TAR(p)/ARCH(1) Process: Guaranteed Parameter Estimation and Change-Point Detection

Yulia Burkatovskaya, Ekaterina Sergeeva, and Sergei Vorobeychikov

***Abstract*—A sequential method of unknown autoregressive parameters estimation of TAR(p)/ARCH(1) model, which all are assumed to be unknown, is presented. This procedure is based on the construction of the special stopping rule and weights for weighted least square estimation method, which allow us to guarantee the prescribe accuracy of the estimation. Also a sequential procedure of change point detection is proposed. Upper bounds for its basic characteristics, such as the probability of false alarm and the delay probability, are obtained.**

***Index Terms*—AR/ARCH, guaranteed parameter estimation, change point detectionAR/ARCH, guaranteed parameter estimation, change point detectionT**

I. Introduction

Threshold autoregressive (TAR) models proposed by Tong [1] definitely are one of the most popular classes of nonlinear time series models for conditional mean, because they not only provide a better fit than linear models, but also reveal strictly nonlinear behavior (e.g. limit cycles, jump resonance, harmonic distortion) which linear models cannot duplicate [2]. But sometimes such models have to be completed by a specification of the conditional variance. ARCH/GARCH type models first introduced by Engle [3] are often considered for the conditional variance.

Estimators of the unknown parameters based on the idea of usage of a special stopping rule in order to guarantee precisely their quality in a special sense were first proposed in [4] and are also very popular. So, in [5] a sequential procedure for estimation of parameters in TAR(1) model, which allows one to construct LSE asymptotically risk efficient estimator, was proposed. In [6] sequential sampling methods was used for construction of confidence intervals for unknown parameters with fixed size and prescribed coverage probability. Konev and Galtchouk [7] proposed sequential least square estimator with stopping rule determined by the trace of the observed Fisher information matrix, which is asymptotically normally distributed in the stability region. The sequential procedure for estimation of unknown parameters of TAR(1)/ARCH(1) process, which can guarantee precise accuracy of estimators was considered in [8]. The problem of change point detection for autoregressive processes with conditional heteroskedasticity is well known and extremely interesting. With different assumption and for different types of models such problem was recently considered for example in [8], [9], [10]. Properties of commonly used algorithms are studied asymptotically or by simulations, because theoretical investigation is extremely difficult or even impossible. Usage of the special stopping rule for construction of the LSE with guaranteed accuracy of unknown parameters allows us to investigate both asymptotic and non-asymptotic properties of algorithms, such as false alarm and delay probabilities. In this paper the guaranteed weighted least square estimators of unknown autoregressive parameters of TAR(p)/ARCH(1) process are proposed. Asymptotic properties for the estimators are considered and the upper bounds for standard deviation (asymptotic and non-asymptotic) are constructed. The procedure of change point detection with guaranteed characteristics for this process is presented.

Manuscript received December 8, 2015; revised February 7, 2016. This paper is supported by The National Research Tomsk State University Academic D.I. Mendeleev Fund Program (NU 8.1.55.2015 L) in 2014–2015 and by RFBR Grant 16-01-00121.

Yu. Burkatovskaya is with the Institute of Cybernetics, Tomsk Polytechnical University, 30 Lenin Prospekt, 634050 Tomsk, Russia; Department of Applied Mathematics and Cybernetics and International Laboratory of Statistics of Stochastic Processes and Quantitative Finance, Tomsk State University, 36 Lenin Prospekt, 634050 Tomsk, Russia, e-mail: tracey@tpu.ru.

E. Sergeeva is with the Institute of Cybernetics, Tomsk Polytechnical University, 30 Lenin Prospekt, 634050 Tomsk, Russia; e-mail: sergeeva_e_e@mail.ru.

S. Vorobeychikov is with the Institute of Cybernetics, Tomsk Polytechnical University, 30 Lenin Prospekt, 634050 Tomsk, Russia; Department of Applied Mathematics and Cybernetics and International Laboratory of Statistics of Stochastic Processes and Quantitative Finance, Tomsk State University, 36 Lenin Prospekt, 634050 Tomsk, Russia, e-mail: sev@mail.tsu.ru.

