应用概率统计 第二十六卷
第二期 2010年4月

Chinese Journal of Applied Probability
and Statistics Vol.26 No.2 Apr. 2010

Wavelet Identification of Structural Change Points in Volatility Models for Time Series

Wang Jingle Liu Weiqi

(*School of Mathematical Sciences, Shanxi University, Taiyuan, 030006*)

Abstract

We propose two estimators, an integral estimator and a discretized estimator, for the wavelet coefficient of volatility in time series models. These estimators can be used to detect the changes of volatility in time series models. The location estimators of the jump points, we proposed, have been shown to have the minimax convergence rate, which is the optimal rate for the estimation of change points. The jump sizes and locations of change points can be consistently estimated by wavelet coefficients. The convergency rates of these estimators are derived and the asymptotic distributions of the statistics are established.

Keywords: Change points in volatility, wavelet coefficient, kernel estimation, local polynomial smoother.

AMS Subject Classification: 62N02.

§1. Introduction

The analysis of change points in nonparametric models has attracted increasing interests. Many attempts have been contribute to test and estimate the jumps and their sizes in volatility model for time series, because of its importance in hedging strategies and risk management. However, the continuous-time model has a severe limitation when jumps exist in underlying sample data. This paper develops a theory of estimating change points in the volatility of a nonparametric model. Wang (1995) first employed the wavelet method to detect jumps in a continuous-time model with a constant volatility. Wong, Ip, Li and Xie (1999) have shown that the wavelet coefficient has significantly large absolute values near the jump points across fine levels, while having relatively small values when the location shifts away from the jump points.

In this paper, we propose two estimators for the wavelet coefficient of volatility in time series models. The first one is an integral estimator. This estimator is very simple and intuitive in constructing the estimator for the wavelet coefficient. The second one is

Received September 24, 2008.

a discretized estimator for the wavelet coefficient. These estimators can be used to detect the change of volatility in time series models. The locations and sizes of change points of volatility are proposed and the convergency rates of these estimators are derived. Wong, Ip and Li (2001) have only proposed a simple empirical estimator of the wavelet coefficient, and have not given the asymptotic distribution of the estimator. In our paper, we proposed the asymptotic distributions of the statistic. In addition, the location estimators of the jump points, we proposed, have been shown to have the minimax convergence rate, which is the optimal rate for the estimation of change points, even if the observations are not a sequence of i.i.d. random variables.

The paper is organizes as follows. Section 2 introduce the nonparametric model and wavelet method. Section 3 provides the estimation of the wavelet coefficient. The estimations of jump sizes and change points are considered in Section 4. Section 5 discusses the asymptotic distribution of the estimators and the estimation of unknown variance. Proofs are collected in the Appendix.

§2. Model and Notations

A nonparametric model is defined as follows:

$$Y_i = T(X_i) + \sigma(X_i)\varepsilon_i, \tag{2.1}$$

where we assume that $\mathsf{E}(\varepsilon_i|X_i) = 0$ and $\mathsf{E}(\varepsilon_i^2|X_i) = 1$ and $\{\varepsilon_i, i = 1, 2, \cdots\}$ is a sequence of random variables. $T(x) = \mathsf{E}(Y|X = x)$, $\sigma^2(x) = \mathsf{Var}\,(Y|X = x)$. $\{(X_i, Y_i), i = 1, 2, \cdots\}$ is a sequence of random vectors which satisfying some mixing dependent conditions. Here we relax $\{(X_i, Y_i), i = 1, 2, \cdots\}$ to allow for dependent observations in a time series.

In fact, we can re-write the model (2.1) as follows:

$$(Y_i - T(X_i))^2 = \sigma^2(X_i) + \sigma^2(X_i)(\varepsilon_i^2 - 1), \tag{2.2}$$

where $\mathsf{E}((\varepsilon_i^2 - 1)|X_i) = 0$, and we assume that $\mathsf{E}((\varepsilon_i^2 - 1)^2|X_i) = v$.

In this paper, we assume that $\sigma^2(x)$ has an α-cusp at t_i, $i = 1, 2, \cdots, p$. $\sigma^2(x)$ is smooth except at those discontinuous points. When $\sigma^2(x)$ have p discontinuous points in $[a, b]$. We can re-write $\sigma^2(x)$ as follows:

$$\sigma^2(x) = C(x) + D(x),$$

where $D(x) = \sum\limits_{i=1}^{p} d_i I_{[t_i,b]}(x)$ with $a < t_1 < t_2 < \cdots < t_p < b$, and $C(x)$ is twice continuously differentiable on (a, b). Let $d_i = \sigma^2(t_i+) - \sigma^2(t_i-)$ denote the size of a jump of the function $\sigma^2(x)$ at point t_i. Our interest is to estimate p, d_i and t_i, $i = 1, 2, \cdots, p$.

