Journal of Statistics: Advances in Theory and Applications
Volume 2, Number 1, 2009, Pages 1-10

ASYMPTOTIC NORMALITY OF WAVELET ESTIMATORS IN SEMIPARAMETRIC REGRESSION MODELS WITH MARTINGALE DIFFERENCE SEQUENCE ERRORS

BIWEN LI[1,2], HONGCHANG HU[2] and BOSHAN CHEN[2]

[1]Department of Control Science and Engineering
Huazhong University of Science and Technology
Wuhan 430074
P. R. China

[2]School of Mathematics and Statistics
Hubei Normal University
Huangshi 435002
P. R. China
e-mail: lbw1818@163.com

Abstract

The paper studies a semiparametric regression model

$$y_i = X_i^T \beta + g(t_i) + \varepsilon_i,\ 1 \le i \le n,$$

where the error $\{\varepsilon_i, F_i, 1 \le i \le n\}$ is a martingale difference sequence. The wavelet estimators of parameter and non-parameter are given and asymptotic normality is investigated under general conditions.

2000 Mathematics Subject Classification: 62E20.

Keywords and phrases: semiparametric regression model, wavelet estimate, martingale difference sequence, asymptotic normality.

Supported by Scientific Research Item of Education Office, Hubei (No. Q200722001), the Natural Science Foundation of Hubei (No. 2007ABA124), Scientific Research Item of Department of Education, Hubei (No. D20092207) and the Key Project of Chinese Ministry of Education (No. 209078), the Major Foundation of Educational Committee of Hubei Province (Z200622002).

Received July 27, 2009

1. Introduction

Consider a semiparametric regression model

$$y_i = X_i^T \beta + g(t_i) + \varepsilon_i, \; 1 \le i \le n, \tag{1}$$

where y_i is real-valued relating to the observation at t_i, the first part of the model is parametric with parameter $\beta \in R^d$ with d denoting the number of parameters, $X = (X_{ir})_{n \times d}$ are random carriers. The second part $g(t)(t \in [0, 1])$ denotes the nonparametric signal, $\{t_i\}$ is a deterministic sequence in interval $[0, 1]$. $\{\varepsilon_i, F_i, i \le n\}$ is a martingale difference sequence (The martingale difference sequence can be found in Stout [15], or Akio [1]).

Following Speckman (See [14]), denote

$$x_{ir} = f_r(t_i) + \eta_{ir}, \; 1 \le i \le n, \; 1 \le r \le d, \tag{2}$$

where $f_r(\cdot)$ is some unknown function on $[0, 1]$, $\{\overline{\eta}_i, i \ge 1\}$ are stochastic sequence with

$$\overline{\eta}_i = (\eta_{i1}, \cdots, \eta_{id})^T \text{ i.i.d.},$$

and

$$E\,\overline{\eta}_i = 0, \; \mathrm{Var}(\overline{\eta}_i) = V, \tag{3}$$

where $V = (V_{ij})\,(j = 1, 2, \cdots, d)$ is a positive definite matrix with d-order. Moreover, $\{\eta_{ir}\}$ and $\{\varepsilon_i\}$ are independent.

Since, the semiparametric regression model contains linear components and a nonparametric component, it is more flexible than the usual standard linear models and attractive in some applications. When the $\{\varepsilon_i\}$ is a strictly stationary error process, the semiparametric regression model has been discussed by Gao and Anh [6]. When the error sequence $\{\varepsilon_i\}$ is independent and identically distributed, many significant results are obtained. See, for example, Chen [5], Bianco and Boente [3], Shi and Teng [12]. In the case, using wavelet method, the model has been studied by Qian and Cai [10], Qian et al. [11], Chai and Xu [4]. When the error sequence $\{\varepsilon_i\}$ is a martingale difference, the model (where $\beta \in R^1, X_i \in R^1$) is studied by using near neighbour method (See Yan et al. [16]). Using wavelet method, Hu and Hu [8] investigate strong consistency in model (1)-(3).

In this paper, using wavelet method, the semiparametric regression model is discussed, while error $\{\varepsilon_i\}$ is a martingale difference sequence. The organization of this paper is as follows. The wavelet estimators of β and $g(t)$ are given in Section 2. Under general conditions, the asymptotic normalities of $\hat{\beta}_n$ and $\hat{g}_n(t)$ are obtained in Section 3. The main proofs are presented in Section 4.

