AIMS Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 9(9): 24776–24795.
DOI: 10.3934/math.20241206
Received: 05 June 2024
Revised: 08 August 2024
Accepted: 14 August 2024
Published: 23 August 2024

Research article

Empirical likelihood method for detecting change points in network autoregressive models

Jingjing Yang[1], Weizhong Tian[2,*], Chengliang Tian[3], Sha Li[4] and Wei Ning[5]

[1] Department of Applied Mathematics, Xi'an University of Technology, Xi'an 710048, China
[2] College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518118, China
[3] College of Computer Science and Technology, Qingdao University, Qingdao 266071, China
[4] School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China
[5] Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403, USA

* **Correspondence:** Email: tianweizhong@sztu.edu.cn.

Abstract: The network autoregressive model is a super high-dimensional time series model that can fully explain social relationships. This model can fully reflect the complex relationships in reality. Therefore, it plays a vital role in detecting the inflection point problem of this network autoregressive model for economics and finance. In this paper, we proposed the change-point problem of detecting network autoregressive models using empirical likelihood statistics based on the expected error term of the switching rule being 0, using the empirical likelihood method. Moreover, the asymptotic null distribution of the proposed empirical likelihood statistic was investigated. Simulation studies based on different settings were considered, and the results showed that the power of test statistics is significant. In the end, the Chinese stock market was investigated to demonstrate the significance of the proposed method.

Keywords: network autoregressive models; change point; empirical likelihood; Gumbel distribution; Chinese stock market
Mathematics Subject Classification: 62C12, 62J05, 62G10

1. Introduction

Nowadays, network research has become increasingly important with the development of science and technology. A network system is composed of intricate relationships. For example, in a large social network, the subjects may be connected by more than one type of relationship. Instead, they can be complexly connected by multiple types of relationships, named the network autoregressive models studied by some scholars. Zhu et al. [1] established a network autoregressive model to study social relationships, which can fully consider all time-series information. However, subjects within a studied network may not be connected only through one type of association; instead, they could be complexly connected through multiple types of relationships. Therefore, Wei and Tian [2] proposed a network regression model with multiple types of connections and considered the impact of multiple connection nodes on the heterogeneity of continuous response variables. To express individual heterogeneity, Zhu and Pan [3] developed a grouped network vector autoregressive (GNAR) model based on the network autoregressive model. Zhu et al. [1] considered both non-specific variables and various connectivity relationships. Subsequently, Tian et al. [4] proposed a mixed network regression model that considers non-specific variable dependence on time based on their foundation. Huang et al. [5] proposed a network autoregressive model with the GARCH effect to describe the return dynamics of stock market indices. Due to different nodes having different effects on others, Tang et al. [6] proposed a penalty method for estimating the network vector autoregressive (NAR) models with different individual effects. Wang et al. [7] proposed a network binary segmentation method for detecting change points, which relies on a weighted average adjacency matrix. Xiao et al. [8] proposed the Huber estimator for estimating the parameters of network autoregressive models. Zhao and Liu [9] proposed the classifier-regularized approach for homogeneous analysis of network effects under the network autoregressive model.

The change-point problem was first proposed by Page [10] and has been widely applied in many fields. It mainly detects whether a change has occurred and determines its position. There are many methods for detecting change points, such as the cumulative sum (CUSUM), likelihood ratio test (LRT), Schwarz information criterion (SIC), Akaike information criterion (AIC), Bayesian information criterion (BIC), and modified information approach (MIC). Page [11] considered the problem of parameter changes in time series based on the CUSUM test. Kim [12] considered a test for a change point in linear regression by using the likelihood ratio statistic and studied the asymptotic behavior of the LRT statistic. Chen et al. [13] proposed the MIC for detecting change points in linear regression models. Jie [14] proposed the SIC to locate change points in simple linear regression models and multiple linear regression models. Basalamah et al. [15] proposed the MIC method to detect parameter changes in linear regression models with normally distributed error terms. Recently, Horváth et al. [16] proposed a weighted function based on the CUSUM of linear model residuals processed to detect the change points of linear regression models. Lee et al. [17] proposed a detector based on CUSUM of score vectors and residuals to investigate the sequential process of early detection of parameter changes in conditional heteroscedasticity time series models.

Empirical likelihood was proposed by Owen [18,19] and has been widely used due to its robustness to non-parametric properties and the efficiency of likelihood construction. Without knowing the distribution, the empirical likelihood method can be used to solve the problem, but it has some limitations in terms of computation. Specifically, many scholars use empirical likelihood methods to detect change points in regression models. Liu et al. [20] proposed a non-parametric method based on empirical likelihood to detect coefficient changes in linear regression models. Ning [21] considered the linear model of mean and proposed a non-parametric method for empirical likelihood testing to detect and estimate the position of change points. Zhao et al. [22] proposed an improved empirical likelihood ratio statistic to test for the presence of change points in long-term experiments and observations, which constructed empirical likelihood ratio statistics based on fitting residuals. Wu et al. [23] proposed a non-parametric method based on jackknife empirical likelihood (JEL) to detect changes in regression coefficients. Because empirical likelihood was initially proposed for independent data, applying it to related data such as time series data is difficult. Some scholars have conducted research on the transformation of dependent data into independent data. Akashi et al. [24] proposed empirical likelihood ratio statistics to detect change points when the position of the change point is unknown in autoregressive models. Gamage and Ning [25] used a powerful non-parametric method to propose empirical likelihood ratio statistics to detect changes in the parameter structure of autoregressive models. Yu et al. [26] proposed the empirical likelihood ratio test to detect structural changes in integer autoregressive (INAR) processes.

To the best of our knowledge, scholars have mainly focused on parameter estimation of network autoregressive models, and no one has used empirical likelihood methods to study the change-point problem. Detecting change points in network autoregressive models has great practical significance. For example, we can construct a network autoregressive model due to the intricate connections between stocks. Since some events may cause changes at a particular moment, detecting the location of these changes can provide better evidence for experts to study the stock market. According to some scholars, the idea of empirical likelihood detection of change points was adopted for time series data. Therefore, it is of great significance to perform change-point detection on network autoregressive models. The network autoregressive model is a high-dimensional time series model that can explain the natural world well by constructing adjacency matrices to clarify the relationships between subjects. This paper considers a non-parametric method to perform change-point detection on complex autoregressive models without knowing the distribution. In this paper, we propose an empirical likelihood method based on network autoregressive models to detect structural changes in the model. The structure of this article is as follows. Hypothesis testing and parameter estimation methods are proposed in Section 2. The empirical likelihood method is presented in Section 3. The simulation studies are considered in Section 4. Actual data application is given in Section 5. The conclusion of the paper is presented in Section 6.

2. The change-point problem in the network vector autoregression model

In the following, we introduce the network autoregressive model proposed by Wei and Tian [2],

$$Y_{i,t} = \beta_0 + \sum_{k=1}^{p} \beta_k n_{k,i}^{-1} \sum_{j=1}^{N} a_{i,j}^k Y_{j,t-1} + \beta_{p+1} Y_{i,t-1} + \epsilon_{i,t}, \tag{2.1}$$

where $i = 1, ..., N$ are the network nodes, $t = 1, ..., T$ represents the observation times, $Y_{i,t}$ represents the response variable of the ith node at time t, β_0 is the intercept term, β_k characterizes the kth connection effect for $k = 1, ..., p$, β_{p+1} characterizes the momentum effect, and $\epsilon_{i,t}$ are the independent random variables with mean 0 and variance σ^2. To describe the network structure composed of N nodes through the kth relationship, an adjacency matrix is defined: $A^k = (a_{i,j}^k) \in R^{N\times N}$, $k = 1, ..., p$. If node i and node j have the kth relation, then $a_{i,j}^k = 1$, otherwise $a_{i,j}^k = 0$. Since the autocorrelation between nodes is not considered, the diagonal of the adjacency matrix A^k is set to 0. $n_{k,i} = \sum_{j\neq i} a_{i,j}^k$ is the total number of nodes connected by the ith node through the kth relationship. $n_{k,i}^{-1} \sum_{j=1}^{N} a_{i,j}^k Y_{j,t-1}$ is the average impact of the kth connection of the ith node's neighbors. At the same time, it can be noted that $Y_t \in R^N$ with $0 \leq t \leq T$ is an ultra-high dimensional time series.

