GUARANTEED CHANGE POINT DETECTION OF LINEAR AUTOREGRESSIVE PROCESSES WITH UNKNOWN NOISE VARIANCE[1]

S. E. Vorobeychikov[2], Y. B. Burkatovskaya[3]
[2] *Tomsk State University*
[3] *Tomsk Polytechnic University*
Tomsk, RUSSIA
e-mail: [2]sev@mail.tsu.ru, [3]tracey@tpu.ru

Abstract

The problem of change point detection of autoregressive processes with unknown parameters is considered. A sequential procedure with guaranteed quality is proposed and both asymptotic and non-asymptotic properties of the algorithm are studied.

1 Introduction and problem statement

The problem of sequential change point detection for autoregressive processes often arises in different applications connected with time series analysis. The most difficult case is the case when all the process parameters are unknown. Theoretical properties of the procedures are commonly studied asymptotically when the number of observations before a change point tends to infinity. For small samples as a rule simulation study is conducted.

This paper develops an alternative approach in the frame of guaranteed sequential methods. It is based on the method of change point detection for AR(p) process proposed in [1]. In this study such an approach is applied to a general autoregressive model with unknown parameters. Using a special stopping rule we construct statistics which variances are bounded from above by a known constant. Hence, we can estimate the probabilities of false alarm and delay non-asymptotically, but asymptotic properties of the statistics are also investigated and more precise results are obtained.

We consider the scalar autoregressive process to be specified by the equation

$$x_{k+1} = A_k\lambda + B_k\xi_{k+1}, \tag{1}$$

where $\{\xi_k\}_{k\geq 0}$ is a sequence of independent identically distributed random variables with zero mean and unit variance. The density distribution function $f_\xi(x)$ of $\{\xi_k\}_{k\geq 0}$ is strictly positive for any x. The value $m > 1$ defines the order of the process; $\lambda = [\lambda_1, \ldots, \lambda_m]$ is the parameter vector of dimension $m \times 1$; A_k is the known $1 \times m$ matrix, the unknown noise variance B_k is bounded from above, i.e., $B_k^2 \leq D^2 < \infty$, $\mathcal{F}_k = \sigma\{\xi_1, ..., \xi_k\}$ is the σ-algebra generated by variables $\{\xi_1, ..., \xi_k\}$, A_k and B_k are $\mathcal{F}_k$-measurable. The value of the parameter vector λ changes at the change point θ:

$$\lambda = \lambda(k) = \begin{cases} \mu_0, & \text{if } k < \theta; \\ \mu_1, & \text{if } k \geq \theta. \end{cases}$$

[1]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 Russian Foundation for Basic Research Grant 16-01-00121 A.

Values of the parameters before and after θ are supposed to be unknown. The difference between μ_0 and μ_1, for some known Δ, satisfies the condition

$$(\mu_0 - \mu_1)'(\mu_0 - \mu_1) \geq \Delta. \tag{2}$$

The problem is to detect the change point θ from observations x_k.

2 Guaranteed parameter estimator

Let $N_1 \geq m$ be the instant of the estimating procedure start, $n > 0$ is the volume of the initial sample used to estimate the noise variance. The estimator is constructed in the form

$$\begin{gathered}\tilde{\lambda}^*(H) = C^{-1}(N_1 + n, \tau) \sum_{k=N_1+n}^{\tau} v_k A_k' x_{k+1}; \\ C(N_1 + n, \tau) = \sum_{k=N_1+n}^{\tau} v_k A_k' A_k.\end{gathered} \tag{3}$$

We choose the value $\Gamma(N_1, n)$ from the following condition:

$$E\left(D^2/\Gamma(N_1, n)\right) \leq 1. \tag{4}$$

The weights on the interval $[N_1 + n, N_1 + n + \sigma]$ are taken in the form

$$v_k = \begin{cases} (\Gamma(N_1, n) A_k A_k')^{-1/2}, & \text{if } A_{N_1}, \ldots, A_k \text{ are linearly independent;} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The weights v_k on the interval $[N_1 + n + \sigma + 1, \tau - 1]$ are found from the following condition:

$$\nu_{\min}(N_1 + n, k)/\Gamma(N_1, n) = \sum_{l=N_1+n+\sigma}^{k} v_l^2 A_l A_l', \tag{6}$$

where $\nu_{\min}(N_1 + n, k)$ is the minimal eigenvalue of the matrix $C(N_1 + n, k)$.

Choosing a positive parameter H, we define the stopping time $\tau = \tau(H)$ as

$$\tau = \inf\left(N > N_1 + n : \ \nu_{\min}(N_1 + n, N) \geq H\right). \tag{7}$$

At the instant τ, the weight is found from the condition:

$$\nu_{\min}(N_1, \tau)/\Gamma(N_1, n) \geq \sum_{l=N_1+n+\sigma}^{\tau} v_l^2 A_l A_l', \quad \nu_{\min}(N_1, \tau) = H. \tag{8}$$

The parameter H defines the accuracy of the estimator. The choice of the weights v_k allows us to establish a non-asymptotic upper bound for the accuracy.

Theorem 1. *Let the parameter λ in (1) be constant, the compensating factor $\Gamma(N_1, n)$ satisfy condition (4) and the weights v_k determined in (5–6) be such that*

$$\sum_{k=0}^{\infty} v_k^2 A_k A_k' = \infty \ \text{a.s.} \tag{9}$$

Then the stopping time τ (7) is finite with probability one and the mean square accuracy of estimator (3) is bounded from above

$$E||\lambda^*(H) - \lambda||^2 \leq P(H)/H^2, \quad P(H) = H + m - 1. \tag{10}$$

Condition (9) hold true for the process AR(p).

