VILNIUS UNIVERSITY

Irma Rastenė

TESTING AND ESTIMATING CHANGED SEGMENT IN AUTOREGRESSIVE MODEL

Summary of Doctoral Disertation

Physical Sciences, Mathematics (01 P)

Vilnius, 2011

The scientific work was carried out during 2005-2011 at Vilnius University.

Scientific supervisor:
Prof. Dr. Habil. Alfredas Račkauskas (Vilnius University, Physical Sciences, Mathematics – 01 P)

The Council:
Chairman:
Prof. Dr. Habil. Vygantas Paulauskas (Vilnius University, Physical Sciences, Mathematics – 01 P)
Members:
Prof. Dr. Habil. Algimantas Jonas Bikelis (Vytautas Magnus University, Physical Sciences, Mathematics – 01 P)
Prof. Dr. Habil. Vydas Čekanavičius (Vilnius University, Physical Sciences, Mathematics – 01 P)
Prof. Dr. Habil. Raimondas Čiegis (Vilnius Gediminas Technical University, Physical Sciences, Mathematics – 01 P)
Prof. Dr. Habil. Remigijus Leipus (Vilnius University, Physical Sciences, Mathematics – 01 P)

Opponents:
Doc. Dr. Danutė Krapavickaitė (Vilnius Gediminas Technical University, Physical Sciences, Mathematics – 01 P)
Prof. Dr. Habil. Jonas Kazys Sunklodas (Institute of Mathematics and Informatics, Physical Sciences, Mathematics – 01 P)

The dissertation will be defended at the public meeting of the Council on June 22, 2011, in The Faculty of Mathematics and Informatics Zigmas Žemaitis (101) auditorium at 4:30 p.m.
Adress: Naugarduko 24, LT-03225 Vilnius, Lithuania.

The summury of the dissertation was distributed on May ......, 2011.
The dissertation is available at the library of Vilnius University.

VILNIAUS UNIVERSITETAS

Irma Rastenė

AUTOREGRESINIO MODELIO PASIKEITUSIO SEGMENTO TESTAVIMAS IR VERTINIMAS

Daktaro disertacijos santrauka

Fiziniai mokslai, matematika (01 P)

Vilnius, 2011

Disertacija rengta 2005-2011 metais Vilniaus universitete

Mokslinis vadovas:
Prof. habil. dr. Alfredas Račkauskas (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P)

Disertacija ginama Vilniaus universiteto Matematikos mokslo krypties taryboje:
Pirmininkas:
Prof. habil. dr. Vygantas Paulauskas (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P)
Nariai:
Prof. habil. dr. Algimantas Jonas Bikelis (Vytauto Didžiojo universitetas, fiziniai mokslai, matematika – 01 P)
Prof. habil. dr. Vydas Čekanavičius (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P)
Prof. habil. dr. Raimondas Čiegis (Vilniaus Gedimino technikos universitetas, fiziniai mokslai, matematika – 01 P)
Prof. habil. dr. Remigijus Leipus (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P)

Oponentai:
Doc. dr. Danutė Krapavickaitė (Vilniaus Gedimino technikos universitetas, fiziniai mokslai, matematika – 01 P)
Prof. habil. dr. Jonas Kazys Sunklodas (Matematikos ir informatikos institutas, fiziniai mokslai, matematika – 01 P)

Disertacija bus ginama viešame Matematikos mokslo krypties tarybos posėdyje 2011 m. birželio 22 d. 16 val. 30 min., Matematikos ir informatikos fakulteto Zigmo Žemaičio (101) auditorijoje.
Adresas: Naugarduko g. 24, LT-03225 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2011m. gegužės mėn. …… d.
Su disertacija galima susipažinti Vilniaus universiteto bibliotekoje.

Summary of Doctoral Disertation

Scientific questions

For each $n \in \mathbb{N}$, we consider the following changed-segment autoregressive model:

$$y_k = \mathbf{1}_{k \in I^{*c}} \rho y_{k-1} + \mathbf{1}_{k \in I^*} \rho^* y_{k-1} + \varepsilon_k, \quad k = 1, 2, \ldots, n,$$

where (ε_k) are independent identically distributed random variables with mean zero and finite variance, and the subset of observation indexes

$$I^* = I^*(n) = \{k^* + 1, \ldots, k^* + \ell^*\}$$

represents a changed segment, $I^{*c} = \{1, 2, \ldots, n\} \backslash I^*$. The model describes an epidemic deviation of an AR(1) process from the ordinary state, and k^* is called the *starting*, whereas ℓ^* is called the *duration of the epidemic state.* In our model, neither k^* nor ℓ^* is known. We focus not only on the stationary case ($|\rho| < 1$, $|\rho^*| < 1$), but also include switches from a stationary state to a nonstationary one ($|\rho| < 1$, $|\rho^*| = 1$) and vice versa ($|\rho| = 1$, $|\rho^*| < 1$). How to identify from given data whether there is a changed segment or not; if yes, where is the starting and what are the duration of the epidemic state and the size of this possible change? Doctoral dissertation gives answers namely to these questions.

In order to construct criteria for testing changed segment, we consider polygonal line processes constructed from estimates $\widehat{\varepsilon}_k$ of the model residuals and partial estimators $\widehat{\rho}_k$ of the model parameter. Typical paths of these processes are presented in Fig. 1.

