Journal of Mathematical Research & Exposition
Mar., 2009, Vol. 29, No. 2, pp. 247–256
DOI:10.3770/j.issn:1000-341X.2009.02.007
Http://jmre.dlut.edu.cn

Wavelet Detection and Estimation of Change Points in Nonparametric Regression Models under Random Design

ZHAO Wen Zhi[1], *TIAN Zheng*[1], *XIA Zhi Ming*[2]
(1. Department of Applied Mathematics, Northwestern Polytechnical University, Shaanxi 710072, China;
2. Department of Mathematics, Northwest University, Shaanxi 710069, China)
(E-mail: zhaowenzhi1979@126.com)

Abstract A wavelet method of detection and estimation of change points in nonparametric regression models under random design is proposed. The confidence bound of our test is derived by using the test statistics based on empirical wavelet coefficients as obtained by wavelet transformation of the data which is observed with noise. Moreover, the consistence of the test is proved while the rate of convergence is given. The method turns out to be effective after being tested on simulated examples and applied to IBM stock market data.

Keywords random design; nonparametric regression model; change point; wavelet transformation; consistent test; rate of convergence.

Document code A
MR(2000) Subject Classification 62G08; 62G10
Chinese Library Classification O212.7

1. Introduction

Nonparametric regression methods are usually applied in order to obtain a smooth fit of a regression curve without having to specify a parametric class of regression function[1]. However, sometimes a generally smooth regression function might contain one or more discontinuities, which are called change points. Thus, it is very necessary to detect and estimate these change points to ensure the accuracy of modeling.

A remarkable property of wavelet coefficients is to reflect the local regularity of the original function, being large where the function is irregular and small where the function is smooth. This property is very useful to detect discontinuities or change points in a regression function observed with noise[2]. Wang[3] proposed a test statistics based on the optimization of the absolute value of the wavelet coefficients; Odgen and Parzen[4] presented a method based on the cumulative sum of squared wavelet coefficients; Raimondo[5] studied the minimax estimation of change point in nonparametric regression model; Antoniadis and Gijbels[6] considered the wavelet detection and estimation problem of the location of discontinuities in piecewise smooth regression function.

All the above results are based on the assumption of i.i.d. Gaussian noise and the data are obtained in a fixed design manner. However, in many practical situations the observed data are

Received date: 2007-06-22; **Accepted date**: 2007-11-22
Foundation item: the National Natural Science Foundation of China (No. 60375003); the Astronautics Basal Science Foundation of China (No. 03I53059).

obtained through a random design manner and the noise process may not be a Gaussian process but an i.i.d. sequence. In this paper, we study the change point problem in nonparametric function when noise process is i.i.d. sequence and the data are obtained in a random design manner.

This paper is organized as follows: Section 2 presents our model and assumptions. Section 3 details the wavelet detection problem and Section 4 establishes the consistency estimators of change points. A simulated study and a data example are discussed in Section 5. Conclusions are summarized in Section 6.

2. Model and assumptions

The model considered in this paper is as follows:

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

where ε_i are i.i.d random variables with zero mean and unit variance and are independent of the X_i's. The design point X_i's are assumed to be supported in the interval [0,1].

Consider the following testing problem:

H_0: f is smooth (at least continuously differentiable on [0,1]);

$H_1(m)$: f has "at-least 1 and at-most m" jump points and is otherwise smooth.

We assume that the number and the location of the change points in the regression function f are unknown. However, we suppose that a realistic upper bound to the number of change-points to be tested is known. In our assumption $H_1(m)$, m denotes such an upper bound, it is supposed to be known. The purpose of this paper is to test H_0 against $H_1(m)$ and to estimate the number and the location of the change points, if H_0 is rejected.

Suppose that a sample of n data pairs $\chi = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ is observed, generated from model (1). Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be the order statistics of the sample $X_1, \ldots, X_n$. Then (1) becomes

$$Y_{[i]} = f(X_{(i)}) + e_i, i = 1, \ldots, n, \tag{2}$$

where $Y_{[i]}$ and e_i are rearrangement of Y_i and ε_i, respectively. In fact, (1) can be rewritten as follows:

$$Y_{[i]} = f(i/n) + [f(X_{(i)}) - f(i/n)] + e_i = f(i/n) + \xi_i + e_i, \tag{3}$$

where $\xi_i = f(X_{(i)}) - f(i/n)$.