2. Estimation Method

Suppose that there exists a scaling function $\phi(x)$ in the Schwartz space S_l and a multiresolution analysis $\{V_m\}$ in the concomitant Hilbert space $L^2(R)$, with its reproducing kernel $E_m(t, s)$ given by

$$E_m(t, s) = 2^m E_0\left(2^m t, 2^m s\right) = 2^m \sum_{k \in z} \phi\left(2^m t - k\right) \phi\left(2^m s - k\right).$$

Let $A_i = [s_{i-1}, s_i]$ denote intervals that partition $[0, 1]$ with $t_i \in A_i$, and $1 \leq i \leq n$. The estimation method will be introduced as following:

Firstly, suppose that β is known, we define estimator of $g(\cdot)$ by

$$\hat{g}_0(t) = \hat{g}_0(t, \beta) = \sum_{i=1}^{n} \left(y_i - X_i^T \beta\right) \int_{A_i} E_m(t, s) ds.$$

In succession, we define wavelet estimator $\hat{\beta}_n$ by minimizing

$$\sum_{i=1}^{n} \left(y_i - X_i^T \beta - \hat{g}_0(t_i, \beta)\right)^2.$$

Finally, we define linear wavelet estimator of $g(\cdot)$ by

$$\hat{g}_n(t) = \hat{g}_0\left(t, \hat{\beta}_n\right) = \sum_{i=1}^{n}\left(y_i - X_i^T\hat{\beta}_n\right)\int_{A_i} E_m(t, s)ds.$$

Let

$$X = (X_{ir})_{n\times d}, Y = (y_1, \cdots, y_n)^T, g = (g(t_1), \cdots, g(t_n))^T, S_{ij} = \int_{A_j} E_m(t_i, s)ds,$$

$$S = (S_{ij})_{n\times n}, \varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)^T, \tilde{X} = (I - S)X, \tilde{Y} = (I - S)Y.$$

Then, we obtain that

$$\hat{\beta}_n = \left(\tilde{X}^T\tilde{X}\right)^{-1}\tilde{X}^T\tilde{Y}.$$

To obtain our results, the following four conditions are sufficient.

(A_1) $g(\cdot), f_r(\cdot) \in H^\alpha$ (Sobolev space, see Chai and Xu [4]), for some $\alpha > 1/2, 1 \le r \le d$;

(A_2) $g(\cdot)$ and $f_r(\cdot)$ are Lipschitz functions of order $\gamma > 0, 1 \le r \le d$;

(A_3) $\phi(\cdot)$ belongs to S_l, which is a Schwartz space for $l \ge \alpha$. ϕ is a Lipschitz function of order 1 and has compact support, in addition to $\left|\hat{\phi}(\xi) - 1\right| = O(\xi)$ as $\xi \to 0$, where $\hat{\phi}$ denotes Fourier transform of ϕ;

(A_4) $s_i (i = 1, \cdots, n)$ and m satisfy $\max_{1\le i\le n}(s_i - s_{i-1}) = O\left(n^{-1}\right)$ and $2^m = O\left(n^{1/3}\right)$, respectively.

3. Statement of the Results

Now, we state the following results of this paper.

Theorem 3.1. *Assume that conditions* (A_1)-(A_4) *hold,* $\{\overline{\eta}_i, 1 \le i \le n\}$ *is a measurable random sequence on* $\bigcap_{k=1}^{n} F_{k-1}, \{\overline{\eta}_i, 1 \le i \le n\}$ *and*

$\{\varepsilon_i, 1 \le i \le n\}$ *are a.s. bounded. If there exists some* $u \in (1/2 + 1/q, 1)$ $(q > 2)$, *such that* $\sup_i \int_{A_i} |E_m(t, s)| ds = O(n^{-u})$, *then for* $\gamma > 1/4$,

$$\sqrt{n}\left(\hat{\beta}_n - \beta\right) \xrightarrow{L} N\left(0, \sigma^2 V^{-1}\right) \quad (n \to \infty).$$

Remark. When the error sequence $\{\varepsilon_i, i = 1, \cdots, n\}$ is i.i.d., by the theorem, we can easily to obtain the corresponding result, which is discussed by Chai and Xu [4].