II. Problem Statement

We consider TAR(p)/ARCH(1) autoregressive process specified by the equation

$$\begin{gathered} x_k = X_k \Lambda^1 \mathbf{1}_{\{x_{k-1}\geq 0\}} + X_k^T \Lambda^2 \mathbf{1}_{\{x_{k-1}<0\}} \\ +\sqrt{\omega + \alpha^2 x_k^2}\xi_k; \\ X_k = [x_{k-1}, ..., x_{k-p}]; \\ \Lambda^j = [\lambda_1^j, ..., \lambda_p^j]^T, \quad j = 1, 2, \end{gathered} \tag{1}$$

where $\{\xi_k\}_{k\geq 0}$ is a sequence of independent identically distributed random variables with zero mean and unit variance, $\omega > 0$, $0 < \alpha^2 < 1$, $\mathbf{1}_A$ is the indicator of the set A, T is the transposition sign. One can see that the process under consideration is the p-order autoregressive process with ARCH noise and the autoregressive parameters depending on the previous value of the process. The value of the parameter vector $\Lambda = [\Lambda^1, \Lambda^2]$ changes from $\mu = [\mu_1^1, ..., \mu_p^1, \mu_1^2, ..., \mu_p^2]$ to $\beta = [\beta_1^1, ..., \beta_p^1, \beta_1^2, ..., \beta_p^2]$ at the change point θ. Values of the parameters before and after θ are supposed to be unknown. The difference between μ and β satisfies the condition

$$(\mu^j - \beta^j)^T(\mu^j - \beta^j) \geq \Delta, \quad j = 1, 2, \tag{2}$$

where Δ is the known value defining the minimum difference between the parameters before and after the change point. The problem is to detect the change point θ from observations x_k.

In [16] and [17] we proposed to detect the instant of parameters change in autoregressive process by making use of guaranteed sequential estimators. The sequence of estimators

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

is constructed and the estimators obtained on different time intervals are compared. In this paper we apply this approach to more complicated model with unbounded noise variance.

III. Ergodicity of the Process

For investigation of asymptotic properties of estimators of unknown parameters of given models it is important to obtain necessary and sufficient conditions for ergodicity or even strongly for geometric ergodicity of such models . There are three main approaches to establish geometrical ergodicity in nonlinear conditionally heterockedastic autoregression [2]. The first approach is based on the assumption that linear regression part becomes main part for the stability research due to usage of infinite number of values of the process considered and the assumption that the radius of the companion matrix of this linear regression part is less than one [11], [12] for AR/ARCH model. The second one uses the concept of the Lyapunov exponent for AR/ARCH [13] and TAR/ARCH [14] and allows one to obtain geometric ergodicity within more general assumptions in much larger parameters space than in [11], [12] but the assumptions appear much more difficult to validate. The last one is approach, which first was proposed by Liebscher [15] and then extended for AR/ARCH model [3], based on concept of the joint spectral radius of a set of matrices and also allows to obtain geometric ergodicity in lager regions of parameter space than [11], [12].

IV. Guaranteed parameter estimator

Since the parameters both before and after the change point are unknown, it is logical to use estimators of the unknown parameters in the change point detection procedure. In this section we construct guaranteed sequential parameter estimators for the parameter vectors $\Lambda^j, j = 1, 2$. Such estimators were proposed in [17] for an autoregressive process. The main advantage of the estimators is their preassigned mean square accuracy depending on the parameter of the estimation procedure.

It should be noted that if parameters ω and α are unknown then process (1) has unknown and unbounded from above noise variance. To obtain a process with bounded noise variance we denote $\max\{1, |x_{k-1}|\}$ as m_k and rewrite the process in the form

$$\begin{gathered} y_k = Y_k^1\Lambda^1 + Y_k^2\Lambda^2 + \gamma_k\xi_k; \\ Y_k^1 = \frac{1}{m_k}X_k\mathbf{1}_{\{x_{k-1}\geq 0\}}. \quad Y_k^2 = \frac{1}{m_k}X_k\mathbf{1}_{\{x_{k-1}<0\}}; \\ y_k = \frac{x_k}{m_k}, \quad \gamma_k = \frac{\sqrt{w + \alpha^2 x_{k-1}^2}}{m_k}. \end{gathered} \tag{3}$$

The noise variance of the process $\{y_k\}$ is bounded from above by the unknown value $(\omega + \alpha^2)$. To eliminate the influence of the unknown constant in [8] we proposed to use the special factor Γ_N constructed by first N observations in the following form

$$\begin{gathered} \Gamma_N = C_N \sum_{k=1}^{N} \left(\frac{x_k}{\min\{1, |x_k|\}}\right)^2; \\ C_N = E\left(\sum_{k=1}^{N}\xi_k^2\right)^{-1}, \end{gathered} \tag{4}$$

where N observations are taken at the interval where all the values $|x_k|$ are sufficiently large. It was proved in [8] that for $p = 1$ the compensating factor satisfies the following condition

$$E\frac{1}{\Gamma_N} \leq \frac{1}{\omega + \alpha^2}. \tag{5}$$

This proof can be generalized for the case $p > 1$ with minimum changes so we omit it.