For the convenience of estimating unknown parameters, let $\boldsymbol{Y_t} = (Y_{1,t}, ..., Y_{N,t})' \in R^N$, $W^k = \text{diag}(n_{k,1}^{-1}, ..., n_{k,N}^{-1})A^k \in R^{N\times N}$, $\mathbf{1} = (1, ..., 1)' \in R^N$, and $\boldsymbol{\epsilon_t} = (\epsilon_{1,t}, ..., \epsilon_{N,t})' \in R^N$. The model (2.1) can be rewritten as:

$$\boldsymbol{Y_t} = \beta_0\mathbf{1} + \beta_1 W^1 \boldsymbol{Y_{t-1}} + \beta_2 W^2 \boldsymbol{Y_{t-1}} + \cdots + \beta_p W^p \boldsymbol{Y_{t-1}} + \beta_{p+1}\boldsymbol{Y_{t-1}} + \boldsymbol{\epsilon_t}. \tag{2.2}$$

We consider one change point k^* in the model (2.2), represented as:

$$\boldsymbol{Y_t} = \begin{cases} \beta_0\mathbf{1} + \beta_1 W^1 \boldsymbol{Y_{t-1}} + \beta_2 W^2 \boldsymbol{Y_{t-1}} + \cdots + \beta_p W^p \boldsymbol{Y_{t-1}} + \beta_{p+1}\boldsymbol{Y_{t-1}} + \boldsymbol{\epsilon_t}, & 1 \leq t \leq k^*, \\ \beta_0^*\mathbf{1} + \beta_1^* W^1 \boldsymbol{Y_{t-1}} + \beta_2^* W^2 \boldsymbol{Y_{t-1}} + \cdots + \beta_p^* W^p \boldsymbol{Y_{t-1}} + \beta_{p+1}^*\boldsymbol{Y_{t-1}} + \boldsymbol{\epsilon_t}, & k^*+1 \leq t \leq T, \end{cases} \tag{2.3}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_p, \beta_{p+1})'$ and $\boldsymbol{\beta^*} = (\beta_0^*, \beta_1^*, ..., \beta_p^*, \beta_{p+1}^*)'$ are the unknown parameters, and k^* is the unknown change-point position the needs to be estimated. To estimate the coefficient vectors $\boldsymbol{\beta}$ and $\boldsymbol{\beta^*}$, we rewrite model (2.3) as:

$$\boldsymbol{Y_t} = \begin{cases} \beta_0\mathbf{1} + \beta_1 W^1 \boldsymbol{Y_{t-1}} + \beta_2 W^2 \boldsymbol{Y_{t-1}} + \cdots + \beta_p W^p \boldsymbol{Y_{t-1}} + \beta_{p+1}\boldsymbol{Y_{t-1}} + \boldsymbol{\epsilon_t} = \boldsymbol{X_{t-1}}\beta + \boldsymbol{\epsilon_t}, & 1 \leq t \leq k^*, \\ \beta_0^*\mathbf{1} + \beta_1^* W^1 \boldsymbol{Y_{t-1}} + \beta_2^* W^2 \boldsymbol{Y_{t-1}} + \cdots + \beta_p^* W^p \boldsymbol{Y_{t-1}} + \beta_{p+1}^*\boldsymbol{Y_{t-1}} + \boldsymbol{\epsilon_t} = \boldsymbol{Z_{t-1}}\beta^* + \boldsymbol{\epsilon_t}, & k^*+1 \leq t \leq T, \end{cases}$$

where $\boldsymbol{X_{t-1}} = (X_{1,t-1}, ..., X_{N,t-1})' \in R^{N\times(p+2)}$, $X_{i,t-1} = (1, w_i^1 Y_{i,t-1}, ..., w_i^p Y_{i,t-1}, Y_{i,t-1})' \in R^{p+2}$, w_i^k is the ith row of W^k, for $1 \leq t \leq k^*$, $\boldsymbol{Z_{t-1}} = (Z_{1,t-1}, ..., Z_{N,t-1})' \in R^{N\times(p+2)}$, $Z_{i,t-1} = (1, w_i^1 Y_{i,t-1}, ..., w_i^p Y_{i,t-1}, Y_{i,t-1})' \in R^{p+2}$, and w_i^k is the ith row of W^k for $k^*+1 \leq t \leq T$. Thus, the estimated coefficients can be obtained as follows:

$$\hat{\boldsymbol{\beta}} = (\sum_{t=1}^{k^*} \boldsymbol{X'_{t-1}}\boldsymbol{X_{t-1}})^{-1} \sum_{t=1}^{k^*} \boldsymbol{X'_{t-1}}\boldsymbol{Y_t}, 1 \leq t \leq k^*, \tag{2.4}$$

$$\hat{\boldsymbol{\beta}}^{\boldsymbol{*}} = (\sum_{t=k^*+1}^{T} \boldsymbol{Z'_{t-1}}\boldsymbol{Z_{t-1}})^{-1} \sum_{t=k^*+1}^{T} \boldsymbol{Z'_{t-1}}\boldsymbol{Y_t}, k^*+1 \leq t \leq T. \tag{2.5}$$

Therefore, model (2.3) can be rewritten as:

$$Y_{i,t} = \begin{cases} \beta_0 1 + \beta_1 w_i^1 Y_{i,t-1} + \beta_2 w_i^2 Y_{i,t-1} + \cdots + \beta_p w_i^p Y_{i,t-1} + \beta_{p+1} Y_{i,t-1} + \epsilon_{i,t}, & 1 \leq t \leq k^*, \\ \beta_0^* 1 + \beta_1^* w_i^1 Y_{i,t-1} + \beta_2^* w_i^2 Y_{i,t-1} + \cdots + \beta_p^* w_i^p Y_{i,t-1} + \beta_{p+1}^* Y_{i,t-1} + \epsilon_{i,t}, & k^*+1 \leq t \leq T, \end{cases} \tag{2.6}$$

and the errors are estimated in the following,

$$\hat{\epsilon}_{i,t}(k^*) = \begin{cases} Y_{i,t} - [\beta_0 1 + \beta_1 w_i^1 Y_{i,t-1} + \beta_2 w_i^2 Y_{i,t-1} + \cdots + \beta_p w_i^p Y_{i,t-1} + \beta_{p+1} Y_{i,t-1}], & 1 \leq t \leq k^*, \\ Y_{i,t} - [\beta_0^* 1 + \beta_1^* w_i^1 Y_{i,t-1} + \beta_2^* w_i^2 Y_{i,t-1} + \cdots + \beta_p^* w_i^p Y_{i,t-1} + \beta_{p+1}^* Y_{i,t-1}], & k^*+1 \leq t \leq T. \end{cases}$$

According to the switching rule suggested by Liu and Qian [27], the estimated error can be expressed as follows,

$$\tilde{\epsilon}_{i,t}(k^*) = \begin{cases} Y_{i,t} - [\beta_0^* 1 + \beta_1^* w_i^1 Y_{i,t-1} + \beta_2^* w_i^2 Y_{i,t-1} + \cdots + \beta_p^* w_i^p Y_{i,t-1} + \beta_{p+1}^* Y_{i,t-1}], & 1 \leq t \leq k^*, \\ Y_{i,t} - [\beta_0 1 + \beta_1 w_i^1 Y_{i,t-1} + \beta_2 w_i^2 Y_{i,t-1} + \cdots + \beta_p w_i^p Y_{i,t-1} + \beta_{p+1} Y_{i,t-1}], & k^*+1 \leq t \leq T. \end{cases} \tag{2.7}$$