To obtain the desired results, we make the following assumptions about the model. Let U be a open neighborhood of the origin of $\mathcal{R}$ and $[a,b]\subset U$.

(A1) The p.d.f. $f(x)$ of X_1 is bounded away from zero and infinity on some open subset U. That is, there exists some positive constant such that $M^{-1}\leqslant f(x)\leqslant M$, $x\in U$.

(A2) The conditional p.d.f. $f(y|x)$ of X_i, given $X_1=x$, is also bounded away from zero and infinity on U.

(A3) $\{X_i, i=1,2,\cdots\}$ is strictly stationary and α-mixing. Its mixing coefficient $\alpha(\mu)=o(\rho^{-\mu})$ for some large $\rho>0$.

(A4) We assume that $f(x)$ is a twice bounded derivative function, and $T(x)$ and $C(x)$ are continuous third order differentiable on U.

(A5) Let $\{(\varepsilon_i^2-1), i=1,2,\cdots\}$ be a sequence of i.i.d. random variables and for each i, ε_i^2-1 is independent of $\{(X_i, Y_{j-1}), j\leqslant i\}$.

Assumptions (A1) and (A2) are necessary for the kernel estimation with dependent data. (A3) is to simplify proofs. (A4) is to meet the continuity requirement for kernel smoothing. (A5) is also made for simplicity of proofs.

Before given the estimators of wavelet coefficient of the volatility, we need to propose some assumptions for wavelet and introduce some notations. In this paper, we mainly use the notations in Chen, Choi and Zhou (2008).

Assumption that $\{X_i, 1\leqslant i\leqslant n\}$ is a realization from model (2.1), Denote

$$I_n(x_0)=\{i: 1\leqslant i\leqslant n, |X_i-x_0|\leqslant\delta_n\},$$
$$I(s,\delta_n)=\Big\{k: \Big|a+\frac{k}{2^J}(b-a)-s\Big|\leqslant\delta_n\Big\},$$

where $\delta_n=2^{-J}$ and $k=[2^J\theta]$, $(0<\theta<1)$, $J=J(n)$ is often a sequence with $J\to\infty$ as $n\to\infty$. let $N_n(x_0)=\#I_n(x_0)$ denote the number of points in $I_n(x_0)$. $D_n=\{0,1,2,\cdots,2^J-1\}$.

We choose wavelet $\psi(x)$ and scale function $\phi(x)$ satisfying the following conditions.

(B1) Both $\psi(x)$ and $\phi(x)$ have finite supports, say, $[-A,A]$, $A>1$, and $\psi(x)=0$, $x\in[-1,1]$. And both have derivatives with bounded variation.

(B2) The wavelet function $\psi(x)$ has the following properties

$$\int_{-A}^{A}\psi(x)\mathrm{d}x=0,\quad \int_{-A}^{A}x\psi(x)\mathrm{d}x=0,\quad \int_{1}^{A}\psi(x)\mathrm{d}x\neq 0,\quad \int_{1}^{A}x\psi(x)\mathrm{d}x\neq 0,$$
$$0<\Big|\int_{y}^{A}\psi(x)\mathrm{d}x\Big|<\Big|\int_{1}^{A}\psi(x)\mathrm{d}x\Big|,\qquad 0<\Big|\int_{-A}^{-y}\psi(x)\mathrm{d}x\Big|<\Big|\int_{-A}^{-1}\psi(x)\mathrm{d}x\Big|,$$

where $1<y<A$.

From the wavelet $\psi(x)$ and scale function $\phi(x)$, we can obtain the orthogonal wavelet basis on $L^2[a,b]$.

$$\{\phi_{i,k}^{\rm per}(x), k\in I_i; \psi_{J,k}^{\rm per}(x), k\in I_J, J\geqslant i\},$$

where

$$\phi_{i,k}^{\rm per}(x)=\sum_{n=-\infty}^{\infty}\frac{1}{\sqrt{b-a}}\phi_{i,k}\Big(\frac{x-a}{b-a}+n\Big),\qquad \psi_{J,k}^{\rm per}(x)=\sum_{n=-\infty}^{\infty}\frac{1}{\sqrt{b-a}}\psi_{J,k}\Big(\frac{x-a}{b-a}+n\Big)$$

with $\phi_{i,k}(x)=2^{i/2}\phi(2^ix-k)$, $\psi_{J,k}(x)=2^{J/2}\psi(2^Jx-k)$ and $I_J=\{0,1,2,\cdots,2^J-1\}$.