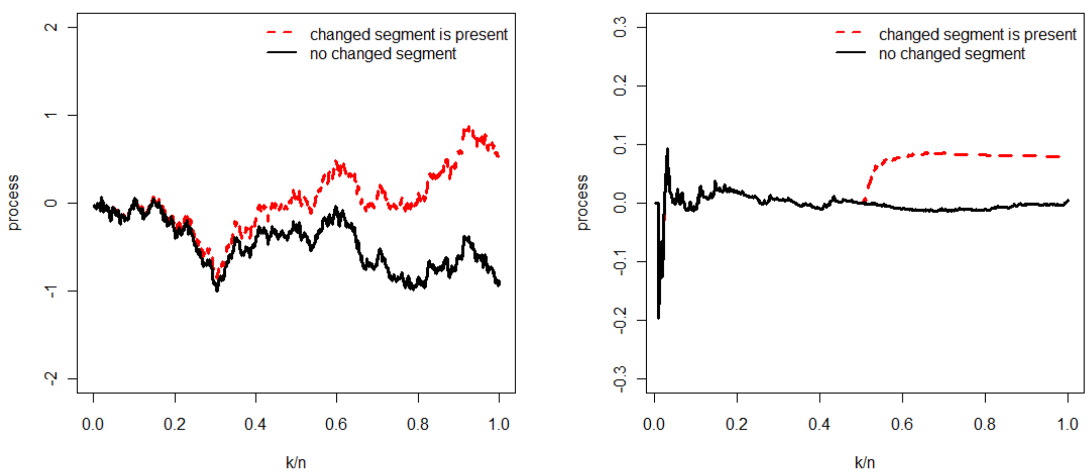

Figure 1. Partial-sum process of normalized residuals $\widehat{\varepsilon}_k$ (left) and the process $\widehat{\rho}_k - \rho$ (right) ($k^*/n = 0.5$, $\ell^* = 0.2$, $\rho = 0.9$, $\rho^* = 1$).

If a changed segment is present, the path of polygonal line process changes. This feature enables us to construct changed-segment testing criteria. Using the invariance principle with respect to Hölderian topology established by Račkauskas and Suquet [38], we get the asymptotic behavior of polygonal line processes under no change. These results provide us with a wider class of functionals that can be used as test statistics for testing stability of AR(1) models. Applying the continuous mapping theorem to polygonal line processes, we construct tests for a changed segment. Under the null hypothesis of no change, the test statistic converges to some asymptotic process and, under the alternative, tends to infinity. When estimating the changed segment, it is important to get consistent estimates of the changed segment and model parameters ρ, ρ^*. In this work, the consistency of least squares estimators is proved, and convergence rates are given.

Relevancy

Various autoregressive processes are widely used in econometric models. In order to get forecasts as accurate as possible, we need to test the structural change and, if it is present, to estimate parameters of the change. Other relevant question is related with the stationarity of an autoregressive model. In this work, we consider testing and estimating questions for one type of possible structural changes, the changed segment. Different scenarios of stationarity–nonstationarity are included.

Aims and problems

The main objective of the thesis is to investigate the first-order changed-segment autoregressive model with unknown starting and duration of the epidemic state, including different scenarios of stationarity. This work deals with following problems:

* convergence of polygonal line processes in Hölder spaces;
* criteria related to polygonal line processes for testing the changed segment;
* consistency of the changed segment and model parameters and of their convergence rates.

Novelty

We give changed segment tests for autoregressive first-order model. They are based on the convergence of polygonal line processes in Hölder spaces, which is proved in this work. We establish the consistency of least squares estimators of the changed segment and model parameters, and provide their convergence rates.

Main statements

The asymptotic behavior of residual partial sums and partial estimator processes of model parameter ρ are given in Theorems 1 and 4, respectively. These results enable us to construct criteria for testing the changed segment in an autoregressive process. Test statistics are defined by (13) and (11). Propositions 7 and 5 describe the convergence of test statistics under the null hypothesis of no change. The behavior of one test statistic under the alternative is given in Proposition 6. The consistency of the least square changed segment estimators is proved, and their convergence rates are given in Theorems 8, 9, and 10.

Methods

We use the invariance principle in Hölder spaces, continues mapping theorem, small ball probabilities. We also apply methods of general probability theory and mathematical statistics.

Main results

In the literature on structural changes, numerous studies have been performed in different directions: testing the presence of structural changes, estimation of change points location and their number, forecasting from models under structural breaks, etc. Models also differ and are linear or nonlinear, with known or unknown change points, with single, epidemic, or multiple changes. We refer to Csörgő and Horváth [13], Chen and Gupta [11], Basseville and Nikiforov [9], Banerjee and Urga [8], Perron [32], and Sánchez [39] for a corresponding bibliography.

Times series stability is very important question in econometrics. Estimators based on unstable models can be shifted, and forecasts loose accuracy. So tests for structural stability are needed. We refer to the papers by Wichern et al. [42], Bagshaw and Johnson [2], Pickard [35], Krämer et al. [25], Tang and MacNeil [41], Davies et al. [15] Kim et al. [24], Lee and Park [30], Lee [29], Gombay and Serban [18], Gombay [17], among others.

