WAVELET ESTIMATION FOR DERIVATIVE OF A DENSITY IN THE PRESENCE OF ADDITIVE NOISE

B.L.S. PRAKASA RAO
CR Rao Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad, India

Abstract: We construct a wavelet estimator for the derivative of a probability density function in the presence of an additive noise and study its L_p-consistency property.

Keywords : Additive noise; Derivative of a probability density function; Estimation; Mean integrated squared error; Nonparametric inference; Wavelets.

Mathematics Subject Classification: 62G07, 62G20.

1 Introduction

Methods of nonparametric estimation of a density function and regression function are widely discussed in the literature (cf. Prakasa Rao (1983, 1999a)). It is known that the estimation of derivatives of a density are also of importance and interest to detect possible bumps and to detect monotonicity, concavity or convexity properties of the density function. Asymptotic properties of the kernel type estimators for the derivatives of density have been investigated earlier (cf. Prakasa Rao (1983)).

Our aim in this paper is to discuss wavelet linear estimators for the derivative of a probability density function in the presence of an additive noise. Estimators of density using wavelets was studied for independent and identically distributed random variables in Antoniadis et al. (1994), for some stationary dependent random variables in Leblanc (1996) and for stationary associated sequences in Prakasa Rao (2003). Chaubey et al. (2006, 2008) extended these results to derivatives of density estimators for associated sequences and for negatively associated processes. The advantages and disadvantages of the use of wavelet based probability density estimators are discussed in Walter and Ghorai (1992) in the case of independent and identically distributed observations. However it was shown in Prakasa Rao (1996, 1999b) that one can obtain precise limits on the asymptotic mean integrated squared error for a wavelet based linear estimator for the density function and its derivatives as well as some other functionals of the density (cf. Theorem 3.1, Prakasa Rao

(1996)). By "precise limit", we mean that the mean integrated squared error, after suitable scaling, converges to a finite limit as the sample size tends to infinity and this limit can be computed explicitly. Tribouley (1995) studied estimation of multivariate densities using wavelet methods. Prakasa Rao (2000) investigated nonparametric estimation of the partial derivatives of a multivariate probability density. Donoho et al. (1996) investigated density estimation by wavelet thresholding. For a discussion on statistical modeling by wavelets, see Vidakovic (1999).

In recent papers, Chesneau and Doosti (2012) studied wavelet estimation of density for a GARCH model under various dependence structures and Chesneau (2013) investigated wavelet estimation of a density in a GARCH-type model leading to upper bounds on the mean integrated squared error. Shirazi et al. (2012) obtained wavelet based estimation of the derivative of a density by blockthresholding under random censorship. We studied estimation of the derivative of a density in GARCH-type model, which can be considered as a generalization of multiplicative censoring model, in Prakasa Rao (2017). Vardi (1989) (cf. Vardi and Zhang (1992)) introduced the multiplicative censoring model which unifies several models including nonparametric inference for renewal processes, non-parametric deconvolution problems and estimation of decreasing density functions. Chaubey et al. (2014) studied adaptive wavelet estimation of a density from mixtures under multiplicative censoring model generalizing the results in Prakasa Rao (2010). Asgharian et al. (2012) investigated asymptotic properties of the kernel density estimators under multiplicative censoring model. Andersen and Hansen (2001) studied density estimation for multiplicative censoring model using a series expansion approach. Chaubey et al. (2011) give a survey of recent results on linear wavelet density estimation.

Estimation of a probability density function, in the presence of an additive noise, via wavelets has been recently investgated in Li and Liu (2014), Geng and Wang (2015) and Hosseinioun (2016). Density estimation for a statistical model with additive noise plays an important role in statistics and econometrics (cf. Li and Racine (2007)). For earlier work on this problem, see Fan and Koo (2002) and Lounici and Nicki (2011). In practical situations, it is not possible to observe data directly . Suppose we have observed data consisting of independent and identically distributed observations $Y_1, \ldots, Y_n$ based on the model

$$Y = X + \epsilon$$

where X is a real valued random variable with *unknown* probability density function f_X and

ϵ is an independent random noise with a *known* probabilty density function g. The problem of estimation of the density f_X based on the observed data $Y_1, \ldots Y_n$ has been investigated by the authors cited earlier among others. Our aim is to investgate the problem of estimation of the derivatives of the density f_X, whenever they exist, based on the observed data $Y_1, \ldots Y_n$. As we mentioned earlier, this problem is also of importance and interest to detect possible bumps of the unknown density function f_X and to detect monotonicity, concavity or convexity properties of the density function f_X. Let f_Y denote the probability density function of the random variable Y. Note that f_Y is the convolution of the probability density functions f_X and g, i.e., $f_Y = f_X * g$ in the standard notation for convolution.

