Journal of Sciences, Islamic Republic of Iran 16(3): 255-259 (2005) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Wavelets for Nonparametric Stochastic Regression with Pairwise Negative Quadrant Dependent Random Variables

H. Doosti* and H.A. Niroumand

Department of Statistics, School of Mathematical Sciences, Ferdowsi University, P.O. Box 1159-91775, Mashhad, Islamic Republic of Iran

Abstract

We propose a wavelet based stochastic regression function estimator for the estimation of the regression function for a sequence of pairwise negative quadrant dependent random variables with a common one-dimensional probability density function. Some asymptotic properties of the proposed estimator are investigated. It is found that the estimators have similar properties to their counterparts studied earlier in literature.

Keywords: Pairwise negative quadrant dependent; Random design; wavelets; Non-parametric curve estimation

1. Introduction

A sequence of random variables $\{X_n, n \geq 1\}$ is said to be pairwise negative quadrant dependent (NQD) if

$$P(X_i > x_i, X_j > x_j) \leq P(X_i > x_i)P(X_j > x_j)$$

for all $x_i, x_j \in \mathbb{R}$ and for all $i, j \geq 1, i \neq j$.

This definition was introduced by Lehmann (1966).

Suppose that $\{X_n, n \geq 1\}$ is a sequence of pairwise NQD random variables with a common one-dimensional marginal probability density function *f*. The problem of interest is the estimation of Nonparametric regression function based on the observations $\{(X_1,Y_1),...,(X_n,Y_n)\}$. There are many interesting examples where applications of regression smoothing methods have yielded analysis essentially unobtainable by other techniques, for example the monographs by Eubank (1988) and Muller (1988) [11,17]. Nonparametric curve estimation by wavelets has been treated in numerous articles in various setups. These range from the simple Gaussian *iid* error situation to more complicated data structures that often call for a specific algorithm tailored to the problem at hand. In the fixed design case, and for sample size that are a power of 2, wavelet methods offer an appealing method for adaptation of nonparametric curve estimation [1,2,4,8,12]. The objective of this article is to extend above results for a sequence of pairwise NQD random variables with a common one-dimensional marginal probability density function *f*. The organization of the paper is as follows. After introducing wavelet density function estimation given in section 2, we introduce our proposed estimator in section 3 and study its asymptotic properties.

* *E-mail: doosti@math.um.ac.ir*

2. Wavelet Linear Density Estimator

Let $\{X_n, n \geq 1\}$ be a sequence of pairwise NQD random variables on the probability space $(\Omega, \aleph, P)$. We suppose that X_i has a bounded and compactly supported marginal density $f(.)$, with respect to Lebesgue measure, which does not depend on *i*. We estimate this density from *n* observations X_i, $i = 1,\ldots,n$. For any function $f \in L_2(R)$, we can write a formal expansion [5]:

$$f = \sum_{k \in Z} \alpha_{j_0,k}\phi_{j_0,k} + \sum_{j \geq j_0}\sum_{k \in Z} \delta_{j,k}\psi_{j,k} = P_{j_0}f + \sum_{j \geq j_0} D_j f$$

where the functions

$$\phi_{j_0,k}(x) = 2^{j_0/2}\phi(2^{j_0}x - k)$$

and

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$$

constitute an (inhomogeneous) orthonormal basis of $L_2(R)$ Here $\phi(x)$ and $\psi(x)$ are the scale function and the orthogonal wavelet, respectively. Wavelet coefficients are given by the integrals

$$\alpha_{j_0,k} = \int f(x)\phi_{j_0,k}(x)dx, \delta_{j,k} = \int f(x)\psi_{j,k}dx$$

We suppose that both ϕ and $\psi \in C^{r+1}$, $r \in N$, have compact supports included in $[-\delta, \delta]$. We construct the density estimator

$$\hat{f} = \sum_{k \in K_{j_0}} \hat{\alpha}_{j_0,k}\phi_{j_0,k}, \, with, \hat{\alpha}_{j_0,k} = \frac{1}{n}\sum_{i=1}^{n}\phi_{j_0,k}(X_i), \quad (2.1)$$

where K_{j_0} is the set of k such that $\sup p(f) \cap \sup p\phi_{j_0,k} \neq 0$. The fact that ϕ has a compact support implies that K_{j_0} is finite and $\mathbf{card}K_{j_0} = O(2^{j_0})$. Wavelet density estimators arouse much interest in the recent literature [7,10]. In the case of independent samples the properties of the linear estimator (2.1) have been studied for a variety of error measures and density classes [14,16,19]. Doosti *et al.* [9] obtained upper bounds on L_p-losses for the linear estimators (2.1) for negative associated random variables. The estimator in Equation (2.1) is a special case of a kernel density estimator with kernel $K(x,y) = \sum_k \phi_{j_0,k}(x)\phi_{j_0,k}(y)$. In terms of this kernel, this can be expressed as

