \geq 1$ and $\alpha = \{\alpha_k\}_{k\in\mathbb{Z}}$.

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

Similar for $B^s_{\infty,\infty}$ (see [12])

$$B^s_{\infty,\infty} = \{f \in C_b(\mathbb{R}) : \{\alpha_k(f)\} \in l^\infty, \quad \sup_{j\geq 0}(2^{j(s+1/2)}\|\beta_{j\cdot}\|_\infty) < \infty\}.$$

Note that if $f \in L^p(\mathbb{R})$ (or $f \in C_b(\mathbb{R})$) and $s^*_p > 0$ then by (1.1)

$$\min\{s^*_p, r\} = \sup\{s < r : \sup_{j\geq 0} 2^{j(s-1/p+1/2)}\|\beta_{j\cdot}\|_p < \infty\}. \tag{2.2}$$

Now we will prove the following theorem (conf. [11] Thm. 1.1).

Theorem 2.1. *Let be given r-regular multiresolution analysis, $r \geq 1$. Let $1 \leq p \leq \infty$. If $0 < s^*_p < r$ then the set of indices*

$$J := \{j \geq 1 : \|\beta_{j\cdot}\|_p \neq 0\}$$

has infinity many elements and

$$\liminf_{j\to\infty,\ j\in J} \frac{-\log_2 \|\beta_{j\cdot}\|_p}{j} = s^*_p - \frac{1}{p} + \frac{1}{2}. \tag{2.3}$$

Remark 2.2. If we use standard notation $\log_2 0 = -\infty$ then we state that

$$\liminf_{j\to\infty} \frac{-\log_2 \|\beta_{j\cdot}\|_p}{j} = s^*_p - \frac{1}{p} + \frac{1}{2}.$$

Proof. Let us fix p. The proof is similar to the one given in [11]. By (2.2) if the set J has finite number of elements then $s^*_p \geq r$ which contradicts the assumption $0 < s^*_p < r$.

Let us denote

$$\tilde{s} = \liminf_{j\to\infty,\ j\in J} \frac{-\log_2 \|\beta_{j\cdot}\|_p}{j} + \frac{1}{p} - \frac{1}{2}.$$

If $s < \tilde{s}$ then there exists N_s such that for all $j \geq N_s$ (recall that $j \in J$)

$$s - \frac{1}{p} + \frac{1}{2} < \frac{-\log_2 \|\beta_{j\cdot}\|_p}{j}.$$

Consequently for all $j \geq N_s$

$$2^{j(s-\frac{1}{p}+\frac{1}{2})}\|\beta_{j\cdot}\|_p < 1.$$

This shows that $\tilde{s} \leq s^*_p$. On the other hand from the definition of $\tilde{s}$, for every $\epsilon > 0$ there exists subsequence $j_n \in J$ and N_ϵ such that for $j_n > N_\epsilon$

$$\frac{-\log_2 \|\beta_{j_n\cdot}\|_p}{j_n} \leq \tilde{s} - \frac{1}{p} + \frac{1}{2} + \epsilon.$$

Now if we take any $s > \tilde{s}$ and $2\epsilon = s - \tilde{s}$ then for $j_n > N_\epsilon$

$$2^{j_n(s-\frac{1}{p}+\frac{1}{2})}\|\beta_{j_n\cdot}\|_p > 2^{j_n(s-\tilde{s}-\epsilon)} = 2^{j_n\epsilon},$$

which implies that $\tilde{s} \leq s^*_p$ and completes the proof of Theorem 2.1. □

This theorem shows, how we can determine the smoothness of a given function. Of course, there is a problem with calculating the lower limit of a sequence. In practice we are only able to calculate a finite number of some sequence values so it is impossible to effectively approximate the lower limit of that sequence. It can be shown, however, that in some cases we can change "lim inf" to "lim" in Theorem 2.1. For example, if function f is a spline of degree m, then the sequence in Theorem 2.1 is convergent and the smoothness parameter is equal to $s^*_p(f) = m + 1/p$ (we will show this in the next section). Even if we cannot change "lim inf" to "lim" (the sequence is not convergent), then we can always evaluate smoothness parameter subjectively by observing the behaviour of that sequence.

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

3. Piecewise-smooth function case

Let us recall that supp $\psi = [0; S(r)]$. It is an essential assumption that a wavelet ψ satisfies following condition: there exist $0 < \delta_1 < 1$ such that

$$|\psi(x)| > 0, \quad x \in (0, 1 + \delta_1]. \tag{3.1}$$

Let us denote

$$\Psi_1 := \min_{0 \le n \le r} \left\{ \left| \int_0^{\delta_1} (\delta_1 - u)^n \psi(u) \mathrm{d}u \right| \right\}.$$

It is easy to see that if $\psi \in C(\mathbb{R})$ satisfies (3.1) then $\Psi_1 > 0$ and for all $\delta_1 < \eta \le 1 + \delta_1$ and for all $0 \le n \le r$ we have

$$\left| \int_0^{\eta} (\eta - u)^n \psi(u) \mathrm{d}u \right| \ge \Psi_1. \tag{3.2}$$

By the moment condition (2.1) we have also equivalent form of (3.2) *i.e.* for all $\delta_1 < \eta \le 1 + \delta_1$

$$\left| \int_0^{\eta} (\eta - u)^n \psi(u) \mathrm{d}u \right| = \left| \int_{\eta}^{S(r)} (\eta - u)^n \psi(u) \mathrm{d}u \right| \ge \Psi_1. \tag{3.3}$$

We will show that all Daubechies wavelets ${}_N\psi$, where $N = 2, \ldots, 20$ (see [8], Chapt. 6) satisfy the condition (3.1). We have similar result for a function ${}_N\phi$, (3.9). For Daubechies wavelets we will also use the notation DB2, DB3, ..., DB20.