Under the null hypothesis H_0, if no change occurs, we can obtain $\boldsymbol{\beta} = \boldsymbol{\beta^*}$. Therefore, we rewrite the hypothesis test as

$$H_0 : E(\tilde{\epsilon}_{i,t}(k^*)) = 0, \text{ for all } k^*,$$
$$H_1 : \exists \text{ a } k^*, \; E(\tilde{\epsilon}_{i,t}(k^*)) \neq 0.$$

3. Empirical likelihood

Notice that under H_0, for any $k^* \in 1, ..., T$, there is $E(\tilde{\epsilon}_{i,t}(k^*)) = 0$. For a fixed k^*, let $p_{i,t}(k^*)$ be the mass probability at the value $\tilde{\epsilon}_{i,t}(k^*)$, and for $1 \le i \le N$, the constraint is $\sum_{t=1}^{T} p_{i,t}(k^*) = 1$. When $p_{i,t}(k^*) = \frac{1}{T}$, for any $t = 1, ...T$ and $1 \le i \le N$, the empirical likelihood $\prod_{t=1}^{T} p_{i,t}(k^*)$ reaches its maximum T^{-T}. The empirical likelihood ratio is $\prod_{t=1}^{T} Tp_{i,t}(k^*)$. Hence, for $1 \le i \le N$, the empirical log-likelihood ratio (ELR) statistic is defined as:

$$\Re(i, k^*) = \sup\left\{\sum_{t=1}^{T} \log(Tp_{i,t}(k^*)) | \sum_{t=1}^{T} p_{i,t}(k^*) = 1, p_{i,t}(k^*) \ge 0, \sum_{t=1}^{T} p_{i,t}(k^*)\tilde{\epsilon}_{i,t}(k^*) = 0\right\}.$$

Owen [19] showed, similar to the log-likelihood ratio test statistic in a parametric model, with mild regular conditions, for $1 \le i \le N$, $-2\Re(i, k^*) \to \chi^2_d$ in distribution as $T \to \infty$, where d is the rank of $Var(\tilde{\epsilon}_{i,t}(k^*))$.

Therefore, for $1 \le i \le N$, we propose an ELR test statistic for the change-point detection in the following,

$$Z_{i,k^*} = -2\Re(i, k^*) = -2\sup\left\{\sum_{t=1}^{T} \log(Tp_{i,t}(k^*)) | \sum_{t=1}^{T} p_{i,t}(k^*) = 1, p_{i,t}(k^*) \ge 0, \sum_{t=1}^{T} p_{i,t}(k^*)\tilde{\epsilon}_{i,t}(k^*) = 0\right\}. \tag{3.1}$$

A Lagrangian argument gives

$$p_{i,t}(k^*) = \frac{1}{T[1 + \lambda_i \tilde{\epsilon}_{i,t}(k^*)]},$$

where λ_i is chosen such that $\sum_{t=1}^{T} p_{i,t}(k^*)\tilde{\epsilon}_{i,t}(k^*) = 0$ for $1 \le i \le N$. After plugging back $p_{i,t}(k^*)$ in (3.1), Z_{i,k^*} can be rewritten as:

$$Z_{i,k^*} = 2\left\{\sum_{t=1}^{T} \log[1 + \lambda_i \tilde{\epsilon}_{i,t}(k^*)]\right\}. \tag{3.2}$$

Define the score function

$$\Phi(\lambda_i') = \frac{\partial Z_{i,k^*}}{2\partial \lambda_i'} = \sum_{t=1}^{T} \frac{\tilde{\epsilon}_{i,t}(k^*)}{1 + \lambda_i \tilde{\epsilon}_{i,t}(k^*)}.$$

Then, $\hat{\lambda}_i$ are determined by

$$\Phi(\hat{\lambda}_i) = 0. \tag{3.3}$$

Therefore, for $1 \le i \le N$, (3.2) can be rewritten as:

$$Z_{i,k^*} = 2\left\{\sum_{t=1}^{T} \log[1 + \hat{\lambda}_i \tilde{\epsilon}_{i,t}(k^*)]\right\}, \tag{3.4}$$

and we propose the following EL test statistic:

$$M_{N,T} = \max_{1 \le i \le N, 1 \le k^* \le T} \{Z_{i,k^*}\}.$$

However, when the position of the change point is close to 1 and T, the estimation efficiency will be poor because there is little information available for parameter estimation, in part with fewer samples, which leads to inaccurate estimation. Therefore, we suggest the trimmed likelihood ratio statistic as

$$M'_{N,T} = \max_{1 \le i \le N, k_0 \le k^* \le T - k_0} \{Z_{i,k^*}\}. \tag{3.5}$$

To choose suitable values for k_0, Csörgő and Horváth [28] gave the conditions of k_0 such that the trimmed test statistic follows an extreme distribution asymptotically. In this paper, we choose $k_0 = 2[T^{\frac{1}{2}}]$ for simplicity and convenience, where $[x]$ means the largest integer not larger than x.

Next, we will outline the main results of the asymptotic distribution under the null hypothesis and the consistency under the alternative hypothesis.

Theorem 1. *Assume that H_0 holds and C.1–C.3 are satisfied. Then under the null model,*

$$\lim_{N,T \to \infty} Pr\left\{A(\log u(T))(M'_{N,T})^{\frac{1}{2}} \le x + D_r(\log u(T))\right\} = \exp(-e^{-x}),$$

for all x, where $A(x) = (2\log x)^{\frac{1}{2}}$, $D_r(x) = 2\log x + (\frac{r}{2})\log\log x - \log\Gamma(\frac{r}{2})$, $u(T) = \frac{T^2 + (2[T^{\frac{1}{2}}])^2 - 2T[T^{\frac{1}{2}}]}{(2[T^{\frac{1}{2}}])^2}$, and r is the dimension of the parameter space.

Theorem 2. *Under the conditions of Theorem 1, for $1 \le i \le N$, and if there exists a positive constant c_0 satisfying $0 < c_0 \le supE\log(1 + \lambda_i \tilde{\epsilon}_{i,t}(k^*)) < \infty$, if change point k^* satisfies $\frac{k^*}{T} \to c \in (0,1)$ as* $\min(N,T) \to \infty$*, then the ELR test is consistent. That is, under the alternative hypothesis,*

$$Z_{i,k^*} \to \infty$$

in probability.

Proofs of the above two theorems are given in the Supplementary Materials.

Theorem 1 indicates that, under the null hypothesis, the asymptotic distribution of the EL test statistic is the Gumbel extreme value distribution.

For any given r,α, N, and T, if $M'_{N,T} < c_{r,\alpha,T,}$ we fail to reject H_0, where $c_{r,\alpha,T}$ are the critical values for r, α, and T. Applying the above Theorem 1, $c_{r,\alpha,T}$ is derived as follows:

$$\begin{aligned}
1 - \alpha &= P[M'_{N,T} < c_{r,\alpha,T}|H_0] = P[0 < M'_{N,T} < c_{r,\alpha,T}|H_0] \\
&= P[0 < (M'_{N,T})^{\frac{1}{2}} < (c_{r,\alpha,T})^{\frac{1}{2}}|H_0] \\
&= P[-D_r(\log u(T)) < A(\log u(T))(M'_{N,T})^{\frac{1}{2}} - D_r(\log u(T)) < A(\log u(T))(c_{r,\alpha,T})^{\frac{1}{2}} - D_r(\log u(T))|H_0] \\
&= P[A(\log u(T))(M'_{N,T})^{\frac{1}{2}} - D_r(\log u(T)) < A(\log u(T))(c_{r,\alpha,T})^{\frac{1}{2}} - D_r(\log u(T))] \\
&- P[A(\log u(T))(M'_{N,T})^{\frac{1}{2}} - D_r(\log u(T)) < -D_r(\log u(T))] \\
&\cong \exp\{-\exp\{D_r(\log u(T)) - A(\log u(T))(c_{r,\alpha,T})^{\frac{1}{2}}\}\} - \exp\{-\exp\{D_r(\log u(T))\}\}.
\end{aligned}$$

Therefore,

$$c_{r,\alpha,T} \cong \left[\frac{\log[-\log(1 - \alpha + \exp\{-\exp\{D_r(\log u(T))\}\})] - D_r(\log u(T))}{-A(\log u(T))}\right]^2.$$

Tables 1 and 2 provide the critical values $c_{r,\alpha,T}$ for different r, α, and T.