2 Preliminaries on wavelets

A wavelet system is an infinite collection of translated and scaled versions of functions $\phi(.)$ and $\psi(.)$ called the *scaling function* and the *primary wavelet function* respectively. In the following discussion, we assume that $\phi(.)$ is real-valued. The function $\phi(x)$ is a solution of the equation

(2. 1) $$\phi(x) = \sum_{k=-\infty}^{\infty} C_k \phi(2x - k)$$

with

(2. 2) $$\int_{-\infty}^{\infty} \phi(x)dx = 1$$

and the function $\psi(x)$ is defined by

(2. 3) $$\psi(x) = \sum_{k=-\infty}^{\infty} (-1)^k C_{-k+1} \phi(2x - k).$$

The choice of the sequence $\{C_k\}$ determines the wavelet system. It is easy to see that

(2. 4) $$\sum_{k=-\infty}^{\infty} C_k = 2.$$

Define

(2. 5) $$\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k), -\infty < j, k < \infty$$

and

(2. 6) $$\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k), -\infty < j, k < \infty.$$

Suppose the coefficients $\{C_k\}$ satisfy the condition

(2. 7) $$\begin{aligned} \sum_{k=-\infty}^{\infty} C_k C_{k+2\ell} &= 2 \text{ if } \ell = 0 \\ &= 0 \text{ if } \ell \neq 0. \end{aligned}$$

It is known that, under some additional conditions on $\phi(.)$, the collection $\{\psi_{j,k}, -\infty < j, k < \infty\}$ is an orthonormal basis for $L^2(R)$, and $\{\phi_{j,k}, -\infty < k < \infty\}$ is an orthonormal system in $L^2(R)$, for each $-\infty < j < \infty$ (cf. Daubechies (1988, 1992)).

Definition 2.1: The scaling function ϕ is said to be *r-regular* for an integer $r \geq 1$, if for every nonnegative integer $\ell \leq r$, and for any integer $k \geq 1$,

(2. 8) $$|\phi^{(\ell)}(x)| \leq c_k(1+|x|)^{-k}, -\infty < x < \infty$$

for some $c_k \geq 0$ depending only on k. Here $\phi^{(\ell)}(.)$ denotes the ℓ-th derivative of $\phi(.)$.

Definition 2.2: A *multiresolution analysis* of $L^2(R)$ consists of an increasing sequence of closed subspaces $\{V_j\}$ of $L^2(R)$ such that
(i) $\cap_{j=-\infty}^{\infty} V_j = \{0\}$;
(ii)$\bar{\cup}_{j=-\infty}^{\infty} V_j = L^2(R)$;
(iii) there is a scaling function $\phi \in V_0$ such that $\{\phi(x-k), -\infty < k < \infty\}$ is an orthonormal basis for V_0;
(iv) for all $h(.) \in L^2(R), -\infty < k < \infty, h(x) \in V_0 \Rightarrow h(x-k) \in V_0$; and
(v) $h(.) \in V_j \Rightarrow h(2x) \in V_{j+1}$.

Mallat (1989) has shown that, given any multiresolution analysis, it is possible to find a function $\psi(.)$ (called primary wavelet function) such that , for any fixed $j, -\infty < j < \infty$, the family $\{\psi_{j,k}, -\infty < k < \infty\}$ is an orthonormal basis of the orthogonal complement W_j of V_j in V_{j+1} so that $\{\psi_{j,k}, -\infty < j, k < \infty\}$ is an orthonormal basis of $L^2(R)$ (cf. Daubechies (1988, 1992)). When the scaling function $\phi(.)$ is r-regular, the corresponding multiresolution analysis is said to be r-regular.

Let $f \in L_2(R)$. The function f can be expanded in the form (cf. Daubechies (1992)):

(2. 9) $$f = \sum_{k=-\infty}^{\infty} a_{s,k}\phi_{s,k} + \sum_{j=s}^{\infty}\sum_{k=-\infty}^{\infty} b_{j,k}\psi_{j,k}$$

for any integer $-\infty < s < \infty$. Observe that the wavelet coefficients are given by

(2. 10) $$a_{s,k} = \int_{-\infty}^{\infty} f(x)\phi_{s,k}(x)dx$$

and

(2. 11) $$b_{j,k} = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx.$$

Suppose that the functions ϕ and ψ belong to C^r, the space of functions with r continuous derivatives for some $r \geq 1$, and have compact support contained in an interval $[-\delta, \delta]$ for some $\delta > 0$. It follows, from the Corollary 5.5.2 in Daubechies (1992), that the function $\psi(.)$ is orthogonal to polynomials of degree less than or equal to r. In particular

$$\int_{-\infty}^{\infty} \psi(x) x^{\ell} dx = 0, \ell = 0, 1, \ldots, r.$$

This brief discussion on wavelets is based on Antoniades et al. (1994). For a more details, see Daubechies (1992) and Strang (1989).

3 More on wavelets

Let $\phi(.)$ be a scaling function as defined earlier. Define

$$\theta_{\phi}(x) = \sum_{k=-\infty}^{\infty} |\phi(x-k)|.$$

Suppose the following conditions hold:

($\mathcal{C}1$) The $ess\sup_x \theta_{\phi}(x) < \infty$ where

$$ess\sup_x g(x) = \inf\{y : \lambda([x : g(x) > y]) = 0\}$$

and λ is the Lebesgue measure on the real line.
($\mathcal{C}2$) There exists a bounded nonincreasing function $\Phi(.)$ such that $|\phi(u)| \leq \Phi(|u|)$ almost every where and

$$\int_0^{\infty} |u|^r \Phi(|u|) du < \infty.$$

for some integer $r \geq 0$.

The following Lemmas 3.1 to 3.3 follow from the results in Hardle et al. (1998).
Lemma 3.1: Suppose the condition ($\mathcal{C}1$) holds. Then, for any sequence $\{\lambda_s, s \in \mathcal{Z}\} \in \ell_p$,

$$C_1 ||\lambda||_{\ell_p} 2^{\frac{s}{2}-\frac{s}{p}} \leq ||\sum_k \lambda_k \phi_{s,k}||_p \leq C_2 ||\lambda||_{\ell_p} 2^{\frac{s}{2}-\frac{s}{p}}$$

where

$$C_1 = (||\theta_{\phi}||_{\infty}^{1/p} ||\phi||_1^{1/q})^{-1}$$

and

$$C_2 = (||\theta_{\phi}||_{\infty}^{1/q} ||\phi||_1^{1/p})^{-1}$$

where $1 \leq p \leq \infty, \frac{1}{p} + \frac{1}{q} = 1$ with suitable interpretation for p and q in the boundary case.