Now, the wavelet coefficient of the volatility function $\sigma^2(X_i)$ of model (2.1) is defined as follow

$$\beta_{J,k}=\int_a^b\sigma^2(x)\psi_{J,k}^{\rm per}(x){\rm d}x. \tag{2.3}$$

§3. Estimation of the Wavelet Coefficient

Wong, Ip and Li (2001) proposed a simple empirical estimator of $\beta_{J,k}$, which is defined as

$$W_{J,k}=\frac{b-a}{N}\sum_{i=1}^{N}\psi_{J,k}^{\rm per}(w_i)\frac{1}{N_n(x)}\sum_{i\in I_n(w_i)}\Big(y_i-\frac{1}{N_n(x)}\sum_{m\in I_n(w_i)}y_m\Big)^2, \tag{3.1}$$

where $N\to\infty$, w_i are those points to divide the interval $[a,b]$ into $N+1$ sub-intervals, that is $w_i=a+i(b-a)/N$.

Let $K(x)$ be a probability density function with bounded support $[-C,C]$, for some constant $C>0$. When $\sigma^2(x)$ is smooth, we have the following estimator:

$$\sigma_n^2(x)=\frac{\sum\limits_{i=1}^{n}\mathcal{K}_{n,h}(X_i-x)(Y_i-T(X_i))^2}{\sum\limits_{i=1}^{n}\mathcal{K}_{n,h}(X_i-x)},$$

where $h=h_n$ is a sequence of bandwidths, with $h\to 0$ and $nh\to\infty$ as $n\to\infty$. When

$$\mathcal{K}_{n,h}(X_i-x)=K_h(X_i-x)=K\Big(\frac{(X_i-x)}{h}\Big),$$

$\sigma_n^2(x)$ is the kernel estimation (see Nadaraya (1964)) and when

$$\begin{aligned}\mathcal{K}_{n,h}(X_i-x) \;&=\; K_h(X_i-x)\sum_{j=1}^{n}K_h(X_j-x)(X_j-x)^2\\ &\quad -K_h(X_i-x)(X_i-x)\sum_{j=1}^{n}K_h(X_j-x)(X_j-x),\end{aligned}$$

$\sigma_n^2(x)$ is the local linear smoothers(see Fan and Gijbels (1996)). We only consider the kernel estimator.

Now, we can propose two estimators of wavelet coefficient of the volatility in model (2.1), as follows

$$U_J(k) = \int_a^b \psi_{J,k}^{\text{per}}(x) \frac{\sum\limits_{i=1}^n \mathcal{K}_{n,h}(X_i - x)(Y_i - T(X_i))^2}{\sum\limits_{i=1}^n \mathcal{K}_{n,h}(X_i - x)} \mathrm{d}x, \tag{3.2}$$

$$W_J(k) = \frac{b-a}{N} \sum_{i=1}^N \psi_{J,k}^{\text{per}}(w_i) \frac{\sum\limits_{j=1}^n \mathcal{K}_{n,h}(X_j - w_i)(Y_j - T(X_j))^2}{\sum\limits_{j=1}^n \mathcal{K}_{n,h}(X_j - w_i)}, \tag{3.3}$$

where N and w_i are the same as those in (3.1).

The first estimator is integral estimator. This estimator is simple and intuitive. The second one is descretized estimator. This one is simple in computation. We can obtain the estimator (3.1) from (3.3) by choosing the kernel function

$$K(x) = \begin{cases} \dfrac{1}{2}, & \|x\| \leqslant 1; \\ 0, & \|x\| > 1. \end{cases}$$

The estimator (3.1) has some drawbacks, because the bandwidth h selected in this estimator cannot reach the optimal value, that is, $h_{\text{opt}} = Cn^{-1/5}$ for some constant C. The estimator (3.1) has a larger mean integration square error than the estimators (3.2) and (3.3), which have the optimal bandwidths.

To obtain the properties of estimators of change points in volatility, we propose some assumptions as follows.

(C1) $\lim\limits_{n\to\infty} 2^{2J}(\log n)^3/n = 0$, $\lim\limits_{n\to\infty} 2^{5J}/n = \infty$, $\lim\limits_{n\to\infty} 2^J h^2 \log n = 0$.

(C2) $\lim\limits_{n\to\infty} n2^J/(Nh)^2 = 0$.

(C3) $\lim\limits_{n\to\infty} 2^J/n = 0$, $\lim\limits_{n\to\infty} 2^{3J}/n = 0$.

Theorem 3.1 Assume (A1)-(A5) are true, Let t_i, $i = 1, 2, \cdots, p$ be p jump points of $\sigma^2(x)$, and the corresponding jump sizes be denoted d_i, $i = 1, 2, \cdots, p$.