Table 1. Approximate critical values with $r = 3$, α, and T.

T	$\alpha = 0.01$	$\alpha = 0.05$	T	$\alpha = 0.01$	$\alpha = 0.05$
100	22.6973	13.4807	160	22.8862	14.2271
110	22.7598	13.7558	170	22.8737	14.1838
120	22.8177	13.9812	180	22.9103	14.3091
130	22.8076	13.9432	200	22.9301	14.3748
140	22.8562	14.1218	300	23.0595	14.7747
150	22.8445	14.0796	400	23.1329	14.9818

Table 2. Approximate critical values with $r = 4$, α, and T.

T	$\alpha = 0.01$	$\alpha = 0.05$	T	$\alpha = 0.01$	$\alpha = 0.05$
100	22.27445	13.15312	160	22.94089	14.27021
110	22.5194	13.56895	170	22.90245	14.20635
120	22.72195	13.9062	180	23.01337	14.39055
130	22.68795	13.84963	200	23.0713	14.48663
140	22.84742	14.11488	300	23.42141	15.06468
150	22.8099	14.05249	400	23.60132	15.35929

4. Simulation study

In this article, we consider $p = 1$ and $p = 2$ in the model (2.1). For the construction of the adjacency matrix A, we follow the same steps described in Zhu et al. [1].

4.1. *The network autoregressive model with $p = 1$*

The adjacency matrix A is constructed from a power-law distribution model: First, we generate for each node its in-degree $d_i = \sum_j a_{ji}$ according to the discrete power-law distribution, that is, $P(d_i = l) = ql^{-b}$ for a normalizing constant q and the exponent parameter $b = 1.2$. A smaller b value implies a heavier distribution tail. Next, for the ith node, we randomly select d_i nodes to be its followers.

To study the detection performance of our proposed method, we study the values of the power of the empirical likelihood statistic. Consider the following model:

$$Y_{i,t} = \begin{cases} \beta_0 1 + \beta_1 w_i^1 Y_{i,t-1} + \beta_2 Y_{i,t-1} + \epsilon_{i,t}, & 1 \le t \le k^*, \\ \beta_0^* 1 + \beta_1^* w_i^1 Y_{i,t-1} + \beta_2^* Y_{i,t-1} + \epsilon_{i,t}, & k^* + 1 \le t \le T. \end{cases}$$

In our simulation, we set parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)' = (0, 0.6, 0.1)'$ and $\boldsymbol{\beta}^* = (\beta_0^*, \beta_1^*, \beta_2^*)' = (0, -0.7, 0.8)'$. Sample sizes $N = 25, 30$ and $T = 100, 200, 400$ are considered. At the same time, we choose the position of the change point $k^* = 0.25T, 0.5T, 0.75T$. According to Theorem 1, we can obtain the p-value and reject the H_0 when it is below the significance level $\alpha = 0.01, 0.05$. At the same time, we computer the value of the power, which is the probability that the H_0 was rejected in 1000 simulations. Moreover, the values of the power are obtained by considering four different distributions of error terms that satisfy $E(\epsilon_{i,t}) = 0$ and $Var(\epsilon_{i,t}) = 1$. The four distributions are the: (i) Standard normal distribution $N(0,1)$, (ii) exponential distribution $\exp(1) - 1$, (iii) chi-square distribution $\frac{1}{2\sqrt{2}}(\chi^2(4) - 4)$, and (iv) student-$t$ distribution $\frac{1}{\sqrt{2}}t(4)$. Next, Tables 3–6 show the values of the power of parameter changes in the network autoregressive model $p = 1$ detected under four different distributions.

Table 3. Power of the EL test when $p = 1$, $\alpha = 0.05$, and $N = 25$.

T	k^*	$N(0,1)$	$\exp(1) - 1$	$\frac{1}{2\sqrt{2}}(\chi^2(4) - 4)$	$\frac{1}{\sqrt{2}}t(4)$
	$k^* = 0$	0.032	0.037	0.030	0.045
	$k^* = 25$	0.978	0.979	0.974	0.982
$T = 100$	$k^* = 50$	0.991	0.992	0.986	0.992
	$k^* = 75$	0.997	0.996	0.998	0.996
	$k^* = 0$	0.038	0.039	0.031	0.046
	$k^* = 50$	0.966	0.975	0.968	0.969
$T = 200$	$k^* = 100$	0.969	0.972	0.976	0.97
	$k^* = 150$	0.974	0.987	0.99	0.993
	$k^* = 0$	0.040	0.040	0.038	0.048
	$k^* = 100$	0.965	0.952	0.952	0.956
$T = 400$	$k^* = 200$	0.967	0.963	0.965	0.961
	$k^* = 300$	0.973	0.969	0.975	0.974

Table 4. Power of the EL test when $p = 1$, $\alpha = 0.05$, and $N = 30$.

T	k^*	$N(0,1)$	$\exp(1) - 1$	$\frac{1}{2\sqrt{2}}(\chi^2(4) - 4)$	$\frac{1}{\sqrt{2}}t(4)$
	$k^* = 0$	0.018	0.035	0.040	0.044
	$k^* = 25$	0.985	0.981	0.987	0.988
$T = 100$	$k^* = 50$	0.992	0.994	0.994	0.995
	$k^* = 75$	0.998	0.999	1	0.999
	$k^* = 0$	0.022	0.043	0.040	0.020
	$k^* = 50$	0.974	0.979	0.975	0.984
$T = 200$	$k^* = 100$	0.985	0.982	0.98	0.99
	$k^* = 150$	0.993	0.991	0.991	0.991
	$k^* = 0$	0.010	0.035	0.041	0.025
	$k^* = 100$	0.972	0.973	0.972	0.971
$T = 400$	$k^* = 200$	0.974	0.977	0.977	0.974
	$k^* = 300$	0.986	0.985	0.98	0.988

Table 5. Power of the EL test when $p = 1$, $\alpha = 0.01$, and $N = 25$.

T	k^*	$N(0,1)$	$\exp(1)-1$	$\frac{1}{2\sqrt{2}}(\chi^2(4)-4)$	$\frac{1}{\sqrt{2}}t(4)$
	$k^* = 0$	0.002	0.010	0.008	0.008
	$k^* = 25$	0.96	0.956	0.95	0.963
$T = 100$	$k^* = 50$	0.975	0.974	0.971	0.974
	$k^* = 75$	0.981	0.988	0.985	0.982
	$k^* = 0$	0.005	0.013	0.008	0.008
	$k^* = 50$	0.949	0.957	0.946	0.947
$T = 200$	$k^* = 100$	0.957	0.962	0.954	0.95
	$k^* = 150$	0.964	0.971	0.978	0.975
	$k^* = 0$	0.011	0.008	0.010	0.011
	$k^* = 100$	0.945	0.93	0.932	0.934
$T = 400$	$k^* = 200$	0.949	0.936	0.934	0.941
	$k^* = 300$	0.953	0.962	0.956	0.953

Table 6. Power of the EL test when $p = 1$, $\alpha = 0.01$, and $N = 30$.