Since the scaling function ϕ satisfies the condition $(\mathcal{C}1)$, the kernel function

$$K(x, y) = \sum_k \phi(x - k)\phi(y - k)$$

is well defined and it is called the orthonormal projection associated with the function ϕ. Let

$$K_s(x, y) = 2^s K(2^s x, 2^s y)$$

for any integer $s \geq 0$. For any function $h \in L_p(R), 1 \leq p \leq \infty$, define

$$K_s h(x) = \int_{-\infty}^{\infty} K_s(x, y)h(y)dy = \sum_s \alpha_{s,k}\phi_{s,k}(x) \tag{3. 1}$$

where

$$\alpha_{s,k} = \int_{-\infty}^{\infty} \phi_{s,k}(x)h(x)dx.$$

Lemma 3.2: Suppose the condition $(\mathcal{C}2)$ holds. Then

$$(i) \int_{-\infty}^{\infty} K(x, y)dy = 1 \text{ a.e.}$$

and

$$(ii)|K(x, y)| \leq C_1\Phi(\frac{|x - y|}{C_2}) \text{ a.e}$$

where C_1 and C_2 are positive constants depending on Φ.

Let $F(x) = C_1\Phi(\frac{|x|}{C_2})$. Then the function $F \in L_1(R) \cap L_\infty(R)$ and $|K(x, y)| \leq F(x-y)$ a.e

Lemma 3.3: Suppose the condition $(\mathcal{C}2)$ holds and $h \in L_p(R), 1 \leq p < \infty$. Then

$$\lim_{n\to\infty} ||K_s h - h||_p = 0.$$

Suppose the function $h^{(d)}$ exists and $h^{(d)} \in L_p(R)$ for some $1 \leq p < \infty$. As a consequence of Lemma 3.3, it follows that

$$\lim_{n\to\infty} ||K_s h^{(d)} - h^{(d)}||_p = 0. \tag{3. 2}$$

It can be shown that Lemma 3.3 holds for $h \in L_\infty(R)$ if the function $h(.)$ is uniformly continuous. We will now state another result known as Rosenthal's inequality (Rosenthal (1970)) which will be used in the sequel.

Lemma 3.4: Let $X_1, \ldots, X_n$ be independent random variables with mean zero and further suppose that $|X_i| \leq M < \infty, 1 \leq i \leq n$. Then there exists a constant $C_p > 0$, such that

$$(i) E(|\sum_{i=1}^{n} X_i|^p) \leq C_p(M^{p-2}\sum_{i=1}^{n} E(X_i^2) + (\sum_{i=1}^{n} E(X_i^2))^{p/2}), p > 2,$$

and

$$(ii) E(|\sum_{i=1}^{n} X_i|^p) \leq C_p(\sum_{i=1}^{n} E(X_i^2))^{p/2}, 0 < p \leq 2.$$

4 Estimation of the d-th derivative of a probability density function

For any function $h(.) \in L_1(R)$, define the Fourier transform

$$\tilde{h}(t) = \int_{-\infty}^{\infty} h(x)e^{-itx}dx, -\infty < t < \infty.$$

It is known that $\tilde{f}_Y(t) = \tilde{f}_X(t)\tilde{g}(t), t \in R$. Suppose that the Fourier transform $\tilde{g}(t)$ of the probability density function g is non-vanishing for all $t \in R$.

Let $\{X_i, 1 \leq i \leq n\}$ be independent and identically distributed random variables with probability density function f_X which is d-times differentiable. Suppose that the derivative $f_X^{(d)}$ of f_X exists, bounded, has compact support and $f_X^{(d)} \in L_2(R)$. Let us first consider the estimation of the probability density function f_X. A wavelet based density estimator of the density function f_X can be motivated in the following way from the expansion given in the equation (2.9) (cf. Prakasa Rao (2003)). We can estimate $f_X(x)$ by $\hat{f}_X(x)$ where

(4. 1)
$$\hat{f}_X(x) = \sum_{k \in N_s} \alpha_{s,k}\phi_{s,k}(x)$$

where

(4. 2)
$$\alpha_{s,k} = \frac{1}{n}\sum_{i=1}^{n} \phi_{s,k}(X_i).$$

Here N_s is the set of integers k such that $supp(f_X) \cap supp(\phi_{s,k})$ is nonempty. Since the functions f_X and ϕ have compact supports, the cardinality of the set N_s is finite and it is of the order $O(2^s)$.