(a) If (C1) is satisfied, then for all $k \in I(t_i, 2^{-J}(b-a))$, we have

$$U_J(k) = 2^{-J/2}(b-a)^{1/2} d_i \int_1^A \psi(x)\mathrm{d}x + O_p(n^{-1/2}), \tag{3.4}$$

where $a_n = O_p(b_n)$ denotes $\lim\limits_{n\to\infty} a_n/b_n = C$ in probability for some constant C and for $k \notin \bigcup\limits_{i=1}^p I(t_i, 2^{-J}(b-a))$, we have

$$U_J(k) = O_p(n^{-1/2}). \tag{3.5}$$

(b) If (C1) and (C2) are satisfied, the (3.4) and (3.5) hold for the discretized estimator $W_J(k)$ of the wavelet coefficient.

§4. Estimation of Jump Size and Change Points

Suppose $\sigma^2(x)$ has p jump points, where p is a finite integer. We assume that $\sigma^2(x)$ is differentiable at all points except for these jump points. Without loss of generality, that $|d_{i+1}| < |d_i|$, $i = 1, 2, \cdots, p-1$. The estimators of t_i are denoted as t_i^U or t_i^W and the estimators of d_i are denoted as d_i^U or d_i^W. The estimators t_i and d_i are constructed as follows:

(1) Find the first change point t_1: $t_1^U = a + k_1(b-a)/2^J$, where $k_1 = \arg\max\limits_{k} |U_J(k)|$.

(2) Find the second change point t_2: $t_2^U = a + k_2(b-a)/2^J$, where $k_2 = \arg\max\limits_{k\in Q_2} |U_J(k)|$, $Q_2 = [a, b] - I(t_1, 2^{-J}A(b-a))$.

(3) Continue the procedure until the location t_p: $t_p^U = a + k_p(b-a)/2^J$, where $k_p = \arg\max\limits_{k\in Q_p} |U_J(k)|$, $Q_p = [a, b] - \bigcup\limits_{i=1}^{p-1} I(t_i, 2^{-J}A(b-a))$.

Making the procedure we can obtain all jump points of $\sigma^2(x)$ by using the integral estimator $U_J(k)$. These procedures also hold for the discretized estimator $W_J(k)$. The estimators of jump sizes of change points t_i, $i = 1, 2, \cdots, p$ can be defined as follows:

$$d_i^U = \frac{2^{J/2} U_J(k_i)}{(b-a)^{1/2} \displaystyle\int_1^A \psi(x) \mathrm{d}x}, \qquad d_i^W = \frac{2^{J/2} W_J(k_i)}{(b-a)^{1/2} \displaystyle\int_1^A \psi(x) \mathrm{d}x}.$$

The following theorems establish convergence rates for above methods.

Theorem 4.1 Assume (A1)-(A5) are true.

(a) If (C3) is satisfied,

$$|d_i^U - d_i| = O_p((2^{-J}n)^{-1/2}), \qquad i = 1, 2, \cdots, p.$$

(b) If (C2) and (C3) is satisfied,

$$|d_i^W - d_i| = O_p((2^{-J}n)^{-1/2}), \qquad i = 1, 2, \cdots, p.$$

Theorem 4.2 Assume (A1)-(A5) are true.

(a) If (C3) is satisfied,

$$|t_i^U - t_i| = O_p(2^{-J}), \qquad i = 1, 2, \cdots, p.$$

(b) If (C2) and (C3) is satisfied,

$$|t_i^W - t_i| = O_p(2^{-J}), \qquad i = 1, 2, \cdots, p.$$

§5. Detection and Unknown Variance

In the former chapter, we assume p is known. Now, we consider the case that p is unknown. We define some elements as follows

$$M_n(x)=\frac{f^{0.5}(a+x(b-a))U_J([2^Jx])}{\sigma^2(a+x(b-a))},\qquad M_n^*(x)=\frac{f^{0.5}(a+x(b-a))W_J([2^Jx])}{\sigma^2(a+x(b-a))},$$

where $0\leqslant x\leqslant 1$. The following results play a role in detecting the change points of $\sigma^2(x)$.

Theorem 5.1 Assume (A1)-(A5) are true.