Throughout this paper, we adopt the following assumptions:

(A): $\varepsilon_1 \stackrel{d}{=} -\varepsilon_1$;

(B): $\varepsilon_i, i = 1, \ldots, n$ satisfy the Cramer's condition, that is, $Ee^{t\varepsilon_i} < \infty$ when $|t| < H$ for some $H > 0$;

(C): f is bounded on the interval $[0, 1]$;

(D): X_i are i.i.d. and $X_i \sim U[0, 1]$.

3. Detection of change point

We detect and estimate change point using Haar wavelet. Given Haar "mother" wavelet function $\psi(u) = I_{[0,1/2)}(u) - I_{[1/2,1]}(u)$, a doubly indexed family of wavelets basis is generated by dilating and translating ψ:

$$\psi_{j,k}(u) = 2^{j/2}\psi(2^j u - k), \quad j \in N, k \in Z.$$

The wavelet transform of a given function f is defined by

$$\int f(u)\psi_{j,k}(u)\mathrm{d}u, \quad j \in N, k \in Z.$$

Taking into account the L^2-normalisation of the empirical wavelet transform, Raimondo[2] and Härdle[7] approximate the above integrals by sums

$$w_{j,k}(f) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{j,k}(i/n)f(i/n).$$

Hence, for any resolution level j, $0 \le j \le J$, $2^J = n$ and indexed $k = 0, 1, \ldots, 2^j$, in the wavelet domain (3) becomes

$$w_k = w_k(f) + w_k(\xi) + w_k(e), \tag{4}$$

where

$$w_k(f) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{j,k}(i/n)f(i/n), \quad w_k(\xi) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{j,k}(i/n)\xi_i, \quad w_k(e) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{j,k}(i/n)e_i.$$

We now present some properties of $w_k(f)$, $w_k(\xi)$ and $w_k(e)$, which are stated as two lemmas. The first one is borrowed from Raimondo[2].

Lemma 1[2] *Under H_0 and H_1 respectively, the following two properties hold:*

1) Under H_0, the function f is differentiable so that for all resolution levels $j \ge j_0 = 3$ and all $k = 0, 1, \ldots, 2^j - 1$, $|w_k(f)| \le c_1(n2^{-3j})^{1/2}$.

2) Under H_1, there exists at least a point $x \in [k/2^j, (k+1)/2^j)$ where f has a jump so that $|w_k(f)| \ge c_2(n2^{-j})^{1/2}$.

The constants c_1, c_2 depend only on f. To simply the exposition we take $c_1 = c_2 = 1$.

Lemma 2 *For any resolution level j and indexed $k = 0, 1, \ldots, 2^j$, we have the following two properties:*

1) $w_k(\xi) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{j,k}(i/n)(f(X_{(i)}) - f(i/n)) = o_P(1)$;

2) $w_k(e) \xrightarrow{d} N(0,1)$, as $n \longrightarrow \infty$.

Proof 1) Since $X_i \sim U[0,1]$ by assumption, we have $X_{(i)} \sim \beta(i, n-i+1)$, thus $|X_{(i)} - i/n| = O_P(1/n)$. Hence using assumption (C) we have

$$w_k(\xi) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{j,k}(i/n)(f(X_{(i)}) - f(i/n)) = O_P(1/\sqrt{n}) = o_P(1).$$

2) Because Haar wavelet is the step function $\psi(u) = I_{[0,1/2)}(u) - I_{[1/2,1]}(u)$, the support of the $\psi_{j,k}(u)$ is exactly the dyadic interval $[k/2^j, (k+1)/2^j)$. The number of points