Let us now consider the problem of estimation of the derivative $f_X^{(d)}$ of f_X. As in Prakasa Rao (1996), we assume that $f_X^d \in L_2(R)$ and that there exist $D_j \geq 0, \beta_j \geq 0$, such that

$$|f_X^{(j)}(x)| \leq D_j|x|^{-\beta_j}, |x| \geq 1, 0 \leq j \leq d$$

where $\beta_0 > 4d+1$. Suppose the multiresolution analysis generated by the scaling function ϕ is r-regular for some $r \geq d$. Then, by definition, $\phi \in C^{(r)}, \phi$ and its derivatives $\phi^{(j)}$ up to order r are rapidly decreasing, i.e., for every integer $m \geq 1$, there exists a constant $A_m > 0$, such that

$$|\phi^{(j)}(x)| \leq A_m(1+|x|)^{-m}, \ \ 0 \leq j \leq r.$$

If $d \geq 1$, then it is clear that

$$\lim_{|x|\to\infty} \phi_{s,k}^{(j)}(x) f^{(d-j-1)}(x) = 0, 0 \leq j \leq d-1$$

for any fixed s and k. The projection of $f_X^{(d)}$ on V_s is

$$f_{X,s}^{(d)}(x) = \sum_{k\in N_s} a_{s,k}\phi_{s,k}(x) \quad (4.\ 3)$$

where

$$\begin{aligned} a_{s,k} &= \int_{-\infty}^{\infty} f_X^{(d)}(x)\phi_{s,k}(x)dx \\ &= (-1)^d \int_{-\infty}^{\infty} f_X(x)\phi_{s,k}^{(d)}(x)dx. \end{aligned} \quad (4.\ 4)$$

The last equality given above can be justified by using integration by parts since the function $\phi(.)$ is r-regular (cf. Prakasa Rao (1996)). This expression motivates the following estimator for $f_X^{(d)}(x)$:

$$\tilde{f}_{X,s}^{(d)}(x) = \sum_{k\in N_s} a_{s,k}^*\phi_{s,k}(x) \quad (4.\ 5)$$

where

$$a_{s,k}^* = \frac{(-1)^d}{n}\sum_{i=1}^{n} \phi_{s,k}^{(d)}(X_i).$$

Note that the estimator defined above in the equation (4.5) reduces to the density estimator given in (4.1) for $d = 0$. Since the random sample $X_i, 1 \leq i \leq n$ is unobservable and the observed data is $Y_i = X_i + \epsilon_i, 1 \leq i \leq n$, we now modify the estimator $\tilde{f}_{X,s}^{(d)}(x)$.

By Plancherel formula,

$$\begin{aligned} a_{s,k} &= (-1)^d \int_{-\infty}^{\infty} f_X(x)\phi_{s,k}^{(d)}(x)dx \\ &= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty} \tilde{f}_X(t)\overline{\tilde{\phi}_{s,k}^{(d)}(t)}dt \\ &= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty} \frac{\tilde{f}_Y(t)}{\tilde{g}(t)}\tilde{\phi}_{s,k}^{(d)}(-t)dt. \end{aligned}$$

For any function $\psi(.)$ which is d-times differentiable, define $\psi_{s,k}(x) = 2^{s/2}\psi(2^s x - k)$ for integers s, k and let $\psi^{(d)}_{s,k}(x)$ denote the d-th derivative of the function $\psi_{s,k}(x)$. Define the operator H_s by the transformation

$$(H_s\psi^{(d)})_{s,k}(y) = \frac{1}{2\pi}\int_R e^{ity}\frac{\tilde{\psi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)}dt, y \in R$$

for all integers $-\infty < s, k < \infty$. It can be checked that

$$\tilde{\psi}^{(d)}_{s,k}(u) = e^{iku2^s}2^{-ds+\frac{s}{2}}\tilde{\psi}^{(d)}_{s,k}(u2^s)$$

which we will use in the computations later. Let

$$\hat{a}_{s,k} = \frac{(-1)^d}{n}\sum_{i=1}^{n}(H_s\phi^{(d)})_{s,k}(Y_i). \tag{4. 6}$$

We will consider the modified estimator $\hat{f}^{(d)}_{X,s}(x)$ defined by

$$\begin{aligned} \hat{f}^{(d)}_{X,s}(x) &= \sum_{k\in N_s}\hat{a}_{s,k}\phi_{s,k}(x) \\ &= \sum_{k\in N_s}[\frac{(-1)^d}{n}\sum_{i=1}^{n}(H_s\phi^{(d)})_{s,k}(Y_i)]\phi_{s,k}(x) \\ &= \frac{(-1)^d}{n}\sum_{i=1}^{n}[\sum_{k\in N_s}(H_s\phi^{(d)})_{s,k}(Y_i)\phi_{s,k}(x)]. \end{aligned} \tag{4. 7}$$

as an estimator of $f^{(d)}_X(x)$.

Lemma 4.1: If the function $f^{(d)}_X \in L_2(R)$, then the estimator $\hat{a}_{s,k}$ defined by the equation (4.6) is an unbiased estimator of the wavelet coefficient $a_{s,k}$ given by the equation (4.4).

Proof : Note that

$$\begin{aligned} E[\hat{a}_{s,k}] &= E[\frac{(-1)^d}{n}\sum_{i=1}^{n}(H_s\phi^{(d)})_{s,k}(Y_i)] \\ &= (-1)^d E[(H_s\phi^{(d)})_{s,k}(Y_1)] \\ &= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty}[\int_{-\infty}^{\infty}e^{ity}\frac{\tilde{\phi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)}dt]f_Y(y)dy \\ &= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty}[\int_{-\infty}^{\infty}e^{ity}f_Y(y)dy]\frac{\tilde{\phi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)}dt \end{aligned}$$

$$
\begin{aligned}
&= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty}\tilde{f}_Y(-t)\frac{\tilde{\phi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)}dt \\
&= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty}\tilde{f}_X(-t)\tilde{\phi}^{(d)}_{s,k}(t)dt \\
&= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty}\tilde{f}_X(t)\tilde{\phi}^{(d)}_{s,k}(-t)dt \\
&= \frac{(-1)^d}{2\pi}\int_{-\infty}^{\infty}\tilde{f}_X(t)\overline{\tilde{\phi}^{(d)}_{s,k}(t)}dt \\
&= (-1)^d\int_{-\infty}^{\infty}f_X(x)\phi^{(d)}_{s,k}(x)dx \\
&= \int_{-\infty}^{\infty}f^{(d)}_X(x)\phi_{s,k}(x)dx \\
&= a_{s,k}.\diamondsuit
\end{aligned}
$$

We will now discuss L_p-consistency of the estimator $\hat{f}^{(d)}_{X,s}(x)$ for estimating of the function $f^{(d)}_X(x)$ following the techniques in Geng and Wang (2015). For any function $f \in L_p(R)$, we write $||f||^p_p$ for $\int_R |f(x)|^p dx$.