T	k^*	$N(0,1)$	$\exp(1)-1$	$\frac{1}{2\sqrt{2}}(\chi^2(4)-4)$	$\frac{1}{\sqrt{2}}t(4)$
	$k^* = 0$	0.001	0.002	0.004	0.004
	$k^* = 25$	0.97	0.963	0.976	0.975
$T = 100$	$k^* = 50$	0.982	0.983	0.977	0.989
	$k^* = 75$	0.992	0.996	0.997	0.993
	$k^* = 0$	0.004	0.002	0.007	0.008
	$k^* = 50$	0.961	0.96	0.957	0.968
$T = 200$	$k^* = 100$	0.972	0.977	0.966	0.977
	$k^* = 150$	0.985	0.979	0.98	0.985
	$k^* = 0$	0.007	0.006	0.009	0.010
	$k^* = 100$	0.954	0.958	0.95	0.948
$T = 400$	$k^* = 200$	0.966	0.959	0.961	0.968
	$k^* = 300$	0.972	0.974	0.966	0.971

It can be seen from Tables 3–6 that the values of the power are approximately similar for four different error distributions under the same conditions. For $\alpha = 0.05$, almost all of the values of the power are greater than 95%. For $\alpha = 0.01$, the values of the power reach over 90%. Specifically, when $N = 30$, the values of the power of the four distributions at the $0.75T$ position are generally close to 1. At the same time, it can effectively control the error of type I when $\alpha = 0.01$. Moreover, the values of the power increase as N increases and decrease as T increases. Perhaps due to the relatively small value of N, there is increasing bias in estimated parameter values, which in turn affects the test statistic constructed based on the error term. This ultimately leads to a decrease in the values of the power as T increases. The values of the power change with the position of the change point for all error distributions and also increase with the increase of the change-point position.

4.2. The network autoregressive model with $p = 2$

The constructions method of adjacency matrices A^1 and A^2 are as follows. First, the adjacency matrix A^1 construction is the same as that of $p = 1$. Next, the adjacency matrix A^2 is simulated from the stochastic block model with 5 blocks. We randomly assign a block label to each node with equal probability, set $P(a_{i,j} = 1) = 0.3N^{-0.3}$ if nodes i and j belong to the same block, and $P(a_{i,j} = 1) = 0.3N^{-1}$ otherwise.

Then, we study the values of the power of the empirical likelihood statistic. Consider the following model:

$$Y_{i,t} = \begin{cases} \beta_0 1 + \beta_1 w_i^1 Y_{i,t-1} + \beta_2 w_i^2 Y_{i,t-1} + \beta_3 Y_{i,t-1} + \epsilon_{i,t}, & 1 \le t \le k^*, \\ \beta_0^* 1 + \beta_1^* w_i^1 Y_{i,t-1} + \beta_2^* w_i^2 Y_{i,t-1} + \beta_3^* Y_{i,t-1} + \epsilon_{i,t}, & k^* + 1 \le t \le T. \end{cases}$$

The parameters are set to $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)' = (0, 0.1, 0.1, 0.1)'$ and $\boldsymbol{\beta^*} = (\beta_0^*, \beta_1^*, \beta_2^*, \beta_3^*)' = (0, -1, -0.2, 0.6)'$. Consider the same sample size and change-point position as $p = 1$. The p-value based on 1000 simulation runs based on $\alpha = 0.01, 0.05$ are calculated. Similarly, like the $p = 1$ model, we consider the same four distributions for the error term to obtain the values of the power. Next, Tables 7–10 show the values of the power of parameter changes in the network autoregressive model $p = 2$ detected under four different distributions.

Table 7. Power of the EL test when $p = 2$, $\alpha = 0.05$, and $N = 25$.

T	k^*	$N(0,1)$	$\exp(1) - 1$	$\frac{1}{2\sqrt{2}}(\chi^2(4) - 4)$	$\frac{1}{\sqrt{2}}t(4)$
$T = 100$	$k^* = 0$	0.020	0.057	0.044	0.032
	$k^* = 25$	0.906	0.898	0.905	0.882
	$k^* = 50$	0.916	0.927	0.924	0.916
	$k^* = 75$	0.945	0.944	0.951	0.957
$T = 200$	$k^* = 0$	0.033	0.060	0.052	0.044
	$k^* = 50$	0.861	0.865	0.877	0.872
	$k^* = 100$	0.88	0.891	0.878	0.878
	$k^* = 150$	0.923	0.923	0.909	0.911
$T = 400$	$k^* = 0$	0.040	0.070	0.060	0.050
	$k^* = 100$	0.852	0.856	0.862	0.852
	$k^* = 200$	0.877	0.868	0.876	0.862
	$k^* = 300$	0.885	0.889	0.89	0.883

Table 8. Power of the EL test when $p = 2$, $\alpha = 0.05$, and $N = 30$.

T	k^*	$N(0,1)$	$\exp(1)-1$	$\frac{1}{2\sqrt{2}}(\chi^2(4)-4)$	$\frac{1}{\sqrt{2}}t(4)$
$T = 100$	$k^* = 0$	0.017	0.050	0.037	0.041
	$k^* = 25$	0.924	0.932	0.926	0.929
	$k^* = 50$	0.928	0.948	0.943	0.946
	$k^* = 75$	0.958	0.973	0.975	0.97
$T = 200$	$k^* = 0$	0.027	0.044	0.047	0.040
	$k^* = 50$	0.899	0.895	0.9	0.88
	$k^* = 100$	0.909	0.911	0.912	0.927
	$k^* = 150$	0.936	0.935	0.931	0.939
$T = 400$	$k^* = 0$	0.020	0.045	0.045	0.047
	$k^* = 100$	0.859	0.881	0.876	0.875
	$k^* = 200$	0.888	0.882	0.893	0.885
	$k^* = 300$	0.916	0.91	0.909	0.895

Table 9. Power of the EL test when $p = 2$, $\alpha = 0.01$, and $N = 25$.

T	k^*	$N(0,1)$	$\exp(1)-1$	$\frac{1}{2\sqrt{2}}(\chi^2(4)-4)$	$\frac{1}{\sqrt{2}}t(4)$
$T = 100$	$k^* = 0$	0.005	0.008	0.009	0.006
	$k^* = 25$	0.872	0.862	0.85	0.842
	$k^* = 50$	0.873	0.89	0.873	0.869
	$k^* = 75$	0.922	0.913	0.915	0.926
$T = 200$	$k^* = 0$	0.010	0.015	0.013	0.011
	$k^* = 50$	0.825	0.815	0.826	0.833
	$k^* = 100$	0.84	0.853	0.843	0.843
	$k^* = 150$	0.878	0.892	0.884	0.875
$T = 400$	$k^* = 0$	0.018	0.020	0.022	0.020
	$k^* = 100$	0.822	0.813	0.824	0.809
	$k^* = 200$	0.829	0.828	0.839	0.825
	$k^* = 300$	0.843	0.856	0.843	0.843

Table 10. Power of the EL test when $p = 2$, $\alpha = 0.01$, and $N = 30$.

T	k^*	$N(0,1)$	$\exp(1)-1$	$\frac{1}{2\sqrt{2}}(\chi^2(4)-4)$	$\frac{1}{\sqrt{2}}t(4)$
	$k^*=0$	0.004	0.006	0.006	0.008
	$k^*=25$	0.89	0.888	0.879	0.882
$T=100$	$k^*=50$	0.897	0.912	0.91	0.913
	$k^*=75$	0.929	0.944	0.949	0.933
	$k^*=0$	0.009	0.013	0.014	0.011
	$k^*=50$	0.844	0.862	0.866	0.844
$T=200$	$k^*=100$	0.87	0.878	0.868	0.885
	$k^*=150$	0.904	0.901	0.902	0.902
	$k^*=0$	0.011	0.016	0.016	0.018
	$k^*=100$	0.837	0.846	0.842	0.839
$T=400$	$k^*=200$	0.858	0.863	0.853	0.845
	$k^*=300$	0.872	0.871	0.861	0.861

It can be seen from Tables 7–10 that the results for the case of $p = 2$ have the same pattern as the case of $p = 1$. Although the values of the power for all four distributions have all decreased compared to the case of $p = 1$, they are all above 80% for $\alpha = 0.05$ or $\alpha = 0.01$. Moreover, the biggest value of the power appears at the $0.75T$ position of $\alpha = 0.05$, $N = 30$, and $T = 100$; and the smallest value of the power appears at the $0.25T$ position of $\alpha = 0.01$, $N = 25$, and $T = 400$. Meanwhile, it can be concluded that the proposed method can effectively control type 1 errors. This indicates that the proposed ELR is sensitive and robust.

