Ann. Inst. Statist. Math.
Vol. 55, No. 3, 467–485 (2003)


# ON THE CUSUM OF SQUARES TEST FOR VARIANCE CHANGE IN NONSTATIONARY AND NONPARAMETRIC TIME SERIES MODELS

SANGYEOL LEE[1], OKYOUNG NA[1] AND SEONGRYONG NA[2]

[1] *Department of Statistics, Seoul National University, Seoul 151-742, Korea*
[2] *Department of Information and Statistics, Yonsei University, Won-Ju, Gangwon-Do, 220-710, Korea*

(Received September 17, 2001; revised September 20, 2002)

**Abstract.** In this paper we consider the problem of testing for a variance change in nonstationary and nonparametric time series models. The models under consideration are the unstable AR($q$) model and the fixed design nonparametric regression model with a strong mixing error process. In order to perform a test, we employ the cusum of squares test introduced by Inclán and Tiao (1994, *J. Amer. Statist. Assoc.*, **89**, 913–923). It is shown that the limiting distribution of the test statistic is the sup of a standard Brownian bridge as seen in iid random samples. Simulation results are provided for illustration.

*Key words and phrases*: Cusum of squares test, variance change, autoregressive model with unit roots, nonparametric regression model, strong mixing process, weak convergence, Brownian bridge.

## 1. Introduction

The problem of testing for a parameter change has attracted much attention from many researchers since the parameter change in the underlying model is occasionally observed in actual practice. Since the paper of Page (1955), a vast amount of relevant articles have appeared in the literature; for example, see Hinkley (1971), Brown *et al.* (1975), Zacks (1983), Csörgő and Horváth (1988, 1997), Krishnaiah and Miao (1988), Wichern *et al.* (1976), Picard (1985), and the articles cited therein. The problem of testing for a variance change has became an important issue in time series analysis since the variance is often interpreted as a risk in econometrics. Inclán and Tiao (1994) considered the cusum of squares test for testing for a variance change. Their method has abundant merits since it is essentially a nonparametric test (distribution free), applicable to detecting multiple change points, and easy to understand and implement under a variety of circumstances; for instance, their test has been extended to GARCH $(1, 1)$ models (cf. Kim *et al.* (2000)) and linear processes (cf. Lee and Park (2001)).

In linear processes a variance change in the observations implies a change in one of the errors and the converse is also true. Thus a test for a variance change can be performed based on the errors rather than the observations themselves. Furthermore, the test based on the errors outperforms the one based on observations since the latter is subject to serious power losses when the data is highly correlated. Thus, if the time series under investigation is stationary and invertible (see Brockwell and Davis (1991), for the definition), then the former is naturally preferred (cf. Park *et al.* (2000)). In fact, the ease of application of the cusum of squares test lies in the fact that the limiting distribution of the test statistic is the sup of a standard Brownian bridge. It has been shown by Lee and Park (2001) that this result holds for stationary processes, but so far no attempt has been made to investigate its extension to nonstationary processes. The issue is intrinsically interesting and the result, if it turned out to be true, would merit special attention. Motivated by this, we considered the variance change problem in unstable processes (cf. Chan and Wei (1988)).

In this paper, we also deal with a nonparametric time series model taking into consideration its practical importance. The nonparametric approach in time series analysis has been advocated by many authors due to its flexibility and robust features when no parametric models are easy to apply to data (see, for example, Truong and Stone (1992), Neumann and Kreiss (1998) and Hafner (1998)). In fact, the nonparametric time series approach has been well appreciated by practitioners as a preliminary search method aimed at establishing a final parametric model. Needless to say, the task of correct modeling requires an analyst to be informed of the possibility of a variance change when she/he speculates as to its presence in given data set. Here we particularly concentrate on the variance change problem in a nonparametric regression model with a strong mixing error process.

The organization of this paper is as follows. In Section 2, we deal with the variance change test for the errors in AR($q$) models, which cover both stationary and nonstationary models. In view of the result of Lee and Wei (1999), which shows that the residual empirical process from the AR($q$) model with unit root 1 has a non-Gaussian process as its limiting process, one would likely guess that the same phenomenon might occur in this case. However, on the contrary, the Brownian bridge result is shown to remain the same as in Inclán and Tiao, and the cusum of squares test is still valid in this case. In light of this result, we discuss a goodness of fit test using the empirical process based on the squares of residuals. It is shown that the empirical process in this case converges weakly to a standard Brownian bridge as long as the error distribution has a symmetric density, which is immediately applicable to a Gaussian test.

In Section 3, we consider the variance change problem in a fixed design nonparametric regression model whose error process is geometrically strong mixing. We show that under regularity conditions the cusum of squares test statistic behaves asymptotically the same as with iid random variables.

Finally in Section 4, we report simulation results for our cusum tests introduced in Sections 2 and 3.

## 2. Test in AR($q$) model

In this section we consider the problem of testing for a variance change in the unstable AR($q$) model:

$$X_t - \beta_1 X_{t-1} - \cdots - \beta_q X_{t-q} = \epsilon_t, \tag{2.1}$$

where $\epsilon_t$ are iid random variables with $E\epsilon_1 = 0$, $E\epsilon_1^2 = \sigma^2$ and $E\epsilon_1^4 < \infty$. We assume that the corresponding characteristic polynomial $\phi$ has a decomposition

$$\begin{aligned}\phi(z) &= 1 - \beta_1 z - \cdots - \beta_q z^q \\ &= (1-z)^a (1+z)^b \prod_{k=1}^{l} (1 - 2\cos\theta_k z + z^2)^{d_k} \psi(z),\end{aligned}$$

where $a, b, l, d_k$ are nonnegative integers, $\theta_k$ belongs to $(0, \pi)$ and $\psi(z)$ is the polynomial of order $r = q - (a + b + 2d_1 + \cdots + 2d_l)$ that has no zeros on the unit disk in the complex plane.

Let $\boldsymbol{X}_t = (X_t, \ldots, X_{t-q+1})'$, where $X_t = 0$ for all $t \leq 0$. Let

$$\hat{\boldsymbol{\beta}}_n = \left( \sum_{t=1}^{n} \boldsymbol{X}_{t-1} \boldsymbol{X}'_{t-1} \right)^{-1} \sum_{t=1}^{n} \boldsymbol{X}_{t-1} X_t, \qquad n > q,$$

be the least squares estimate of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)'$ based on $X_1, \ldots, X_n$. Then the residuals are

$$\hat{\epsilon}_t = X_t - \hat{\boldsymbol{\beta}}'_n \boldsymbol{X}_{t-1}, \qquad t = 1, \ldots, n.$$

As mentioned earlier, our goal is to test the following hypotheses:

$H_0$: the $\epsilon_t$ have the same variance $\sigma^2$ vs.

$H_1$: not $H_0$.

In order to perform a test, we employ the cusum of squares test statistic $T_n$ based on the residuals:

$$T_n = \frac{1}{\sqrt{n}\hat{\kappa}_n} \max_{1 \leq k \leq n} \left| \sum_{t=1}^{k} \hat{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^{n} \hat{\epsilon}_t^2 \right|, \tag{2.2}$$

where $\hat{\kappa}_n^2 = n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^4 - (n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^2)^2$. Then we have the following result.