Theorem 4.1: Suppose that $\tilde{g}(t) \simeq (1+|t^2|)^{-\beta/2}, t \in R$ for some $\beta \geq 0$ and the function $f^{(d)}_X \in L_p(R)$ for some $2 \leq p < \infty$. Further suppose that $f_Y \in L_{p/2}(R)$. Choose the positive integer s such that $2^s \simeq n^{\frac{1-\epsilon}{1+2\beta+4d\frac{2p-1}{p}}}$ for some $0 < \epsilon < 1$. Define the estimator $\hat{f}^{(d)}_{X,s}(x)$ as an estimator of the function $f^{(d)}_X(x)$. Then the estimator $\hat{f}^{(d)}_{X,s}(x)$ is L_p-consistent, i.e.,

$$\lim_{n\to\infty} E||\hat{f}^{(d)}_{X,s} - f^{(d)}_X||_p = 0.$$

Proof : Note that

$$
\begin{aligned}
E[\hat{f}^{(d)}_{X,s}(x)] &= E[\sum_{k\in N_s}\hat{a}_{s,k}\phi_{s,k}(x)] \\
&= E[\sum_{k\in N_s}[\frac{(-1)^d}{n}\sum_{i=1}^{n}(H_s\phi^{(d)})_{s,k}(Y_i)]\phi_{s,k}(x)] \\
&= E[(-1)^d\sum_{k\in N_s}(H_s\phi^{(d)})_{s,k}(Y_1)]\phi_{s,k}(x)] \\
&= (-1)^d\sum_{k\in N_s}E[(H_s\phi^{(d)})_{s,k}(Y_1)]\phi_{s,k}(x) \\
&= \sum_{k\in N_s}a_{s,k}\phi_{s,k}(x)
\end{aligned}
$$

$$= \quad K_s f_X^{(d)}(x)$$

where the operator K_s is as defined by the equation (3.1).

As a consequence of the equation (3.2) following Lemma 3.3 (cf. Hardle et al. (1998)), it follows that

$$\lim_{n\to\infty} ||f_X^{(d)} - E[\hat{f}_{X,s}^{(d)}]||_p = \lim_{n\to\infty} ||f_X^{(d)} - K_s f_X^{(d)}||_p = 0. \tag{4. 8}$$

In the following discussion, we will denote $A_s \simeq B_s$ if there exist positive constants C_1 and C_2 such that

$$C_1 B_s \leq A_s \leq C_2 B_s$$

as $s \to \infty$. We will now estimate the term

$$||\hat{f}_{X,s}^{(d)} - E[\hat{f}_{X,s}^{(d)}]||_p.$$

Note that

$$\begin{aligned} ||\hat{f}_{X,s}^{(d)} - E[\hat{f}_{X,s}^{(d)}]||_p^p &= ||\sum_{k\in N_s} \hat{a}_{s,k}\phi_{s,k}(x) - \sum_{k\in N_s} a_{s,k}\phi_{s,k}(x)]||_p^p \\ &= ||\sum_{k\in N_s}(\hat{a}_{s,k} - a_{s,k})\phi_{s,k}(x)||_p^p \\ &\simeq 2^{s(\frac{p}{2}-1)}[\sum_{k\in N_s} |\hat{a}_{s,k} - a_{s,k}|^p] \ \text{ (by Lemma 3.1)} \end{aligned}$$

and hence

$$\begin{aligned} E[||\hat{f}_{X,s}^{(d)} - E[\hat{f}_{X,s}^{(d)}]||_p^p] &\simeq 2^{s(\frac{p}{2}-1)}E[\sum_{k\in N_s} |\hat{a}_{s,k} - a_{s,k}|^p] \\ &= 2^{s(\frac{p}{2}-1)}[\sum_{k\in N_s} E|\hat{a}_{s,k} - a_{s,k}|^p]. \end{aligned}$$

Observe that

$$\begin{aligned} |\hat{a}_{s,k} - a_{s,k}| &= |\frac{1}{n}\sum_{i=1}^n (H_s\phi^{(d)})_{s,k}(Y_i) - \frac{1}{n}\sum_{i=1}^n E[(H_s\phi^{(d)})_{s,k}(Y_i)]| \\ &= \frac{1}{n}|\sum_{i=1}^n Z_{ik}| \end{aligned}$$

where

$$Z_{ik} = (H_s\phi^{(d)})_{s,k}(Y_i) - E[(H_s\phi^{(d)})_{s,k}(Y_i).$$

Therefore

$$\begin{aligned}
|(H_s\phi^{(d)})_{s,k}(y)| &= |\frac{1}{2\pi}\int_R e^{ity}\frac{\tilde{\phi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)}dt| \\
&\leq \frac{1}{2\pi}\int_R |\frac{\tilde{\phi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)}|dt \\
&\simeq \frac{1}{2\pi}\int_R |\tilde{\phi}^{(d)}_{s,k}(t)|(1+|t|)^{\beta/2}dt \\
&\simeq 2^{ds-\frac{s}{2}}\int_R |\tilde{\phi}^{(d)}(u)|(1+|u|2^s)^{\beta/2}2^s du \\
&\simeq 2^{ds+\frac{s}{2}}2^{\beta s}.
\end{aligned}$$