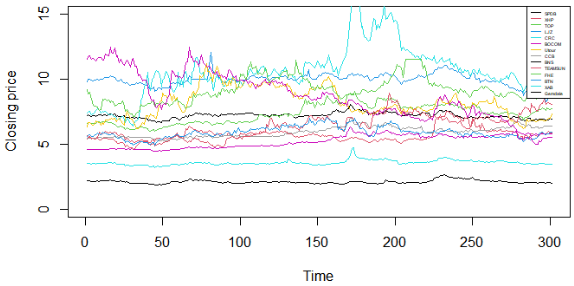
5. Application

We apply our method to the Chinese stock market dataset, which contains the daily closing prices of 30 stocks from August 18, 2022 to November 15, 2023. The detailed information for these 30 stocks is given in the Supplementary Materials. Furthermore, a time series chart of the daily closing prices of 30 stocks with different prices from 0 to 230 (CNY) is displayed in Figure 1. In order to make the time series chart more aesthetically pleasing, we divided them into four charts for display. The stock prices range from 0–15, as shown in Figure 1(a); 10–40, as shown in Figure 1(b); 20–60, as shown in Figure 1(c); and 70–230, as shown in Figure 1(d).

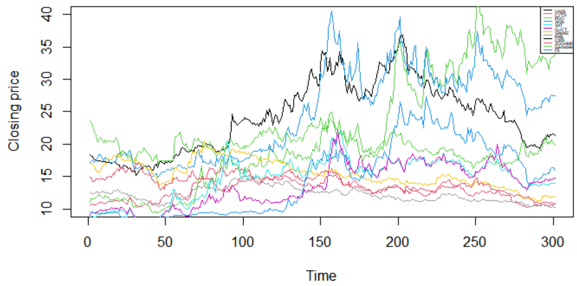
The response variables of this dataset ($Y_{i,t}, i = 1, ..., 30, t = 1, ..., 300$) take into account the daily closing prices of 30 stocks over 300 days. We consider the following two different connections: A^1 is composed of four regional sectors of stocks: Shanghai, Beijing, Shaanxi, and Shenzhen, and A^2 is constructed by five industry sectors: real estate, financial, tourism, media, and technology. Table 11 shows the construction of adjacency matrices. From Formula (3.5), $k^* = 251$ with the corresponding $p_{value} = 5.6488 \times 10^{-7}$ can be obtained. Therefore, the change point is detected at position 251, corresponding to August 29, 2023. In addition, Figure 2 shows the values of Z_{i,k^*} for 30 stocks under different k^* conditions. Figure 2 shows the maximum value of Z_{i,k^*} at position $k^* = 251$, with $M'_{N,T} = 112.1537$.

Table 11. Regions and industry sectors of 30 stocks.

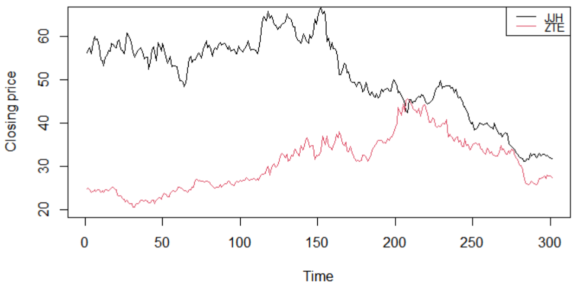
Name of stock	Regional sector	Industry sector
Shanghai Pudong Development Bank (SPDB)	Shanghai	Financial
Xinhuangpu (XHP)	Shanghai	Real estate
The Oriental Pearl (TOP)	Shanghai	Media
Jinjiang Hotel (JJH)	Shanghai	Tourism
Lujiazui (LJZ)	Shanghai	Real estate
China Reform Culture (CRC)	Shanghai	Media
ACM Research ShanghaiH (ACMSH)	Shanghai	Technology
Bank of Communications (BOCOM)	Shanghai	Financial
Unisplendour (UNIS)	Beijing	Technology
UTour Group (Utour)	Beijing	Tourism
China Youth Travel Service (CYTS)	Beijing	Tourism
China Duty Free Group (CDFG)	Beijing	Tourism
China Construction Bank (CCB)	Beijing	Financial
People's Daily Online (PDO)	Beijing	Media
Xinhuanet (XHN)	Beijing	Media
Beijing North Star (BNS)	Beijing	Real estate
Beijing Teamsun (TEAMSUN)	Beijing	Technology
Fenghuo Electronics (FHE)	Shaanxi	Technology
Xi'an Tourism (XAT)	Shaanxi	Tourism
Broadcast and TV Network (BTN)	Shaanxi	Media
Xi'an Bank (XAB)	Shaanxi	Financial
Qujiang Cultural Tourism (QJCT)	Shaanxi	Tourism
China Vanke (VANKE)	Shenzhen	Real estate
Ping An Bank (PAB)	Shenzhen	Financial
Zhongxing Telecommunication Equipment (ZTE)	Shenzhen	Technology
China Merchants Shekou (CMS)	Shenzhen	Real estate
Shenzhen ZQGAME (ZQGAME)	Shenzhen	Technology
Gemdale Corporation (Gemdale)	Shenzhen	Real estate
Foxconn Industrial Internet (FII)	Shenzhen	Technology
China Merchants Bank (CMB)	Shenzhen	Financial



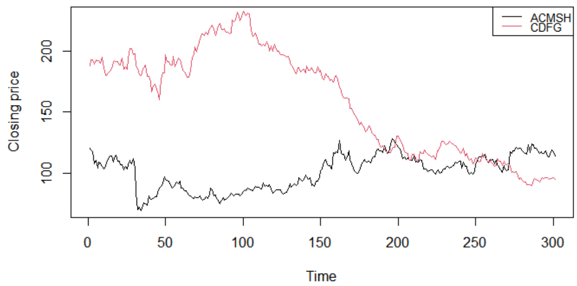
(a) The closing price of stocks is between 0 and 15.



(b) The closing price of stocks is between 10 and 40.



(c) The closing price of stocks is between 20 and 65.



(d) The closing price of stocks is between 70 and 230.

Figure 1. Time series chart of the daily closing prices of 30 stocks.

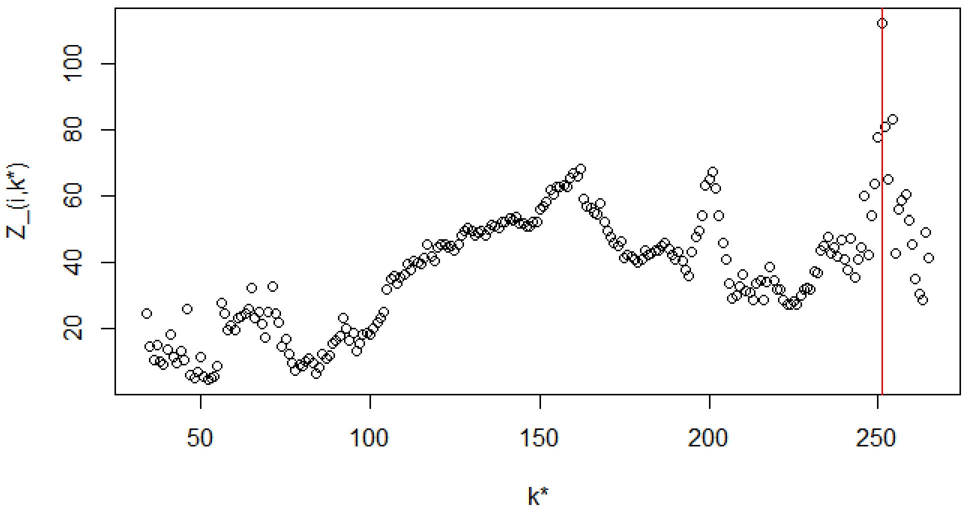


Figure 2. The values of Z_{i,k^*} in the data.