In doctoral dissertation, we use CUSUM (cumulative-sum) method for testing structural change. This method is used commonly for testing change in mean, variance, and other parameters of regression-type models. Usually, functional limit theorems for model residual-based processes are first proved. Then appropriate functionals of these processes are used as test statistics. The weak limits of partial sums of regression residuals were established by MacNeil and Jandhyala [22], Tang and MacNeil [41], whereas Kulperger [26] and Bai [3] studied AR and ARMA models, respectively. Horvath [19] and Bai [3] proposed some applications to change-point problems. Some authors have established weak limits for more general processes based on residuals. For example, Jandhyala and MacNeil [21] studied iterated partial sum sequences of regression residuals, Yu [43] and Kulperger and Yu [27] constructed high-moment partial-sum processes based on residuals of ARMA and GARCH models, respectively. All mentioned studies of structural changes focus on stationary processes. Nonstationary AR models where investigated by Shin [40]. He proved that residual partial-sum processes of the AR(1) model $y_k = \rho y_{k-1} + e_k$, $k = 1, 2, \ldots, n$, with independent identically distributed finite variance errors (e_k) weakly converge in the Skorokhod space D[0, 1]. The limiting process is a standard Brownian motion when $|\rho| < 1$, whereas it is a randomly shifted Brownian motion when $|\rho| = 1$. We extend these results establishing the convergence in Hölder spaces of the polygonal line processes constructed from partial sums of residuals of AR(1) model.

The literature on estimating the location of changes and models parameters is sparser. Bai [4] proposed least square estimate for change point in the mean of a linear process of general type under weak regularity assumptions. These estimates do not require specifying the distribution of the error process (like maximum likelihood methods) and are computationally efficient. Therefore, this method became most frequently used. Linear regression models with multiple breaks were studied by Bai and Perron [6], [7], Perron and Qu [34], Kejriwal and Perron [23], and Bai et al. [5]. Lavielle and Moulines [28] proposed a multiple change-point estimation procedure that can be applied to a large class of dependent processes. Qu and Perron [33] cover the more general case of multiple structural changes in a system of equations allowing arbitrary restrictions on the parameters.

Literature on change location estimation for a specific case of epidemic-type change addresses mainly to independent observations (see Fu and Curnow [16], Hušková [20], and Račkauskas and Suquet [36], [37]).

Jouini and Boutahar [10] call attention to relatively sparse number of works focussing on the changes in nonstationary time series. Chong [12] proposed single change-point location estimator for AR(1) model including the cases with change from a nonstationary state to a stationary one and vice versa. In this paper we extended his results concerning AR(1) model with single break to AR(1) model with an epidemic change. In this work, we prove the consistency of the changed segment estimators for different scenarios and provide the convergence rates.

Some asymptotic results for polygonal line processes

We consider the following AR(1) model:

$$y_k = \rho y_{k-1} + \varepsilon_k, \tag{1}$$

$k = 1, 2, \ldots, n$, where $\varepsilon_1, \ldots, \varepsilon_n$ are independent identically distributed random variables with mean zero and finite variance $\sigma^2 = E\varepsilon_1^2$. Throughout we assume that $\sigma^2 = 1$ and set $y_0 = 0$ for simplicity. The residuals $(\widehat{\varepsilon}_k, k = 1, \ldots, n)$ are defined by

$$\widehat{\varepsilon}_k = y_k - \widehat{\rho}\, y_{k-1} = (\rho - \widehat{\rho})\, y_{k-1} + \varepsilon_k, \tag{2}$$

$k = 1, 2, \ldots, n$, where $\widehat{\rho}$ is the least squares estimator of ρ given by

$$\widehat{\rho} = \frac{\sum_{k=1}^{n} y_k y_{k-1}}{\sum_{k=1}^{n} y_{k-1}^2}. \tag{3}$$

Throughout we set $\sum_{k\in\emptyset} \equiv 0$. Shin [40] showed that residual partial-sum processes

$$\widehat{W}_n(t) = \sum_{k=1}^{[nt]} \widehat{\varepsilon}_k,\ \ t \in [0,\ 1],$$

normalized by $n^{-1/2}\sigma^{-1}$, as $n \to \infty$, have a weak limit in the Skorokhod space $D[0,1]$. If $|\rho| \neq 1$, the processes $(\sigma^{-1}n^{-1/2}\widehat{W}_n)$ converge in distribution to a standard Brownian motion on $[0,\ 1]$. If $|\rho| = 1$, then the limit is a randomly shifted Brownian motion.

We investigate the polygonal line process

$$\widehat{V}_n(t) = \sum_{k=1}^{[nt]} \widehat{\varepsilon}_k + (nt - [nt])\widehat{\varepsilon}_{[nt]+1}, \quad t \in [0,\ 1], \tag{4}$$

$n = 1, 2, \dots,$ in a framework of the separable Hölder spaces

$$H_\alpha^o[0,\ 1] := \{f \in C[0,\ 1] : \lim_{\delta \to 0} \omega_\alpha(f,\ \delta) = 0\},\ (0 < \alpha < 1)$$

endowed with the norm

$$\|f\|_\alpha := |f(0)| + \omega_\alpha(f,\ 1),$$

where

$$\omega_\alpha(f,\ \delta) := \sup_{\substack{s,t \in [0,1] \\ 0 < t-s < \delta}} \frac{|f(t) - f(s)|}{|t-s|^\alpha}.$$

As usual, $C[0,1]$ denotes the space of continuous functions endowed with the uniform norm $||f|| = \sup_{0 \le t \le 1} |f(t)|$, $f \in C[0,1]$.

Clearly, the processes $\widehat{V}_n, n \geq 1$, belong to $H_\alpha^o[0,1]$ for each $0 < \alpha < 1$. However, its limiting process has a version in $H_\alpha^o[0,1]$ for $0 < \alpha < 1/2$ only.

Let $W = (W(t), t \in [0,1])$ be a standard Wiener process.