Hence

$$\begin{aligned}
|Z_{ik}| &= |(H_s\phi^{(d)})_{s,k}(Y_i) - E[(H_s\phi^{(d)})_{s,k}(Y_i)]| \\
&\leq |(H_s\phi^{(d)})_{s,k}(Y_i)| + E|(H_s\phi^{(d)})_{s,k}(Y_i)| \\
&= |\frac{1}{2\pi}\int_{-\infty}^{\infty} 2^{ds+\frac{s}{2}}e^{it(2_i^Y-k)}\frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)}dt| \\
&\quad +\frac{1}{2\pi}\int_{-\infty}^{\infty}|\int_{-\infty}^{\infty} 2^{ds+\frac{s}{2}}e^{it(2^y-k)}\frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)}dt|f_y(y)dy \\
&\simeq 2^{s(\frac{1}{2}+\beta+d)}.
\end{aligned}$$

Applying Rosenthal's inequality (Lemma 3.4), it follows that

$$\begin{aligned}
(4.\ 9)\qquad E[|\hat{a}_{s,k} - a_{s,k}|^p] &= \frac{1}{n^p}E|\sum_{i=1}^n Z_{ik}|^p \\
&\simeq \frac{1}{n^p}[2^{s(\frac{1}{2}+\beta+d)(p-2)}\sum_{i=1}^n E|Z_{ik}|^2 + (\sum_{i=1}^n E|Z_{ik}|^2)^{p/2}] \\
&= \frac{2^{s(\frac{1}{2}+\beta+d)(p-2)}}{n^{p-1}}E|Z_{1k}|^2 + \frac{1}{n^{p/2}}(E|Z_{1k}|^2)^{p/2}.
\end{aligned}$$

We will now estimate $\sum_k (E|Z_{1k}|^2)^{p/2}$. Observe that

$$A = \int_R |(H_s\phi^{(d)})_{s,k}(y)|^2 dy$$

$$
\begin{aligned}
&= 2\pi \int_R |\frac{\tilde{\phi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)}|^2 dt \\
&\simeq 2^{4ds-s} \int_R |\frac{\tilde{\phi}^{(d)}(t2^{-s})}{\tilde{g}(-t)}|^2 dt \\
&\simeq 2^{s(4d-1)} \int_R |\frac{\tilde{\phi}^{(d)}(u)}{\tilde{g}(-u2^s)}|^2 2^s du \\
&\simeq 2^{4ds} \int_R |(1+|u^2 2^{2s}|)^{\beta/2} \tilde{\phi}^{(d)}(u)|^2 du \\
&\simeq 2^{4ds+2\beta s}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
(E|Z_{1k}|^2)^{p/2} &= (E|(H_s\phi^{(d)})_{s,k}(Y_1) - E[(H_s\phi^{(d)})_{s,k}(Y_1)]|^2)^{p/2} \\
&\leq (E|(H_s\phi^{(d)})_{s,k}(Y_1)|^2)^{p/2} \\
&= (\int_R |(H_s\phi^{(d)})_{s,k}(y)|^2 f_Y(y) dy)^{p/2} \\
&= A^{p/2} (\int_R |\frac{(H_s\phi^{(d)})_{s,k}(y)|^2}{A} f_Y(y) dy)^{p/2} \\
&\leq A^{\frac{p}{2}-1} (\int_R |(H_s\phi^{(d)})_{s,k}(y)|^2)(f_Y(y))^{p/2} dy.
\end{aligned}
$$

Furthermore

$$
\begin{aligned}
\sum_k |(H_s\phi^{(d)})_{s,k}(y)|^2 &= \sum_k |\frac{1}{2\pi} \int_R e^{ity} \frac{\tilde{\phi}^{(d)}_{s,k}(t)}{\tilde{g}(-t)} dt|^2 \\
&\simeq \sum_k (2^{ds-s/2} | \int_{-4\pi/3}^{4\pi/3} e^{it(y-k)} \frac{\tilde{\phi}^{(d)}(t2^{-s})}{\tilde{g}(-t)} dt|)^2 \\
&\simeq 2^{2ds-s} \sum_k (| \int_{-4\pi/3}^{4\pi/3} e^{it(y-k)} \frac{\tilde{\phi}^{(d)}(t2^{-s})}{\tilde{g}(-t)} dt|)^2 \\
&= 2^{2ds-s} \sum_k (| \int_{-4\pi/3}^{4\pi/3} e^{i(y-k)u2^s} \frac{\tilde{\phi}^{(d)}(u)}{\tilde{g}(-u2^s)} 2^s du|)^2 \\
&= 2^{2ds+s} \sum_k (| \int_{-4\pi/3}^{4\pi/3} e^{i(y-k)u2^s} \frac{\tilde{\phi}^{(d)}(u)}{\tilde{g}(-u2^s)} du|)^2 \\
&= 2^{2ds+s} \sum_k (| \int_0^{4\pi/3} e^{it2^s y} \frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)} e^{-it2^s k} dt| \\
&\quad + | \int_{-4\pi/3}^0 e^{it2^s y} \frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)} e^{-it2^s k} dt|)^2
\end{aligned}
$$

$$
\begin{aligned}
&\simeq \quad 2^{2ds+s}[\sum_k (|\int_0^{4\pi/3} e^{it2^s y} \frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)} e^{-it2^s k} dt|^2 \\
&\quad + \sum_k (|\int_{-4\pi/3}^{0} e^{it2^s y} \frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)} e^{-it2^s k} dt|^2]
\end{aligned}
$$