$i/n, i = 1, 2, \ldots$ in the dyadic interval $[k/2^j, (k+1)/2^j)$ is $n/2^j$. Since e_i are the rearrangement of ε_i and ε_i are i.i.d. random variables, we know from the independence between ε_i and X_i and assumption (A) that e_i are i.i.d. random variables and $e_i \stackrel{d}{=} -e_i$. Hence

$$w_k(e) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_{j,k}(i/n) e_i \stackrel{d}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2^j} 2^{j/2}(\pm 1) e_i \stackrel{d}{=} \frac{1}{\sqrt{n/2^j}} \sum_{i=1}^{n/2^j} e_i. \tag{5}$$

Thus by the central limit theorem, as $n \longrightarrow \infty$,

$$w_k(e) \xrightarrow{d} N(0,1).$$

3.1 Critical region and confidence bounds

In the next theorem, we choose the first m maximum values of $|w_k|, k = 1, 2, \ldots, 2^j$ as the statistics. We give a condition on the resolution j which ensures that there are at least a finite number of exceedances over the threshold v_n. Our condition depends on the sample size, and $a_n \asymp b_n$ means that there exist some positive constants c_1 and c_2, such that, for n large enough, $c_1 b_n \le a_n \le c_2 b_n$.

Theorem 1 *Let $|w_{(k)}|$ be the ordered (absolute value of) wavelet coefficients at the resolution level j so $|w_{(1)}| > |w_{(2)}| > \cdots > |w_{(2^j)}|$. Let $v_n = \sqrt{n 2^{-3j}}$. Under assumptions (A)-(D) and H_0, if the resolution level $j = j(n)$ satisfies*

$$2^j \asymp \frac{n}{(\log n)^\delta}, 1 < \delta < 2, \tag{6}$$

then for any arbitrary constant $m \ge 1$,

$$P(|w_{(m)}| > v_n) \longrightarrow 1, \text{ as } n \longrightarrow \infty.$$

Proof By (4), there is an indexed k_1 such that $|w_{(m)}| = |w_{k_1}(f) + w_{k_1}(\xi) + w_{k_1}(e)|$. Using the first result of Lemma 1 and the triangle inequality

$$|w_{(m)}| \ge |w_{k_1}(\xi) + w_{k_1}(e)| - |w_{k_1}(f)|,$$

we know that

$$P\{|w_{(m)}| > v_n\} \ge P\{|w_{k_1}(\xi) + w_{k_1}(e)| \ge 2v_n\}.$$

Since $v_n = o(\sqrt{n/2^j})$ and e_i are i.i.d. sequence, by Lemma 2, assumption (B) and Cramer Large Deviation Theorem[8] we have that

$$\begin{aligned}
\lim_{n\to\infty} P\{|w_{(m)}| > v_n\} &\ge \lim_{n\to\infty} P\{|w_{k_1}(\xi) + w_{k_1}(e)| \ge 2v_n\} \\
&= \lim_{n\to\infty} P\{|w_{k_1}(e)| \ge 2v_n\} \\
&= \lim_{n\to\infty} P\{|N| \ge 2v_n\} \quad (\text{where } N \sim N(0,1)).
\end{aligned}$$

Observe that the resolution level j satisfies (6) and $v_n \longrightarrow 0$, as $n \longrightarrow \infty$, so that

$$\lim_{n\to\infty} P\{|N| \ge 2v_n\} = 1 \quad (\text{where } N \sim N(0,1)),$$

which completes the proof. $\square$

Theorem 1 shows that the number of the wavelet coefficients $|w_k|$ which exceed the threshold v_n is m at least, which means that v_n may be too small in application. Instead, we consider the statistics

$$T_i = \left|w_{(i)}\right| - \left|w_{(m+1)}\right|, i = 1, 2, \ldots, m. \tag{7}$$

In the theorem below we use result of Theorem 1 to derive $100(1-\beta)/\%$ confidence bounds for the combined m-exceedances (7).