THEOREM 2.1. *Under $H_0$, as $n \to \infty$,*

$$T_n \xrightarrow{w} \sup_{0 \leq u \leq 1} |W^\circ(u)|, \tag{2.3}$$

*where $W^\circ$ denotes a standard Brownian bridge. We reject $H_0$ if $T_n$ is large.*

PROOF OF THEOREM 2.1. Since

$$\begin{aligned} \frac{1}{\sqrt{n}} \left( \sum_{t=1}^{k} \hat{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^{n} \hat{\epsilon}_t^2 \right) &= \frac{1}{\sqrt{n}} \left( \sum_{t=1}^{k} \epsilon_t^2 - \frac{k}{n} \sum_{t=1}^{n} \epsilon_t^2 \right) \\ &\quad + \frac{1}{\sqrt{n}} \left( \sum_{t=1}^{k} \hat{\epsilon}_t^2 - \sum_{t=1}^{k} \epsilon_t^2 \right) - \frac{k}{n} \frac{1}{\sqrt{n}} \left( \sum_{t=1}^{n} \hat{\epsilon}_t^2 - \sum_{t=1}^{n} \epsilon_t^2 \right) \end{aligned}$$

and

$$\max_{1 \leq k \leq n} \frac{1}{\sqrt{n}\, \mathrm{var}(\epsilon_1^2)} \left| \sum_{t=1}^{k} \epsilon_t^2 - \frac{k}{n} \sum_{t=1}^{n} \epsilon_t^2 \right| \xrightarrow{w} \sup_{0 \leq u \leq 1} |W^\circ(u)|,$$

it suffices to show that

$$\max_{1 \leq k \leq n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^{k} \hat{\epsilon}_t^2 - \sum_{t=1}^{k} \epsilon_t^2 \right| = o_P(1) \tag{2.4}$$

and

$$\hat{\kappa}_n^2 \xrightarrow{P} \mathrm{var}(\epsilon_1^2). \tag{2.5}$$

Note that

$$\max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}\hat{\epsilon}_t^2-\sum_{t=1}^{k}\epsilon_t^2\right| = \max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}(\hat{\epsilon}_t-\epsilon_t)^2+2\sum_{t=1}^{k}(\hat{\epsilon}_t-\epsilon_t)\epsilon_t\right| \le I_n+2II_n,$$

where

$$\begin{aligned} I_n &= \max_{1\le k\le n}\frac{1}{\sqrt{n}}\sum_{t=1}^{k}(\hat{\epsilon}_t-\epsilon_t)^2 = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\hat{\epsilon}_t-\epsilon_t)^2 = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{(\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta})'\boldsymbol{X}_{t-1}\}^2, \\ II_n &= \max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}(\hat{\epsilon}_t-\epsilon_t)\epsilon_t\right| = \max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}(\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta})'\boldsymbol{X}_{t-1}\epsilon_t\right|. \end{aligned}$$

As in Lee and Wei (1999), in order to show the negligibility of $I_n$ and $II_n$, we decompose the time series into several components so that each component has its own distinct characteristic roots.

Let

$$\begin{aligned} u_t &= \phi(B)(1-B)^{-a}X_t, \\ v_t &= \phi(B)(1+B)^{-b}X_t, \\ x_t(k) &= \phi(B)(1-2\cos\theta_k B+B^2)^{-d_k}X_t, \qquad k=1,\ldots,l, \\ z_t &= \phi(B)\psi^{-1}(B)X_t, \end{aligned}$$

where $B$ denotes the back-shift operator. For convenience, set

$$\begin{aligned} \boldsymbol{u}_t &= (u_t,\ldots,u_{t-a+1})', \qquad \boldsymbol{v}_t = (v_t,\ldots,v_{t-b+1})', \\ \boldsymbol{x}_t(k) &= (x_t(k),\ldots,x_{t-2d_k+1}(k))', \qquad \boldsymbol{z}_t = (z_t,\ldots,z_{t-r+1})'. \end{aligned}$$

Since $\boldsymbol{X}_0=\boldsymbol{0}$, we have $\boldsymbol{u}_0=\boldsymbol{v}_0=\boldsymbol{x}_0(1)=\cdots=\boldsymbol{x}_0(l)=\boldsymbol{z}_0=\boldsymbol{0}$.

According to Chan and Wei (1988), there exists a $q\times q$ nonsingular matrix $Q$ such that

$$Q\boldsymbol{X}_t=(\boldsymbol{u}_t',\boldsymbol{v}_t',\boldsymbol{x}_t'(1),\ldots,\boldsymbol{x}_t'(l),\boldsymbol{z}_t')'$$

and there exist block diagonal matrices $S_n=\operatorname{diag}(J_n,K_n,L_n(1),\ldots,L_n(l),M_n)$ such that

$$\begin{aligned} &S_nQ\sum_{t=1}^{n}\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-1}'Q'S_n' \\ &\quad\sim_P \operatorname{diag}\left(J_n\sum_{t=1}^{n}\boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}'J_n',\ldots,M_n\sum_{t=1}^{n}\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}'M_n'\right) \\ &\quad= O_P(1), \end{aligned} \tag{2.6}$$

where $J_n,K_n,L_n(1),\ldots,L_n(l),M_n$ are $a\times a, b\times b, 2d_1\times 2d_1,\ldots,2d_l\times 2d_l$ and $r\times r$ matrices. Moreover, it holds that

$$(Q'S_n')^{-1}(\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}) \sim_P \begin{pmatrix} (J_n')^{-1}\left(\sum_{t=1}^{n}\boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}'\right)^{-1}\sum_{t=1}^{n}\boldsymbol{u}_{t-1}\epsilon_t \\ \vdots \\ (M_n')^{-1}\left(\sum_{t=1}^{n}\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}'\right)^{-1}\sum_{t=1}^{n}\boldsymbol{z}_{t-1}\epsilon_t \end{pmatrix} \tag{2.7}$$

$$= O_P(1).$$

Here, for any sequences of r.v.'s $\{X_n\}$ and $\{Y_n\}$, $X_n \sim_P Y_n$ means that $X_n - Y_n = o_P(1)$.

First, note that $I_n = O_P(1/\sqrt{n})$, since

$$\sum_{t=1}^{n}\{(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})'\boldsymbol{X}_{t-1}\}^2 = O_P(1), \tag{2.8}$$

which is due to (2.6) and (2.7).

In order to deal with $II_n$, note that

$$S_n Q \sum_{t=1}^{k} \boldsymbol{X}_{t-1}\epsilon_t = \begin{pmatrix} J_n \sum_{t=1}^{k} \boldsymbol{u}_{t-1}\epsilon_t \\ \vdots \\ M_n \sum_{t=1}^{k} \boldsymbol{z}_{t-1}\epsilon_t \end{pmatrix}. \tag{2.9}$$

Then, in view of (2.7) and (2.9), we have that

$$\begin{aligned} II_n \sim_P \max_{1\le k\le n} \frac{1}{\sqrt{n}} \Bigg| & \left\{ (J_n')^{-1} \left( \sum_{t=1}^{n} \boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}' \right)^{-1} \sum_{t=1}^{n} \boldsymbol{u}_{t-1}\epsilon_t \right\}' J_n \sum_{t=1}^{k} \boldsymbol{u}_{t-1}\epsilon_t \\ & + \cdots + \left\{ (M_n')^{-1} \left( \sum_{t=1}^{n} \boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}' \right)^{-1} \sum_{t=1}^{n} \boldsymbol{z}_{t-1}\epsilon_t \right\}' M_n \sum_{t=1}^{k} \boldsymbol{z}_{t-1}\epsilon_t \Bigg|, \end{aligned}$$

which is no more than

$$\begin{aligned} \frac{1}{\sqrt{n}} & \left\| (J_n')^{-1} \left( \sum_{t=1}^{n} \boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1} \right)' \sum_{t=1}^{n} \boldsymbol{u}_{t-1}\epsilon_t \right\| \max_{1\le k\le n} \left\| J_n \sum_{t=1}^{k} \boldsymbol{u}_{t-1}\epsilon_t \right\| \\ & + \cdots + \frac{1}{\sqrt{n}} \left\| (M_n')^{-1} \left( \sum_{t=1}^{n} \boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1} \right)' \sum_{t=1}^{n} \boldsymbol{z}_{t-1}\epsilon_t \right\| \max_{1\le k\le n} \left\| M_n \sum_{t=1}^{k} \boldsymbol{z}_{t-1}\epsilon_t \right\|, \end{aligned} \tag{2.10}$$

where $\|\cdot\|$ denotes the Euclidean norm. Since the first term in each summand in (2.10) is $O_P(1)$ by Chan and Wei (1988), we only have to deal with the second terms.