Since we assume that supp ${}_N\psi = [0; S(r) = 2N - 1]$ it means that we use Daubechies wavelets translated by $N + 1$ and hence supp ${}_N\phi = [N + 1; 3N]$

Lemma 3.1. *Daubechies wavelets ${}_N\psi$ satisfy the condition (3.1), where $N = 2, \ldots, 20$.*

Proof. The proof is based on numerical results for a cascade algorithm for Daubechies father wavelets ${}_N\phi$, $N = 2, 3, \ldots$ (we assume that numerical error is negligible). We know that a difference in L^∞ norm between a numerical father wavelet $\widetilde{{}_N\phi_j}$, for fixed $j \ge 0$ given in dyadic points $k2^{-j}$, $k \in \mathbb{Z}$, interpolate by piecewise constant or piecewise linear and a father wavelet $N\phi$ is given by

$$\|{}_N\phi - \widetilde{{}_N\phi_j}\|_{L^\infty} \le C 2^{-\alpha j}$$

(see [8], Prop. 6.5.2). On the other hand the main property of the wavelets ([8], Chapt. 6, Prop. 6.3.6) states that there is a sequence b_k such that

$${}_N\psi(x) = \sum_{k \in Z} (-1)^k b_{-k+1} {}_N\phi(2x - k) \tag{3.4}$$

and

$${}_N\phi(x) = \sum_{k \in Z} b_k {}_N\phi(2x - k), \quad x \in \mathbb{R}. \tag{3.5}$$

This formula is used to calculate the numerical mother wavelet $\widetilde{{}_N\psi_j}$ using the numerical father wavelet $\widetilde{{}_N\phi_j}$. Consequently we know a difference in L^∞ norm between the numerical mother wavelet $\widetilde{{}_N\psi_j}$ and the mother wavelet ${}_N\psi$. If $j \ge 0$ is large enough one can observe that for numerical mother wavelet $\widetilde{{}_N\psi_j}$ ($N = 2, 3, \ldots 20$), there exists $\beta_{N1} < 1/2 < 1 < \beta_{N2}$ and $c_N > 0$ such that

$$|\widetilde{{}_N\psi_j}(x)| \ge c_N > 0, \quad \text{for} \quad x \in [\beta_{N1}, \beta_{N2}]. \tag{3.6}$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

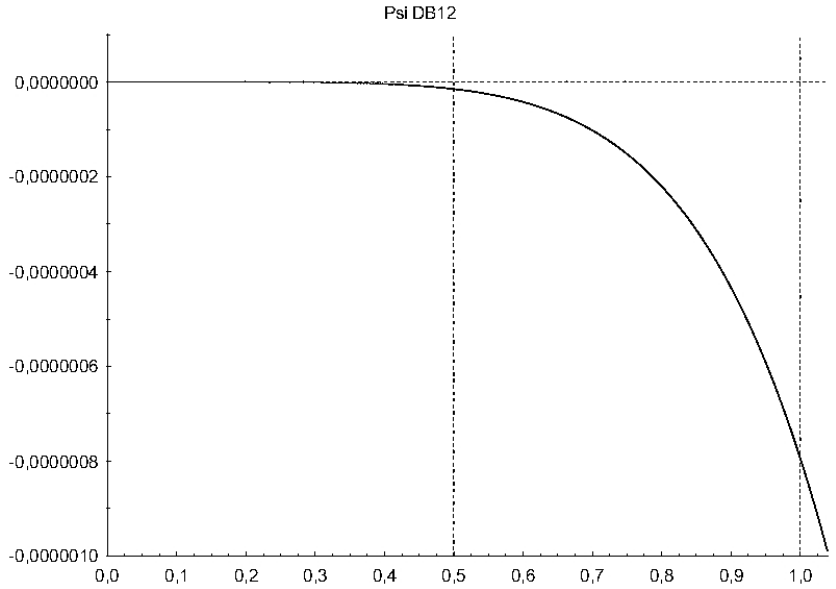


FIGURE 1. The example of Daubechies, numerical mother wavelet for $N = 12$ ($S(r) = 23$) on the interval $[0; 1]$, in a dyadic points $k2^{-10}$.

By (3.6) and the numerical accuracy

$$|{}_N\psi(x)| \geq c_N/2 > 0, \quad \text{for} \quad x \in [\beta_{N1}, \beta_{N2}]. \tag{3.7}$$

Hence to show (3.1) it is sufficient to prove it only for $x \in (0, 1/2]$. Since ${}_N\psi$ and ${}_N\phi$ have compact supports by (3.4) we have

$${}_N\psi(x) = b_{-N-1\,N}\phi(2x + N + 1), \quad \text{for} \quad x \in [0, 1/2]. \tag{3.8}$$

We know that $b_{-N-1} \neq 0$ so we have to check if

$$|{}_N\phi(x)| > 0, \quad \text{for} \quad x \in (N + 1, N + 1 + 1]. \tag{3.9}$$

Calculating the values of the numerical father wavelets in the interval $[N + 1 + 1/2, N + 1 + 1]$ (recall that we shift the wavelet by $N + 1$) and taking to account the numerical precision we get

$$|{}_N\phi(x)| > C_N > 0, \quad x \in [N + 1 + 1/2, N + 1 + 1]. \tag{3.10}$$

Now by (3.5) we get

$$|{}_N\phi(x)| = |b_{N+1}||{}_N\phi(2x - (N + 1))|, \quad x \in [N + 1, N + 1 + 1/2],$$

and by (3.10) we have

$$|{}_N\phi(x)| > C_N|b_{N+1}| > 0, \quad x \in [N + 1 + 1/4, N + 1 + 1/2],$$

and so on. Finally we get (3.9) which finishes the proof. □

Now let us introduce for $m \in \mathbb{N}$ and for arbitrary $a \in \mathbb{R}$ two functions

$$V_a^1(x) = V_{m,a}^1(x) = (x - a)_+^m$$
$$V_a^2(x) = V_{m,a}^2(x) = (a - x)_+^m.$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