(a) If (C1) is satisfied. When there is no change point in $\sigma^2(x)$,

$$\mathsf{P}\Big\{A(\delta_n)\Big(\frac{n}{k_2v}\Big)^{1/2}\sup_{0\leqslant x\leqslant 1}|M_n(x)|-a(\delta_n)<z\Big\}\longrightarrow\exp(-2\exp(-z)).$$

(b) If (C1) and (C2) is satisfied. When there is no change point in $\sigma^2(x)$,

$$\mathsf{P}\Big\{A(\delta_n)\Big(\frac{n}{k_2v}\Big)^{1/2}\sup_{0\leqslant x\leqslant 1}|M_n^*(x)|-a(\delta_n)<z\Big\}\longrightarrow\exp(-2\exp(-z)),$$

where

$$A(x)=|2\log x|^{1/2},\qquad a(x)=|2\log x|^{1/2}+|2\log x|^{-1/2}\log\Big(\frac{k_1^{0.5}}{2\pi k_2^{0.5}}\Big),$$

$$k_1=\int_{-A}^{A}(\psi'(x))^2\mathrm{d}x,\qquad k_2=\int_{-A}^{A}\psi^2(x)\mathrm{d}x,\qquad v=\mathsf{E}((\varepsilon_i^2-1)^2|X_i).$$

Now, Let us estimate $T(x)$ and $f(x)$ by a nonparametric technique. As we known, we can estimate the density function by nonparametric kernel estimate. The kernel estimator of $f(x)$ is defined as follow

$$f_n(x)=\frac{1}{nh}\sum_{i=1}^{n}K\Big(\frac{X_i-x}{h}\Big).$$

Similar to $\sigma^2(x)$, the estimate of $T(x)$ can be estimated by kernel estimator or local linear estimator. We can write as follow

$$T_n(x)=\frac{\sum\limits_{i=1}^{n}\mathcal{K}_{n,h}(X_i-x)Y_i}{\sum\limits_{i=1}^{n}\mathcal{K}_{n,h}(X_i-x)}.$$

From Theorem 5.1, we can obtain the following corollary.

Corollary 5.1 If we use $f_n(x)$ instead of $f(x)$ and $T_n(x)$ instead of $T(x)$ in $M_n(x)$ and $M_n^*(x)$, the consequences of Theorem 5.1 hold for $M_n(x)$ and $M_n^*(x)$, respectively.

Appendix

We denote

$$W_j(x)=\frac{K_h(X_j-x)}{\sum\limits_{j=1}^{n}K_h(X_j-x)},\qquad K_h(X_j-x)=K\Big(\frac{X_j-x}{h}\Big)$$

in the following proofs. For simplicity, without loss of generality, let

$$\phi_{i,k}^{\rm per}(x)=\frac{1}{\sqrt{b-a}}\phi_{i,k}\Big(\frac{x-a}{b-a}\Big),\qquad \psi_{J,k}^{\rm per}(x)=\frac{1}{\sqrt{b-a}}\psi_{J,k}\Big(\frac{x-a}{b-a}\Big).$$

Lemma 1 Assume that (A1)-(A3) are true. $K(x)$ is a continuously differentiable kernel function with finite support $[-C,C]$, and $\int K(x){\rm d}x=0$, $\int xK(x){\rm d}x=0$. Let $h\to 0$, $nh\to\infty$ as $n\to\infty$, then

(a) for any positive integer i, we have

$$\sum_{t=1}^{n}K_h(X_t-x)(X_t-x)^i=nh^{i+1}q_if(x)+nh^{i+2}q_{i+1}f'(x)+O(nh^{i+1}c_n)\qquad \text{a.s.}$$

uniformly for $x\in[a,b]$, where

$$K_h(x)=K\Big(\frac{x}{h}\Big),\qquad c_n=h^2+\Big(\frac{\log n}{nh}\Big)^{1/2},\qquad q_i=\int x^iK(x){\rm d}x.$$

(b) assume that $\{(\varepsilon_t^2-1),t=1,2,\cdots\}$ satisfies (A1)-(A5), then

$$\sum_{t=1}^{n}K_h(X_t-x)(\varepsilon_t^2-1)=O_p((nh)^{1/2}),$$
$$\sum_{t=1}^{n}K_h(X_t-x)\Big(\frac{X_t-x}{h}\Big)(\varepsilon_t^2-1)=O_p((nh)^{1/2})$$

uniformly for $x\in[a,b]$.

Lemma 2 (a) Assume that $\psi(x)$ satisfies the assumptions (B1)-(B3), and that $C(x)$ is a continuously differentiable function in the order of two. Denote

$$W_j(x)=\frac{K_h(X_j-x)}{\sum\limits_{j=1}^{n}K_h(X_j-x)},\qquad K_h(X_j-x)=K\Big(\frac{X_j-x}{h}\Big).$$