Now, we show that

$$\max_{1\le k\le n} \left\| J_n \sum_{t=1}^{k} \boldsymbol{u}_{t-1}\epsilon_t \right\| = O_P(1). \tag{2.11}$$

Recall that $(1-B)^a u_t = \epsilon_t$ and $\boldsymbol{u}_0 = \boldsymbol{0}$. Set $u_t(j) = (1-B)^{a-j}u_t$, $j = 0, \ldots, a-1$, and $U_t = (u_t(a), \ldots, u_t(1))'$. By (3.13) of Chan and Wei (1988), there exists an $a \times a$ matrix $M$ such that $M\boldsymbol{u}_t = U_t$ and $J_n = N_n^{-1}M$ where $N_n = \mathrm{diag}(n^a, \ldots, n)$. In this case, we have

$$J_n \sum_{t=1}^{k} \boldsymbol{u}_{t-1}\epsilon_t = \left( \sum_{t=1}^{k} n^a u_{t-1}(a)\epsilon_t, \ldots, \sum_{t=1}^{k} n^{-1} u_{t-1}(1)\epsilon_t \right)',$$

so that

$$\max_{1\le k\le n}\left\|J_n\sum_{t=1}^{k}\boldsymbol{u}_{t-1}\epsilon_t\right\|^2\le\sum_{j=1}^{a}n^{-2j}\max_{1\le k\le n}\left(\sum_{t=1}^{k}u_{t-1}(j)\epsilon_t\right)^2. \tag{2.12}$$

Since $\{\sum_{t=1}^{k}u_{t-1}(j)\epsilon_t, \mathcal{F}_k=\sigma(\epsilon_k:t\le k)\}$ forms a martingale, by using the first submartingale inequality of Theorem 3.8 in Karatzas and Shreve (1988), p. 13, we have that for any $\delta>0$,

$$P\left[\max_{1\le k\le n}\left(\sum_{t=1}^{k}u_{t-1}(j)\epsilon_t\right)^2\ge\delta\right]\le\frac{1}{\delta}E\left(\sum_{t=1}^{n}u_{t-1}(j)\epsilon_t\right)^2.$$

Note that for all $t$ and $j$,

$$Eu_t^2(j)\le t^{2j-1}\sigma^2,$$

and

$$E\left(\sum_{t=1}^{n}u_{t-1}(j)\epsilon_t\right)^2=\sum_{t=1}^{n}Eu_{t-1}^2(j)\sigma^2\le\sum_{t=1}^{n}(t-1)^{2j-1}\sigma^2\le n^{2j}\sigma^4.$$

Hence,

$$n^{-2j}\max_{1\le k\le n}\left(\sum_{t=1}^{k}u_{t-1}(j)\epsilon_t\right)^2=O_P(1)\quad\text{for all}\quad j,$$

and the right hand side of (2.12) is $O_P(1)$. This proves (2.11).

Meanwhile, in a similar fashion, we can show that

$$\max_{1\le k\le n}\left\|K_n\sum_{t=1}^{k}\boldsymbol{v}_{t-1}\epsilon_t\right\|=O_P(1); \tag{2.13}$$

$$\max_{1\le k\le n}\left\|L_n(i)\sum_{t=1}^{k}\boldsymbol{x}_{t-1}(i)\epsilon_t\right\|=O_P(1),\qquad i=1,\ldots,l; \tag{2.14}$$

$$\max_{1\le k\le n}\left\|M_n\sum_{t=1}^{k}\boldsymbol{z}_{t-1}\epsilon_t\right\|=O_P(1), \tag{2.15}$$

the proofs of which are omitted for brevity. Combining (2.11), (2.13)–(2.15), we obtain the argument in (2.10) is $O_P(1)$, which entails $II_n=o_P(1)$. This proves (2.4). Since (2.5) follows from (2.4) and the fact $I_n=O_P(1/\sqrt{n})$, we establish the theorem. □

So far, we have seen that the test for the variance change can be performed based on the least squares residuals. Since the approach based on the squares of residuals works adequately, it is natural to ask if an analogous phenomenon happens in the SREP (empirical process based on the squares of residuals). Recall that the REP (residual empirical process) converges to a non-Gaussian process in the presence of unit root 1 (cf. Lee and Wei (1999)). Also, it was shown that the residual based Bickel-Rosenblatt test using a smoothing technique can fail contrary to intuition (cf. Lee and Na (2002)). In short, the following derivation shows us that the SREP can be used as a basic process for a goodness of fit test so long as the underlying density is symmetric and satisfies the regularity conditions in Lee and Wei, Section 3.2. The result is directly applicable to a Gaussian test.

Let $F$ be the distribution of $\epsilon_1$ with symmetric density $f$, and let $G$ be the distribution of $\epsilon_1^2$, namely, $G(x^2) = F(x) - F(-x)$ for $x \geq 0$. Define

$$\mathcal{E}_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{I(\epsilon_t^2 \leq x^2) - G(x^2)\}, \qquad x \geq 0,$$

$$\hat{\mathcal{E}}_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{I(\hat{\epsilon}_t^2 \leq x^2) - G(x^2)\}, \qquad x \geq 0.$$

Then,

$$\begin{aligned} \hat{\mathcal{E}}_n(x) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{I(-x \leq \hat{\epsilon}_t \leq x) - (F(x) - F(-x))\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{I(\hat{\epsilon}_t \leq x) - F(x)\} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{I(\hat{\epsilon}_t \leq -x) - F(-x)\}. \end{aligned}$$

Since in view of Lee and Wei (1999),

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{I(\hat{\epsilon}_t \leq x) - F(x)\} &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{I(\epsilon_t \leq x) - F(x)\} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \boldsymbol{X}_{t-1} f(x) + \eta_n(x) \end{aligned}$$

with $\sup_x |\eta_n(x)| = o_P(1)$, and $f(x) = f(-x)$, we can see that

$$\sup_x |\hat{\mathcal{E}}_n(x) - \mathcal{E}_n(x)| = o_P(1).$$

Therefore,

$$\tilde{\mathcal{E}}_n(u) := \hat{\mathcal{E}}_n((G^{-1}(u))^{1/2}) \xrightarrow{w} W^\circ(u), \qquad 0 \leq u \leq 1.$$

The above result suggests that a goodness of fit test for a symmetric density, including a Gaussian test, can be accomplished based on $\tilde{\mathcal{E}}_n$. In fact, the Gaussian test (when the variance is known) is converted into a chi-square distribution test. In actual practice, one should keep in mind that, if an estimate of variance is plugged into the empirical process, the limiting distribution is no longer a Brownian bridge, but a Gaussian process as we usually observe in the empirical process context (cf. Lee and Wei (1999), Section 3.2). Besides the goodness of fit test, we can reason that the sequential SREP (cf. Bai (1994)) can be employed to detect a distributional change in autoregressive models under the same conditions; it is well-known that this result does not hold when using the sequential REP (cf. Ling (1998)). All these facts support the usefulness of the method employing the squares of residuals in autoregressive models.

## 3. Test in nonparametric regression model

In this section, we develop the variance change test procedure in a fixed design nonparametric regression model with a strong mixing error process. The nonparametric regression model under consideration is as follows:

$$Y_t = g(x_t) + \epsilon_t, \qquad t = 1, \dots, n, \tag{3.1}$$

where $n$ denotes the sample size, $x_t = t/n$, $t = 1, \dots, n$, are the equally-spaced design points, $g$ is the regression function, and $\{\epsilon_t\}$ is a stationary strong mixing process with zero mean and finite variance. The model in (3.1) has been studied by several authors: see Hall and Hart (1990), Hart (1991) and Wu and Chu (1994). Their concern was the estimation of the regression function $g$ rather than the change detection problem itself.