If $\{\xi_k\}$ have standard normal distribution then the sum $\sum_{k=1}^{N}\xi_k^2$ has χ^2 distribution with N degrees of freedom. In this case one has

$$C_N = \frac{1}{2^{N/2}\Gamma(N/2)}\int_0^{+\infty} x^{N/2-3}e^{-x/2}dx = \frac{1}{(N-2)(N-4)}.$$

This constant is defined for $N \geq 5$.

Consider now a weighted least squares estimator for process (3). The process can be rewritten in the form

$$\begin{gathered} y_k = Y_k\Lambda + \gamma_k\xi_k; \\ Y_k = [Y_k^1, Y_k^2]. \end{gathered} \tag{6}$$

So a weighted least squares estimator has the following form

$$\begin{gathered} \hat{\Lambda} = C^{-1}(m)\sum_{k=N+1}^{m} v_k(Y_k)^T y_k; \\ C(m) = \sum_{k=N+1}^{m} v_k(Y_k)^T Y_k, \end{gathered} \tag{7}$$

where $0 < v_k \leq 1$. According to definition (3) $(Y_k^j)^T Y_k^i = O_p$ for $i \neq j$ (here O_p stands for a zero matrix of the order p). Hence, taking into account (6) one obtains that the matrix $C(m)$ has a block structure

$$\begin{gathered} C(m) = \begin{bmatrix} C(1,m) & O_p \\ O_p & C(2,m) \end{bmatrix} \\ C^{-1}(m) = \begin{bmatrix} C^{-1}(1,m) & O_p \\ O_p & C^{-1}(2,m) \end{bmatrix} \\ C(j,m) = \sum_{k=N+1}^{m} v_{j,k}(Y_k^j)^T Y_k^j, \quad j = 1, 2. \end{gathered}$$

Using this result and (3), (6) one can obtain

$$\begin{aligned} \hat{\Lambda}^1 &= [C^{-1}(1,m)\ O_p]\sum_{k=N+1}^{m} v_k(Y_k)^T y_k \\ &= [C^{-1}(1,m)\ O_p]\sum_{k=N+1}^{m} v_k(Y_k)^T(Y_k\Lambda + \gamma_k\xi_k) \\ &= [C^{-1}(1,m)\ O_p]\begin{bmatrix} C(1,m)\Lambda_1 \\ C(2,m)\Lambda_2 \end{bmatrix} \\ &+ [C^{-1}(1,m)\ O_p]\sum_{k=N+1}^{m} v_k(Y_k)^T\gamma_k\xi_k. \end{aligned}$$

Hence,

$$\begin{gathered} \hat{\Lambda}^1 = \Lambda_1 + C^{-1}(1,m)\eta(1,m); \\ \eta(1,m) = \sum_{k=N+1}^{m} v_k(Y_k^1)^T\gamma_k\xi_k. \end{gathered} \tag{8}$$

The same result can be obtained for $\hat{\Lambda}^2$. It allows us to construct estimators for Λ^1 and Λ^2 separately, i.e.

$$\hat{\Lambda}^j = C^{-1}(j;m)\sum_{k=N+1}^{m} v_k(Y_k^j)^T y_k, \quad j = 1, 2. \tag{9}$$

The obtained estimator can be modified in order to bound the standard deviation of the estimator from above. It can be done if we change the sample size m for a special random stopping time τ^j and if we use special weights $v_{j,k}$ for every estimator. Let $H > 0$ be a parameter of the estimation

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

procedure. Further we prove that it defines the accuracy of the proposed parameter estimators. Then the estimators are constructed using the weighted least squares method and hence can be written in the following form

$$\begin{aligned}\hat{\Lambda}^j = \hat{\Lambda}^j(H) = C^{-1}(j;\tau^j) \sum_{k=N+1}^{\tau^j} v_{j,k}(Y_k^j)^T y_k;\\ C(j,m) = \sum_{k=N+1}^{m} v_{j,k}(Y_k^j)^T Y_k^j, \quad j=1,2.\end{aligned} \tag{10}$$

Let $\nu_{\min}(j,m)$ be the minimum eigenvalue of the matrix $C(j,m)$. Then the stopping time $\tau^j = \tau^j(H)$ are defined by the following conditions

$$\tau^j = \inf\left(m > N : \ \nu_{\min}(j,m) \geq H\right). \tag{11}$$

Now we consider the choice of the weights $v(j,k)$. Let for $m = N+1, ..., N+\sigma$ the matrix $C(j,m)$ is degenerate and $C(j,m+1)$ is not degenerate. The weights on the interval $[N+1, N+\sigma]$ are taken in the following form

$$v(j,k) = \begin{cases} \dfrac{1}{\sqrt{\Gamma_N Y_k^j (Y_k^j)^T}}, & \begin{array}{l}\text{if } Y_N^j, \dots, Y_k^j\\ \text{are linearly}\\ \text{independent;}\end{array}\\ 0, & \text{otherwise.}\end{cases} \tag{12}$$