Theorem 2 *Let β be an arbitrary number, $0 < \beta < 1$, and put $c_i = \sqrt{-2\log\frac{\beta}{m(m-i+1)}}$,*

$$R_n(\beta) = \bigcup_{i=1,\ldots,m} \{T_i > c_i\}. \tag{8}$$

Under assumptions (A)–(D) and H_0, if the level j satisfies (6) and $n \to \infty$, then

$$P\{R_n(\beta)\} \le \beta.$$

Proof Let A_n be the event that $\left|w_{(m+1)}\right| > v_n$. Then we have that $\lim_{n\to\infty} P(A_n) = 1$ and $P\{T_i > x|\, A_n\} \le P\{\left|w_{(i)}\right| - v_n > x\right|\, A_n\}$ for any constant x. Writing $|w_i|, |w_{i+1}|, \ldots, |w_m|$ for the unordered set of coefficients whose ordered sequences are $\left|w_{(i)}\right| \ge \left|w_{(i+1)}\right| \ge \cdots \ge \left|w_{(m)}\right|$, we obtain from (4) and Lemma 1 that

$$\begin{aligned} P\{\left|w_{(i)}\right| - v_n > x\right|\, A_n\} &= P\{\max_{k=i,\ldots,m} |w_k| - v_n > x|\, A_n\} \le \sum_{k=i}^{m} P\{|w_k| - v_n > x|\, A_n\} \\ &\le \sum_{k=i}^{m} P\{|w_k(\xi) + w_k(e)| > x|\, A_n\}. \end{aligned}$$

By Theorem 1, Lemma 2 and Feller's inequality, we have

$$\begin{aligned} \lim_{n\to\infty} P\{T_i > x|\, A_n\} &\le \lim_{n\to\infty} \sum_{k=i}^{m} P\{|w_k(\xi) + w_k(e)| > x|\, A_n\} \\ &\le \sum_{k=i}^{m} P\{|N| > x\} \quad (\text{where } N \sim N(0,1)) \\ &\le (m-i+1)\frac{2}{\sqrt{2\pi}} \cdot \frac{1}{x} \cdot \exp(-\frac{x^2}{2}). \end{aligned}$$

Putting $x = c_i$, then we obtain

$$\begin{aligned} \lim_{n\to\infty} P\{R_n(\beta)\} &= \lim_{n\to\infty} P\{R_n(\beta)|\, A_n\} \\ &\le \lim_{n\to\infty} \sum_{i=1}^{m} P\{T_i > c_i|\, A_n\} \\ &\le m(m-i+1)\frac{2}{\sqrt{2\pi}} \cdot \frac{1}{c_i} \cdot \exp(-\frac{c_i^2}{2}) \\ &= m(m-i+1) \cdot \frac{2}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{-2\log\frac{\beta}{m(m-i+1)}}} \cdot \frac{\beta}{m(m-i+1)} \\ &\le \beta. \end{aligned}$$

□

Remark 1 If the model considered is $Y_i = f(X_i) + \sigma\varepsilon_i, i = 1, 2, \ldots, n$, then the critical value in (8) is σc_i.

3.2 The consistence of detection

We discuss the consistence of the detection in this section.

Theorem 3 *Let $R_n(\beta)$ be defined by Theorem 2. Under $H_1(m)$, if the resolution level j satisfies (6), then*

$$\lim_{n\to\infty} P\{R_n(\beta)\} = 1.$$

Proof To simplify the exposition we suppose that $m = 1$ (extension to other cases is straightforward). The theorem will follow if we prove that

$$T_1 \longrightarrow \infty, \text{ as } n \longrightarrow \infty. \tag{9}$$

Let $x_n = \sqrt{2\log n}$, $B_n = \{\max_{k=0,\ldots,2^j-1} |w_k(\xi) + w_k(e)| \le x_n\}$. We prove the theorem in the following two steps. Firstly, we prove (9) on condition B_n. Secondly, we prove that