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{j_0}(x, X_i)$$

where the orthogonal projection kernels are $K_{j_0}(x,y) = 2^{j_0}K(2^{j_0}x, 2^{j_0}y)$. Huang (1999) studied asymptotic bias and variance of linear wavelet density estimator [13]. Define

$$b_m(x) = x^m - \int_{-\infty}^{\infty} K(x,y)y^m dy .$$

The functions $b_m(x)$ are important in expressing the asymptotic bias of linear estimators and finding their efficiencies with respect to the standard kernel density estimators. Theorem 2.1 gives the bias for the density function estimator (2.1).

Theorem 2.1. *[13] Assume that the density f belongs to the Holder space* $C^{m+\alpha}$, $0 \leq \alpha \leq 1$, *and the wavelet-kernel K(x,y) satisfies the localization property:* $\int_{-\infty}^{\infty} K(x,y)(y-x)^{m+\alpha}dy \leq C$ *, for some positive C. Let* $j \to \infty$ *and* $n2^{-j} \to \infty$, *as* $n \to \infty$. *Then, for x fixed,*

$$E\hat{f}(x) - f(x) = -\frac{1}{m!}f^{(m)}(x)b_m(2^j x)2^{-mj}$$
$$+ O(2^{-j(m+\alpha)}).$$

The asymptotic variance of $\hat{f}$ is given in following Theorem 2.2. This theorem is a generalization of a theorem proved by Huang [13]. In the following theorem we suppose scale function be a monotone function.

Theorem 2.2. *Let* $f \in C^1$, f' *is the first derivative of f and f and* f' *be uniformly bounded. Then, for x fixed,*

$$Var\hat{f}(x) = \frac{2^j}{n}f(x)V(2^j x) + O(n^{-1}),$$

where $V(x) = \int_{-\infty}^{\infty} K^2(x,y)dy = K(x,x)$.

Proof.

$$Var\hat{f}(x) = Var\{\frac{1}{n}\sum_{i=1}^{n} K_h(x, X_i)\}$$
$$= \frac{1}{n^2}\sum_{i=1}^{n} Var\{K_h(x, X_i)\}$$
$$+ \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov(K_h(x, X_i), K_h(x, X_j))$$
$$\leq \frac{Var\{K_h(x, X_i)\}}{n}.$$

Because of pairwise NQD property and monotonicity of scale function, we know that $\{K_h(x,X_i), i\geq 1\}$ remains a sequence of pairwise NQD random variables, therefore $Cov(K_h(x,X_i), K_h(x,X_j))\leq 0$. Now, we have

$$Var\hat{f}(x)\leq \frac{1}{n}\int_{-\infty}^{\infty}K_h^2(x,y)f(y)dy$$
$$-\frac{1}{n}(\int_{-\infty}^{\infty}K_h(x,y)f(y)dy)^2$$
$$=\frac{1}{n}f(x)\int_{-\infty}^{\infty}K_h^2(x,y)dy$$
$$+\frac{1}{n}\int_{-\infty}^{\infty}K_h^2(x,y)(f(y)-f(x))dy$$
$$-\frac{1}{n}(\int_{-\infty}^{\infty}K_h(x,y)f(y)dy)^2$$
$$=\frac{1}{nh}f(x)V(x/h)$$
$$+\frac{1}{n}\int_{-\infty}^{\infty}K_h^2(x,y)(f(y)-f(x))dy$$
$$-\frac{1}{n}(\int_{-\infty}^{\infty}K_h(x,y)f(y)dy)^2.$$