Theorem 3.2. *Assume that* $0 < c \le \sigma_i^2 = E\left((\varepsilon_i)^2 | F_{i-1}\right) < \infty$, $\{\varepsilon_i, 1 \le i \le n\}$ *are a.s. bounded and there exists some* $u \in (1/2 + 1/q, 1)$ $(q > 2)$, *such that*

$$\sup_i \int_{A_i} |E_m(t, s)| ds = O(n^{-u}).$$

If condition (A3) *holds, then*

$$\frac{\left\{\hat{g}_n(t) - \sum_{i=1}^{n}\left[X_i^T\left(\beta - \hat{\beta}_n\right) + g(t_i)\int_{A_i} E_m(t, s) ds\right]\right\}}{\sqrt{\sum_{i=1}^{n}\left[\sigma_i \int_{A_i} E_m(t, s) ds\right]^2}} \xrightarrow{D} N(0, 1).$$

4. Proofs of Theorems

Before the proofs of the theorems, we introduce some preliminary results. For simplicity, C is an arbitrary positive constant, which could take different value at each occurrence.

Lemma 4.1 (Antoniads et al. [2]). *If condition* (A3) *holds, then*

(I) $|E_0(t, s)| \le \frac{C_k}{(1 + |t - s|)^k}$ *and* $|E_m(t, s)| \le \frac{2^m C_k}{(1 + 2^m |t - s|)^k}$ *for* $k \in N$, *where* C_k *is a real constant depending only on* k;

(II) $\sup_{0 \leq s \leq 1} | E_m(t, s)| = O(2^m)$;

(III) $\sup_t \int_0^1 |E_m(t, s)| ds \leq C.$

Lemma 4.2 (Yang [17]). *Let* $S_i = \sum_{j=1}^{i} X_j$. *If* $\{S_i, F_i, i \geq 1\}$ *is a martingale sequence, then*

$$E \max_{1 \leq i \leq n} |S_i|^r \leq 2C_{r_2} \left(\sum_{i=1}^{n} \left(E|X_i|^r \right)^{2/r} \right)^{r/2} \quad (\forall r > 2).$$

Further, if there exists some α *with* $0 < \alpha \leq 2$, *and a positive constant sequence* $\{M_n, n \geq 1\}$, *such that* $\sum_{i=1}^{n} E\left(|X_i|^{\alpha} | F_{i-1} \right) \leq M_n$ *a.s. holds, then*

$$E \max_{1 \leq i \leq n} |S_i|^r \leq \frac{2 + r^2}{\alpha} C_{r_1} \sum_{i=1}^{n} E|X_i|^r + C_{r_2}^{\frac{2}{r}} M_n^{\frac{r}{\alpha}} \quad (\forall r > 2),$$

where $C_{r_1} = (r / (r-1))^r 2^{r-3} r$, and $C_{r_2} = \left(C_{r_1} r \right)^{r/2}$.

Lemma 4.3. *Suppose that* $\{\varepsilon_i, F_i, i \leq n\}$ *is a martingale difference sequence with* $E(|\varepsilon_i||F_{i-1}) \leq C$ *a.s. and* $\sup_i E|\varepsilon_i|^q < \infty$ *for some* $q > 2$, *and there exists* $u \in (1/2 + 1/q, 1)$, *such that* $\sup_i \int_{A_i} |E_m(t, s)| ds = O\left(n^{-u}\right)$, *for all* $t \in [0, 1]$. *Then for all* $t \in [0, 1]$,

$$\sum_{i=1}^{n} \varepsilon_i \int_{A_i} E_m(t, s) ds \xrightarrow{a.s.} 0, \quad n \to \infty.$$

Proof. By Lemma 4.2, we have that

$$E\left|\sum_{i=1}^{n}\varepsilon_i\int_{A_i}E_m(t,s)ds\right|^q \le C\left(\sum_{i=1}^{n}\left|\int_{A_i}E_m(t,s)ds\right|^q E|\varepsilon_i|^q + \left(\sum_{i=1}^{n}\left(\int_{A_i}E_m(t,s)ds\right)^2\left(E|\varepsilon_i|^q\right)^{2/q}\right)^{q/2}\right)$$

$$\le C\left(\sum_{i=1}^{n}\left|\int_{A_i}E_m(t,s)ds\right|^q + \left(\sum_{i=1}^{n}\left(\int_{A_i}E_m(t,s)ds\right)^2\right)^{q/2}\right)$$

$$\le C\left(n^{1-uq} + n^{(1-2u)q/2}\right). \tag{4}$$

Therefore, $\sum_{n=1}^{\infty}E\left|\sum_{i=1}^{n}\varepsilon_i\int_{A_i}E_m(t,s)ds\right| < \infty$ follows from (4). Thus, using Borel-Cantelli lemma, we can obtain

$$\sum_{i=1}^{n}\varepsilon_i\int_{A_i}E_m(t,s)ds \xrightarrow{a.s.} 0, \quad n\to\infty. \qquad \blacksquare$$

Lemma 4.4. *Let* $\{\xi_{nk}, F_k^n, n, k \ge 1\}$ *be a martingale difference sequence with* $E\,\xi_{nk}^2 < \infty$. *Assume that the following conditions hold.*

(I) $\sum_{h=1}^{n}E\left(\xi_{nk}^2\, I(|\xi_{nk}| > \delta)\middle|F_{k-1}^n\right)\xrightarrow{p} 0$, *for* $\forall\delta > 0$ *as* $n\to\infty$,

(II) $\sum_{h=1}^{n}E\left(\xi_{nk}^2\middle|F_{k-1}^n\right)\xrightarrow{p}\sigma_1^2$.