Then, uniformly for $k\in D_n$,

$$\int_a^b\psi_{J,k}^{\rm per}(x)\sum_{j=1}^{n}W_j(x)[C(X_j)-C(x)]{\rm d}x=O_p(2^{-J/2}hc_n),$$
$$\int_a^b\psi_{J,k}^{\rm per}(x)C(x){\rm d}x=O(2^{-5J/2}),$$

where

$$c_n = h^2 + \Big(\frac{\log n}{nh}\Big)^{1/2}.$$

(b) Furthermore, we have uniformly for $k \in D_n$

$$\frac{b-a}{N}\sum_{i=1}^{N}\psi_{J,k}^{\mathrm{per}}(w_i)\sum_{j=1}^{n}W_j(w_i)[C(X_j)-C(w_i)] = O_p\Big(\frac{2^{J/2}}{N}\Big) + O_p(2^J h c_n),$$

$$\frac{b-a}{N}\sum_{i=1}^{N}\psi_{J,k}^{\mathrm{per}}(w_i)C(w_i) = O\Big(\frac{2^{J/2}}{N}\Big) + O(2^{-5J/2}).$$

Lemma 3 (a) Assume that $K(x)$ is a kernel function with finite support $[-C, C]$, and $h \to 0$ as $n \to \infty$. Denote

$$W_j(x) = \frac{K_h(X_j - x)}{\sum\limits_{j=1}^{n} K_h(X_j - x)}, \qquad K_h(X_j - x) = K\Big(\frac{X_j - x}{h}\Big).$$

We have

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x)\sum_{j=1}^{n}W_j(x)\Big[\sum_{i=1}^{p} d_i I(t_i \leqslant X_j \leqslant b)\Big]\mathrm{d}x = 2^{-J/2}(b-a)^{1/2} d_i \int_1^A \psi(x)\mathrm{d}x$$

uniformly for $k \in I(t_i, 2^{-J}(b-a))$, and

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x)\sum_{j=1}^{n}W_j(x)\Big[\sum_{i=1}^{p} d_i I(t_i \leqslant X_j \leqslant b)\Big]\mathrm{d}x = 0$$

uniformly for $k \notin \bigcup\limits_{i=1}^{p} D(A)$. Where

$$D(A) = \Big\{k : \Big|a + \frac{k(b-a)}{2^J} - t_i\Big| < 2^{-J}A(b-a)\Big\}$$

in which A is the support point.

(b) Furthermore, we have

$$\begin{aligned}&\frac{b-a}{N}\sum_{i=1}^{N}\psi_{J,k}^{\mathrm{per}}(w_i)\sum_{j=1}^{n}W_j(w_i)\Big[\sum_{i=1}^{p} d_i I(t_i \leqslant X_j \leqslant b)\Big]\\ =\ & 2^{-J/2}(b-a)^{1/2} d_i \int_1^A \psi(x)\mathrm{d}x + O_p\Big(\frac{2^{J/2}}{N}\Big)\end{aligned}$$

uniformly for $k \in I(t_i, 2^{-J}(b-a))$, and

$$\frac{b-a}{N}\sum_{i=1}^{N}\psi_{J,k}^{\mathrm{per}}(w_i)\sum_{j=1}^{n}W_j(w_i)\Big[\sum_{i=1}^{p} d_i I(t_i \leqslant X_j \leqslant b)\Big] = O_p\Big(\frac{2^{J/2}}{N}\Big)$$

uniformly for $k \notin \bigcup_{i=1}^{p} D(A)$.

The proofs of Lemma 1, Lemma 2 and Lemma 3 are to see the Lemma B.2, Lemma A.2 and Lemma A.3 of Chen, Choi and Zhou (2008).

Lemma 4 Suppose that $K(x)$ is a kernel function with finite support $[-C, C]$ and $h \to 0$ as $n \to \infty$. Denote

$$W_j(x) = \frac{K_h(X_j - x)}{\sum\limits_{j=1}^{n} K_h(X_j - x)}, \qquad K_h(X_j - x) = K\Big(\frac{X_j - x}{h}\Big).$$

Then

$$\sum_{j=1}^{n} W_j I(t_i \leqslant X_j \leqslant b)(\varepsilon_j^2 - 1) = \sum_{j=1}^{n} W_j I(t_i \leqslant x \leqslant b)(\varepsilon_j^2 - 1).$$