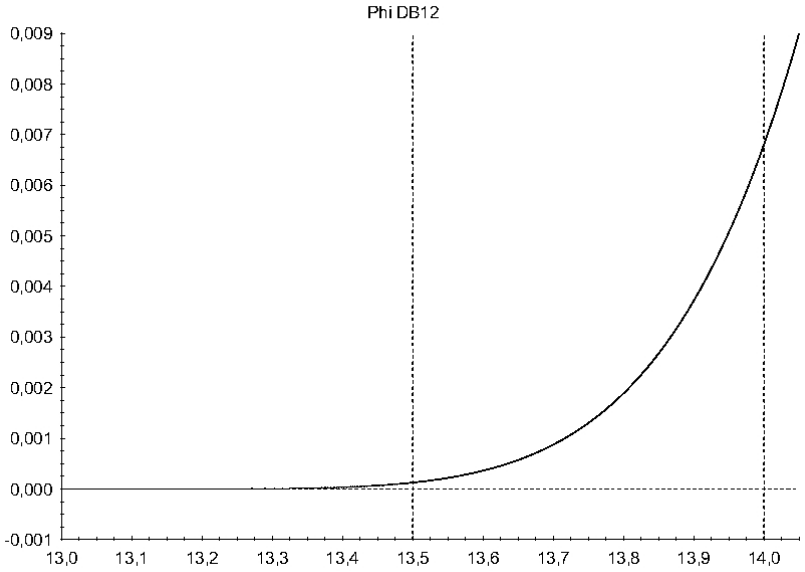


FIGURE 2. The example of Daubechies, numerical father wavelet for $N = 12$ on the interval $[13; 14]$, in a dyadic points $k2^{-10}$.

This functions do not belong to $L^p(\mathbb{R})$ or $C_b(\mathbb{R})$ but it makes sense to use them in the following lemma:

Lemma 3.2. *Let be given r-regular multiresolution analysis, where $r \geq 1$ and $m \leq r$. If $\psi(x)$ satisfies (3.2) and $1 \leq p \leq \infty$ then*

$$\lim_{j\to\infty} \frac{-\log_2 \|\beta^i_{j\cdot}\|_p}{j} = m + \frac{1}{2},$$

where $\beta^i_{jk} := \int_{\mathbb{R}} \psi_{jk}(x) V^i_a(x) \mathrm{d}x$, for $i = 1, 2$.

Proof. Let $1 \leq p < \infty$. It is sufficient to prove that for all j large enough

$$\frac{\Psi_1^p}{2^{jp(1/2+m)}} \leq \|\beta^i_{j\cdot}\|_p^p \leq \frac{\Psi_2^p (S(r))^{p(m+1)}}{2^{jp(1/2+m)}(m+1)^p}, \tag{3.11}$$

for $i = 1, 2$ where

$$\Psi_2 = \sup |\psi(u)|.$$

Note that if $\lfloor x \rfloor = \sup\{n \in \mathbb{Z} : n < x\}$ then

$$\|\beta^1_{j\cdot}\|_p^p = \frac{1}{2^{jp(1/2+m)}} \sum_{\lfloor 2^j a \rfloor - S(r)+1 \leq k \leq \lfloor 2^j a \rfloor} \left| \int_{2^j a - k}^{+\infty} (u - (2^j a - k))^m \psi(u) \mathrm{d}u \right|^p,$$

$$\|\beta^2_{j\cdot}\|_p^p = \frac{1}{2^{jp(1/2+m)}} \sum_{\lfloor 2^j a \rfloor - S(r)+1 \leq k \leq \lfloor 2^j a \rfloor} \left| \int_{-\infty}^{2^j a - k} ((2^j a - k) - u)^m \psi(u) \mathrm{d}u \right|^p.$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

The right side of the inequality (3.11) is obvious. To prove the left side let us take such k that $\eta = 2^j a - k \in [\delta_1, 1+\delta_1]$ (see (3.2)). Then

$$\|\beta^1_{j\cdot}\|_p^p \geq \frac{1}{2^{jp(1/2+m)}} \left| \int_{\eta}^{S(r)} (u-\eta)^m \psi(u) \mathrm{d}u \right|^p \geq \frac{\Psi_1^p}{2^{jp(1/2+m)}}. \tag{3.12}$$

We proceed analogously for $i = 2$. A case $p = \infty$ is a consequence of the following inequality: for all $k \in \mathbb{Z}$, $i = 1, 2$ and $j \geq 0$

$$|\beta^i_{jk}| \leq \|\beta^i_{j\cdot}\|_\infty \leq \|\beta^i_{j\cdot}\|_1. \qquad \square$$

Using the lemma above we will prove the following corollary (by C^{-1} we denote measurable functions)

Corollary 3.3. *Let $p > 1$. Let $f \in C^{m-1}(\mathbb{R})$ be a bounded function with a compact support, where $m \in \mathbb{N} \cup \{0\}$, $f \in C^{m+1}((-\infty, a])$ and $f \in C^{m+1}([a, \infty))$ but $f^{(m)}(a^-) \neq f^{(m)}(a^+)$. Let be given r-regular multiresolution analysis, where $r \geq 1$ and $m + 1 \leq r$. If $\psi(x)$ satisfies (3.2) then*

$$\lim_{j\to\infty} \frac{-\log_2 \|\beta_{j\cdot}(f)\|_p}{j} + \frac{1}{p} - \frac{1}{2} = s_p^*(f) = m + \frac{1}{p},$$

where $\beta_{jk}(f) := \int_{\mathbb{R}} \psi_{jk}(x) f(x) \mathrm{d}x$.

Remark 3.4. We will prove Corollary 3.3 in a special case for $p = 2$. The prove in the other cases is a simple analogy.