Theorem 1 *Let $0 < \alpha < 1/2$ and assume that*

$$\lim_{t \to \infty} tP(|\varepsilon_1| \geq t^{1/2-\alpha}) = 0. \tag{5}$$

Then, in the Hölder space $H_\alpha^o[0,1]$, we have:

(*a*) *if $|\rho| \neq 1$, then*

$$n^{-1/2}\widehat{V}_n \xrightarrow[n \to \infty]{\mathcal{D}} W;$$

(*b*) *if $\rho = 1$, then*

$$n^{-1/2}\widehat{V}_n \xrightarrow[n \to \infty]{\mathcal{D}} W - A^{-1}BD;$$

(*c*) *if $\rho = -1$, then*

$$n^{-1/2}\widehat{V}_n \xrightarrow[n \to \infty]{\mathcal{D}} W + \tilde{A}^{-1}\tilde{B}\tilde{D};$$

where

$$A = 2\int_0^1 W^2(r)dr,\ B = W^2(1) - 1,\ D(t) = \int_0^t W(r)dr,\ t \in [0,1],$$

and the random vector $(\tilde{A}, \tilde{B}, (\tilde{D}(t), t \in [0,1]))$ has the same distribution as $(A, B, (D(t), t \in [0,1]))$ and is independent of W.

Remark 2 *Via the continuous mapping theorem, Theorem 1 implies that, for each continuous function $F : H_\alpha^o[0,1] \to R$:*

(i) if $|\rho| \neq 1$*, then*

$$F(n^{-1/2}\widehat{V}_n) \xrightarrow[n\to\infty]{\mathcal{D}} F(W);$$

(ii) if $\rho = 1$*, then*

$$F(n^{-1/2}\widehat{V}_n) \xrightarrow[n\to\infty]{\mathcal{D}} F(W - A^{-1}BD);$$

(iii) if $\rho = -1$*, then*

$$F(n^{-1/2}\widehat{V}_n) \xrightarrow[n\to\infty]{\mathcal{D}} F(W + \tilde{A}^{-1}\tilde{B}\tilde{D}).$$

Now the statistics $T_n = F(n^{-1/2}\widehat{V}_n)$ *can be used to test stability of* AR(1) *model (1) by choosing a continuous function* $F : H_\alpha^o[0,1] \to R$*.*

Let $\widehat{\rho}_k$ denote the least square estimator of the parameter ρ based on the first k observations $y_1, \dots, y_k$:

$$\widehat{\rho}_k = \frac{\sum_{j=1}^k y_j y_{j-1}}{\sum_{j=1}^k y_{j-1}^2}, \quad k = 2, 3, \dots, n.$$

For $t \in [0,1]$ and $n \geq 1$, define the polygonal line process

$$r_n(t) = \begin{cases} \widehat{\rho}_k + (nt-k)(\widehat{\rho}_{k+1} - \widehat{\rho}_k) & \text{for} \quad k/n \leq t < (k+1)/n, \quad k = 2, \dots, n, \\ 0, & \text{for} \quad 0 \leq t < 2/n. \end{cases}$$

The classical asymptotic normality results for $\widehat{\rho}_n$ (see, e.g., Mann and Wald [31] and Anderson [1]) yields

$$\sqrt{n}(1-\rho^2)^{-1/2}(\widehat{\rho}_{[nt]} - \rho) \xrightarrow[n\to\infty]{\mathcal{D}} t^{-1/2}\mathcal{N}(0,1).$$

Hence, for small $t \in (0,1)$, the least square estimator $\widehat{\rho}_{[nt]}$ performs very inaccurately. Therefore, in order to establish the functional behavior of the process r_n, we restrict it to the interval $[\delta, 1]$ with $\delta \in (0,1)$.

Theorem 3 *For each* $\delta \in (0,1)$*,*

$$\sqrt{n}(1-\rho^2)^{-1/2}(r_n - \rho) \xrightarrow[n\to\infty]{\mathcal{D}} \left\{\frac{W(t)}{t}, \quad \delta \leq t \leq 1\right\} \tag{6}$$

in the space $C[\delta, 1]$*. If for* $0 < \alpha < 1/2$*, the condition*

$$E|\varepsilon_1|^{p(\alpha)} < \infty, \quad \textit{where} \quad p(\alpha) = \frac{1}{0.5 - \alpha}, \tag{7}$$

is satisfied, then the convergence (6) *holds also in the space* $H_\alpha^o[\delta, 1]$*.*

For practical purposes, the convergence (6) is not interesting since ρ is unknown. The following result uses another standardization.

Theorem 4 *For each $\delta \in (0,1)$,*

$$\Big(\sum_{k=1}^{n} y_{k-1}^2\Big)^{1/2} (r_n - \rho) \xrightarrow[n\to\infty]{\mathcal{D}} \Big\{\frac{W(t)}{t}, \quad \delta \le t \le 1\Big\} \tag{8}$$

in the space $C[\delta,1]$. If for an $0 < \alpha < 1/2$, the condition

$$E|\varepsilon_1|^{p(\alpha)} < \infty, \quad \textit{where} \quad p(\alpha) = \frac{1}{0.5-\alpha}, \tag{9}$$

is satisfied, then the convergence (8) holds also in the space $H_\alpha^o[\delta,1]$.