The weights on the intervals $[N+\sigma+1, \tau^j - 1]$ are found from the following condition:

$$\frac{\nu_{\min}(j,k)}{\Gamma_N} = \sum_{l=N+\sigma}^{k} v^2(j,l) Y_l^j (Y_l^j)^T. \tag{13}$$

At the instants τ^j, the weights are determined by the condition:

$$\begin{aligned}\frac{\nu_{\min}(j,\tau^j)}{\Gamma_N} \geq \sum_{l=N_1+n+\sigma}^{\tau^j} v^2(j,l) Y_l^j (Y_l^j)^T;\\ \nu_{\min}(j,\tau^j) = H.\end{aligned} \tag{14}$$

Theorem 1. Let the parameter vector Λ^j in (1) be constant. Then the stopping time τ^j (11) is finite with probability one and the mean square accuracy of estimator (10) is bounded from above

$$E||\hat{\Lambda}^j(H) - \Lambda^j||^2 \leq \frac{H+p-1}{H^2}. \tag{15}$$

Proof. According to the definition of the instant τ^j (11) it is finite with probability one if

$$\sum_{l=N+\sigma}^{k} v^2(j,l) Y_l^j (Y_l^j)^T \to \infty \text{ as } k \to \infty. \tag{16}$$

The series converges if and only if $\forall \varepsilon > 0$ as $M \to \infty$ (see [18])

$$P\left\{ \sum_{l \geq M} v^2(j,l) Y_l^j (Y_l^j)^T \geq \varepsilon \right\} \to 0. \tag{17}$$

The factor $Y_l^j (Y_l^j)^T$ does not tend to zero because the absolute value of first component is equal to $|x_{l-1}|/\max\{1, |x_{l-1}|\}$ and, according to equation (1), $|x_{l-1}|$ exceeds unity with non-zero probability and can be both negative and positive. So condition (17) can hold true only because of the choice of the weights $v(j,l)$.

Suppose that matrix $C(j,M-1)$ is not diagonal. According to the definition of the minimum eigenvalue of matrix

$$\nu_{\min}(j,M) = \min_{x:||x||=1} (x, C(j,M)x),$$

where (x,y) is the scalar product of the vectors x and y. Then using (10), we obtain

$$\begin{aligned}&\nu_{\min}(N_1, N)\\ &= \min_{x:||x||=1} (x, ((C(j,M-1) + v(j,M)(Y_M^j)^T Y_M^j)x)\\ &= \min_{x:||x||=1} ((x, C(j,M-1)x) + v_N(Y_M^j x)^2).\end{aligned}$$

Let z_M be the argument of the minimum in the last equation. According to (13), we obtain

$$\begin{aligned}&(z_M, C(j,M-1)z_M) + v(j,M)(Y_M^j z_M)^2\\ &= \nu_{\min}(j,M-1) + v(j,M)^2 Y_M^j (Y_M^j)^T.\end{aligned}$$

So we have derived the quadratic equation for $v(j,M)$ with roots in the form

$$\begin{aligned}v_{1,2} = \frac{1}{2Y_M^j(Y_M^j)^T} \Big[(Y_M^j z_M)^2 \pm \Big((Y_M^j z_M)^4\\ +4Y_M^j(Y_M^j)^T((z_M, C(j,M-1)z_M) - \nu_{\min}(j,M-1))\Big)^{1/2}\Big].\end{aligned}$$

It is obvious that

$$(z, C(j,M-1)z) - \nu_{\min}(j,M-1) \geq 0.$$

Thus the following two cases are possible.

Case 1. The equation has two zero roots: $v_1 = v_2 = 0$. This is possible if and only if z is the eigenvector of the matrix $C(j,M-1)$ corresponding to $\nu_{\min}(j,M-1)$ and $Y_M^j z = 0$. However, the first component of $Y_M^j z$ depends on the random variable ξ_M, which is independent on the $\{x_k\}_{k<N}$. Hence the vector Y_M^j is orthogonal to the given eigenvector of the matrix $C(j,M-1)$ with zero probability.