Proof. Let us consider a function $h \in C^{m+1}(\mathbb{R} \setminus \{a\})$ with a compact support such that in a neighbourhood of a

$$\begin{aligned} h(x) = {} & \frac{(-1)^{m+1}}{(m+1)!} f^{(m+1)}(a^-)[(a-x)_+]^{m+1} + \frac{1}{(m+1)!} f^{(m+1)}(a^+)[(x-a)_+]^{m+1} \\ & + \frac{(-1)^m}{m!} f^{(m)}(a^-)[(a-x)_+]^m + \frac{1}{m!} f^{(m)}(a^+)[(x-a)_+]^m. \end{aligned}$$

Note that $f - h$ belongs to $C^{m+1}(\mathbb{R})$ and have a compact support. In fact,

$$(f-h)^{(m+1)}(a^+) = (f-h)^{(m+1)}(a^-) = 0$$

and

$$(f-h)^{(m)}(a^+) = (f-h)^{(m)}(a^-) = 0.$$

By C_c^{m+1} let us denote a space of $m+1$ times differentiable functions with compact supports. We have an embedding $C_c^{m+1} \subset W_2^{m+1} \subset B^s_{2,\infty}$ for all $s < m+1$, where W_2^{m+1} denotes the Sobolev space. Hence we conclude that

$$\|\beta_{j\cdot}(f-h)\|_2 \leq C_1 \cdot 2^{-(m+3/4)j}$$

for some constant C_1 and j large enough.

Let

$$h_1(x) = \frac{(-1)^m}{m!} f^{(m)}(a^-)[(a-x)_+]^m + \frac{1}{m!} f^{(m)}(a^+)[(x-a)_+]^m.$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

From the proof of Lemma 3.2 it follows for the function h_1 that we have

$$\begin{aligned}\|\beta_{j\cdot}(h_1)\|_2^2 = \frac{1}{2^{2j(1/2+m)}} \sum_{\lfloor 2^j a\rfloor - S(r)+1 \le k \le \lfloor 2^j a\rfloor} \Bigg| \frac{1}{m!} f^{(m)}(a^+) \int_{2^j a-k}^{+\infty} (u-(2^j a-k))^m \psi(u)\mathrm{d}u \\ + \frac{(-1)^m}{m!} f^{(m)}(a^-) \int_{-\infty}^{2^j a-k} ((2^j a-k)-u)^m \psi(u)\mathrm{d}u \Bigg|^2 \\ = \frac{1}{2^{2j(1/2+m)}} \left(\frac{1}{m!}\right)^2 \sum_{\lfloor 2^j a\rfloor - S(r)+1 \le k \le \lfloor 2^j a\rfloor} \Bigg| f^{(m)}(a^+) \int_{2^j a-k}^{+\infty} (u-(2^j a-k))^m \psi(u)\mathrm{d}u \\ + f^{(m)}(a^-) \int_{-\infty}^{2^j a-k} ((u-2^j a-k))^m \psi(u)\mathrm{d}u \Bigg|^2 .\end{aligned}$$

Hence using (3.12) in Lemma 3.2 and the moment condition (2.1)

$$\|\beta_{j\cdot}(h_1)\|_2^2 \ge \frac{\Psi_1^2}{2^{2j(1/2+m)}} \left(\frac{1}{m!}\right)^2 |f^{(m)}(a^+) - f^{(m)}(a^-)|^2.$$

Once again from Lemma 3.2 it follows that for h there are constants C_2 and C_3 such that

$$C_2 \cdot 2^{-(m+1/2)j} \le \|\beta_{j\cdot}(h)\|_2 \le C_3 \cdot 2^{-(m+1/2)j}.$$

From the triangle inequality we have

$$\|\beta_{j\cdot}(f)\|_2 \le \|\beta_{j\cdot}(f-h)\|_2 + \|\beta_{j\cdot}(h)\|_2 \le C_1 \cdot 2^{-(m+3/4)j} + C_3 \cdot 2^{-(m+1/2)j} \le 2C_3 \cdot 2^{-(m+1/2)j}.$$

Since $\|\beta_{j\cdot}(h)\|_2 \ge C_2 \cdot 2^{-(m+1/2)j}$ and $\|\beta_{j\cdot}(f-h)\|_2 \le C_1 \cdot 2^{-(m+3/4)j}$ then for j large enough we have

$$\|\beta_{j\cdot}(f)\|_2 \ge |\ \|\beta_{j\cdot}(h)\|_2 - \|\beta_{j\cdot}(f-h)\|_2\ | = \|\beta_{j\cdot}(h)\|_2 - \|\beta_{j\cdot}(f-h)\|_2 \ge$$

$$\ge C_3 \cdot 2^{-(m+1/2)j} - C_1 \cdot 2^{-(m+3/4)j} \ge \frac{1}{2} C_3 \cdot 2^{-(m+1/2)j}$$

and the proof is complete. □

Remark 3.5. It is easy to see that the corollary above is true for a piecewise-smooth function f *i.e.* f is a finite combination of functions from Corollary 3.3 with different points a.

Let us note that in Corollary 3.3 the lower bound for $\|\beta_{j\cdot}\|_2^2$ for the piecewise constant function depends on the function

$$F_\psi(x) = \sum_{\lfloor x\rfloor - S(r)+1 \le k \le \lfloor x\rfloor} \left(\int_{x-k}^{\infty} \psi(u)\mathrm{d}u\right)^2$$

The graphs of functions F_ψ in case of Daubechies wavelets DB2-DB20 are presented below.