Testing changed segment

In this section, we consider a problem of changed segment in an autoregressive model. Assume that

$$y_k = \rho^* y_{k-1}\mathbf{1}_{I^*}(k) + \rho y_{k-1}\mathbf{1}_{I^{*c}}(k) + \varepsilon_k, k = 1, \dots, n, \tag{10}$$

where $I^* = \{k^*+1, \dots, k^*+\ell^*\}$ and $I^{*c} = \{1,\dots,n\} \setminus I^*$, $|\rho| < 1$ $(\varepsilon_k, k = 0,1,2,\dots)$ are independent identically distributed random variables with mean zero and finite variance $\sigma^2 = E\varepsilon_1^2$. The model (10) describes an epidemic deviation of an AR(1) process from the ordinary state, and k^* is called the *starting*, whereas ℓ^* is called the *duration of the epidemic state.* Usually, neither k^* nor ℓ^* is known. The aim is to test the null hypothesis

$$H_0: \quad \ell^* = 0$$

that there is no epidemic deviation from ordinary state against the alternative

$$H_A: \quad \ell^* > 0$$

that, on the contrary, such deviation exists. To this aim, we consider the following statistic for some fixed $\delta \in (0,1)$:

$$T_n := T_n^{(\delta)} := \Big(\sum_{i=1}^{n} y_{i-1}^2\Big)^{1/2} \max_{1\le \ell \le (1-\delta)n} \ell^{-\alpha} \max_{\delta n \le k \le n} |\widehat{\rho}_{k+\ell} - \widehat{\rho}_k|. \tag{11}$$

Proposition 5 *Under H_0, assume that $E|\varepsilon_1|^{p(\alpha)} < \infty$, $\quad p(\alpha) = \frac{1}{0.5-\alpha}$. Then*

$$n^\alpha T_n^{(\delta)} \xrightarrow[n\to\infty]{\mathcal{D}} \max_{0<h<1-\delta} h^{-\alpha} \max_{\delta \le t \le 1} \Big|\frac{W(t+h)}{t+h} - \frac{W(t)}{t}\Big|. \tag{12}$$

Proposition 6 *Under alternative H_A, assume that $k^* \ge \delta n$ for some $\delta \in (0,1)$ and that $\ell^* \to \infty$ and $\ell^*/n \to 0$ as $n \to \infty$. Then*

$$n^\alpha T_n^{(\delta)} \to \infty,$$

provided that

$$\Big(\frac{\ell^*}{n}\Big)^{1-\alpha} \sqrt{n}|\rho - \rho^*| \to \infty \quad \textit{as} \quad n \to \infty.$$

For power analysis, some experiments were made, and their results were presented by size-power curves (see [14]). On the x-axis, we have set the values of the empirical p-value distribution function under the null hypothesis, whereas on the y-axis, we have set the values of the empirical p-value distribution function under the alternative (empirical power function).

The results show that the test statistic gains more power as the number of observations n, the duration of the epidemic state ℓ^*, and δ increases (the prehistory of epidemic change should be of the length proportional to the number of observations). The power decreases when start position increases. From here we get that the test is more powerful when we have a longer post-history after the epidemic change. Analysis of the influence of parameters $\rho*$ and ρ to the test power shows that the test is more powerful in detecting changes of the parameter ρ when the absolute values of this parameter are quite large or autoregressive process changes from a stationary state to a nonstationary one.

Let us define the statistic based on partial residual sums $\hat{S}_0 = 0$, $\hat{S}_k = \hat{\varepsilon}_1 + \cdots + \hat{\varepsilon}_k$, $k = 1, \ldots, n$, of model (10):

$$T(n;\alpha) = \max_{1<l<n} \frac{1}{l^\alpha} \max_{0\leqslant k\leqslant n-l} \left| \hat{S}(k+l) - \hat{S}(k) - \frac{l}{n}\hat{S}(n) \right|, \tag{13}$$

where $0 < \alpha < 1/2$.

Proposition 7 *Under H_0, assume that* $\lim_{t\to\infty} t^{1/(0.5-\alpha)} P(|\varepsilon_0| > t) = 0$. *Then*

$$n^{-1/2+\alpha}\sigma^{-1}T(n;\alpha) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{0<h<1} h^{-\alpha} \max_{0<t<1-h} |W(t+h) - W(t) - hW(1)| . \tag{14}$$

Under alternative H_A, $n^{-1/2+\alpha}\sigma^{-1}T(n;\alpha) \to \infty$.

Empirical test power analysis has shown similar results as those for the test based on the partial-estimator process of parameter ρ. Only changed segment starting position has no significant influence on the test power, and as the parameter α goes closer to 1/2, the test power increases. Criteria based on partial-estimator process of parameter ρ is more powerful, except the cases where the changed-segment prehistory is short.

Changed-segment estimators: consistency and convergence rates

For each $n \in \mathbb{N}$, we consider the following changed-segment autoregressive model:

$$y_i^{(n)} = \rho_n y_{i-1}^{(n)} \mathbf{1}_{I^{*c}}(i) + \rho_n^* y_{i-1}^{(n)} \mathbf{1}_{I^*}(i) + \varepsilon_i^{(n)}, \tag{15}$$

$i \in N_n := \{1, 2, \ldots, n\}$, where the subset of observations indexes

$$I^* = I^*(n) = \{k_n^* + 1, \ldots, k_n^* + \ell_n^*\} \in I_n$$

represents a changed segment, $I^{*c} = N_n \backslash I^*$, and

$$I_n := \{\{k+1, \ldots, k+\ell\}, \quad \ell = 1, \ldots, n, \quad k = 0, \ldots, n-\ell\} .$$

The model (15) describes an epidemic deviation of an AR(1) process from the ordinary state, and k_n^* is called the *starting*, whereas ℓ_n^* is called the *duration of the epidemic state.* In our model (15), neither k_n^* nor ℓ_n^* is known. Our aim is to estimate them from given observations.