The connecting effects before and after August 29, 2023 can be calculated as $\hat{\beta} = (5.6314 \times$

$10^{-2}, -6.9384\times10^{-4}, -5.1843\times10^{-5}, 9.9816\times10^{-1})'$ and $\hat{\beta}^{*} = (9.9786\times10^{-2}, -4.022\times10^{-3}, -9.6017\times 10^{-4}, 9.9765\times10^{-1})'$. In fact, in late August 2023, the Chinese government issued some home purchase policies, such as recognizing a house but not a loan, lowering mortgage interest rates, and driving people's consumption through implementation in different regions, greatly promoting the development of the real estate and financial industries. The change in connectivity effect from $\hat{\beta}_1 = -6.9384 \times 10^{-4}$ to $\hat{\beta}_1^{*} = -4.022\times10^{-3}$ may be due to the different implementation times of government-issued housing policies in different regions. The change in connectivity effect from $\hat{\beta}_2 = -5.1843 \times 10^{-5}$ to $\hat{\beta}_2^{*} = -9.6017\times10^{-4}$ may be due to the degree to which the government's purchasing policies affect different industry sectors.

6. Conclusions

In this paper, we considered the EL method to detect structural changes for the network autoregressive models. The asymptotic null distribution of the test statistic is the Gumbel extreme value distribution, which was also studied. Through the simulation studies, different error distributions were illustrated, and the results proved that the proposed test statistic has good performance in detecting the change points. The simulation experimental results show that our method can effectively identify changes in the given network autoregressive model. The final application of the Chinese stock market further demonstrated the practical significance of the proposed method. In the future, we will extend the network model to the spatiotemporal data and consider the potential change-point problem. On the other hand, due to the EL method's computational limitations, we can adopt other methods, such as JEL and Adjusted Empirical Likelihood (AEL), in the future to improve the computational problem. Finally, we can also apply the method to more updated network regression models to detect change points in the future.

Author contributions

W. T., C. T., and W. N.: Conceptualization, Methodology, Validation, Investigation, Resources, Supervision, Project administration, Visualization, Writing-review and editing; J. Y. and S. L.: Software, Formal analysis, Data curation, Writing-original draft preparation, Visualization. All authors have read and agreed to the published version of the manuscript.

Use of AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors thank the editor and three anonymous referees for their valuable comments and suggestions that helped to improve this article significantly.

The research was supported by the Natural Science Foundation of Top Talent of SZTU (GDRC202214).

Conflict of interest

The authors declare no conflicts of interest.

References

1. X. Zhu, R. Pan, G. Li, Y. Liu, H. Wang, Network vector autoregression, *Ann. Statist.*, **45** (2017), 1096–1123. http://doi.org/10.1214/16-AOS1476
2. F. Wei, W. Tian, Heterogeneous connection effects, *Stat. Prob. Lett.*, **133** (2018), 9–14. http://doi.org/10.1016/j.spl.2017.09.015
3. X. Zhu, R. Pan, Grouped network vector autoregression, *Statistica Sinica.*, **30** (2020), 1437–1462. http://doi.org/10.5705/ss.202017.0533
4. W. Tian, F. Wei, T. Brown, Mixture network autoregressive model with application on students' successes, *Front. Math. China.*, **15** (2020), 141–154. http://doi.org/10.1007/s11464-020-0813-5
5. S. Huang, H. Chiang, Y. Lin, A network autoregressive model with GARCH effects and its applications, *Plo. One*, **16** (2021), e0255422. http://doi.org/10.1371/JOURNAL.PONE.0255422
6. Y. Tang, Y. Bai, T. Huang, Network vector autoregression with individual effects, *Metrika*, **84** (2021), 1–19. http://doi.org/10.1007/S00184-020-00805-Y
7. D. Wang, Y. Yu, A. Rinaldo, Optimal change point detection and localization in sparse dynamic networks, *Annal. Stat.*, 2021. http://doi.org/10.1214/20-AOS1953
8. X. Xiao, X. Xu, W. Zhong, Huber estimation for the network autoregressive model, *Stat. Prob. Lett.*, **203** (2023). http://doi.org/10.1016/J.SPL.2023.109917
9. J. Zhao, J. Liu, Homogeneous analysis on network effects in network autoregressive model, *Finance Research Lett.*, **58** (2023). http://doi.org/10.1016/J.FRL.2023.104671
10. E. S. Page, Continuous inspection schemes, *Biometrika*, **41** (1954), 100–115. https://doi.org/10.1093/biomet/41.1-2.100.
11. E. S. Page, A test for a change in a parameter occurring at an unknown point, *Biometrika*, **42** (1955), 523–527. http://doi.org/10.2307/2333401
12. H. J. Kim, Tests for a change-point in linear regression, *IMS Lect. Notes-Monogr. Series*, **23** (1994), 170–176. Available from: `http://www.jstor.org/stable/4355772`
13. J. Chen, A. K. Gupta, J. Pan, Information criterion and change point problem for regular models, *Sankhyā*, **68** (2006), 252–282. Available from: `https://www.jstor.org/stable/25053496`
14. C. Jie, Testing for a change point in linear regression models, *Commun. Stat.-Simul. Comput.*, **27** (2007), 2481–2493. http://doi.org/10.1080/03610929808832238
15. D. Basalamah, K. K. Said, W. Ning, Y. Tian, Modified information criterion for linear regression change-point model with its applications, *Commun. Stat.-Simul. Comput.*, 2019, 1–18. http://doi.org/10.1080/03610918.2018.1554109

16. L. Horváth, G. Rice, Y. Zhao, Testing for changes in linear models using weighted residuals, *J. Multivar. Anal.*, **198** (2023), 105210. http://doi.org/10.1016/J.JMVA.2023.105210
17. Y. Lee, S. Kim, H. Oh, Sequential change-point detection in time series models with conditional heteroscedasticity, *Economics Lett.*, **236** (2024), 111597. http://doi.org/10.1016/J.ECONLET.2024.111597
18. A. B. Owen, Empirical likelihood ratio confidence Intervals for a single functional, *Biometrika*, **75** (1988), 237–249. http://doi.org/10.1093/biomet/75.2.237
19. A. B. Owen, A. B. Empirical likelihood ratio confidence regions, *Annal.Statist.*, **18** (1990), 90–120. http://doi.org/10.1214/aos/1176347494
20. Y. Liu, C. Zou, R. Zhang, Empirical likelihood ratio test for a change-point in linear regression model, *Commun. Stat.-Theory Methods*, **37** (2008), 2551–2563. http://doi.org/10.1080/03610920802040373
21. W. Ning, Empirical likelihood ratio test for a mean change point model with a linear trend followed by an abrupt change, *J. Appl. Stat.*, **39** (2012), 947–961. http://doi.org/10.1080/02664763.2011.628647
22. H. Zhao, H. Chen, W. Ning, Changepoint analysis by modified empirical likelihood method in two-phase linear regression models, *Open J. Appl. Sci.*, **3** (2013), 1–6.
23. X. Wu, S. Zhang, Q. Zhang, S. Ma, Detecting change point in linear regression using jackknife empirical likelihood, *Stats. Interf.*, **9** (2015), 113–122. http://doi.org/10.4310/SII.2016.V9.N1.A11
24. F. Akashi, H. Dette, Y. Liu, Change point detection in autoregressive models with no moment assmptions, *J. Time Series Anal.*, **5** (2018), 763–786. https://doi.org/10.1111/jtsa.12405
25. R. D. P. Gamage, W. Ning, Empirical likelihood for change point detection in autoregressive models, *J. Korean Statist. Soci.*, 2020, 1–29. http://doi.org/10.1007/s42952-020-00061-w
26. K. Yu, H. Wang, C. H. Weiß, An empirical-likelihood-based structural-change test for INAR processes, *J. Statist. Comput. Simul.*, **93** (2023), 442–458. http://doi.org/10.1080/00949655.2022.2109635
27. Z. Liu, L. Qian, Changepoint estimation in a segmented linear regression via empirical likelihood, *Commun. Statist.-Simul. Comput.*, **39** (2010), 85–100. http://doi.org/10.1080/03610910903312193
28. M. Csörgő, L. Horváth, *Limit theorems in Change-Point analysis*, New York: Wiley and Sons, 1971.