$$P(B_n) \longrightarrow 1, \text{ as } n \longrightarrow \infty. \tag{10}$$

Under $H_1(1)$, by Lemma 1, there exists a unique index k_2 such that $|w_{k_2}(f)| \ge \sqrt{n2^{-j}}$ where $|w_k(f)| \le \sqrt{n2^{-3j}} = v_n$ for any $k \ne k_2$. Working conditionally on B_n, using (6) and the definition of x_n, it is not hard to check that $v_n = o(x_n), x_n = o(\sqrt{n2^{-j}})$. It follows that $|w_k| \le |w_k(f)| + |w_k(\xi) + w_k(e)| \le v_n + x_n = O(x_n)$, $k \ne k_2$ and $|w_{k_2}| \ge |w_k(f)| - |w_k(\xi) + w_k(e)| \ge \sqrt{n2^{-j}} - x_n = \sqrt{n2^{-j}}(1 + o(1))$. This shows that for n large enough,

$$\left|w_{(1)}\right| = \max_{k=0,\ldots,2^j-1} |w_k| = |w_{k_2}| \ge \sqrt{n2^{-j}}(1 + o(1)).$$

Since $\left|w_{(2)}\right| = |w_{k_3}|$ for some $k_3 \ne k_2$, we have

$$T_1 = \left|w_{(1)}\right| - \left|w_{(2)}\right| \ge \sqrt{n2^{-j}} - O(x_n) = \sqrt{n2^{-j}}(1 + o(1)),$$

thus $T_1 \longrightarrow \infty$, as $n \longrightarrow \infty$. Now we prove (10). Note that

$$P(B_n^c) = P(\max_{k=0,\ldots,2^j-1} |w_k(\xi) + w_k(e)| > x_n) \le \sum_{k=1}^{2^j} P\{|w_k(\xi) + w_k(e)| > x_n\}.$$

Recalling that $x_n = o(\sqrt{n2^{-j}})$, e_i are i.i.d. random variables, and applying Lemma 2 and Cramer Large Deviation Theorem[8], we have that

$$\begin{aligned}
\lim_{n\to\infty} P(B_n^c) &\le \lim_{n\to\infty} \sum_{k=1}^{2^j} P\{|w_k(\xi) + w_k(e)| > x_n\} \\
&\le \lim_{n\to\infty} \sum_{k=1}^{2^j} P\{|w_k(e)| > x_n\} = \lim_{n\to\infty} \sum_{k=1}^{2^j} P\{|N| > x_n\} \\
&= 2 \lim_{n\to\infty} \sum_{k=1}^{2^j} P\{N) > x_n\} \quad (\text{where } N \sim N(0,1))
\end{aligned}$$

$$\leq 2 \lim_{n\to\infty}(2^j \cdot \frac{1}{x_n} \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{n}) \leq C \lim_{n\to\infty} \frac{1}{(\log n)^{1/2+\delta}}$$
$$= 0.$$

□

In the next section, we will use statistics (7) to estimate the location and the number of the change points.

4. Estimation of change points

Theorem 4 *Under $H_1(m)$, denote by q and $0 < \theta_1 < \cdots < \theta_q < 1$ the number and locations of the change points of the function f, $1 \leq q \leq m$, respectively. Take $\hat{q} = \sup\{i : 1 \leq i \leq m, T_i > c_i\}$. Let $k_1 \leq k_2 \cdots \leq k_{\hat{q}}$ index the wavelet coefficients $|w_{(1)}| \geq |w_{(2)}| \geq \cdots \geq |w_{(\hat{q})}|$, i.e., $|w_{k_1}|, |w_{k_2}|, \ldots, |w_{k_{\hat{q}}}|$ are the unordered coefficients whose ordered sequences are $|w_{(1)}| \geq |w_{(2)}| \geq \cdots \geq |w_{(\hat{q})}|$. Let $(\hat{\theta}_1, \ldots, \hat{\theta}_{\hat{q}}) = (k_1/2^j, \ldots, k_{\hat{q}}/2^j)$. Then under the same assumptions as in Theorem 3,*

$$P(\hat{q} \to q) = 1, \text{ as } n \longrightarrow \infty, \tag{11}$$

$$\sum_{i=1}^{\hat{q}} (\hat{\theta}_i - \theta_i)^2 = O_p(\frac{1}{2^{2j}}). \tag{12}$$