Let the following assumptions be satisfied.

(A1) For each n, the random variables $\varepsilon_1^{(n)}, \dots, \varepsilon_n^{(n)}$ are independent identically distributed with mean zero and finite variance $\sigma^2 = E(\varepsilon_1^{(n)})^2$;

(A2) $y_0^{(n)} = 0$;

(A3) There exist $\theta_0, \theta_1 \in (0,1)$ such that $\ell_n^* \in [\theta_0 n, \theta_1 n]$.

The strongest assumption (A3) means that a changed segment is of length proportional to the number of observations. The case where $\ell_n^* = o(n)$ requires different approach than suggested in this paper.

In what follows. we skip an extra index n in the notation just for simplification purposes. We define

$$\hat{\rho}(A) = \frac{\sum_{i\in A} y_i y_{i-1}}{\sum_{i\in A} y_{i-1}^2},$$
$$RSS_n(A) = \sum_{i\in A^c} (y_i - \hat{\rho}(A^c) y_{i-1})^2 + \sum_{i\in A} (y_i - \hat{\rho}(A) y_{i-1})^2,$$

where $A \in I_n$ and $A^c = N_n \setminus A$. The changed segment in the model (15) is then estimated by

$$(\hat{k^*}, \hat{k^*} + \hat{\ell^*}) = \hat{I}^* = \arg\min_{I\in I_n} RSS_n(I). \tag{16}$$

Set, for $0 \le \beta < \theta_0/3$,

$$I_\beta = (\hat{k^*} + \beta n, \hat{k^*} + \hat{\ell^*} - \beta n), \quad J_\beta = [1, \hat{k^*} - \beta n] \cup [\hat{k^*} + \hat{\ell^*} + \beta n, n].$$

When the consistency of the estimator $\hat{I}^*$ is established, the choice of β will ensure that, with high probability, both intervals I_β and J_β are nonempty. We shall consider the estimators of ρ and ρ^* defined by

$$\hat{\rho}^* = \hat{\rho}(I_\beta), \quad \hat{\rho} = \hat{\rho}(J_\beta). \tag{17}$$

In the case where the autoregressive process changes from one stationary state to another, we have the following result.

Theorem 8 *For the model* (15), *assume that* $|\rho| < 1$, $|\rho^*| < 1$, *and that conditions (A1)–(A3) are fulfilled. Then*

(*i*) *for the estimators of* (k^*, ℓ^*) *given by* (16),

$$|\hat{k^*} - k^*| + |\hat{\ell^*} - \ell^*| = O_P\Big(\frac{\sqrt{n}}{|\rho - \rho^*|}\Big);$$

(*ii*) *for the estimators of* ρ *and* ρ^* *given by* (17) *with* $\beta = 0$,

$$|\rho^* - \hat{\rho^*}| = O_P\Big(n^{-1/2}\Big), \quad |\rho - \hat{\rho}| = O_P\Big(n^{-1/2}\Big),$$

provided that $\sqrt{n}|\rho - \rho^*| \to \infty$ *as* $n \to \infty$.

If the changed segment represents a random walk, we have the following results.

Theorem 9 *For the model* (15), *assume that* $|\rho| < 1$ *and* $\rho^* = 1$. *Assume also that conditions* (*A*1)*–*(*A*3) *are fulfilled. Then*

(*i*) *for the estimators of* (k^*, ℓ^*) *given by* (16),

$$|\hat{k^*} - k^*| + |\hat{\ell^*} - \ell^*| = o_P(n);$$

(*ii*) *for the estimators of* ρ *and* ρ^* *given by* (17) *with any* $0 < \beta < \theta_0/3$,

$$|1 - \hat{\rho^*}| = O_P(n^{-1}), \quad |\rho - \hat{\rho}| = O_P(n^{-1/2})$$

as $n \to \infty$.

Finally, in the case where a stationary autoregressive segment is inserted into a random walk, the result is similar to the previous one.

Theorem 10 *For the model* (15), *assume that* $\rho = 1$ *and* $|\rho^*| < 1$. *Assume also that conditions* (*A*1)*–*(*A*3) *are fulfilled. Then*

(*i*) *for the estimators of* (k^*, ℓ^*) *given by* (16),

$$|\hat{k^*} - k^*| + |\hat{\ell^*} - \ell^*| = o_P(n);$$

(*ii*) *for the estimators of* $\hat{\rho}$ *and* $\hat{\rho^*}$ *given by* (17) *with any* $0 < \beta < \theta_0/3$,

$$|1 - \hat{\rho}| = O_P(n^{-1}), \quad |\rho^* - \hat{\rho^*}| = O_P(n^{-1/2})$$

as $n \to \infty$.

Conclusions

In the doctoral dissertation, we consider problems of testing and estimating changed segment with unknown starting position and duration of epidemic state in the autoregressive first-order model. The proposed tests are based on partial sums of model residuals and model-parameter partial-estimator polygonal line processes. We derive asymptotic results for these processes in Hölder spaces. The behavior of test statistics under the null hypothesis of no change and alternative is provided. Empirical power analysis has shown that tests are more powerful when absolute values of model parameter are quite large or autoregressive process changes from a stationary state to a nonstationary one. We prove the consistency of the least square changed-segment estimators and provide their convergence rates.