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

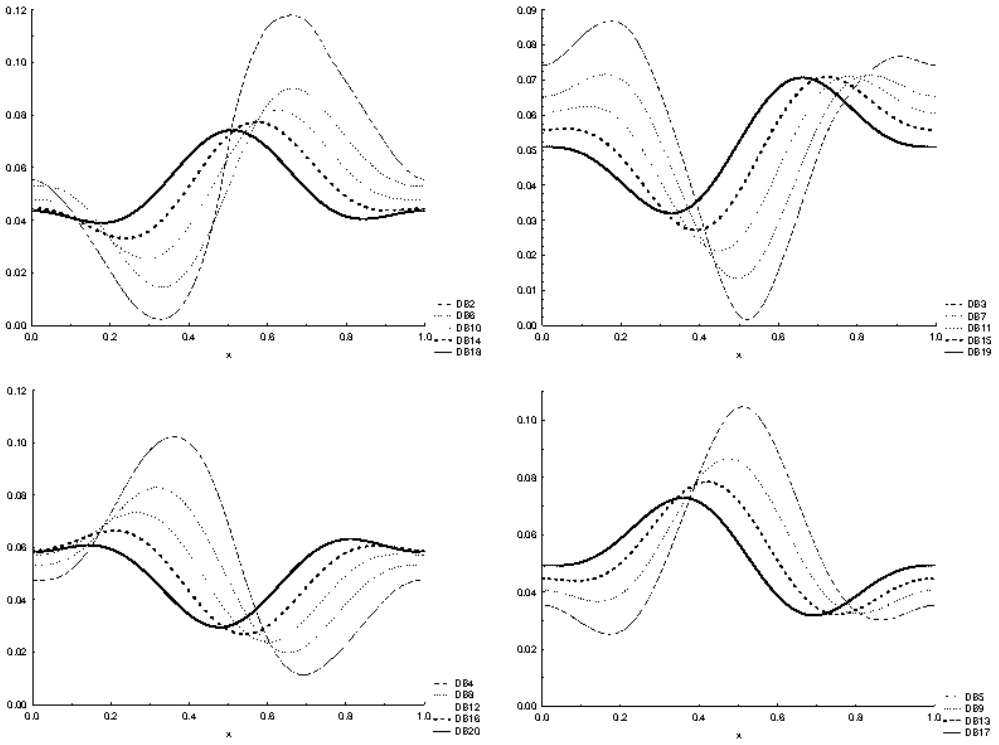


FIGURE 3. Function F_ψ for Daubechies, numerical mother wavelets DB2-DB20, divided on four groups due to the graphs similarities.

Since we assume that $r \geq 1$ for r-multiresolution analysis then we can use only DB3-DB20. The smoothness of ${}_N\phi$ is given in [8], Chapt. 7.1. One can see from the above pictures that for higher N Daubechies wavelets perform better evaluation in Corollary 3.3 (the higher $\min_{x\in[0;1]} F_\psi(x)$ the better lower bound for $\|\beta_{j\cdot}\|_2^2$).

4. ESTIMATION

We will prove the following theorem:

Theorem 4.1. *Let be given r-regular multiresolution analysis, $r \geq 1$. Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d random variables with density f, where n is the size of the experiment. Assume that for some $a, b \in \mathbb{R}$,* $\mathrm{supp} f \subset [a; b]$ *and* $||f||_\infty < \infty$. *If* $0 < s_p^* = s_p^*(f) < r$, *where* $2 \leq p \leq \infty$ *then*

$$\liminf_{j\to\infty} \frac{-\log_2 \|\hat{\beta}_{j\cdot}\|_p}{j} = s_p^* - \frac{1}{p} + \frac{1}{2}.a.e.,$$

where $\hat{\beta}_{jk} = \frac{1}{n}\sum_{i=1}^{n} \psi_{jk}(X_i)$, *and* $n \asymp 2^{2j(r-1/p+1)}$. *For simplicity let*

$$j = \lfloor \log_2 n/(2(r - 1/p + 1)) \rfloor.$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

Proof. Let $2 \leq p \leq \infty$. Let us evaluate

$$\mathbf{E}(\hat{\beta}_{jk}) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\psi_{jk}(X_i)) = \int_{\mathbb{R}} \psi_{jk}(x)f(x)\mathrm{d}x = \beta_{jk},$$

$$\begin{aligned}\mathbf{E}(\hat{\beta}_{jk})^2 &= \frac{1}{n^2}\mathbf{E}\left(\sum_{i=1}^{n}\psi_{jk}^2(X_i) + 2\sum_{m<l}\psi_{jk}(X_m)\psi_{jk}(X_l)\right)\\ &\leq \frac{1}{n}\int_{\mathbb{R}}\psi_{jk}^2(x)f(x)\mathrm{d}x + \frac{n-1}{n}\beta_{jk}^2 \leq \frac{||f||_\infty}{n} + \beta_{jk}^2.\end{aligned}$$

Let us denote

$$\mathcal{K}_j := \{k \in \mathbb{Z} : \text{supp } \psi_{jk} \cap [a;b] \neq \emptyset\}.$$

It is easy to see that $|\mathcal{K}_j| < L2^j$ for some constant L. Now, using Markov's inequality, we have

$$\mathrm{P}\left(\sum_{k\in\mathbb{Z}}(\hat{\beta}_{jk} - \beta_{jk})^2 > 2^{-2j(s-1/p+1/2)}\right) \leq \frac{\mathbf{E}\left(\sum\limits_{k\in\mathcal{K}_j}(\hat{\beta}_{jk} - \beta_{jk})^2\right)}{2^{-2j(s-1/p+1/2)}} \leq \frac{\|f\|_\infty L2^{2j(s-1/p+1)}}{n}.$$

Moreover since $p \geq 2$ then

$$\mathrm{P}\left(\|\hat{\beta}_{j\cdot} - \beta_{j\cdot}\|_p > 2^{-j(s-1/p+1/2)}\right) \leq \mathrm{P}\left(\|\hat{\beta}_{j\cdot} - \beta_{j\cdot}\|_2 > 2^{-j(s-1/p+1/2)}\right)$$

If we take $n = \lfloor 2^{2j(r-1/p+1)}\rfloor$ then, we obtain

$$\forall\ 0 < s < r \quad \sum_{j=0}^{\infty}\mathrm{P}\left(\|\hat{\beta}_{j\cdot} - \beta_{j\cdot}\|_p > 2^{-j(s-1/p+1/2)}\right) < \infty.$$

By a triangle inequality

$$\left|||\hat{\beta}_{j\cdot}||_p - ||\beta_{j\cdot}||_p\right| \leq \|\hat{\beta}_{j\cdot} - \beta_{j\cdot}\|_p,$$

so we have for all $C0$ and

$$\forall\ 0 < s < r \quad \sum_{j=0}^{\infty}\mathrm{P}\left(\left|||\hat{\beta}_{j\cdot}||_p - ||\beta_{j\cdot}||_p\right| > C2^{-j(s-1/p+1/2)}\right) < \infty. \tag{4.1}$$