Proof We can easily obtain that $I(t_i \leqslant X_j \leqslant b) - I(t_i \leqslant x \leqslant b) = I(t_i \leqslant X_j \leqslant b, x < t_i) + I(t_i \leqslant X_j \leqslant b, x > b) + I(t_i \leqslant x \leqslant b, X_j < t_i) + I(t_i \leqslant x \leqslant b, X_j > b)$. It easy to show that

$$\sum_{j=1}^{n} W_j I(t_i \leqslant X_j \leqslant b, x < t_i) = \sum_{j=1}^{n} W_j I\Big(\frac{t_i - x}{h} \leqslant Y_j \leqslant \frac{b - x}{h}, x < t_i\Big),$$

where $Y_j = (X_j - x)/h$. As $|Y_j| > C$, $K(Y_j) = 0$ and for large enough n, $x < t_i$ implies $(t_i - x)/h \to \infty$. So that, for large n, we have

$$\sum_{j=1}^{n} K(Y_j) I\Big(\frac{t_i - x}{h} \leqslant Y_j \leqslant \frac{b - x}{h}, x < t_i\Big) = 0,$$
$$\sum_{j=1}^{n} K(Y_j) I\Big(\frac{t_i - x}{h} \leqslant Y_j \leqslant \frac{b - x}{h}, x < t_i\Big)\varepsilon_j^2 = 0.$$

Hence,

$$\sum_{j=1}^{n} W_j I(t_i \leqslant X_j \leqslant b, x < t_i)(\varepsilon_j^2 - 1) = 0.$$

Similarly, we can show that

$$\sum_{j=1}^{n} W_j I(t_i \leqslant X_j \leqslant b, x > b)(\varepsilon_j^2 - 1) = 0,$$
$$\sum_{j=1}^{n} W_j I(t_i \leqslant x \leqslant b, X_j < t_i)(\varepsilon_j^2 - 1) = 0,$$
$$\sum_{j=1}^{n} W_j I(t_i \leqslant x \leqslant b, X_j > b)(\varepsilon_j^2 - 1) = 0.$$

This completes the proof of Lemma 4. □

Lemma 5 (a) Denote

$$W_j(x) = \frac{K_h(X_j - x)}{\sum\limits_{j=1}^{n} K_h(X_j - x)}, \qquad K_h(X_j - x) = K\Big(\frac{X_j - x}{h}\Big).$$

Assume that (A1)-(A5) are true, then we have

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x) \sum_{j=1}^{n} W_j(x)\sigma^2(X_j)(\varepsilon_j^2 - 1)\mathrm{d}x = O_p(n^{-1/2}).$$

(b) Furthermore, the (C2) is satisfied, then we have

$$\frac{b-a}{N} \sum_{i=1}^{N} \psi_{J,k}^{\mathrm{per}}(w_i) \sum_{j=1}^{n} W_j(w_i)\sigma^2(X_j)(\varepsilon_j^2 - 1) = O_p(n^{-1/2}).$$

Proof From the decomposition of $\sigma^2(x)$, we have

$$\begin{aligned} &\int_a^b \psi_{J,k}^{\mathrm{per}}(x) \sum_{j=1}^{n} W_j(x)\sigma^2(X_j)(\varepsilon_j^2 - 1)\mathrm{d}x \\ =\ &\int_a^b \psi_{J,k}^{\mathrm{per}}(x) \sum_{j=1}^{n} W_j(x)C(X_j)(\varepsilon_j^2 - 1)\mathrm{d}x + \int_a^b \psi_{J,k}^{\mathrm{per}}(x) \sum_{j=1}^{n} W_j(x)D(X_j)(\varepsilon_j^2 - 1)\mathrm{d}x. \end{aligned}$$

By Taylor's expansion, we obtain $C(X_j) = C(x) + C'(x)(X_j - x) + (1/2)C'(\xi_j)(X_j - x)^2$, where ξ_j lies between X_j and x. From Lemma 1, we have

$$\sup_{x\in\Lambda} \Big| \sum_{j=1}^{n} K_h(X_j - x)(X_j - x)^i(\varepsilon_j^2 - 1) \Big| = O_p((nh)^{1/2}h^i), \qquad i = 0, 1, 2,$$

where $\Lambda = [a - \delta_0, b + \delta_0]$ for some $\delta_0 > 0$. Hence

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x) \sum_{j=1}^{n} W_j(x)C(X_j)(\varepsilon_j^2 - 1)\mathrm{d}x = O_p(n^{-1/2}).$$

From Lemma 4. We have

$$\begin{aligned} &\int_a^b \psi_{J,k}^{\mathrm{per}}(x) \sum_{j=1}^{n} W_j(x)D(X_j)(\varepsilon_j^2 - 1)\mathrm{d}x \\ =\ &\sum_{i=1}^{p} d_i \int_a^b \psi_{J,k}^{\mathrm{per}}(x)I(t_i \leqslant x \leqslant b) \sum_{j=1}^{n} W_j(x)(\varepsilon_j^2 - 1)\mathrm{d}x = O_p(n^{-1/2}). \end{aligned}$$

This implies that Lemma 5 holds for (a). Similarly, we can prove (b) of the Lemma 5. □