Proof Using the arguments in the proof of Theorem 3, we can easily prove (11) and we omit it here. Let x_n and B_n be defined as in Theorem 3. Observe that

$$|w_k(f)| - |w_k(\xi) + w_k(e)| \leq |w_k| \leq |w_k(f)| + |w_k(\xi) + w_k(e)|. \tag{13}$$

Combing (13), Lemma 1, $x_n = o(\sqrt{n2^{-j}})$, $\sqrt{n2^{-3j}} = o(x_n)$ along with (10), we obtain that with probability tending to 1, the following two inequalities hold

$$\max_{k=0,\ldots,2^j}\{|w_k| : \exists\theta_i \in [\frac{k}{2^j}, \frac{k+1}{2^j})\} \geq \sqrt{n2^{-j}} - x_n = \sqrt{n2^{-j}}(1 + o(1)), \tag{14}$$

$$\max_{k=0,\ldots,2^j}\{|w_k| : \forall\theta_i \notin [\frac{k}{2^j}, \frac{k+1}{2^j})\} \leq \sqrt{n2^{-3j}} + x_n \leq x_n(1 + o(1)) \leq \sqrt{n2^{-j}}(1 + o(1)). \tag{15}$$

Thinking of $\max_k |w_k|$ as a function of k, by (14) and (15) we can easily see that, with probability tending to 1, the maximum of $|w_k|$ will be achieved at some k, where $\theta_i \in [\frac{k}{2^j}, \frac{k+1}{2^j})$, $i = 1, 2, \ldots, q$. By the definition of k_i, with probability tending to 1, $\theta_i \in [\frac{k_i}{2^j}, \frac{k_i+1}{2^j}), i = 1, 2, \ldots, \hat{q}$. That is, $\frac{\theta_i - \hat{\theta}_i}{1/2^j} \in [0, 1)$. Thus $\theta_i - \hat{\theta}_i = O_p(\frac{1}{2^j}), i = 1, 2, \ldots, \hat{q}$, which yields (12). □

Remark 2 We know from (6) that $\theta_i - \hat{\theta}_i = O_p(\frac{1}{n}(\log n)^\delta), i = 1, 2, \ldots, \hat{q}, 1 < \delta < 2$. Such rate of convergence is the same as Wang's[3], whose result is based on the assumptions of Gaussian noise and the data are obtained in a fixed design manner.

5. Numerical properties

5.1 Simulation study

Consider the following data generating process:

$$Y_i = f(X_i) + \sigma\varepsilon_i, i = 1, 2, \ldots, n, \tag{16}$$

where $X_i \sim U[0,1]$, $\sigma = 0.2$, $\varepsilon_i \sim N(0,1)$. We carry out the experiment in the following two cases:

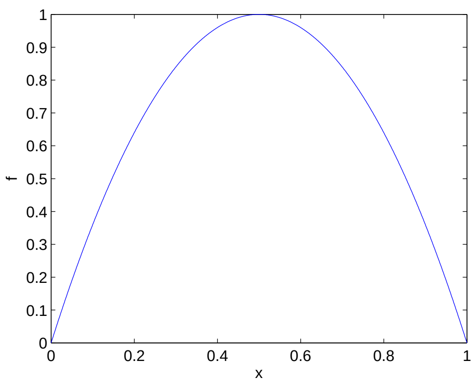
H_0: $f(x) = 4x(1-x)$ (f is continuously differentiable on [0,1]),

H_1: $g(x) = f(x) + 1_{\{x>0.25\}} - 1_{\{x>0.5\}}$ (there are two change points in function g).

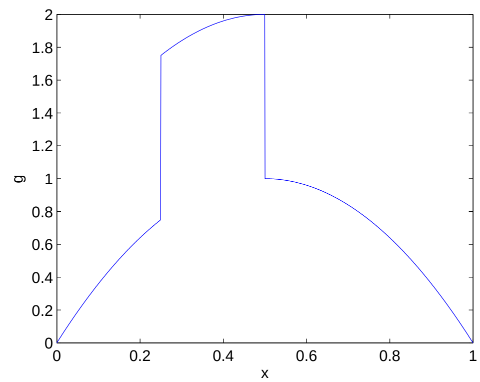
Figure 1(a) and (b) are the noise-free function of f and g, respectively. While Figure 1(c) and (d) are the illustrations of f and g which are observed with Gaussian noise $N(0, 0.2^2)$.