Then $\sum_{h=1}^{n}\xi_{nk}\xrightarrow{L} N\left(0, \sigma_1^2\right)$.

Proof. See Theorem 1.2 of Kundu et al. [9], or Lemma 1.1 of Hu [7]. ■

Proof of Theorem 3.1. Note that

$$\hat{\beta}_n - \beta = \left(n^{-1}\tilde{X}^T\tilde{X}\right)^{-1}\left(n^{-1}\tilde{X}^T\tilde{g} + n^{-1}\tilde{X}^T\tilde{\varepsilon}\right). \tag{5}$$

From proof of Theorem 2.1 in [8], it is easy to see that

$$n^{-1}\widetilde{X}^T\widetilde{X} \xrightarrow{P} V \quad (n \to \infty), \tag{6}$$

and

$$n^{-1/2}\widetilde{X}^T\widetilde{g} \xrightarrow{P} 0 \quad (n \to \infty). \tag{7}$$

We will show that

$$n^{-1/2}\widetilde{X}^T\widetilde{\varepsilon} \xrightarrow{L} N\left(0, \sigma^2 V\right)(n \to \infty). \tag{8}$$

In fact, we can get $n^{-1/2}a^T\widetilde{X}^T\widetilde{\varepsilon} \hat{=} n^{-1/2}b^T\widetilde{\varepsilon} = \sum_{h=1}^{n} n^{-1/2}b_h\widetilde{\varepsilon}_h$ for arbitrary choosing vector a^T, where $b^T \hat{=} a^T\widetilde{X}^T \hat{=} (b_1, \cdots, b_n)$.

Since $\{\varepsilon_h, F_h, 1 \le h \le n\}$ is a martingale difference sequence, $\{\overline{\eta}_k, F_{h-1}, h \ge 1\}$ is a random sequence, it is easy to show that $\left\{n^{-1/2}b_h\widetilde{\varepsilon}_h, F_h, h \ge 1\right\}$ is a martingale difference sequence (see exercise 三(5) of Shi [12]). By the fact, Lemma 4.3, the dominated convergence theorem and (6), we get

$$\sum_{h=1}^{n} E\left(\frac{1}{n} b_h^2 \widetilde{\varepsilon}_h^2 \,|F_{h-1}\right) = \frac{1}{n}\sum_{h=1}^{n} b_h^2\, E\left(\widetilde{\varepsilon}_h^2 \,|F_{h-1}\right) \text{(a.s.)} \xrightarrow{P} \sigma^2 a^T Va.$$

Since, it is easy to see that $b_h\widetilde{\varepsilon}_h$ is a.s. bounded, we obtain

$$\sum_{h=1}^{n} E\left(\frac{1}{n} b_h^2\, \widetilde{\varepsilon}_h^2\, I\left(\left|n^{-1/2}b_h\widetilde{\varepsilon}_h\right| > \delta\right)\middle|F_{h-1}\right) \le \max_{1\le h\le n} E\left(b_h^2\, \widetilde{\varepsilon}_h^2\, I\left(\left|b_h\widetilde{\varepsilon}_h\right| > \sqrt{n}\delta\right)\middle|F_{h-1}\right) \xrightarrow{p} 0.$$

By Lemma 4.4, we can obtain that

$$n^{-1/2}a^T\widetilde{X}^T\widetilde{\varepsilon} = \sum_{h=1}^{n} n^{-1/2}b_h\widetilde{\varepsilon}_h \xrightarrow{L} N\left(0, \sigma^2 a^T Va\right).$$

Therefore, we have that $n^{-1/2}\tilde{X}^T\tilde{\varepsilon} \xrightarrow{L} N\left(0, \sigma^2 V\right)$. Now, the theorem follows from (5)-(8). ■

Proof of Theorem 3.2. Let

$$C_{ni} = \int_{A_i} E_m(t, s)ds \Bigg/ \sqrt{\sum_{i=1}^{n}\left(\sigma_i \int_{A_i} E_m(t, s)ds\right)^2}.$$