Proof of Theorem 3.1 Note that $U_J(k)$ can be decomposed into two parts

$$U_J(k) = U_J^C(k) + U_J^D(k),$$

where

$$U_J^C(k)=\int_a^b \psi_{J,k}^{\rm per}(x)\sum_{j=1}^n W_j(x)\sigma^2(X_j){\rm d}x,$$
$$U_J^D(k)=\int_a^b \psi_{J,k}^{\rm per}(x)\sum_{j=1}^n W_j(x)\sigma^2(X_j)(\varepsilon_j^2-1){\rm d}x.$$

From Lemma 2 and Lemma 3, we have

$$U_J^C(k)=2^{-J/2}(b-a)^{1/2}d_i\int_1^A\psi(x){\rm d}x+O_p(2^{-5J/2}+2^{-J/2}hc_n)$$

uniformly for $k\in I(t_i,2^{-J}A(b-a))$, and $U_J^C(k)=O_p(2^{-5J/2}+2^{-J/2}hc_n)$ for $k\notin\bigcup\limits_{i=1}^p I(t_i,$ $2^{-J}A(b-a))$. From Lemma 5, it follows that

$$U_J^D(k)=O_p(n^{-1/2})$$

for all $k\in D_n$. This implies that Theorem 3.1 hold for the integral estimation of the wavelet coefficient. Similarly, we can prove Theorem 3.1 for the discretized estimation of the wavelet coefficient. □

Proof of Theorem 4.1 The proof is straightforward from Theorem 3.1. □

Proof of Theorem 4.2 Assume that $\sigma^2(x)$ has only a jump point t_i. Note that $U_J(k)$ can be decomposed into two parts

$$U_J(k)=U_J^C(k)+U_J^D(k),$$

where $U_J^C(k)$ and $U_J^D(k)$ are the same as those in the proof of Theorem 3.1. It follows from the similar arguments in Lemma 2 and 3

$$|U_J^C(k)|\leqslant C2^{-3J/2}$$

for all $k\notin I(t_i,2^{-J}A(b-a))$, where C is a generic constant whose value may change from line to line. By Lemma 3 we have

$$|U_J^C(k)|\geqslant C2^{-J/2}$$

for $k\in I(t_i,2^{-J}A(b-a))$. By Lemma 5, we have

$$U_J^D(k)=O_p(n^{-1/2}).$$

From Assumption (C3), we have $\lim\limits_{n\to\infty}2^J/n=0$, $\lim\limits_{n\to\infty}2^{3J}/n=0$.

Hence,

$$\max\{|U_J(k)|, k \in D_n\} = \max\{|U_J(k)|, k \in I(t_i, 2^{-J}A(b-a))\}.$$

Hence,

$$|t_i^U - t_i| = \left|a + \frac{k_i}{2^J}(b-a) - t_i\right| < 2^{-J}A(b-a).$$

When the number of change points of $\sigma^2(x)$ is p, we can similarly prove above formula. Hence, Theorem 4.2 holds for (a). Similarly, we can prove (b). This completes the proof of Theorem 4.2. □

Proof of Theorem 5.1 The proof is straightforward from Corollary 2.1 in Chen, Choi and Zhou (2008). □

References

[1] Wang, Y., Jump and sharp cusp detection by wavelets, *Biometrika*, **82**(1995), 385–397.

[2] Wong, H., Ip, W., Li, Y. and Xie, Z., Threshold variable selection by wavelets in open-loop threshold autoregressive models, *Statistics and Probability Letters*, **42**(1999), 375–392.

[3] Wong, H., Ip, W. and Li, Y., Detection of jumps by wavelets in a heteroscedastic autoregressive model, *Statistics and Probability Letters*, **52**(2001), 365–372.

[4] Chen, G., Choi, Y. and Zhou, Y., Detections of changes in return by a wavelet smoother with conditional heteroscedastic volatility, *Journal of Econometrics*, **143**(2008), 227–262.

[5] Nadaraya, E.A., On estimating regression, *Theory of Probability and Its Applications*, **9**(1964), 141–142.

[6] Fan, J. and Gijbels, I., *Local Polynomial Modeling and Its Applications*, Chapman and Hall, London, 1996.

时间序列中方差的结构变点的小波识别

王景乐 刘维奇

(山西大学数学科学学院, 太原, 030006)

本文给出了时间序列中方差的小波系数的两种估计: 连续估计和离散估计. 这两种估计可以用来检测时间序列中方差的结构变点. 利用这两种估计我们给出了方差变点的位置和跳跃幅度的估计, 并且显示出这些估计可达到最佳收敛速度. 同时, 我们还给出了这些估计的收敛速度以及检验统计量的渐进分布!

关键词: 方差变点, 小波系数, 核估计, 局部线形估计.

学科分类号: O212.7.