Our goal is to test the following hypotheses:

$H_0$: $E\epsilon_t^2 = \sigma^2$ for all $t = 1, \dots, n$ vs.

$H_1$: not $H_0$.

Towards this end, we employ the cusum of squares test based on residuals. For an estimator of $g$, we consider the kernel-type regression estimator introduced by Priestley and Chao (1972). Let $K$ be a kernel function and $h = h_n$ be a bandwidth. Given observations $Y_1, \dots, Y_n$, the kernel regression function estimator $g_n(\cdot)$ is given as

$$g_n(x) = \frac{1}{n} \sum_{t=1}^{n} Y_t K_h(x - x_t), \qquad 0 \le x \le 1,$$

where $K_h(u) = K(u/h)/h$, and the residuals are

$$\hat{\epsilon}_t = Y_t - g_n(x_t), \qquad t = 1, \dots, n.$$

In order to obtain an asymptotic result as in Inclán and Tiao (1994), we assume that $\{\epsilon_t\}$ is geometrically strong mixing, viz., if we put

$$\alpha_k = \alpha(\sigma(\epsilon_s, s \le 0), \sigma(\epsilon_s, s \ge k)), \qquad k = 0, 1, 2, \dots,$$

where $\alpha(\mathcal{F}, \mathcal{G})$ denotes the strong mixing coefficient between $\sigma$-fields $\mathcal{F}$ and $\mathcal{G}$ (cf. Doukhan (1994), p. 3), $\{\alpha_k\}$ satisfies

$$\alpha_k \le C e^{-\rho k} \tag{3.2}$$

for some $C > 0$ and $\rho > 0$.

Before we state the main theorem of this section, we introduce some notation and conditions. We first define the partial sum of $\epsilon_t^2$, namely, $S_k = \sum_{t=1}^{k} \epsilon_t^2$, $k = 1, \dots, n$. In fact, if we set $s_n^2 = \mathrm{Var}(S_n)$, then provided $E\epsilon_1^2 = \sigma^2$ and $E\epsilon_1^4 < \infty$, we have

$$\begin{aligned} s_n^2 &= n \left\{ E(\epsilon_1^2 - \sigma^2)^2 + 2 \sum_{k=1}^{n} \left(1 - \frac{k}{n}\right) E(\epsilon_1^2 - \sigma^2)(\epsilon_{1+k}^2 - \sigma^2) \right\} \\ &= n \left\{ E(\epsilon_1^2 - \sigma^2)^2 + 2 \sum_{k=1}^{n} E(\epsilon_1^2 - \sigma^2)(\epsilon_{1+k}^2 - \sigma^2) \right\} \\ &\quad - 2 \sum_{k=1}^{n} k E(\epsilon_1^2 - \sigma^2)(\epsilon_{1+k}^2 - \sigma^2), \end{aligned} \tag{3.3}$$

which is useful for later work. Since we have to deal with the residuals, we also define the partial sum of residuals:

$$\hat{S}_k = \sum_{t=[nh]+1}^{k} \hat{\epsilon}_t^2 \quad \text{for} \quad k = [nh]+1, \ldots, n-[nh],$$

where the truncations for $k$ are concerned with the kernel function $K$ satisfying $\mathrm{supp}(K) \subset [-1,1]$ which is assumed in (A3) below. Moreover, $\gamma(k) := E(\epsilon_1^2 - \sigma^2)(\epsilon_{i+k}^2 - \sigma^2)$ is estimated by

$$\hat{\gamma}(k) = \frac{1}{n-2[nh]} \sum_{t=[nh]+1}^{n-[nh]-k} (\hat{\epsilon}_t^2 - \hat{\mu}_2)(\hat{\epsilon}_{t+k}^2 - \hat{\mu}_2),$$

where

$$\hat{\mu}_2 = \frac{1}{n-2[nh]} \sum_{t=[nh]+1}^{n-[nh]} \hat{\epsilon}_t^2.$$

Then the estimator $\hat{s}_n^2$ of $s_n^2$ is given by

$$\hat{s}_n^2 = n\left\{\hat{\gamma}(0) + 2\sum_{k=1}^{l_n} \hat{\gamma}(k)\right\},$$

for a sequence of positive integers $\{l_n\}$ satisfying $l_n \to \infty$ and $l_n/n \to 0$ as $n \to \infty$ as will be explained in more detail shortly. Note that the residuals near end points are discarded to avoid the boundary effect in nonparametric regression.

Below are the conditions imposed in this section.

(A1) $\{\alpha_k\}$ satisfies (3.2), and $E|\epsilon_1^2 - \sigma^2|^r < \infty$ for some $r > 2$.

(A2) The regression function $g$ satisfies the Lipschitz condition, viz.,

$$|g(x) - g(y)| \le D_1|x-y|, \qquad 0 \le x, y \le 1$$

for some constant $0 < D_1 < \infty$.

(A3) The kernel function $K$ vanishes outside $[-1,1]$ and is Lipschitz continuous on $[-1,1]$, viz.,

$$|K(x) - K(y)| \le D_2|x-y|, \qquad -1 \le x, y \le 1$$

for some constant $0 < D_2 < \infty$. And $K$ satisfies $\int K(x)dx = 1$.

(A4) The bandwidth $h = h_n$ satisfies $nh^2 \to \infty$ and $nh^4 \to 0$ as $n \to \infty$.

(A5) $l_n$ satisfies $l_n \to \infty$, $l_n/\sqrt{nh} \to 0$ and $l_n h \to 0$ as $n \to \infty$.

*Remark.* A broad class of processes, including invertible stationary ARMA$(p,q)$ processes with innovations having a continuous distribution, satisfy Condition (A1) (cf. Gorodetskii (1977)). Conditions (A2)–(A4) are the usual conditions assumed in nonparametric regression estimation.

Here is the main result of this section.

THEOREM 3.1. *Assume that $H_0$ holds. Under Conditions* (A1)–(A5),

$$(3.4) \qquad T_n^* := \max_{[nh]+1\le k\le n-[nh]} \left| \frac{1}{\hat{s}_n}\left(\hat{S}_k - \frac{k-[nh]}{n-2[nh]}\hat{S}_{n-[nh]}\right)\right| \xrightarrow{w} \sup_{0\le u\le 1}|W^0(u)|.$$

*We reject $H_0$ if $T_n^*$ is large.*

Now we prove Theorem 3.1. We start with a lemma which can be found in Doukhan ((1994), p. 46).

LEMMA 3.1. *Assume that $H_0$ holds. If $E|\epsilon_1^2-\sigma^2|^r<\infty$ for some $r>2$ and* (3.2) *is true for some $C>0$ and $\rho>0$, then*

$$(3.5) \qquad \frac{1}{s_n}\left(S_{[n\cdot]} - \frac{[n\cdot]}{n}S_n\right) \xrightarrow{w} W^0.$$

LEMMA 3.2. *Under $H_0$ and* (A1)–(A4), *as $n\to\infty$,*

$$\max_{[nh]+1\le k\le n-[nh]} \frac{1}{\sqrt{n}}|\hat{S}_k - \tilde{S}_k| \xrightarrow{P} 0,$$

*where $\tilde{S}_k = \sum_{t=[nh]+1}^{k}\epsilon_t^2$.*

PROOF. We write

$$\begin{aligned}
\frac{1}{\sqrt{n}}(\hat{S}_k - \tilde{S}_k) &= \frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{k}(\hat{\epsilon}_t^2 - \epsilon_t^2)\\
&= \frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{k}(g_n(x_t)-g(x_t))^2 + 2\frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{k}(g(x_t)-g_n(x_t))\epsilon_t\\
&= I_k + II_k.
\end{aligned}$$