Case 2. The equation has one non-positive and one positive root. Taking the major root as $v(j,M)$, one obtains

$$\begin{aligned}v(j,M)^2 Y_M^j (Y_M^j)^T \geq \frac{(Y_M^j z_M)^4}{Y_M^j (Y_M^j)^T}\\ +(z_M, C(j,M-1)z_M) - \nu_{\min}(j,M-1).\end{aligned} \tag{18}$$

The first term in (18) is equal to $Y_M^j (Y_M^j)^T \cos^4(\alpha_M)/2$, where α_M is the angle between Y_M^j and z_M. Since $Y_M^j (Y_M^j)^T$ does not converge to zero, the first term converges to zero if and only if $\cos(\alpha_M) \to 0$ when $M \to \infty$. On the other hand, if the second term in (18) converges to zero then z_M converges to the eigenvector of the matrix $C(j,M-1)$ corresponding to $\nu_{\min}(j,M-1)$. If $v(j,M) \to 0$, then the matrix $C(j,M)$ changes slightly with increasing M. Hence, the eigenvectors of the matrix change slightly too, and z_M converges to a certain vector z^*. Therefore, the right side of (18) converges to zero if the cosine of the angle between Y_M^j and z^* converges to zero. However, the first component of Y_M^j depends on x_{M-1} which can take any value, this cosine can be sufficiently large with non-zero probability.

Note that condition (17) can hold true if all eigenvalues of the matrix $C(j,M-1)$ are equal for certain M. It is possible if and only if the matrix $C(j,M-1)$ is diagonal. The matrix $C(j,N+k) = v_{j,N+k}(Y_{N+k}^j)^T Y_{N+k}^j$, where k is the least number such as Y_{N+k}^j is non-zero, is not diagonal. It can be easily proved that if the matrix $C(j,M-1)$ is not

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

diagonal, then the next matrix $C(j, M)$ is diagonal with zero probability.

Hence, condition (17) does not hold true for AR process (1), and this implies (16).

According to (8) one has

$$\hat{\Lambda}^j = \Lambda^j + C^{-1}(j, \tau^j)\eta(j, \tau^j);$$
$$\eta(j, \tau^j) = \sum_{k=N+1}^{\tau^j} v_{j,k}(Y_k^j)^T \gamma_k \xi_k.$$

Using the norm properties and (13) one has

$$||\hat{\Lambda}^j(H) - \Lambda^j||^2 \le (\nu_{\min}(j, \tau^j))^{-2}||\eta(j, \tau^j)||^2 \le \frac{||\eta(j, \tau^j)||^2}{H^2}. \tag{19}$$

Let $F_k = \sigma\{\xi_1, ..., \xi_k\}$ be a sigma-algebra generated by the random variables $\{\xi_1, ..., \xi_k\}$ and $\tau^j(M) = \min\{\tau^j, M\}$ is a truncated stopping instant. According to (11) the instant $\{\tau^j(M) = k\} \in F_{k-1}$. Using the properties of conditional expectations one obtains

$$\begin{aligned} &E||\eta(j, \tau^j(M))||^2 \\ &= E\sum_{k=N+1}^{M} E[v_{j,k}^2 Y_k^j (Y_k^j)^T \gamma_k^2 \xi_k^2 \mathbf{1}_{\tau^j \le k}|F_{k-1}] \\ &+2E\sum_{k=N+2}^{M}\sum_{l=N+1}^{k-1} E[v_{j,k}v_{j,l}Y_k^j(Y_l^j)^T\gamma_k\gamma_l\xi_k\xi_l\mathbf{1}_{\tau^j\le k}|F_{k-1}] \\ &= E\sum_{k=N+1}^{M} v_{j,k}^2 Y_k^j (Y_k^j)^T \gamma_k^2 \mathbf{1}_{\tau^j \le k} E[\xi_k^2|F_{k-1}] \\ &+2E\sum_{k=N+2}^{M}\sum_{l=N+1}^{k-1} v_{j,k}v_{j,l}Y_k^j(Y_l^j)^T\gamma_k\gamma_l\xi_l\mathbf{1}_{\tau^j\le k}E[\xi_k|F_{k-1}]. \end{aligned}$$

Since ξ_k does not depend on F_{k-1} the second summand is equal to zero and one obtains

$$E||\eta(j, \tau^j(M))||^2 = E\sum_{k=N+1}^{M} v_{j,k}^2 Y_k^j (Y_k^j)^T \gamma_k^2.$$

Due to the choice of the weights $v_{j,k}$ (12–13) one obtains

$$\begin{aligned} E\sum_{k=N+1}^{\tau^j} v_{j,k}^2 Y_k^j (Y_k^j)^T &= E\sum_{k=N+1}^{N_\sigma} v_{j,k}^2 Y_k^j (Y_k^j)^T \\ +E\sum_{k=N+\sigma+1}^{\tau^j} v_{j,k}^2 Y_k^j (Y_k^j)^T &\le \frac{p-1}{\Gamma_N} + \frac{H}{\Gamma_N}. \end{aligned}$$

According to (3) one can see that $\gamma_k^2 \le \omega + \alpha^2$. Note that $\tau^j(M) \to \tau^j$ as $M \to \infty$, hence

$$E||\eta(j, \tau^j)||^2 \le (\omega + \alpha^2)(H + p - 1)E\frac{1}{\Gamma_N}.$$

Due to property (5) of the factor Γ_N and inequality (19) one obtains

$$||\hat{\Lambda}^j(H) - \Lambda^j||^2 \le E\frac{H + p - 1}{H^2}.$$

The theorem has been proved.