Take the sample size $n = 2^{10} = 1024$, the upper bound of the number of change points $m = 3$ and the significance level $\beta = 0.05$. At resolution level $j = 5$, the wavelet coefficient w_k is obtained by the Haar wavelet transformation. Figure 1(e) and (f) depict the absolute values of wavelet coefficients ($|w_k|$) under H_0 and H_1, respectively. The threshold line is drawn at $|w_{(4)}|$. The critical region is drawn according to (8).

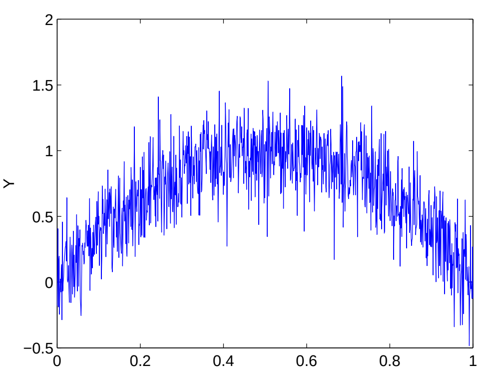
From Figure 1(e), we can find that the number of the wavelet coefficients that exceed the critical region line is 0. Thus we accept H_0. This implies that there is no change in function f. While from Figure 1(f) we can see that, the wavelet coefficients exceed the critical region line nearby 0.25 and 0.5 where the function g has change points. The simulation results show that our method is effective.