Observe that

$$\begin{aligned}
(3.6) \qquad I_k &\le \frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{n-[nh]}\left\{\frac{1}{n}\sum_{j=1}^{n}\epsilon_j K_h(x_t-x_j)\right.\\
&\qquad\qquad \left.+\frac{1}{n}\sum_{j=1}^{n}g(x_j)K_h(x_t-x_j) - g(x_t)\right\}^2\\
&\le 3\left[\frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{n-[nh]}\left\{\frac{1}{n}\sum_{j=1}^{n}\epsilon_j K_h(x_t-x_j)\right\}^2\right.\\
&\qquad + \frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{n-[nh]}\left\{\frac{1}{n}\sum_{j=1}^{n}(g(x_j)-g(x_t))K_h(x_t-x_j)\right\}^2\\
&\qquad \left.+ \frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{n-[nh]}\left\{g(x_t)\left(\frac{1}{n}\sum_{j=1}^{n}K_h(x_t-x_j)-1\right)\right\}^2\right].
\end{aligned}$$

It is easy to check that the first term in (3.6) is $O_P(n^{-1/2}h^{-1})$ since

$$E\left\{\frac{1}{n}\sum_{j=1}^{n}\epsilon_j K_h(x_t-x_j)\right\}^2 = O\left(\frac{1}{nh}\right),$$

due to (A1) and (A3). The remaining nonstochastic terms in (3.6) are of order $O(n^{1/2}h^2)$ and $O(n^{-3/2}h^{-2})$, respectively, since

$$\frac{1}{n}\sum_{j=1}^{n}(g(x_j)-g(x_t))K_h(x_t-x_j) = O(h)$$

and

$$g(x_t)\left(\frac{1}{n}\sum_{j=1}^{n}K_h(x_t-x_j)-1\right) = O\left(\frac{1}{nh}\right), \tag{3.7}$$

where we have used (A2) and (A3). Then, by using (A4) we obtain

$$\max_{[nh]+1\le k\le n-[nh]}|I_k| = o_P(1). \tag{3.8}$$

For $II_k$, we decompose it into three terms as follows:

$$\begin{aligned}
II_k &= \frac{2}{\sqrt{n}}\sum_{t=[nh]+1}^{k}\left\{g(x_t)\left(1-\frac{1}{n}\sum_{j=1}^{n}K_h(x_t-x_j)\right)\right\}\epsilon_t \\
&\quad + \frac{2}{\sqrt{n}}\sum_{t=[nh]+1}^{k}\left\{\frac{1}{n}\sum_{j=1}^{n}(g(x_t)-g(x_j))K_h(x_t-x_j)\right\}\epsilon_t \\
&\quad - \frac{2}{\sqrt{n}}\sum_{t=[nh]+1}^{k}\left\{\frac{1}{n}\sum_{j=1}^{n}\epsilon_j K_h(x_t-x_j)\right\}\epsilon_t \\
&= II_{k,1}+II_{k,2}-II_{k,3}.
\end{aligned}$$

First, observe that

$$|II_{k,1}| \le \frac{2}{\sqrt{n}}\sum_{t=[nh]+1}^{n-[nh]}\left|g(x_t)\left(1-\frac{1}{n}\sum_{j=1}^{n}K_h(x_t-x_j)\right)\right||\epsilon_t|.$$

From this and (3.7), we have

$$\max_{[nh]+1\le k\le n-[nh]}|II_{k,1}| = O_P(n^{-1/2}h^{-1}). \tag{3.9}$$

For $II_{k,2}$, set $p=[n^{1/3}]$ and $q=[(\log n)^2]$, and define random variables $V_i$ and $V_i'$ as follows:

$$V_i=\sum_{t=a_i}^{b_i}\eta_t, \qquad V_i'=\sum_{t=b_i+1}^{a_{i+1}-1}\eta_t, \qquad i=1,\dots,r=\left[\frac{n-2[nh]}{p+q}\right],$$

where $\eta_t = (n^{-1}\sum_{j=1}^n (g(x_t) - g(x_j))K_h(x_t - x_j))\epsilon_t$, $a_i = [nh] + 1 + (p+q)(i-1)$ and $b_i = [nh] + 1 + (p+q)(i-1) + p$. Then, we can write

$$II_{k,2} = \frac{2}{\sqrt{n}} \sum_{t=[nh]+1}^{k} \eta_t = \frac{2}{\sqrt{n}} \sum_{i=1}^{u_k} V_i + \frac{2}{\sqrt{n}} \sum_{i=1}^{u_k} V_i' + \frac{2}{\sqrt{n}} \Delta_k, \tag{3.10}$$

where $u_k$ is the largest integer such that $[nh]+1+u_k(p+q) \le k$, and $\Delta_k = \sum_{t=[nh]+1}^k \eta_t - \sum_{i=1}^{u_k} V_i - \sum_{i=1}^{u_k} V_i'$. Let $\zeta$ be any positive real number. By the coupling theorem (cf. Doukhan (1994), p. 8, and Bosq (1996), p. 18), there exist independent random variables $V_1^*, \ldots, V_r^*$ such that $V_i^*$ and $V_i$ have the same distribution and

$$P\left(|V_i - V_i^*| \ge \frac{\zeta}{2}\frac{\sqrt{n}}{r}\right) \le 18\left(\alpha_q^2 \frac{2rE|V_i|}{\zeta\sqrt{n}}\right)^{1/3}.$$

Then, we can see that

$$\max_{[nh]+1 \le k \le n-[nh]} \frac{1}{\sqrt{n}} \left|\sum_{i=1}^{u_k} V_i\right| = o_P(1), \tag{3.11}$$

since

$$\begin{aligned}
&P\left(\max_{[nh]+1 \le k \le n-[nh]} \frac{1}{\sqrt{n}} \left|\sum_{i=1}^{u_k} V_i\right| > \zeta\right) \\
&\quad \le P\left(\max_{1 \le u \le r} \frac{1}{\sqrt{n}} \left|\sum_{i=1}^{u} V_i\right| > \zeta\right) \\
&\quad \le P\left(\max_{1 \le u \le r} \frac{1}{\sqrt{n}} \left|\sum_{i=1}^{u} V_i^*\right| > \frac{\zeta}{2}\right) + P\left(\max_{1 \le u \le r} \frac{1}{\sqrt{n}} \left|\sum_{i=1}^{u} (V_i - V_i^*)\right| > \frac{\zeta}{2}\right) \\
&\quad \le \frac{4}{n\zeta^2} \operatorname{Var}\left(\sum_{i=1}^{r} V_i^*\right) + \sum_{i=1}^{r} P\left(|V_i - V_i^*| > \frac{\zeta\sqrt{n}}{2r}\right) \\
&\quad = O\left(\frac{rph^2}{n}\right) + O\left(r\alpha_q^{2/3}\left(\frac{rph}{\sqrt{n}}\right)^{1/3}\right) \\
&\quad = o(1).
\end{aligned}$$

In the same manner, we also get

$$\max_{[nh]+1 \le k \le n-[nh]} \frac{1}{\sqrt{n}} \left|\sum_{i=1}^{u_k} V_i'\right| = o_P(1). \tag{3.12}$$

Furthermore, applying the coupling theorem again, we can show that for any $\zeta > 0$,

$$\begin{aligned}
&P\left(\max_{[nh]+1 \le k \le n-[nh]} \frac{1}{\sqrt{n}} |\Delta_k| > \zeta\right) \\
&\quad \le \sum_{i=1}^{r} P\left(\max_{a_i \le k \le a_{i+1}-1} \left|\sum_{t=a_i}^{k} \eta_t\right| > \zeta\sqrt{n}\right) + P\left(\max_{a_{r+1} \le k \le n-[nh]} \left|\sum_{t=a_{r+1}}^{k} \eta_t\right| > \zeta\sqrt{n}\right) \\
&\quad = O\left(r\frac{(p+q)h^2}{n\zeta^2}\right),
\end{aligned}$$

the detailed proofs of which are omitted for brevity. This entails that

$$\max_{[nh]+1\leq k\leq n-[nh]} \frac{1}{\sqrt{n}}|\Delta_k| = o_P(1). \tag{3.13}$$