Example. Let the observed process AR(p) be described by equation

$$x_{k+1} = \lambda_1 x_k + \ldots + \lambda_m x_{k-m+1} + B\xi_{k+1}. \tag{11}$$

Then the compensating factor can be chosen in the form proposed in [2]

$$\Gamma(N_1, n) = D(N_1, n) \sum_{l=N_1}^{N_1+n-1} x_l^2. \tag{12}$$

If the noises $\{\varepsilon_k\}_{k\geq 1}$ in (11) are normally distributed with zero mean and unit variance then the multiplier $D(N_1, n) = (n-2)^{-1}$.

3 Change point detection

Consider now the change point detection problem for process (1). We construct a set of sequential estimation plans

$$(\tau_i, \lambda_i^*) = (\tau_i(H), \lambda_i^*(H)),\ i \geq 1,$$

where $\{\tau_i\},\ i \geq 0$ is the increasing sequence of the stopping instances ($\tau_0 = -1$), and λ_i^* is the guaranteed parameter estimator on the interval $[\tau_{i-1}+1, \tau_i]$. The following condition holds true for the estimator

$$E\,||\lambda_i^*(H) - \lambda||^2 \leq P(H)/H^2. \tag{13}$$

Then we choose an integer $l > 1$. We associate the statistic J_i with the $i-$th interval $[\tau_{i-1}+1, \tau_i]$ for all $i > l$

$$J_i = \left(\lambda_i^* - \lambda_{i-l}^*\right)' \left(\lambda_i^* - \lambda_{i-l}^*\right). \tag{14}$$

This statistic is the squared deviation of the estimators with numbers i and $i-l$.

Theorem 2. *The expectation of the statistics J_i (14) satisfies the following inequalities:*

$$E\left[J_i|\,\tau_i < \theta\right] \leq 4P(H)/H^2, \quad E\left[J_i|\,\tau_{i-l} < \theta \leq \tau_{i-1}\right] \geq \Delta - 4\sqrt{\Delta P(H)/H^2}. \tag{15}$$

Hence, the change of the expectation of the statistic J_i allows us to construct the following change point detection algorithm. We choose the values of the parameter H and of the parameter $\delta > 0$ satisfying the following condition

$$4P(H)/H^2 < \delta < \Delta - 4\sqrt{\Delta P(H)/H^2}. \tag{16}$$

The J_i values are compared with the threshold δ. The change point is considered to be detected when the statistic exceeds δ. Due to the application of the guaranteed parameter estimators in the statistics, we can bound the probabilities of false alarm and delay from above.

Theorem 3. *The probability of false alarm P_0^i and the probability of delay P_1^i in any observation cycle $[\tau_{i-1}+1, \tau_i]$ are bounded from above*

$$P_0^i \leq 4P(H)/\delta H^2, \quad P_1^i \leq 4P(H)/\left((\sqrt{\Delta} - \sqrt{\delta})^2 H^2\right). \tag{17}$$

4 Asymptotic properties of the statistics

In the following theorem an asymptotic upper bound for the probability of large values of the standard deviation for the estimator (3) is obtained.

Theorem 4. *If for process (1)* $B_k^2 \leq D^2 < \infty$*, and*

$$\max_{1 \leq k \leq \tau(H)} \frac{v_k^2 D^2 ||A_k||^2}{\Gamma(N_1, N) H} \to^{\mathcal{P}} 0, \text{ as } H \to \infty;$$

and the compensating factor $\Gamma(N_1, N)$ *satisfies the following conditions*

$$N \to \infty, \quad N/H \to 0 \quad \text{as } H \to \infty, \quad \Gamma(N_1, N) \to^{\mathcal{P}} \text{const} \quad \text{as } N \to \infty$$

then for sufficiently large H *in the conditions of Theorem 1*

$$\mathcal{P}\left\{||\lambda^* - \lambda||^2 > x\right\} \leq 2\left(1 - \Phi\left(\sqrt{\tfrac{xH^2}{H+m-1}}\right)\right), \tag{18}$$

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

The following theorem provides the asymptotic inequalities for the probabilities of false alarm and delay for the change point detection procedure.

Theorem 5. *For process (1) in the conditions of Theorem 4 for sufficiently large* H *the probabilities of false alarm* P_0^i *and delay* P_1^i *in any observation cycle* $[\tau_{i-1} + 1, \tau_i]$ *are bounded from above*

$$\begin{gathered} P_0^i = \mathcal{P}\{||\zeta_i - \zeta_{i-l}||^2 > \delta\} \leq 4\left(1 - \Phi\left(\sqrt{\tfrac{\delta H^2}{4(H+m-1)}}\right)\right); \\ P_1^i \leq \mathcal{P}\left\{||\zeta_i - \zeta_{i-l}||^2 > \left(\sqrt{\Delta} - \sqrt{\delta}\right)^2\right\} \leq 4\left(1 - \Phi\left(\sqrt{\tfrac{\left(\sqrt{\Delta}-\sqrt{\delta}\right)H^2}{4(H+m-1)}}\right)\right), \end{gathered} \tag{19}$$

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

The conditions of the theorems hold true for the stable AR(p) process.

References

[1] Burkatovskaya Y.B., Vorobeychikov S.E. (2011). Change point detection of autoregressive process with unknown parameters. *Preprints of the 18th IFAC World Congress.* pp. 13215–13220.

[2] Konev V.V., Dmitrienko A.A. (1994). On guaranteed estimation of autoregression parameters when the noise variance is unknown. *Automatics and Remote Control.* Vol. **2**, pp. 87–99.