Theorem 2. If process (1) is ergodic, and the compensating factor Γ_N satisfies the following conditions

$$N \to \infty, \quad N/H \to 0 \text{ as } H \to \infty,$$

then for sufficiently large H

$$\mathcal{P}\left\{\left|\left|\hat{\Lambda}^j - \Lambda\right|\right|^2 > x\right\} \le 1 - \left(2\Phi\left(\sqrt{\frac{xH^2}{H+p-1}}\right) - 1\right)^p, \tag{20}$$

where $\Phi(\cdot)$ is the standard normal distribution function

An analogous theorem for the AR(p) process with unknown variance has been proved in [17]. Theorem 2 can be proved in much the same way. We omit the proof because of the limited volume of the paper.

V. Change Point Detection Procedure

Consider now the change point detection problem for process (1). At the first stage, we define intervals $[\tau_{n-1}^j + 1, \tau_n^j]$, $n \ge 1$. The estimators $\hat{\Lambda}_n^j$ of the parameters of process (1) are constructed on each interval. Then the estimators on intervals $[\tau_{n-l-1}^j + 1, \tau_{n-l}^j]$ and $[\tau_{n-1}^j + 1, \tau_n^j]$, where $l > 1$ is an integer, are compared. If the interval $[\tau_{n-l-1}^j + 1, \tau_n^j]$ does not include the change point θ, then vector Λ^j on this interval is constant. It can be equal to the initial value μ^j or the final value β^j. Thus for certain n, if $\tau_{n-l}^j < \theta < \tau_{n-1}^j + 1$, the difference between values of the parameters on intervals $[\tau_{n-l-1}^j + 1, \tau_{n-l}^j]$ and $[\tau_{n-1}^j + 1, \tau_n^j]$ is no less then Δ. This is the key property for the change point detection.

We construct a set of sequential estimation plans

$$(\tau_n^j, \hat{\Lambda}_n^j) = (\tau_n^j(H), \hat{\Lambda}_n^j(H)), \quad n \ge 1, \quad j = 1, 2,$$

where $\{\tau_n^j\}$, $n \ge 0$ is the increasing sequence of the stopping instances ($\tau_0 = N$), and $\hat{\Lambda}_n^j$ is the guaranteed parameter estimator on the interval $[\tau_{n-1}^j + 1, \tau_n^j]$. Then we choose an integer $l > 1$ and define the statistics I_n^j

$$I_n^j = ||\hat{\Lambda}_n^j - \hat{\Lambda}_{n-l}^j||^2. \tag{21}$$

This statistic is the squared deviation of the estimators with numbers n and $n - l$. Properties of the statistics are given in the following theorem.

Theorem 3. The expectation of the statistics I_n^j (21) satisfies the following inequality:

$$\begin{aligned} E\left[I_n^j\middle|\,\tau_n^j < \theta\right] &\le \frac{4(H + p - 1)}{H^2}; \\ E\left[I_n^j\middle|\,\tau_{n-l}^j < \theta \le \tau_{n-1}^j\right] &\ge \Delta. \end{aligned} \tag{22}$$

Proof. Denote the deviation of the estimator $\hat{\Lambda}_n^j$ from the true value of the parameter Λ^j as ζ_n^j. Let the parameter value remain unchanged until the instant τ_n^j, i.e., $\theta > \tau_n^j$. In this case, $\hat{\Lambda}_n^j = \mu^j + \zeta_n^j$, $\hat{\lambda}_{n-l}^j = \mu^j + \zeta_{n-l}^j$ and statistic (21) can be written in the form

$$I_n^j = \left|\left|\zeta_n^j - \zeta_{n-l}^j\right|\right|^2.$$

According to the Theorem 1,

$$E||\zeta_n^j||^2 \le \frac{H + p - 1}{H^2} \tag{23}$$

To estimate the expectation of the statistic, we use property (23) and the inequality $||a - b||^2 \le 2||a||^2 + 2||b||^2$:

$$EI_n^j \le E\left(2||\zeta_n^j||^2 + 2||\zeta_{n-l}^j||^2\right) \le 4\frac{P(H)}{H^2}. \tag{24}$$