Combining (3.10) and (3.11)–(3.13), we have

$$\max_{[nh]+1\leq k\leq n-[nh]} |II_{k,2}| = o_P(1). \tag{3.14}$$

Finally, we show that

$$\max_{[nh]+1\leq k\leq n-[nh]} |II_{k,3}| = o_P(1). \tag{3.15}$$

Let $\xi_t = (n^{-1}\sum_{j=1}^n \epsilon_j K_h(x_t - x_j))\epsilon_t$, $p = p_n = [nh\log n]$ and $q = q_n = 2[nh] + [(\log n)^2]$, and define $a_i$, $b_i$, $r$ and $u_k$ as we did before. Set

$$W_i = \sum_{t=a_i}^{b_i}(\xi_t - E\xi_t), \qquad W_i' = \sum_{t=b_i+1}^{a_{i+1}-1}(\xi_t - E\xi_t), \qquad i = 1, \ldots, r.$$

If we put $\gamma_\epsilon(k) = E\epsilon_1\epsilon_{1+k}$, then

$$E\xi_t = \frac{1}{n}\sum_{j=1}^n K_h(x_t - x_j)\gamma_\epsilon(t-j),$$

and

$$II_{k,3} = \frac{2}{\sqrt{n}}\sum_{i=1}^{u_k} W_i + \frac{2}{\sqrt{n}}\sum_{i=1}^{u_k} W_i' + \frac{2}{\sqrt{n}}\Delta_k' + \frac{2}{\sqrt{n}}\sum_{t=[nh]+1}^{k} E\xi_t,$$

where $\Delta_k' = \sum_{t=[nh]+1}^k(\xi_t - E\xi_t) - \sum_{i=1}^{u_k} W_i - \sum_{i=1}^{u_k} W_i'$. Simple algebra shows that if (A1) and (A3) hold, then

$$\frac{1}{\sqrt{n}}\sum_{t=[nh]+1}^{n-[nh]} |E\xi_t| = O\left(\frac{1}{\sqrt{nh}}\right),$$

and for all $i = 1, \ldots, r$,

$$\mathrm{Var}(W_i) = O\left(\frac{p^2(\log n)^2}{n^2h^2}\right).$$

Then, in a manner similar to the derivation of (3.11)–(3.13), we can obtain (3.15) by utilizing the coupling theorem since only $\epsilon_j$ with $j = t - [nh], \ldots, t + [nh]$ are involved in each $\xi_t$. Combining (3.8), (3.9), (3.14) and (3.15), we establish the lemma. □

LEMMA 3.3. *Under $H_0$ and* (A1)–(A5), *as $n \to \infty$,*

$$\frac{\hat{s}_n^2}{n} \xrightarrow{P} \tau^2 := \sum_{k=-\infty}^{\infty}\gamma(k),$$

*where $\gamma(k) = E(\epsilon_1^2 - \sigma^2)(\epsilon_{1+k}^2 - \sigma^2)$.*

PROOF. Let

$$\gamma^*(k) = (n-2[nh])^{-1} \sum_{t=[nh]+1}^{n-[nh]-|k|} (\epsilon_t^2 - \mu_2)(\epsilon_{t+|k|}^2 - \mu_2),$$

where $\mu_2 = (n-2[nh])^{-1} \sum_{t=[nh]+1}^{n-[nh]} \epsilon_t^2$. It suffices to show that

$$\sum_{k=-l_n}^{l_n} \hat{\gamma}(k) - \sum_{k=-l_n}^{l_n} \gamma^*(k) \xrightarrow{P} 0 \tag{3.16}$$

and

$$\sum_{k=-l_n}^{l_n} \gamma^*(k) - \sum_{k=-l_n}^{l_n} \gamma(k) \xrightarrow{P} 0. \tag{3.17}$$

In order to verify (3.16), we write

$$\begin{aligned}
\hat{\gamma}(k) - \gamma^*(k) &= \frac{1}{n-2[nh]} \sum_{t=[nh]+1}^{n-[nh]-k} \{(\hat{\epsilon}_t^2 - \epsilon_t^2)(\hat{\epsilon}_{t+k}^2 - \epsilon_{t+k}^2) + (\hat{\epsilon}_t^2 - \epsilon_t^2)(\epsilon_{t+k}^2 - \mu_2) \\
&\qquad + (\hat{\epsilon}_t^2 - \epsilon_t^2)(\mu_2 - \hat{\mu}_2) + (\epsilon_t^2 - \mu_2)(\hat{\epsilon}_{t+k}^2 - \epsilon_{t+k}^2) \\
&\qquad + (\epsilon_t^2 - \mu_2)(\mu_2 - \hat{\mu}_2) + (\mu_2 - \hat{\mu}_2)(\hat{\epsilon}_{t+k}^2 - \epsilon_{t+k}^2) \\
&\qquad + (\mu_2 - \hat{\mu}_2)(\epsilon_{t+k}^2 - \mu_2) + (\mu_2 - \hat{\mu}_2)^2\} \\
&= R_1(k) + \cdots + R_8(k).
\end{aligned}$$

Observe that by the Schwarz inequality,

$$\begin{aligned}
|R_1(k)| &\le \frac{1}{n-2[nh]} \sum_{t=[nh]+1}^{n-[nh]} (\hat{\epsilon}_t^2 - \epsilon_t^2)^2 \\
&\le \frac{2}{n-2[nh]} \sum_{t=[nh]+1}^{n-[nh]} (g_n(x_t) - g(x_t))^4 + \frac{8}{n-2[nh]} \sum_{t=[nh]+1}^{n-[nh]} (g_n(x_t) - g(x_t))^2 \epsilon_t^2.
\end{aligned}$$

By Jensen's inequality and Condition (A1), we can see that

$$\begin{aligned}
\sum_{t=[nh]+1}^{n-[nh]} (g_n(x_t) - g(x_t))^4 &\le 27 \sum_{t=[nh]+1}^{n-[nh]} \left\{ \left( \frac{1}{n} \sum_{j=1}^{n} \epsilon_j K_h(x_t - x_j) \right)^4 \right. \\
&\qquad + \left( \frac{1}{n} \sum_{j=1}^{n} (g(x_j) - g(x_t)) K_h(x_t - x_j) \right)^4 \\
&\qquad \left. + \left( g(x_t) \left( \frac{1}{n} \sum_{j=1}^{n} K_h(x_t - x_j) - 1 \right) \right)^4 \right\} \\
&= O_P\left( n \frac{1}{n^2 h^2} \right) + O(nh^4) + O\left( n \frac{1}{n^4 h^4} \right).
\end{aligned}$$