Observe that the function

$$
e^{it2^s y} \frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)} I_{[0,2\pi]} \in L_2[0, 2\pi]
$$

and the series $\{e^{-it2^s k}, k \in Z\}$ is an orthonormal basis for $L_2[0, 2\pi]$. An application of the Parseval formula shows that

$$
\sum_k (|\int_0^{4\pi/3} e^{it2^s y} \frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)} e^{-it2^s k} dt|^2 = \int_0^{4\pi/3} |e^{it2^s y} \frac{\tilde{\phi^{(d)}}(t)}{\tilde{g}(-2^s t)}|^2 dt = 2^{2s\beta}.
$$

In a similar way, we get that

$$
\sum_k (|\int_{-4\pi/3}^{0} e^{it2^s y} \frac{\tilde{\phi}^{(d)}(t)}{\tilde{g}(-2^s t)} e^{-it2^s k} dt|^2 = 2^{2s\beta}.
$$

Combining the above bounds, it follows that

$$
\sum_k |(H_s \phi^{(d)})_{s,k}(y)|^2 \leq 2^{s(2\beta+1+2d)}
$$

which in turn implies that

$$
\sum_k (E|Z_{1k}|^2)^{p/2} \leq A^{\frac{p}{2}-1} 2^{s(2\beta+1+2d)} = 2^{s((\beta p+1)+2d(p-1))}.
$$

Hence

$$
\begin{aligned}
\sum_k E|\hat{a}_{s,k} - a_{s,k}|^p &= \frac{2^{s(\frac{1}{2}+\beta+d)(p-2)}}{n^{p-1}} \sum_k E|Z_{1k}|^2 + \frac{1}{n^{p/2}} \sum_k (E|Z_{1k}|^2)^{p/2} \\
&\leq \frac{2^{s(\frac{1}{2}+\beta+d)(p-2)} 2^{s(2\beta+1+2d)}}{n^{p-1}} + \frac{2^{s(\beta p+1+2d(p-1))}}{n^{p/2}} \\
&= \frac{2^{s((\beta p+1)+2d(p-1))}}{n^{p/2}} (1 + \frac{2^{s(\frac{p}{2}-1-d(p-2))}}{n^{\frac{p}{2}-1}}).
\end{aligned}
$$

As a consequence of the bound obtained above, it follows that

$$
\begin{aligned}
E[||\hat{f}_{X,s}^{(d)} - E[\hat{f}_{X,s}^{(d)}]||_p^p] &\leq 2^{s(\frac{p}{2}-1)} \frac{2^{s(\beta p+1)+2d(p-1)}}{n^{p/2}} (1 + \frac{2^{s(\frac{p}{2}-1-d(p-2))}}{n^{\frac{p}{2}-1}}) \\
&\simeq (\frac{2^{s(2\beta+1+4d\frac{(p-1)}{p})}}{n})^{p/2}.
\end{aligned}
$$

Choosing $2^s \simeq n^{\frac{1-\epsilon}{1+2\beta+4d\frac{(p-1)}{p}}}$ for some $0 < \epsilon < 1$, we obtain that

$$\lim_{n\to\infty} E[||\hat{f}^{(d)}_{X,s} - E[\hat{f}^{(d)}_{X,s}]||_p^p] = 0. \tag{4. 10}$$

Combining the relations (4.8) and (4.10), we obtain that

$$\lim_{n\to\infty} E[||\hat{f}^{(d)}_{X,s} - f^{(d)}_{X}||_p^p] = 0 \tag{4. 11}$$

by the inequality

$$||U+V||_p \leq ||U||_p + ||V||_p$$

for $U, V \in L_p(R)$. This proves the L_p-consistency of the estimator $\hat{f}^{(d)}_{X,s}$ for estimating the derivative $f^{(d)}_X$.$\diamondsuit$

Remarks : We have discussed the problem of estimation of derivative of a density function in the presence of independent additive noise whose distribution is known using linear wavelet estimators. It is not clear how to estimate the original density or its derivative if the error density is also unknown. Another problem is to study adaptive nonlinear wavelet estimators of derivative of a density in the presence of known or unknown independent additive noise and to construct shape preserving estimators. These problems need investigation.

Acknowledgement: This work was supported under the scheme "Ramanujan Chair Professor" at the CR Rao Advanced Institute for Mathematics, Statistics and Computer Science, Hyderabad, India.

References :

Andersen, K. and Hansen, M. (2001) Multiplicative censoring: density estimation by a series expansion, *J. Statist. Plan. and Infer.*, **98**, 137-155.

Antoniadis, A., Gregoire, G., and McKeague, I.W. (1994) Wavelet methods for curve estimation, *J. Amer. Statist. Assoc.*, **89**, 1340-1353.

Asgharian, M., Carone, M. and Fakoor, V. (2012) Large-sample study of the kernel density estimation under multiplicative censoring, *Ann. Statist.*, **40**, 159-187.

Chaubey, Y.P., Chesneau, C., and Doosti, H. (2011) On linear wavelet density estimation: Some recent developments, *J. Indian. Soc. Agricult. Statist.*, **65**, 169-179.

Chaubey, Y.P, Chesneau, C. and Doosti, H. (2014) Adaptive wavelet estimation of a density from mixtures under multiplicative censoring, *Statistics: A Journal of Theoretical and Applied Statistics*, DOI:10.1080/02331888.2014.938652.