Publications

1. Rastenė I., Testing AR(1) model, *Lietuvos Matematikos Rinkinys*, **47**(spec. issue), 375–379 (2007).

2. Račkauskas A., Rastenė I., Hölder convergence of autoregression residuals partial sum processes, *Lithuanian Mathematical Journal*, **48**(4), 438–450, (2008).

3. Račkauskas A., Rastenė I., Estimating changed segment in AR(1) model, *Department of Mathematics and Informatics, Vilnius University*, **preprint** (2011).

4. Račkauskas A., Rastenė I., Some asymptotic results for polygonal line processes related to an autoregressive process, *Department of Mathematics and Informatics, Vilnius University*, **preprint** (2011).

Conferences

1. June 27–28, 2007. XLVIII Conference of Lithuanian Mathematical Society. "Research of structural change in AR(1) model."

2. June 18–19, 2009. L Conference of Lithuanian Mathematical Society. "Estimating changed segment in AR(1) model" (stationary case).

3. June 28-July 2, 2010. X International Vilnius Conference on Probability Theory and Mathematical Statistics. "Estimating changed segment in AR(1) model" (stationary and nonstationary cases).

4. June 7–10, 2011. XIV Conference of the Applied Stochastic Models and Data Analysis International Society. "Testing and estimating changed segment in AR(1) model."

Reziumė

Disertacijoje nagrinėjamas pirmos eilės autoregresinio modelio pasikeitusio segmento testavimo ir vertinimo uždavinys. Aprašomo modelio epideminio pasikeitimo pradžia ir ilgis nėra žinomi. Nagrinėjami galimi pasikeitimai ne tik iš stacionarios būklės į stacionarią, bet ir iš stacionarios į nestacionarią bei atvirkščiai.

Įvade skaitytojas supažindinamas su nagrinėjamais uždaviniais, trumpai aprašomi jų sprendimo būdai, pateikiami disertacijos tikslai bei ginamieji teiginiai. Pirmajame skyriuje supažindinama su moksline literatūra, kurioje nagrinėjami artimi šiam darbui uždaviniai. Antrasis disertacijos skyrius skirtas laužčių procesų ribinėms teoremoms Hiolderio erdvėse. Nagrinėjami modelio paklaidų įvertinių dalinių sumų ir modelio parametro ρ dalinių įvertinių laužčių procesai. Trečiajame skyriuje pasiūlyti kriterijai pasikeitusio segmento testavimui. Statistikų ribinis elgesys esant teisingai nulinei (modelio be struktūrinio pasikeitimo) hipotezei gaunamas taikant ankstesniajame skyriuje įrodytas ribines teoremas. Parodoma, kad esant teisingai alternatyvai testo statistikos išsigimsta. Iš empirinio galios tyrimo rezultatų matyti, kad pasiūlytų testų galia didžiausia aptinkant pasikeitimus iš stacionarios būklės į nestacionarią arba esant artimoms vienetui parametro ρ reikšmėms. Ketvirtajame disertacijos skyriuje įrodoma, kad pasikeitusio segmento pradžios ir ilgio įvertiniai bei autoregresinio modelio su pasikeitusiu segmentu parametrų įvertiniai yra suderintieji bei pateikiamas jų konvergavimo greitis. Darbas baigiamas išvadomis ir literatūros sąrašu.

Curriculum Vitae

Date of birth: April 2, 1981.

Education:

2010 finished doctoral studies at Department of Econometric Analysis, Faculty of Mathematics and Informatics, Vilnius University.

2005 achieved MA degree in mathematics at the Faculty of Mathematics and Informatics, Vilnius University.

2003 Erasmus Programme studies at the Faculty of Mathematics and Natural Sciences, Ernst Moritz Arndt University of Greifswald.

2003 achieved BA magna cum laude degree in mathematics at the Faculty of Mathematics and Informatics, Vilnius University.

1999 finished Kaišiadorys 1-st Secondary School.

Work experience:

Since 2010, 2007 Assistant at Department of Econometric Analysis, Faculty of Mathematics and Informatics, Vilnius University.

2006-2007 Data analyst at "Euromonitor International - Eastern Europe".

2004-2006 Non-life insurance actuary at "Lietuvos draudimas".

References

[1] T. W. Anderson, On asymptotic distributions of estimates of parameters of stochastic difference equations, *The Annals of Mathematical Statistics*, **30**, 676–687 (1959).

[2] M. Bagshaw, R. A. Johnson, Sequentional procedures for detecting parameter changes in a time series model, *Journal of American Statistical Association*, **72**, 593–597 (1977).

[3] J. Bai, On the partial sums of residuals in autoregressive and moving average models, *Journal of Time Series Analysis* **14**, 247–260 (1993).

[4] J. Bai, Least square estimation of a shift in linear processes, *Time Series Analysis*, **15**(5), 453–472 (1994).

[5] J. Bai, H. Chen, T.T.L. Chong, S.X. Wang, Generic consistency of the break-point estimators under specification errors in a multiple-break model, *Econometrics Journal*, **11**, 287–307 (2008).

[6] J. Bai, P. Perron, Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, **18**, 1–22 (2003).

[7] J. Bai, P. Perron, Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**(1), 47–78 (1998).

[8] A. Banerjee, G. Urga, Modelling structural breaks, long memory and stock market volatility: an overview, *Journal of Econometrics*, **129**, 1–34 (2005).

[9] M. Basseville, N. Nikiforov, *Detection of abrupt changes: theory and application*, Information and system science series, Prentice Hall, New York (1993).