Since $E\sum_{t=[nh]+1}^{n-[nh]}\epsilon_t^4 = O(n)$ and

$$\sum_{t=[nh]+1}^{n-[nh]}(g_n(x_t)-g(x_t))^2\epsilon_t^2 \le \left(\sum_{t=[nh]+1}^{n-[nh]}(g_n(x_t)-g(x_t))^4\right)^{1/2}\left(\sum_{t=[nh]+1}^{n-[nh]}\epsilon_t^4\right)^{1/2},$$

we obtain, uniformly in $k$,

$$|R_1(k)| = O_P\left(\frac{1}{n^2h^2}\vee h^4\right)+O_P\left(\left(\frac{1}{n^2h^2}\vee h^4\right)^{1/2}\right) = O_P\left(\frac{1}{nh}\vee h^2\right), \tag{3.18}$$

where $a\vee b$ denotes the maximum of $a$ and $b$. For $i=2,4$, we have

$$\begin{aligned}|R_i(k)| &\le \left(\frac{1}{n-2[nh]}\sum_{t=[nh]+1}^{n-[nh]}(\hat{\epsilon}_t^2-\epsilon_t^2)^2\right)^{1/2} \\ &\quad\cdot\left(\frac{1}{n-2[nh]}\sum_{t=[nh]+1}^{n-[nh]}(\epsilon_t^2-\mu_2)^2\right)^{1/2} \\ &= \left(O_P\left(\frac{1}{nh}\vee h^2\right)\right)^{1/2}(O_P(1))^{1/2} \\ &= O_P\left(\frac{1}{\sqrt{nh}}\vee h\right)\end{aligned} \tag{3.19}$$

in view of (3.18). Furthermore, by Lemma 3.2, we have that for $i=3,6$,

$$\begin{aligned}|R_i(k)| &\le \frac{n}{(n-2[nh])^2}\max_{[nh]+1\le j\le n-[nh]}\frac{1}{\sqrt{n}}\left|\sum_{t=[nh]+1}^{j}(\hat{\epsilon}_t^2-\epsilon_t^2)\right| \\ &\quad\cdot\frac{1}{\sqrt{n}}\left|\sum_{t=[nh]+1}^{n-[nh]}(\hat{\epsilon}_t^2-\epsilon_t^2)\right| \\ &= o_P\left(\frac{1}{n}\right),\end{aligned} \tag{3.20}$$

and for $i=5,7$,

$$\begin{aligned}|R_i(k)| &\le \frac{1}{n-2[nh]}\sum_{t=[nh]+1}^{n-[nh]}|\epsilon_t^2-\mu_2|\cdot\frac{\sqrt{n}}{n-2[nh]}\frac{1}{\sqrt{n}}\left|\sum_{t=[nh]+1}^{n-[nh]}(\hat{\epsilon}_t^2-\epsilon_t^2)\right| \\ &= o_P\left(\frac{1}{\sqrt{n}}\right).\end{aligned} \tag{3.21}$$

Also, by Lemma 3.2,

$$\begin{aligned}|R_8(k)| &\le (\mu_2-\hat{\mu}_2)^2 = \frac{n}{(n-2[nh])^2}\left(\frac{1}{\sqrt{n}}\left|\sum_{t=[nh]+1}^{n-[nh]}(\epsilon_t^2-\hat{\epsilon}_t^2)\right|\right)^2 \\ &= o_P\left(\frac{1}{n}\right).\end{aligned} \tag{3.22}$$

Hence, (3.16) is yielded by (3.18)–(3.22) and Condition (A5).

Now, it remains to prove (3.17). Noticing that

$$\begin{aligned}\gamma^*(k)-\gamma(k) = {} & \frac{1}{n-2[nh]}\sum_{t=[nh]+1}^{n-[nh]-k}(\epsilon_t^2\epsilon_{t+k}^2 - E\epsilon_t^2\epsilon_{t+k}^2) - \frac{k}{n-2[nh]}(E\epsilon_1^2\epsilon_{1+k}^2+\sigma^4) \\ & - \frac{n-2[nh]+k}{n-2[nh]}(\mu_2^2-\sigma^4) \\ & + \frac{\mu_2}{n-2[nh]}\left\{\sum_{t=[nh]+1}^{[nh]+k}\epsilon_t^2 + \sum_{t=n-[nh]-k+1}^{n-[nh]}\epsilon_t^2\right\},\end{aligned}$$

we have

$$\left|\sum_{k=-l_n}^{l_n}(\gamma^*(k)-\gamma(k))\right| = \left|\sum_{k=-l_n}^{l_n}\frac{1}{n-2[nh]}\sum_{t=[nh]+1}^{n-[nh]-k}(\epsilon_t^2\epsilon_{t+k}^2 - E\epsilon_t^2\epsilon_{t+k}^2)\right| + o_P(1).$$

Then, using the Minkovski inequality in $L^2$-norm $\|\cdot\|_2$, and Conditions (A1) and (A5), we obtain

$$\begin{aligned}& \left\|\sum_{k=-l_n}^{l_n}\frac{1}{n-2[nh]}\sum_{t=[nh]+1}^{n-[nh]-k}(\epsilon_t^2\epsilon_{t+k}^2 - E\epsilon_t^2\epsilon_{t+k}^2)\right\|_2 \\ & \quad \le \sum_{t=-l_n}^{l_n}\left\|\frac{1}{n-2[nh]}\sum_{t=[nh]+1}^{n-[nh]-k}(\epsilon_t^2\epsilon_{t+k}^2 - E\epsilon_t^2\epsilon_{t+k}^2)\right\|_2 \\ & \quad = O_P\left(l_n\frac{1}{\sqrt{n}}\right) \\ & \quad = o_P(1),\end{aligned}$$

which implies (3.17) and the lemma is established. □

Proof of Theorem 3.1. Let $\tilde{s}_n^2 = \mathrm{Var}(\tilde{S}_{n-[nh]})$. Then, from (3.3) and Lemma 3.3, we can see that $\tilde{s}_n/\hat{s}_n \xrightarrow{P} 1$ and $\hat{s}_n/\sqrt{n} \xrightarrow{P} \tau$. Now, note that

$$\begin{aligned}& \frac{1}{\hat{s}_n}\left(\hat{S}_k - \frac{k-[nh]}{n-2[nh]}\hat{S}_{n-[nh]}\right) \\ & \quad = \frac{\tilde{s}_n}{\hat{s}_n}\frac{1}{\tilde{s}_n}\left\{\tilde{S}_k - \frac{k-[nh]}{n-2[nh]}\tilde{S}_{n-[nh]}\right\} \\ & \qquad + \frac{\sqrt{n}}{\hat{s}_n}\frac{1}{\sqrt{n}}\left\{(\hat{S}_k-\tilde{S}_k) - \frac{k-[nh]}{n-2[nh]}(\hat{S}_{n-[nh]}-\tilde{S}_{n-[nh]})\right\}.\end{aligned}$$

Then, the theorem is a direct result of Lemmas 3.1 and 3.2. □

## 4. Simulation results

In this section we conduct a simulation study to evaluate the tests in Sections 2 and 3. In this simulation we perform a test at a nominal level $\alpha = 0.05$. The empirical sizes and powers are calculated as the rejection number of the null hypothesis out of 2000 repetitions. First, in order to see the performance of $T_n$ in Section 2, we consider the model

$$X_t = \beta X_{t-1} + \epsilon_t, \qquad t = 1, \ldots, n,$$

where $X_0$ is assumed to be 0 and $\epsilon_t$ are iid normal random variables with mean zero and variance $\sigma^2$. Now we consider the problem of testing the following hypotheses:

$H_0$: $\sigma^2$ remains equal to 1 for $t = 1, \ldots, n$, vs.

$H_1$: $\sigma^2 = 1$ for $t = 1, \ldots, [n/2]$ and $\sigma^2 = \Delta$ for $t = [n/2] + 1, \ldots, n$,

where $\Delta$ takes the values 2 and 4. Here we evaluate $T_n$ with the sample size $n = 200$, 300, 500 and $\beta = 0.2$, 0.5, 0.8, 1.0. The empirical sizes and powers are summarized in Table 1. As seen in the table, $T_n$ does not have size distortions and produces good powers. It is manifest that the sizes and powers do not depend upon the values of $\beta$, and the test works well for the unstable case as well as the stationary case.

Now, in order to evaluate the performance of the test $T_n^*$ in Section 3, we consider the nonparametric regression model in (3.1):

$$Y_t = g(x_t) + \epsilon_t, \qquad t = 1, \ldots, n,$$

where $g(x) = 25x^3 - 45x^2 + 24x - 3.6$ and $\{\epsilon_t\}$ satisfies the equation:

$$\epsilon_t = \phi\epsilon_{t-1} + e_t, \quad |\phi| < 1, \qquad t = 1, \ldots, n,$$

where $e_t$ are iid normal random variables with mean zero and variance $\omega^2$. For the estimation of the regression function, we use the kernel function

$$K(x) = \frac{3}{4}(1 - x^2)I_{[-1,1]}(x),$$

where $I(\cdot)$ denotes the indicator function, the bandwidth $h = h_n = n^{-1/3}/3$, and $l_n = [n^{1/4}]$ in estimating $s_n^2$. As before, we assume that the variance change occurs at $t = n/2$ and perform a test for the following hypotheses:

$H_0$: $\omega^2$ remains equal to 1 over $t = 1, \ldots, n$, vs.