Since

$$\mathrm{P}\left(\left|||\hat{\beta}_{j\cdot}||_p - ||\beta_{j\cdot}||_p\right| > 2^{-j(s-1/p+1/2)}\right) \geq \mathrm{P}\left(||\hat{\beta}_{j\cdot}||_p > 2^{-j(s-1/p+1/2)} + ||\beta_{j\cdot}||_p\right),$$

then

$$\forall\ 0 < s < r \quad \sum_{j=0}^{\infty}\mathrm{P}\left(||\hat{\beta}_{j\cdot}||_p > 2^{-j(s-1/p+1/2)} + ||\beta_{j\cdot}||_p\right) < \infty.$$

By Theorem 2.1 we know that, for each $0 < s < s_p^*$, there exists $N_s \in \mathbb{N}$, such that for $j > N_s$ $||\beta_{j\cdot}||_p \leq 2^{-j(s-1/p+1/2)}$, so

$$\forall\ 0 < s < s_p^* \quad \exists\ N \in \mathbb{N} \quad \sum_{j=N}^{\infty}\mathrm{P}\left(||\hat{\beta}_{j\cdot}||_p \geq 2\cdot 2^{-j(s-1/p+1/2)}\right) < \infty,$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

and consequently

$$\forall\ 0 < s < s_p^* \quad \exists\ N \in \mathbb{N} \quad \sum_{j=N}^{\infty} \mathrm{P}\left(\frac{-\log_2 ||\hat{\beta}_{j\cdot}||_p}{j} \leq s - \frac{1}{p} + \frac{1}{2} - \frac{1}{j} \right) < \infty.$$

Now we apply Borel–Cantelli Lemma to any given $s < s_p^*$ and we get that for almost all $\omega \in \Omega$

$$\frac{-\log_2 ||\hat{\beta}_{j\cdot}(\omega)||_p}{j} > s - \frac{1}{p} + \frac{1}{2} - \frac{1}{j}$$

for $j \geq N_s(\omega)$, where Ω is the probability space of the observations and $\hat{\beta}_{jk}(\omega) = \frac{1}{n} \sum_{i=1}^{n} \psi_{jk}(X_i(\omega))$. We conclude that for all $0 < s < s_p^*$

$$\liminf_{j\to\infty} \frac{-\log_2 ||\hat{\beta}_{j\cdot}||_2}{j} \geq s - \frac{1}{p} + \frac{1}{2} \quad a.e. \tag{4.2}$$

On the other hand, by Theorem (2.1) we know that, if $s > s_p^*$, there exists a subsequence $(j_k)_{k=1}^{\infty}$, such that $||\beta_{j_k\cdot}||_2 \geq 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})}$, so

$$\begin{aligned} \mathrm{P}\left(||\hat{\beta}_{j_k\cdot}||_p < \frac{1}{2} \cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})} \right) &= \mathrm{P}\left(||\beta_{j_k\cdot}||_p - ||\hat{\beta}_{j_k\cdot}||_p > ||\beta_{j_k\cdot}||_p - \frac{1}{2} \cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})} \right) \\ &\leq \mathrm{P}\left(\left| ||\beta_{j_k\cdot}||_p - ||\hat{\beta}_{j_k\cdot}||_p \right| > \frac{1}{2} \cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})} \right). \end{aligned}$$

Now using (4.1) we obtain

$$\forall\ s_p^* < s < r \quad \exists\ (j_k)_{k=1}^{\infty} \quad \sum_{k=1}^{\infty} \mathrm{P}\left(||\hat{\beta}_{j_k\cdot}||_2 < \frac{1}{2} \cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})} \right) < \infty, \tag{4.3}$$

so

$$\forall\ s_p^* < s < r \quad \exists\ (j_k)_{k=1}^{\infty} \quad \sum_{k=1}^{\infty} \mathrm{P}\left(\frac{-\log_2 ||\hat{\beta}_{j_k\cdot}||_2}{j_k} > s - \frac{1}{p} + \frac{1}{2} + \frac{1}{j_k} \right) < \infty.$$

From Borel–Cantelli Lemma we conclude, that for $s_p^* < s < r$

$$\liminf_{j\to\infty} \frac{-\log_2 ||\hat{\beta}_{j\cdot}||_p}{j} \leq s - \frac{1}{p} + \frac{1}{2} \quad a.e. \tag{4.4}$$

Now let us denote

$$A_s = \left\{ \omega \in \Omega : \liminf_{j\to\infty} \frac{-\log_2 ||\hat{\beta}_{j\cdot}(\omega)||_p}{j} \geq s - \frac{1}{p} + \frac{1}{2} \right\}.$$

and

$$B_s = \left\{ \omega \in \Omega : \liminf_{j\to\infty} \frac{-\log_2 ||\hat{\beta}_{j\cdot}(\omega)||_p}{j} \leq s - \frac{1}{p} + \frac{1}{2} \right\}.$$

By (4.2) and (4.4) we know that

$$\forall\ 0 < s < s_p^* \ \ P(A_s) = 1, \quad \text{and} \quad \forall\ s_p^* < s < r \ \ P(B_s) = 1.$$

Since

$$\left\{ \omega \in \Omega : \liminf_{j\to\infty} \frac{-\log_2 ||\hat{\beta}_{j\cdot}(\omega)||_p}{j} = s_p^* - \frac{1}{p} + \frac{1}{2} \right\} = \bigcap_{k=1}^{\infty} A_{s_p^* - \frac{1}{k}} \cap \bigcap_{k=1}^{\infty} B_{s_p^* + \frac{1}{k}},$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

we have

$$P\left(\liminf_{j\to\infty}\frac{-\log_2||\hat{\beta}_{j\cdot}||_p}{j}=s_p^*-\frac{1}{p}+\frac{1}{2}\right)=1-P\left(\bigcup_{k=1}^{\infty}A'_{s_p^*-\frac{1}{k}}\cup\bigcup_{k=1}^{\infty}B'_{s_p^*+\frac{1}{k}}\right)$$
$$\geq 1-\sum_{k=1}^{\infty}P\left(A'_{s_p^*-\frac{1}{k}}\right)-\sum_{k=1}^{\infty}P\left(B'_{s_p^*+\frac{1}{k}}\right)=1,$$

which completes the proof of the Theorem 4.1. □

Since $\liminf_{j\to\infty}\hat{s}_p^*=s_p^*$ a.e. where $\hat{s}_p^*=-\log_2\|\hat{\beta}_{j\cdot}\|_p/j+\frac{1}{p}-\frac{1}{2}$, we will call it "pseudo-consistent". Using the same technique one can prove:

Corollary 4.2. *Let $2\leq p\leq\infty$. Let $X_1,X_2,\dots X_n$ be a sequence of i.i.d random variables with density f, where f satisfies conditions from Corollary 3.3 and $\psi(x)$ satisfies (3.2), then*

$$\lim_{j\to\infty}\frac{-\log_2\|\hat{\beta}_{j\cdot}\|_p}{j}=n+\frac{1}{2}=s_p^*-\frac{1}{p}+\frac{1}{2}\quad a.e.,$$

where $\hat{\beta}_{jk}=\frac{1}{n}\sum_{i=1}^{n}\psi_{jk}(X_i)$, and $n\asymp 2^{2j(r-1/p+1)}$.

Proof. We repeat the whole proof of the Theorem 4.1 with only one difference: condition (4.3) is now true for almost all $j\in\mathbb{N}$ not only for some subsequence. Hence using Borel–Cantelli Lemma we obtain

$$\forall\ s^*<s<r\quad\limsup_{j\to\infty}\frac{-\log_2^+||\hat{\beta}_{j\cdot}||_2}{j}\leq s\quad a.e.$$

Combinig this with condition (4.2) we finish the proof of Corollary 4.2. □

Remark 4.3. The Corollary above, says that $\hat{s}_p^*=-\log_2\|\hat{\beta}_{j\cdot}\|_p/j+\frac{1}{p}-\frac{1}{2}$ is a strongly consistent estimator of the smoothness parameter if f is a piecewise-smooth function.

5. Numerical Experiment

In this section, we will check the behaviour of the estimator in a numerical experiment, when $p=2$. For an estimation the Daubechies wavelets "DB12" with the support $[0;S(r)]=[0;23]$ were used. Two values of the experiment size: $n=2^{16}$ and $n=2^{24}$ were used. For the calculation of the wavelets values, a dyadic discretization was used. For obtaining wavelets values between the discretization points, a linear interpolation was used. The distance between discretization points on the resolution level j was $2^{-(j+10)}$.

Data samples was generated from the following density functions:

- Step function:

$$f(x)=0.5\cdot\mathbb{1}_{[0;1/8]}(x)+1.5\cdot\mathbb{1}_{(1/8;2/8]}(x)+1\cdot\mathbb{1}_{(2/8;3,8]}(x)+1.5\cdot\mathbb{1}_{(3/8;4/8]}(x)$$
$$+0.5\cdot\mathbb{1}_{(4/8;5,8]}(x)+1.5\cdot\mathbb{1}_{(5/8;6/8]}(x)+1\cdot\mathbb{1}_{(6/8;7/8]}(x)+0.5\cdot\mathbb{1}_{(7/8;1]}(x)$$

- Polyline function:

$$f(x)=4x\cdot\mathbb{1}_{[0;0.5]}(x)-4(x-1)\cdot\mathbb{1}_{(0.5;1]}(x)$$

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

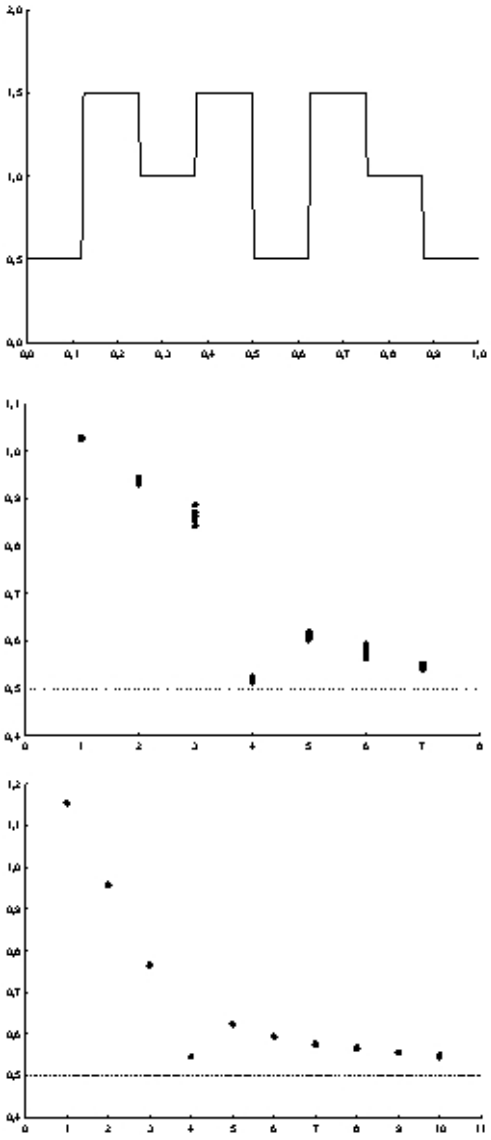


FIGURE 4. Estimation results for the "Step function". From top: Density function, The estimators of the smoothness parameter for 6 generations of the data with $n = 2^{16}$, The estimators of the smoothness parameter for 6 generations of the data with $n = 2^{24}$. On the horizontal axis (for the estimation result pictures) we have the resolution levels.

The real smoothness parameter for the "Step function" is $s_1^* = 0.5$ and for the "Polyline function" is $s_2^* = 1.5$. The estimators were constructed for 6 artificial data samples generated for each function and they are presented in Figures 4 and 5.