[10] M. Boutahar, J. Jouini, Evidence on structural changes in U.S. time series, *Economic Modelling*, **22**, 391–422 (2005).

[11] J. Chen, A. K. Gupta, *Parametric statistical change point analysis*, Birkhäuser Verlag, Boston (2000).

[12] T.T. Chong, Structural change in AR(1) models, *Econometric theory*, **17**, 87–155 (2001).

[13] M. Csörgő, L. Horváth, *Limit Theorems in Change-Point Analysis*, John Wiley & Sons, Baffins Lane, Chichester (1997).

[14] R. Davidson, J. G. MacKinnon, Graphical methods for investigating the size and power of hyphothesis tests, *Manchester School*, **66**(1), 1–26 (1998).

[15] R.A. Davies, D. Huang, Y.-C. Yao, Testing for a change in the parameter value and order of an autoregressive model, *Annals of Statistics*, **23**, 282–304 (1995).

[16] Y. Fu, R.N. Curnow, Locating a changed segment in a sequence of Bernoulli variables, *Biometrika*, **77**(2), 295–304 (1990).

[17] E. Gombay, Change detection in autoregressive time series, *Journal of Multivariate Analysis Volume*, **99**(3), 451–464 (2008).

[18] E. Gombay, D. Serban, Monitoring parameter change in AR(p) time series models, *Journal of Multivariate Analysis*, **100**(4), 715–725 (2009).

[19] L. Horvath, Change in autoregressive process, *Stochastic Processes and their Applications*, **44**, 221–242 (1993).

[20] M. Hušková, Estimators for epidemic alternatives, *Commentationes Mathematicae Universitatis Carolinae*, **36**(2), 281–293 (1995).

[21] V. K. Jandhyala, I. B. MacNeil, Iterated partial sum sequences of regression residuals and tests for changepoints with continuity constraints, *Journal of the Royal Statistical Society*, **59**, 147–156 (1997).

[22] V. K. Jandhyala, I. B. MacNeil, Residuals partial sum limit process for regression models with applications to detecting parameter changes at unknown times, *Stochastic Processes and their Applications*, **33**, 309–323 (1989).

[23] M. Kejrival, P. Perron, The limit distribution of the estimates in cointegrated regression models with multiple structural changes, *Journal of Econometrics*, **146**, 59–73 (2008).

[24] S. Kim, S. Cho, S. Lee, On the cusum test for parameter changes in GARCH(1,1) models, *Communications in Statistics: Theory and Methods*, **29**, 445–462 (2000).

[25] W. Krämer, W. Ploberger, R. Alt, Testing for structural change in dynamic models, *Econometrica*, **56**, 1355–1369 (1988).

[26] R.J. Kulperger, On the residuals of autoregressive processes and polynomial regression, *Stochastic Processes and their Applications*, **21**, 107–118(1985).

[27] R. J. Kulperger, H. Yu, High moment partial sum processes of residuals in GARCH models and their applications, *The Annals of Statistics*, **33**, 2395–2422 (2005).

[28] M. Lavielle, E. Moulines, Least Squares estimation of an unknown number of shifts in a time series, *Journal of Time Series Analysis*, **21**(1), 33-59 (2000).

[29] S. Lee, J. Ha, O. Na, S. Na, The cusum test for parameter change in time series models, *Scandinavian Journal of Statistics*, **30**(4), 78–796 (2003).

[30] S. Lee, S. Park, The cusum of squares test for scale changes in infinite order moving average processes, *Scandinavian Journal of Statistics*, **28**, 625–644 (2001).

[31] H.B. Mann, A. Wald , On the statistical treatment of linear stochastic difference equations, *Econometrica*, **11**, 173–220 (1943).

[32] P. Perron, Dealing with Structural Breaks, in: Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, K. Patterson and T.C. Mills (eds.), Palgrave Macmillan, 278–352 (2006).

[33] P. Perron, Z. Qu. Estimating and testing structural changes in multivariate regressions. *Econometrica*, **75**, 459–502, 2007.

[34] P. Perron, Z. Qu, Estimating restricted structural change models, *Journal of Econometrics*, **134**, 373–399, 2006.

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

[36] A. Račkauskas, Ch. Suquet, Estimation of change points of infinite dimensional parameters in short epidemics, *Lithuanian Mathematical Journal*, **45**(4), 458–474, 2005.

[37] A. Račkauskas, Ch. Suquet, Estimating a changed segment in a sample, *Acta Applicandae Mathematicae*, **97**, 189–210 (2007).

[38] A. Račkauskas, Ch. Suquet, Necessary and sufficient condition for the functional central limit theorem in Hölder spaces, *Journal of Theoretical Probability*, **17**, 221–243 (2004).

[39] P.A. Sánchez, Cambios estructurales en series de tiempo: Una revisión del estrado del arte, *Revista de Ingenierías Universidad de Medellín*, **012**(6), 115–140 (2008).

[40] D.W.Shin, The limiting distribution of the residuals processes in non-stationary autoregressive processes, *Journal of time series analysis*, **06**, 723–736 (1998).

[41] S. M. Tang, I. B. MacNeil, The effect of serial correlation on tests for parameter change at unknown time, *The Annals of Statistics*, **21**, 552–575 (1993).

[42] D. W. Wichern, R. B. Miller, D. A. Hsu, Changes of variance in first-order autoregressive time series models - with an application, *Applied Statistics*, **25**, 248–256 (1976).

[43] H. Yu, High moment partial sum processes of residuals in ARMA models and their applications, *Journal of time series analysis*, **28**, 72–91 (2007).