Below, we show that the second and the third terms in last equality are of order $O(n^{-1})$

$$|\frac{1}{n}\int_{-\infty}^{\infty}K_h^2(x,y)(f(y)-f(x))dy|$$
$$\leq\frac{1}{n}\sup_x|f(x)|\frac{1}{h^2}\int_{-\infty}^{\infty}K^2(x/h,y/h)|y-x|dy$$
$$\leq\frac{1}{n}\sup_x|f(x)|\sup_{s,t\in\mathrm{R}}|K(s,t)|$$
$$\int_{-\infty}^{\infty}|K(x/h,t)(t-x/h)|dt$$
$$=O(n^{-1}).$$

By the uniform boundedness of $f(x)$, it is easy to conclude that

$$\frac{1}{n}(\int_{-\infty}^{\infty}K_h(x,y)f(y)dy)^2=O(\frac{1}{n}).$$

3. Wavelet Regression Estimators

Consider the nonparametric regression model which is given as the following. Let (X_i,Y_i), $i = 1,\ldots n$ be identically distributed as a two-dimensional random vector (X,Y) with $\mathsf{E}(Y^2)<\infty$. The propose here is to estimate the regression function of Y on X, denoted by $r(x)=\mathsf{E}(Y|X=x)$. An alternative way to write the regression model is the following:

$$Y_i=r(X_i)+\varepsilon_i,\quad i=1,\ldots,n \tag{3.1}$$

where the error ε_i, conditionally on X_i, are assumed to be independent with zero expectation and a bounded (conditional) variance. The above setup corresponds to the random design, but we shall also consider the fixed design model. In this case, the X_i are deterministic and the Y_i follow the same relation (3.1), with ε_i being independent random variables with zero mean and bounded variance. Note that, in both cases, the variance of noise is not assumed to be constant, allowing therefore the analysis of data with heteroscedastic noise. When the X_i are random, it is assumed that their common distribution admits a density f. Otherwise, it is assumed that their empirical distribution converges as $n\to\infty$ to a distribution admitting a density f. Similar to the set up in earlier literature [3,4,6], our estimator of r will be obtained by taking the ratio of wavelet estimators of $g = r f$ and f. The estimator proposed in the above papers is given by,

$$\hat{g}(x)=\sum_{k=-\infty}^{\infty}[\frac{1}{n}\sum_{i=1}^{n}Y_i\phi_{j,k}(Xi)]\phi_{j,k}(x)$$
$$=\frac{1}{n}\sum_{i=1}^{n}Y_iK_h(x,X_i). \tag{3.2}$$

We want to extend their result for sequence of pairwise NQD random variables with a common one-dimensional probability density function. The following theorems give the bias and variance of above estimator.

Theorem 3.1. *Assume that the g(x) belongs to the Holder space* $\mathsf{C}^{m+\alpha}$, $0\leq\alpha\leq 1$, *and the wavelet-kernel K(x,y) satisfies the localization property:* $\int_{-\infty}^{\infty}K(x,y)(y-x)^{m+\alpha}dy\leq C$, *for some positive C. Let* $j\to\infty$ *and* $n2^{-j}\to\infty$, *as* $n\to\infty$. *Then, for x fixed,*

$$\mathsf{E}\,\hat{g}(x)-g(x)=-\frac{1}{m!}g^{(m)}(x)b_m(2^jx)2^{-mj}$$
$$+O(2^{-j(m+\alpha)}).$$

Proof. We have

$$\mathsf{E}\,\hat{g}(x)-g(x)=\mathsf{E}\{\frac{1}{n}\sum_{i=1}^{n}Y_iK_h(x,X_i)-g(x)\}$$
$$=\int_{-\infty}^{\infty}K_h(x,y)g(y)dy-g(x)$$

$$=\int_{-\infty}^{\infty}K_h(x,y)\{g(y)-g(x)\}dy$$

$$=\int_{-\infty}^{\infty}K_h(x,y)(\sum_{i=1}^{m-1}\frac{g^{(i)}(x)}{i!}(y-x)^i$$

$$+\frac{g^{(m)}(\xi_{x,y})}{m!}(y-x)^m)dy$$

$$=\int_{-\infty}^{\infty}K_h(x,y)\frac{g^{(m)}(\xi_{x,y})}{m!}(y-x)^m dy$$

$$=\frac{-1}{m!}g^{(m)}(x)b_m(\frac{x}{h})h^m$$

$$+\int_{-\infty}^{\infty}K_h(x,y)\frac{g^{(m)}(\xi_{x,y})-g^{(m)}}{m!}(y-x)^m dy$$