The sequence length should by as long as possible, but if it is too long, then the variances of the $\hat{\beta}_{jk}$ is too high, and the $||\hat{\beta}_{j\cdot}||_2$ is to high, consequently the values of the sequence are false low (can even be negative). Concerning that the number of the resolution levels (sequence length) have been selected subjectively to compromise those

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

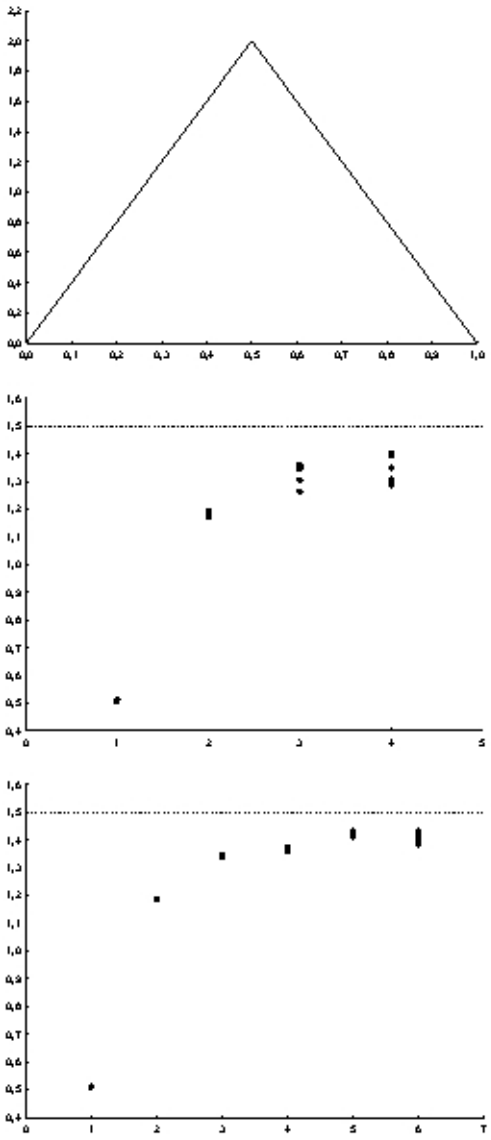


FIGURE 5. Similar to Figure 4 but for the "Polyline function".

requirements. In practice we can always take $j = \lfloor \log_2 n/(2r+1) \rfloor$ but in our experiment we have considered a little bit larger resolution levels to check the estimator behaviour (in the equation $j = \lfloor \log_2 n/(2r+1) \rfloor$ instead of r one can take any number higher than the real smoothness of function f).

Acknowledgements. The first author is supported by Polish Ministry of Science Grant N-N201-609440. The second author is supported in part by the grant The exploitation of white cabbage for phytoremediation and biofumigation of soils (AGROBIOKAP) co-financed by the European Union from European Regional Development Fund in a framework of the Innovative Economy Operational Programme 20072013.

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press

References

[1] E. Belitser and F. Enikeeva, Empirical Bayesian Test of the Smoothness. *Math. Methods Stat.* **17** (2008) 1–18.

[2] A.D. Bull, A Smirnov-Bickel-Rosenblatt theorem for compactly-supported wavelets. *Constructive Approximation* **37** (2013) 295–309.

[3] A.D. Bull, Honest adaptive confidence bands and self-similar functions. *Electron. J. Stat.* **6** (2012) 1490–1516.

[4] T. Cai, Adaptive Wavelet Estimation: A Block Thresholding and Oracle Inequality Approach. *Ann. Stat.* **27** (1999) 898–924.

[5] T. Cai and M.G. Low, An adaptation theory for nonparametric confidence intervals. *Ann. Stat.* 32 **5** (2004) 1805–1840.

[6] T. Cai and M.G. Low, Adaptive confidence balls. *Ann. Stat.* **34** (2006) 202–228.

[7] E. Chicken and T. Cai, Block thresholding for density estimation: local and global adaptivity. *J. Multivariate Anal.* **95** (2005) 76–106.

[8] I. Daubechies, Ten lectures on wavelets. *SIAM Philadelphia* (1992).

[9] D.L. Donoho and I.M. Johnstone, Minimax estimation *via* wavelet shrinkage. *Ann. Stat.* **26** (1996) 879–921.

[10] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian and D. Picard, Density estimation by wavelet thresholding. *Ann. Stat.* **24** (1996) 508–539.

[11] K. Dziedziul, M. Kucharska and B. Wolnik, Estimation of the smoothness parameter. *J. Nonparametric Stat.* **23** (2011) 991–1001.

[12] E. Giné and R. Nickl, Confidence bands in density estimation. *Ann. Stat.* **38** (2010) 1122–1170.

[13] A. Gloter and M. Hoffmann, Nonparametric reconstruction of a multifractal function from noisy data. *Probab. Theory Relat. Fields* **146** (2010) 155187.

[14] P. Hall and M.C. Jones, Adaptive M-Estimation in Nonparametric Regression. *Ann. Stat.* **18** (1990) 1712–1728.

[15] W. Härdle, G. Kerkyacharian, D. Picard and A.B. Tsybakov, Wavelets, *Approximation and Statistical Applications*. Springer-Verlag, New York (1998).

[16] M. Hoffmann and R. Nickl, On adaptive inference and confidence bands. *Ann. Stat.* **39** (2011) 2383–2409.

[17] L. Horvath and P. Kokoszka, Change-point detection with non parametric regression. *Statistics: A J. Theoret. Appl. Stat.* **36** (2002) 9–31.

[18] Y. Ingster and N. Stepanova, Estimation and detection of functions from anisotropic Sobolev classes. *Electron. J. Stat.* **5** (2011) 484–506.

[19] S. Jaffard, Conjecture de Frisch et Parisi et généricité des fonctions multifractales. *C. R. Acad. Sci. Paris Sér. I Math.* 330 **4** (2000) 265–270.

[20] M.G. Low, On nonparametric confidence intervals. *Ann. Stat.* **25** (1997) 2547–2554.

[21] Y. Meyer, Wavelets and operators. In *Cambridge Stud. Advanc. Math.* of vol. 37. Translated from the 1990 French original by D.H. Salinger. Cambridge University Press, Cambridge. (1992).

[22] S. Ropela, Spline bases in Besov spaces. *Bull. Acad. Pol. Sci. Serie Math. astr. Phys.* **24** (1976) 319–325.

[23] S.J. Sheather and M.C. Jones, A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *J. Royal Stat. Soc. Ser. B.* **53** (1991) 683–690.

https://doi.org/10.1051/ps/2013030 Published online by Cambridge University Press