Chaubey, Y.P, Doosti, H. and Prakasa Rao, B.L.S. (2006) Wavelet based estimation of the derivatives of a density with associated variables, *Int. J. Pure and Appl. Math.*, **27**, 97-106.

Chaubey, Y.P, Doosti, H. and Prakasa Rao, B.L.S. (2008) Wavelet based estimation of the derivatives of a density for a negatively associated process, *J. Statistical Theory and Practice*, **2**, 453-463.

Chesneau, C. (2013) Wavelet estimation of a density in a GARCH-type model, *Comm. Stat. Theor. Meth.*, **42**, 98-117.

Chesneau, C. and Doosti, H. (2012) Wavelet linear density estimation for a GARCH model under various dependence structures, *Journal of Iranian Statistical Society*, **11**, 1-21.

Daubecheis, I. (1988) Orthogonal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **41**, 909-996.

Daubecheis, I. (1992) *Ten Lectures on Wavelets*, SIAM, Philadelphia.

Donoho, D., Johnstone, I., Kerkyacharian, G., and Picard, D. (1996) Density estimation by wavelet thresholding, *Ann. Statistics*, **24**, 508-539.

Fan, J and Koo, J. (2002) Wavelet deconvolution, *IEEE Trans. Inform. Theory*, **48**, 734-747.

Geng, Zijuan and Wang, Jinru (2015) The mean consistency of wavelet density estimators, *Journal of Inequalities and Applications* 2015:111. http://dx.doi.org/10.1186/s13660-015-0636-1.

Hardle, W., Kerkycharian, G., Picard, D., and Tsybakov. A (1998) *Wavelets, Approximations, and Statistical Applications*, Lecture Notes in Statistics, Vol. 129, Springer, New York.

Hosseinioun, N. (2016) Wavelet-based density estimation in presence of additive noise under various dependence structures, *Advances in Pure Mathematics*, **6**, 7-15.

Leblanc, F. (1996) Wavelet linear density estimator for a discrete-time stochastic process: L_p-losses, *Statist. Probab. Lett.*, **27**, 71-84.

Li, Rui and Liu, Youming (2014) Wavelet estimations for a density with some additive noises, *Appl. and Comput. Harmon. Anal.*, **36**, 416-433.

Li, Q. and Racine, J.S. (2007) *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, Princeton.

Lounici, K. and Nickl, R. (2011) Global uniform risk bounds for wavelet deconvolutionestimators, *Ann. Statist.*, **39**, 201-231.

Mallat, S.G. (1989) A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 674-693.

Prakasa Rao, B.L.S. (1983) *Nonparametric Functional Estimation*, Academic Press, Orlando.

Prakasa Rao, B.L.S. (1996) Nonparametric estimation of the derivatives of a density by the method of wavelets, *Bull. Inform. Cyb.*, **28**, 91-100.

Prakasa Rao, B.L.S. (1999a) Nonparametric functional estimation: An overview, In *Asymptotics, Nonparametrics and Time Series,* Ed. Subir Ghosh, Marcel Dekker Inc. New York, pp. 461-509.

Prakasa Rao, B.L.S. (1999b) Estimation of the integrated squared density derivative by wavelets, *Bull. Inform.Cyb.*, **31**, 47-65.

Prakasa Rao, B.L.S. (2000) Nonparametric estimation of partial derivatives of a multivariate probability density by the method of wavelets. In *Asymptotics in Statistics and Probability*, Festschrift for George G. Roussas, Ed. M.L. Puri, VSP, The Netherlands, pp 321-330.

Prakasa Rao, B.L.S. (2003) Wavelet linear density estimation for associated sequences, *J. Indian Statist. Assoc.*, **41**, 369-379.

Prakasa Rao, B.L.S. (2010) Wavelet linear estimation for derivatives of a density from observations of mixtures with varying mixing proportions, *Indian J. of Pure & Appl. Math.*, **41**, 275-291.

Prakasa Rao, B.L.S. (2017) Wavelet estimation for derivative of a density in a GARCH-type model, *Comm. Statist. Theory and Methods*, **46**, 2396-2410.

Rosentahl, H.P. (1970) On the subspaces of $L^p, (p > 2)$ spanned by sequences of independent random variables, *Israel J. Math.*, **8**, 273-303.

Shirazi, E., Chaubey, Y.P., Doosti, H. and Nirumand, H.A. (2012) Wavelet based estimation for the derivative of a density by block thresholding under random censorship, *J. Korean Stat. Soc.*, **41**, 199-211.

Strang, G. (1989) Wavelets and dilation equations: a brief introduction, *SIAM Review*, **31**, 614-627.

Tribouley, K. (1995) Practical estimation of multivariate densities using wavelet methods, *Statistica Neerlandica*, **49**, 41-62.

Vardi, Y. (1989) Multiplicative censoring, renewal processes, deconvolution and decreasing density: Nonparmetric estimation, *Biometrika*, **76**, 751-761.

Vardi, Y. and Zhang, C.H. (1992) Large sample study of empirical distributions in a random multiplicative censoring model, *Ann. Statist.*, **20**, 1022-1039.

Vidakovic, B. (1999) *Statistical Modeling by Wavelets*, Wiley, New York.

Walter, G and Ghorai, J. (1992) Advantages and disadvantages of density estimation with wavelets, In *Proceedings of the 24th Symp. on the Interface*, Ed. H. Joseph Newton, Interface FNA,VA 24: 234-243.