Let the change of the parameter take place on the interval $[\tau_{n-l}^j, \tau_{n-1}^j]$ i.e. $\tau_{n-l}^j < \theta \le \tau_{n-1}^j$. In this case, $\hat{\Lambda}_n^j = \beta^j + \zeta_n^j$, $\hat{\Lambda}_{n-l}^j = \mu^j + \zeta_{n-l}^j$, and statistic (21) is

$$I_n^j = ||\beta^j - \mu^j + \zeta_n^j - \zeta_{n-l}^j||^2.$$

To estimate the expectation of the statistics, we take advantage of the inequality $||a-b|| \geq ||a|| - ||b||$ and condition (23)

$$\begin{gathered} EI_n^j \geq E\left(||\beta^j - \mu^j|| - ||\zeta_n^j - \zeta_{n-l}^j||\right)^2 \\ \geq ||\beta^j - \mu^j||^2 - 2||\beta^j - \mu^j||\, E||\zeta_n^j - \zeta_{n-l}^j|| \\ \geq \Delta - 4\sqrt{\Delta \frac{H+p-1}{H^2}}. \end{gathered}$$

The theorem has been proved.

Hence, the change of the expectation of the statistic I_n^j allows us to construct the following change point detection algorithm. The I_n^j values are compared with a certain threshold δ, where $4(H+p-1)/H^2 < \delta < \Delta$. When the value of the statistic exceeds δ then the change point is considered to be detected. If at least one parameter of the vector $\Lambda = [\Lambda_0, \Lambda_1]$ changes then the change point θ can be detected.

The probabilities of false alarm and delay in the change point detection in any observation cycle are important characteristics of any change point detection procedure. Due to the application of the guaranteed parameter estimators in the statistics, we can bound these probabilities from above.

Theorem 4. The probability of false alarm $P_{0,n}$ and the probability of delay $P_{1,n}$ in n-th observation cycle $[\tau_{n-1}^j + 1, \tau_n^j]$ are bounded from above

$$\begin{aligned} P_{0,n} &\leq \frac{2(H+p-1)}{\delta H^2}; \\ P_{1,n} &\leq \frac{2(H+p-1)}{(\sqrt{\Delta}-\sqrt{\delta})^2 H^2}. \end{aligned} \tag{25}$$

Proof. First we consider the false alarm probability, i.e. the probability that the statistic J_i exceeds the threshold before the change point. Using the norm properties and the Chebyshev inequality, we obtain

$$\begin{gathered} P_{0,n} = \mathcal{P}\left\{ I_n^j > \delta \,\middle|\, \tau_n^j < \theta \right\} \\ = \mathcal{P}\left\{ ||\zeta_n^j - \zeta_{n-l}^j||^2 > \delta \right\} \leq \frac{2E\left(||\zeta_n^j||^2 + ||\zeta_{n-l}^j||^2\right)}{\delta}. \end{gathered}$$

This and (23) imply the first inequality from (25).

Then we consider delay probability, i.e., the probability that the statistic I_n^j does not exceed the threshold after the change point

$$\begin{gathered} P_{1,n} = \mathcal{P}\left\{ I_n^j < \delta \,\middle|\, \tau_{n-l}^j < \theta < \tau_{n-1}^j \right\} \\ = \mathcal{P}\left\{ ||\beta^j - \mu^j + \zeta_n^j - \zeta_{n-l}^j||^2 < \delta \right\} \\ = \mathcal{P}\left\{ || + \zeta_n^j - \zeta_{n-l}^j|| < \sqrt{\delta} \right\}. \end{gathered}$$

Taking into account that $||\beta^j - \mu^j||^2 > \Delta$ and using the norm properties and the Chebyshev inequality, one obtains

$$\begin{aligned} P_{1,n} &\leq \mathcal{P}\left\{ ||\beta^j - \mu^j|| - ||\zeta_n^j - \zeta_{n-l}^j|| < \sqrt{\delta} \right\} \\ &\leq \mathcal{P}\left\{ \sqrt{\Delta} - ||\zeta_n^j - \zeta_{n-l}^j|| < \sqrt{\delta} \right\} \\ &= \mathcal{P}\left\{ ||\zeta_n^j - \zeta_{n-l}^j|| > \sqrt{\Delta} - \sqrt{\delta} \right\} \\ &\leq \frac{2E\left(||\zeta_n^j||^2 + ||\zeta_{n-l}^j||^2\right)}{(\sqrt{\Delta}-\sqrt{\delta})^2}. \end{aligned}$$

This and (23) imply the second inequality from (25).

The theorem has been proved.

Then we consider asymptotic properties of the proposed change point detection procedure for $H \to \infty$ if process (1) is ergodic, i.e. the asymptotic inequalities for the probabilities of false alarm and delay.