$H_1$: $\omega^2$ changes from 1 to $\delta$ at $t = [n/2]$,

where $\delta$ takes the values 2, 4 and 9. Here we employ the sample size $n = 200$, 300, 500, and $\phi = 0$, 0.3, 0.5, 0.8 in order to see the correlation effect. The figures in Tables 2 and 3 denote the empirical sizes and powers, respectively. From the results, we can see that the test has no severe size distortions at moderate sample size, say, $n \geq 300$, and it produces good powers under $H_1$. The power depends on the values of $\phi$, which decreases as $\phi$ approaches 1 and when $\delta$ has lower values. As anticipated, it increases as either $\delta$ or $n$ increases. The results obtained in our simulation study enable us to conclude that the cusum of squares test performs adequately for the variance change in nonstationary and nonparametric time series models.

Table 1. Empirical sizes and powers of $T_n$.

| $\beta$ | | 0.2 | | | 0.5 | | | 0.8 | | | 1.0 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Delta$ | | size | 2 | 4 | size | 2 | 4 | size | 2 | 4 | size | 2 | 4 |
| | 200 | .033 | .818 | 1.00 | .030 | .818 | 1.00 | .037 | .788 | 1.00 | .037 | .826 | 1.00 |
| $n$ | 300 | .034 | .953 | 1.00 | .039 | .953 | 1.00 | .032 | .953 | 1.00 | .037 | .957 | 1.00 |
| | 500 | .042 | .998 | 1.00 | .050 | .997 | 1.00 | .042 | 1.00 | 1.00 | .042 | .998 | 1.00 |

Table 2. Empirical sizes of $T_n^*$.

| $\phi$ | | 0.0 | 0.3 | 0.5 | 0.8 |
|---|---|---|---|---|---|
| $n$ | 200 | .028 | .030 | .035 | .030 |
| | 300 | .026 | .034 | .041 | .040 |
| | 500 | .034 | .048 | .042 | .050 |

Table 3. Empirical powers of $T_n^*$.

| $\phi$ | | 0.0 | | | 0.3 | | | 0.5 | | | 0.8 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\delta$ | | 2 | 4 | 9 | 2 | 4 | 9 | 2 | 4 | 9 | 2 | 4 | 9 |
| $n$ | 200 | .559 | .981 | .998 | .511 | .978 | .998 | .466 | .949 | .995 | .318 | .851 | .978 |
| | 300 | .834 | 1.00 | 1.00 | .774 | .999 | 1.00 | .706 | .995 | 1.00 | .474 | .961 | .996 |
| | 500 | .990 | 1.00 | 1.00 | .975 | 1.00 | 1.00 | .940 | 1.00 | 1.00 | .725 | .999 | 1.00 |

## Acknowledgements

We like to thank the referees for their valuable comments. This research was supported (in part) by KOSEF through Statistical Research Center for Complex Systems at Seoul National University.

## References

Bai, J. (1994). Weak convergence of the sequential empirical processes of residuals in ARMA models, *Ann. Statist.*, **22**, 2051–2061.

Bosq, D. (1996). *Nonparametric Statistics for Stochastic Processes*; *Estimation and Prediction*, Lecture Notes in Statist., Springer, New York.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed., Springer, New York.

Brown, R. L., Durbin, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time, *J. Roy. Statist. Soc. Ser. B*, **37**, 149–163.

Chan, N. H. and Wei, C. Z. (1988). Limiting distribution of least squares estimates of unstable autoregressive processes, *Ann. Statist.*, **16**, 367–401.

Csörgő, M. and Horváth, L. (1988). Nonparametric methods for change point problems, *Handbook of Statistics* (eds. P. R. Krishnaiah and C. D. Rao), Vol. 7, 403–425, Elsevier, New York.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-point Analysis*, Wiley, Chichester.

Doukhan, P. (1994). *Mixing: Properties and Examples*, Springer, New York.

Gorodetskii, V. V. (1977). On the strong mixing properties for linear sequences, *Theory Probab. Appl.*, **22**, 411–413.

Hafner, C. M. (1998). *Nonlinear Time Series Analysis with Applications to Foreign Exchange Rate Volatility*, Physica, Heidelberg.

Hall, P. and Hart, J. D. (1990). Nonparametric regression with long range dependence, *Stochastic Process. Appl.*, **36**, 339–351.

Hart, J. D. (1991). Kernel regression estimation with time series errors, *J. Roy. Statist. Soc. Ser. B*, **53**, 173–187.

Hinkley, D. V. (1971). Inference about the change-point from cumulative sums tests, *Biometrika*, **58**, 509–523.

Inclán, C. and Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variances, *J. Amer. Statist. Assoc.*, **89**, 913–923.

Karatzas, I. and Shreve, S. E. (1988). *Brownian Motion and Stochastic Calculus*, Springer, New York.

Kim, S., Cho, S. and Lee, S. (2000). On the cusum test for parameter changes in GARCH(**1**,**1**) models, *Comm. Statist. Theory Methods*, **29**, 445–462.

Krishnaiah, P. R. and Miao, B. Q. (1988). Review about estimation of change points, *Handbook of Statistics* (eds. P. R. Krishnaiah and C. P. Rao), Vol. 7, 375–402, Elsevier, New York.

Lee, S. and Na, S. (2002). On the Bickel-Rosenblatt test for first order autoregressive models, *Statist. Probab. Lett.*, **56**, 21–35.

Lee, S. and Park, S. (2001). The cusum of squares test for scale changes in infinite order moving average processes, *Scand. J. Statist.*, **28**, 625–644.

Lee, S. and Wei, C. Z. (1999). On residual empirical processes of stochastic regression models with applications to time series, *Ann. Statist.*, **27**, 237–261.

Ling, S. (1998). Weak convergence of the sequential empirical processes of residuals in nonstationary autoregressive models, *Ann. Statist.*, **26**, 741–754.

Neumann, M. H. and Kreiss, J. P. (1998). Regression-type inference in nonparametric autoregression, *Ann. Statist.*, **26**, 1570–1613.

Page, E. S. (1955). A test for change in a parameter occurring at an unknown point, *Biometrika*, **42**, 523–527.

Park, S., Lee, S. and Jeon, J. (2000). The cusum of squares test for variance changes in infinite order autoregressive models, *J. Korean Statist. Soc.*, **29**, 351–361.

Picard, D. (1985). Testing and estimating change-points in time series, *Advances in Applied Probability*, **17**, 841–867.

Priestley, M. B. and Chao, M. T. (1972). Non-parametric function fitting, *J. Roy. Statist. Soc. Ser. B*, **34**, 385–392.

Truong, Y. K. and Stone, C. J. (1992). Nonparametric function estimation involving time series, *Ann. Statist.*, **20**, 77–97.

Wichern, D. W., Miller, R. B. and Hsu, D. A. (1976). Changes of variance in first-order autoregressive time series models—with an application, *Appl. Statist.*, **25**, 248–256.

Wu, J. S. and Chu, C. K. (1994). Nonparametric estimation of a regression function with dependent observations, *Stochastic Process. Appl.*, **50**, 149–160.

Zacks, S. (1983). Survey of classical and Bayesian approaches to the change-point problem: Fixed sample and sequential procedures of testing and estimation, *Recent Advances in Statistics* (eds. M. H. Rizvi, J. S. Rustagi and D. Siegmund), 245–269, Academic Press, New York.