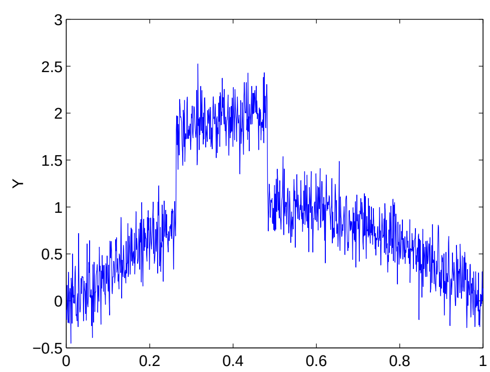
(a) Noise-free function under H_0



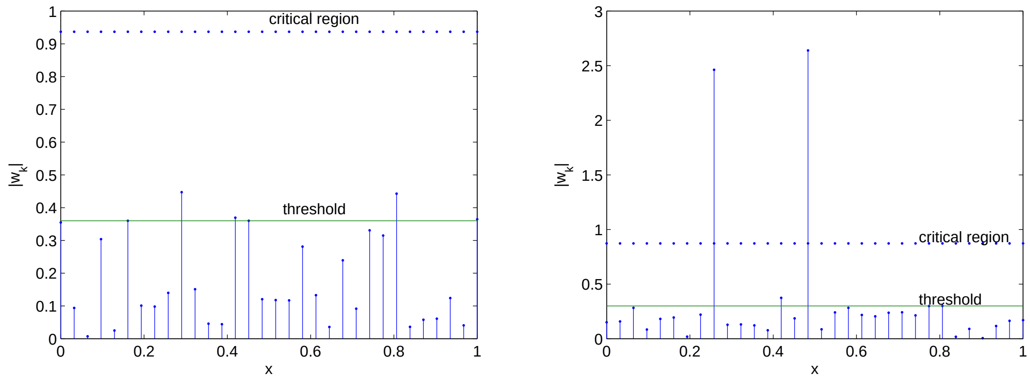
(b) Noise-free function under H_1



(c) Simulated data under H_0



(d) Simulated data under H_1

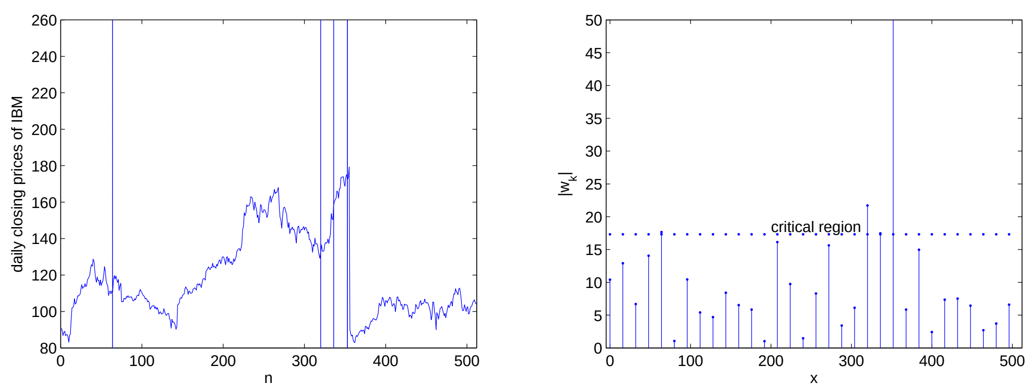


(e) Absolute wavelet coefficients under H_0 (f) Absolute wavelet coefficients under H_1

Figure 1 Simulated data

5.2 Application to IBM data

We applied our method to the IBM data in this section. Figure 2(a) shows the $2^9 = 512$ data points of IBM daily closing prices starting from January 2, 1996 until January 8, 1998. Figure 2(b) is the plot of the absolute wavelet coefficients which are obtained at $j = 4$, where the dashed horizontal line is the critical region which is drawn at significance level $\beta = 0.05$, $m = 6$. In Figure 2(b), the wavelet coefficients exceed the critical region line at four locations. The vertical line in Figure 2(a) shows these locations. The locations of these large wavelet coefficients are observations 64 (April 1, 1996), 320 (April 4, 1997), 336 (April 29, 1997) and 353 (May 22, 1997), so there are local structural changes near the corresponding times. The first change point in 1996 may be caused by Crisis in the Taiwan strait in 1996. While the next three change points in 1997 are caused by the Asian Economic Storms in 1997.



(a) IBM daily closing prices (b) Absolute wavelet coefficients at $j = 4$

Figure 2 IBM data

6. Conclusions

We study the detection and estimation problem of change points in nonparametric regression model under random design. The $100(1-\beta)/\%$ confidence bounds using the test statistics based on the wavelet coefficients is derived, meanwhile, the consistence of test is proved. The test statistics can also be used to estimate the locations and the number of the change points. The convergence rate of the change estimator is the same as Wang's[3]. The simulation study supports our result and the IBM real example shows that our method can be used to solve practical problems.

However, we must point out that, though we succeed in avoiding the assumptions of Gaussian noise and fixed random design, we still assume that the noises are i.i.d.. While in applications, the data may be obtained in dependent cases. At this moment, our method is invalid. This problem has not been solved yet. It will take more time to get it right.

References

[1] MÜLLER H. *Change-points in nanparametric regression analysis* [J]. Ann. Statist., 1992, **20**(2): 737–761.
[2] RAIMONDO M, TAJVIDI N. *A peaks over threshold model for change-point detection by wavelets* [J]. Statist. Sinica, 2004, **14**(2): 395–412.
[3] WANG Yazhen. *Jump and sharp cusp detection by wavelets* [J]. Biometrika, 1995, **82**(2): 385–397.
[4] ODGEN T, PARZEN O. *Change-point approach to data analytic thresholding* [J]. Statist. Comput., 1996, **6**: 93–99.
[5] RAIMONDO M. *Minimax estimation of sharp change points* [J]. Ann. Statist., 1998, **26**(4): 1379–1397.
[6] ANTONIADIS A, GIJBELS I. *Detecting abrupt changes by wavelet methods* [J]. J. Nonparametr. Stat., 2002, **14**(1-2): 7–29.
[7] HÄRDLE W, KERKYACHARIAN G, PICARD D. et al. *Wavelets, Approximation, and Statistical Applications* [M]. Springer-Verlag, New York, 1998.
[8] PETROV V. *Sums of Independent Random Variables* [M]. Springer-Verlag, New York-Heidelberg, 1975.