Theorem 5. If process (1) is ergodic, and the compensating factor Γ_N satisfies the following conditions $N \to \infty$, $N/H \to 0$ as $H \to \infty$, then for sufficiently large H the probabilities of false alarm and delay in n-th observation cycle $[\tau_{n-1}^j + 1, \tau_n^j]$ are bounded from above

$$\begin{aligned} P_{0,n} &\leq 1 - \left(2\Phi\left(\frac{\delta H}{\sqrt{2(H+p-1)}}\right) - 1\right)^p; \\ P_{1,n} &\leq 1 - \left(2\Phi\left(\frac{\left(\sqrt{\Delta}-\sqrt{\delta}\right)H}{\sqrt{2(H+p-1)}}\right) - 1\right)^p. \end{aligned} \tag{26}$$

where $\Phi(\cdot)$ is the standard normal distribution function

These results can be obtained along the lines of the proof of Theorem 4 if instead of the Chebyshev inequality one uses the result of the Theorem 2.

These estimators can be used instead of (25) for sufficiently large H.

VI. Conclusion

The results in this paper were derived with strong mathematical evidence and are theoretical, but it can be very interesting to test them on the real data. So we plan to do simulation experiments in order to demonstrate efficiency of the procedures proposed.

References

[1] H. Tong , "On the Threshold Model," in Chen. C.H.(Ed.), Pattern Recognition and Signal Processing. Sijthoff and Noordhoff, Amsterdam.1978.

[2] M. Meitz, P. Saikkonen, "A note on the geometric ergodicity of nonlinear AR-ARCH model," Statistics and Probability Letters, vol.80,pp. 631-638, 2010.

[3] R.F.Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kindom inflation," Econometrica, vol.50,pp.987-1008,1982.

[4] A.Vald, "Sequential analysis", Phismatgis, Moscow,1960.

[5] S.Lee, T.N.Sriram, "Sequential point estimation of parameters in a threshold AR(1) model," StochasticProcessesand their Applications,vol.84, pp.343-355,1999.

[6] T.N. Sriram, "Fixed size confidence regions for parameters of threshold AR(1) models," Journal of Statistical Planning and Inference, vol.97, pp.293-304, 2001.

[7] V.Konev, L. Galtchouk, "On asymptotic normality of sequential LS-estimate for unstable autoregressive process AR(2)," Journal of Multivariate Analysis, vol.101, pp.2616-2636,2010.

[8] S.Vorobeychikov, J.Burkatovskaya, E.Sergeeva, "On Guaranteed Sequential Change point Detection for TAR(1)/ARCH(1) process," Communications in Computer and Information Science, vol.487, pp.59-68,2014.

[9] P. Fryzlewicz, S.Subba Rao, "Multipue-change-point detection for autoregressive conditional heteroscedastic processes," Journal of the Royal Statistical Society.,Series B,Statistical Methodolog.,pp.1-22,2014.

[10] O.Na, J. Lee, S. Lee, "Change Point Detection in Copula ARMA-GARCH Models," Journal of Time Series Analysis, vol.33,no.4,pp.554-569,2012.

[11] G.Chen, M.Chen,"On a class of nonlinear AR(p) models with nonlinear ARCH errors," Australian and New Zealand Journal of Statistics, vol.43,pp.445-454,2001.

[12] Z.Lu, Z.Jiang,"On the geometric ergodicity of a non-linear autoregressive model with an autoregressive conditional heteroscedastic term," Statistics and Probability Letters, vol.51,pp.121-130,2001.

[13] D.B.H. Cline, "Evaluating the Lyapounov exponent and existence of moments for threshold AR-ARCH models," Journalof Time Series Analysis, vol.28,pp.241-260,2007.

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

[14] D.B.H. Cline, H.H. Pu, ”Stability and the Lyapounov exponent of threshold AR-ARCH models,” Annals of Applied Probability,vol.14,pp.1920-1949,2004.
[15] E.Liebscher,”Towards a unified approach for proving geometrical ergodicity and mixing properties of nonlinear autoregressive processes,” Journal of Time Series Analysis,vol.26, pp.669-689,2005.
[16] Sergeeva, E.E., S. Vorobejchikov, S.E.. An Efficient Algorithm for Detecting a Change Point of Autoregressive Parameters of AR(p)/ARCH(q) Process. In:Proceedings of the 11th International Conference of Pattern Recognition and Information Processing, 156–159 (2011).
[17] Burkatovskaya, Y.B., Vorobeychikov, S.E.: Change Point Detection of Autoregressive Process with Unknown Parameters. In: Preprints of the 18th IFAC World Congress, 13215–13220 (2011).
[18] Shiryaev A.N.: Probability.Moscow: Nauka, Chief Editorial Staff of Physical-Mathematical Literature, (1980).

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)