where

$$b_m(x)=x^m-\int_{-\infty}^{\infty}K(x,y)y^m dy$$

and $\xi_{x,y}$ is some number lying between *x* and *y*. It will be shown below that the term in the last equality is $O(h^{m+\alpha})$. By the localization assumption of $K(x,y)$, we have

$$|\int_{-\infty}^{\infty}K_h(x,y)\frac{g^{(m)}(\xi_{x,y})-g^{(m)}}{m!}(y-x)^m dy|$$

$$\le\frac{A}{m!}\int_{-\infty}^{\infty}|K_h(x,y)(y-x)^{m+\alpha}|dy$$

$$=\frac{Ah^{m+\alpha}}{m!}\int_{-\infty}^{\infty}|K_h(x/h,y)(y-x/h)^{m+\alpha}|dy$$

$$=O(h^{m+\alpha})$$

Theorem 3.2. *Assume that density function f and the regression function r(x) are locally bounded and scale function be a monotone function, Then*

$$Var\{g(x)\}=O(\frac{2^{j_0}}{n}).$$

Proof.

$$Var\,\hat{g}(x)=Var\{\frac{1}{n}\sum_{i=1}^{n}Y_iK_h(x,X_i)\}$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}Var\{Y_iK_h(x,X_i)\}$$

$$+\frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}Cov(Y_iK_h(x,X_i),Y_jK_h(x,X_j)) \quad (3.3)$$

$$=T_1+T_2.$$

Now, we want to find upper bounds for T_1 and T_2. Now, by [3]

$$T_1\le K\frac{2^{j_0}}{n} \quad (3.4)$$

Next, we have

$$Cov(Y_iK_h(x,X_i),Y_jK_h(x,X_j))$$

$$=\mathsf{E}(Y_iK_h(x,X_i)Y_jK_h(x,X_j)) \quad (3.5)$$

$$-\mathsf{E}(Y_iK_h(x,X_i))^2$$

which will be locally bounded by assumptions and because of pairwise NQD property and monotonicity of scale function, we know that $\{K_h(x,X_i),i\ge 1\}$ remains a sequence of pairwise NQD random variables, therefore $Cov(K_h(x,X_i),K_h(x,X_j))\le 0$ so we have

$$T_2=O(\frac{2^{j_0}}{n})$$

These results allow one to control the convergence rate of estimators $\hat{r}=\frac{\hat{g}}{\hat{f}}$. Using [18, p. 13] expansion (2.6), we have

$$\hat{r}(x)=\frac{\mathsf{E}\hat{g}}{\mathsf{E}\hat{f}}+\frac{\hat{g}(x)-\mathsf{E}\hat{g}(x)}{\mathsf{E}\hat{f}(x)}-\frac{\hat{f}(x)-\mathsf{E}\hat{f}(x)}{[\mathsf{E}\hat{f}(x)]^2}$$

$$+O_p([\hat{g}(x)-\mathsf{E}\hat{g}(x)]^2)+O_p([\hat{f}(x)-\mathsf{E}\hat{f}(x)]^2).$$

Then, it follows by using Theorems 2.2 and 3.2 that

$$\mathsf{E}\hat{r}(x)=\frac{\mathsf{E}\hat{g}}{\mathsf{E}\hat{f}}+O(Var\hat{g}(x))+O(Var\hat{f}(x))$$

$$\le\frac{\mathsf{E}\hat{g}}{\mathsf{E}\hat{f}}+O(\frac{2^{j_0}}{n}) \quad (3.6)$$

Now, by using Equation (2.7) of [18], we have

$$\frac{\mathrm{E}\hat{g}}{\mathrm{E}\hat{f}} = r(x) + \frac{\mathrm{E}\hat{g}(x) - g(x)}{f(x)} - \frac{\mathrm{E}\hat{f}(x) - f(x)}{f(x)} r(x)$$

$$+O_p([g(x) - \mathrm{E}\hat{g}(x)]^2) + O_p([f(x) - \mathrm{E}\hat{f}(x)]^2).$$