Then

$$\frac{\left\{\hat{g}_n(t) - \sum_{i=1}^{n}\left[X_i^T\left(\beta - \hat{\beta}_n\right) + g(t_i)\int_{A_i} E_m(t, s)ds\right]\right\}}{\sqrt{\sum_{i=1}^{n}\left[\sigma_i \int_{A_i} E_m(t, s)ds\right]^2}}$$

$$= \frac{\sum_{i=1}^{n}\varepsilon_i \int_{A_i} E_m(t, s)ds}{\sqrt{\sum_{i=1}^{n}\left[\sigma_i \int_{A_i} E_m(t, s)ds\right]^2}} = \sum_{i=1}^{n} C_{ni}\varepsilon_i, \tag{9}$$

$$\sum_{i=1}^{n} C_{ni}^2\sigma_i^2 = 1, \quad \sup_{n\geq 1}\sum_{i=1}^{n} C_{ni}^2 \leq \sup_{n\geq 1}\left(1\Big/\inf_i \sigma_i^2\right) < \infty. \tag{10}$$

By Cauchy equality and Lemma 4.1, we have

$$\max_{1\leq i\leq n}|C_{ni}| \leq \frac{\sup\int_{A_i}|E_m(t, s)|ds}{\inf \sigma_i \cdot \sqrt{\sum_{i=1}^{n}\left(\int_{A_i} E_m(t, s)ds\right)^2}} \leq \frac{\sqrt{n}\cdot \sup\int_{A_i}|E_m(t, s)|ds}{\inf \sigma_i \cdot \left|\int_{0-}^{1} E_m(t, s)ds\right|}$$

$$= O\left(n^{1/2-u}\right) \to 0 \quad (n \to \infty). \tag{11}$$

Therefore, the desired conclusion follows from (9)-(11) and Lemma 4.4.

References

[1] T. Akio, Martingale limit theorem and its application to an ergodic controlled Markov Chain, Systems & Control Letters 26 (1995), 261-266.

[2] A. Antoniads, G. Gregoire and I. W. Mckeague, Wavelet methods for curve estimation, J. Amer. Statist. Assoc. 89 (1994), 1340-1353.

[3] A. Bianco and G. Boente, Robust estimators in semiparametric partly linear regression models, J. Statist. Planning and Inference 122 (2004), 229-252.

[4] G. Chai and K. Xu, Wavelet smoothing in semiparametric regression model, Chinese J. Appl. Prob. Statist. 15 (1999), 97-105.

[5] H. Chen, Convergence rates for parametric components in a partly linear model, Annals of Statist. 16 (1988), 136-146.

[6] J. T. Gao and V. V. Anh, Semiparametric regression model under long-range dependent errors, J. Statist. Planning and Inference 80 (1999), 37-57.

[7] S. Hu, Central limit theorem for weighted sum of martingale difference, Acta Mathematicae Applicatae Sinica 24 (2001), 539-545.

[8] H. C. Hu and D. H. Hu, Strong consistency of wavelet estimation in semiparametric regression models, Acta Math Sinica (Chinese Series) 49(6) (2006), 1417-1424.

[9] S. Kundu, S. Majumder and K. Mukherjee, Central limit theorems revisited, Statist. Prob. Lett. 47 (2000), 265-275.

[10] W. Qian and G. Cai, Strong approximability of wavelet estimate in semiparametric regression model, Science in China (series A) 29 (1999), 233-240.

[11] W. Qian, G. Cai and F. Jiang, Error variance of wavelet estimate in semiparametric regression, Chinese Journal of Annual of Maths 21 (2000), 341-350.

[12] P. Shi and X. Teng, Asymptotic distributions of *M*-estimators of the parametric components of partly linear models with fixed carriers, Advances in Mathematics 28 (1999), 447-461.

[13] J. Shi, Discrete Martingale and its Application, Science Press, Beijing, 1996.

[14] P. Speckman, Kernel smoothing in partial linear models, J. R. Statist. Soc. 50 (1988), 413-436.

[15] W. F. Stout, Almost Sure Convergence, Academic Press, New York, 1974.

[16] Z. Yan, W. Wu and Z. Nie, Near neighbour estimate in semiparametric regression model: The martingale difference error sequence case, Chinese J. Appl. Prob. Statist. 17 (2001), 44-50.

[17] S. Yang, Nonparametric regression weighted function estimator for martingale sequence, J. Sys. Sci. Math. Scis. 19 (1999), 79-85.

■