By the above results, it follows that

$$\frac{\mathrm{E}\hat{g}}{\mathrm{E}\hat{f}} \le r(x) + O(2^{-j_0 m}),$$

hence, from Equation (3.6), order of the bias of the estimators $\hat{r}$, is given by

$$bias(\hat{r}(x)) = O(2^{-j_0 m}) + O(\frac{2^{j_0}}{n}).$$

For variance of $r(x)$, we have

$$Var(\hat{r}(x)) \le \frac{Var\hat{g}(x)}{[\mathrm{E}\hat{f}(x)]^2} + \frac{[\mathrm{E}\hat{g}(x)]^2}{[\mathrm{E}\hat{f}(x)]^4} Var\hat{f}(x)$$

$$+O_p(\mathrm{E}[g(x) - \mathrm{E}\hat{g}(x)]^4)$$

$$+O_p(\mathrm{E}[f(x) - \mathrm{E}\hat{f}(x)]^4)$$

Assuming that $f(x) > 0$ for all x, and using the results on the asymptotic bias and variance of $\hat{g}(x)$ and $\hat{f}(x)$, we conclude that

$$Var(\hat{r}(x)) \le O(\frac{2^{j_0}}{n}).$$

As we see, the convergence rate for bias and variance of our proposed estimator is similar to that of [3]. Numerical studies are required in order to judge the merits of one over the other, which will be pursued further.

References

1. Antoniadis A., Gregoire G., and McKeague I. Wavelet methods for curve estimation. *Journal of the Americal Statistical Association*, **89**: 1340-1353 (1994).
2. Antoniadis A. Smoothing noisy data with coiflets. *Statistica Sinica*, **4**: 651-678 (1994).
3. Antoniadis A. and Pham D.T. Wavelet regression for random or irregular design. Technical Report RT 148, IMAG-LMC, University of Grenoble, France (1995).
4. Antoniadis A., Gregoire G., and Vial P. Random Design Wavelet Curve Smoothing. *Statistics and Probability Letters*, **35**: 225-232 (1997).
5. Daubechies I. Ten Lectures on wavelets. CBMS-NSF regional conferences series in applied mathematics. SIAM, Philadelphia (1992).
6. Delouille V., Simoens J., and von Sachs R. Smooth design-adapted wavelets for nonparametric stochastic regression. *Journal of the Americal Statistical Association*, **99**: 643-659 (2004).
7. Donoho D.L., Johnstone I.M., Kerkyacharian G., and Picard D. Density estimation by wavelet thresholding. *The Annals of Statistics*, **2**: 508-539 (1996).
8. Donoho D.L., Johnstone I.M., Kerkyacharian G., and Picard D. Wavelet shrinkage: asymptopia (with discussion). *Journal of Royal Statistical Society*, Ser. B **57**(2): 301-370 (1995).
9. Doosti H., Fakoor V., and Chaubey P.Y. Wavelet linear density estimation for negative associated sequences. Submitted (2005).
10. Doukhan P. and Leon J.R. Une note sur la deviation quadratique destimateurs de densites par projections orthogonales. *C.R. Acad. Sci.Paris*, t310, serie 1, 425-430 (1990).
11. Eubank R.L. *Spline Smoothing and Nonparametric Regression*. Marcel Dekker, New York (1988).
12. Hrdle W., Kerkyacharian G., Picard D., and Tsybabov A. *Wavelets: Approximation and Statistical Applications.* Springer-Verlag, New York (1998).
13. Huang S.Y. Density estimation by wavelet-based reproducting kernels. *Statistica Sinica*, **9**: 137-151 (1999).
14. Leblanc F. Wavelet linear density estimator for a discrete-time stochastic process: L_p-losses. *Statistics and Probability Letters*, **27**: 71-84 (1996).
15. Lehman E.L. Some concepts of dependence *Annals of Mathematics and Statistics*, **37**: 1137-1153 (1966).
16. Kerkyacharian G. and Picard D. Density estimation in Besov spaces. *Statistics and Probability Letters*, **13**: 15-24 (1992).
17. Muller H.G. *Nonparametric Regression Analysis of Longitudinal Data*. Vol. 46 of Lecture Notes in Statistics. Springer-Verlag, Berlin (1988).
18. Rosenblatt M. *Stochastic Curve Estimation*. NSF-CBMS Regional Conference Series in Probability and Statistics, **3**, Institute of Mathematical Statistics, Hayward, California (1991).
19. Tribouley K. Density estimation by cross-validation with wavelet method. *Statistica Neerlandica*, **45**: 41-62 (1995).