Supplementary Materials

The regular conditions needed are listed as follows. For $1 \leq i \leq N$, assume

C.1. $\mathrm{rank}(X_{i,t-1})=\mathrm{rank}(Z_{i,t-1})=d$ for $k_0 \leq k^* \leq T - k_0$.

C.2. There are some $\delta > 0, \nu > 0, \nu > 2 + 27/\min(1+\delta)$, $\sigma_{i,1}^2 > 0$, and $\sigma_{i,2}^2 > 0$, and positive-definite matrices $\Sigma_{i,1}$, $\Sigma_{i,2}$ such that as $k^* \to \infty$ and $T - k^* \to \infty$,

$$|\frac{1}{k^*}X_{i,t-1}^{'}X_{i,t-1} - \Sigma_{i,1}| = o(r(k^*)),$$

$$|\frac{1}{T-k^*}Z'_{i,t-1}Z_{i,t-1}-\Sigma_{i,2}|=o(r(T-k^*)),\tag{6.1}$$

$$|(\sum_{t=1}^{k^*}q_{i,t-1})'(Z'_{i,t-1}Z_{i,t-1})^{-1}(\sum_{t=1}^{k^*}q_{i,t-1})-\sigma_{i,1}^2|=o(r(T-k^*)),\tag{6.2}$$

$$|(\sum_{t=k^*+1}^{T}q_{i,t-1})'(X'_{i,t-1}X_{i,t-1})^{-1}(\sum_{t=k^*+1}^{T}q_{i,t-1})-\sigma_{i,2}^2|=o(r(k^*)),\tag{6.3}$$

where $r(x)=1/(\log x)^{\gamma}$, and $|\cdot|$ is the ordinary norm: $|(a_{ij})|=(\sum_i\sum_j a_{ij}^2)^{1/2}$.

C.3. There is some $\delta>0$ such that $\max_{1\le k^*\le T}|q_{i,t-1}|=o(T^{1/(2+\delta)})$, and $E|\epsilon_{i,t}|^{2+\delta}<\infty$.

Assumption C.2 is slightly weaker than C.9 in Csörgő and Horváth (1997, page 204) that assumes $\Sigma_{i,1}=\Sigma_{i,2}$. $(1/k^*)X'_{i,t-1}X_{i,t-1}$ and $(1/T-k^*)Z'_{i,t-1}Z_{i,t-1}$ may have different limits if existing. In the commonly adapted regression model that $(y_{i,t},x_{i,t})$'s are an independent and identically distributed sample with $E|(y_{i,t},x_{i,t})|^{2+\delta}<\infty$ for some, it is easily seen that C.2 and C.3 hold in probability one. The first lemma gives an order estimate for $\max(\tilde{\epsilon}_{i,t}(k^*))$. Denote $\bar{\epsilon}_i(k^*)=(1/T)\sum_{t=1}^{T}\tilde{\epsilon}_{i,t}(k^*)$,

$$\tilde{\epsilon}_{i,t}=\begin{cases}\epsilon_{i,t}-q'_{i,t-1}(Z'_{i,t-1}Z_{i,t-1})^{-1}Z'_{i,t-1}\gamma_{i,2}I, & 1\le t\le k^*,\\ \epsilon_{i,t}-q'_{i,t-1}(X'_{i,t-1}X_{i,t-1})^{-1}X'_{i,t-1}\gamma_{i,1}I, & k^*+1\le t\le T,\end{cases}$$

and $s_i^2(k^*)=(1/T)\sum_{t=1}^{T}\tilde{\epsilon}_{i,t}^2(k^*)$.

Lemma 1. Assume that H_0 and C.1–C.3 hold. Then

$$\max_{1\le t\le T}\{\max_{k_0\le k^*\le T-k_0}|\tilde{\epsilon}_{i,t}(k^*)|\}=O_P(T^{1/(2+\delta)}).$$

Lemma 2. Assume that H_0 and C.1–C.3 hold. Then

(a) $\max_{k_0\le k^*\le T-k_0}|\bar{\epsilon}_i(k^*)|=O_P(T^{-1/2}\log\log^{1/2}T)$.

(b) $\max_{k_0\le k^*\le T-k_0}s_i^2(k^*)=O_P(1)$ and in probability,

$$\liminf_{T\to\infty}\max_{k_0\le k^*\le T-k_0}s_i^2(k^*)\ge\sigma_i^2>0.$$

Furthermore, if $k_T\vec{\ }\infty$ as $T\to\infty$, we have $\max_{k_T\le k^*\le T-k_T}|s_i^2(k^*)-\sigma_i^2|=O_P(1)$.

Lemma 3. Assume that H_0 and C.1–C.3 hold. Then for some $\tau>0$,

$$\max_{k_0\le k^*\le T-k_0}|\hat{\lambda}_i-\bar{\epsilon}_i(k^*)/s_i^2(k^*)|=O_P(T^{-1/2-\tau}).$$

The proof process of Lemmas 1–3 is similar to Zhao et al. (2013).

Proof of Theorem 1. First, we use Lemmas 1–3 to obtain a quadratic approximation to $-2\mathfrak{R}(i,k^*)$, uniformly in k^*. Following Owen's (2001, page 221) arguments, denote $\zeta_{i,t}=\hat{\lambda}_i\tilde{\epsilon}_{i,t}(k^*)$. Using Taylor's expansion,

$$\begin{aligned}-2\mathfrak{R}(i,k^*)&=2\sum_{t=1}^{T}\log(1+\zeta_{i,t})\\&=2\sum_{t=1}^{T}\{\zeta_{i,t}-\frac{1}{2}\zeta_{i,t}^2+\frac{1}{3}\frac{\zeta_{i,t}^3}{(1+\zeta_{i,t})^3}\},\end{aligned}\tag{6.4}$$

where $|\xi_{i,t}| \le |\eta_{i,t}| = \hat{\lambda}\tilde{\epsilon}_{i,t}(k^*) = O_P(1)$, uniformly in k^*. By Lemmas 1 and 3, for some $\delta > 0$,

$$\begin{aligned}
&\max_{k_0 \le k^* \le T-k_0} \sum_{t=1}^{T} \{\zeta_{i,t} - \frac{1}{2}\zeta_{i,t}^2 + \frac{1}{3}\frac{\zeta_{i,t}^3}{(1+\zeta_{i,t})^3}\} \\
&\le T\{\max_{k_0 \le k^* \le T-k_0} [|\hat{\lambda}_i^3| s^2(k^*)]\} \max_{1 \le t \le T} |\tilde{\epsilon}_{i,t}(k^*)| \\
&= O_P\{T^{-\frac{3}{2}+1+\frac{1}{(2+\delta)}} \log\log^{2/3} T\} \\
&= O_P\{T^{-\delta/(4+2\delta)} \log\log^{2/3} T\}.
\end{aligned} \tag{6.5}$$

Next, by Lemma 3, for some $\tau > 0$,

$$\begin{aligned}
2\sum_{t=1}^{T} \zeta_{i,t} &= 2T\bar{\epsilon}_i^2(k^*)/s^2(k^*) + T\bar{\epsilon}_i(k^*)o(T^{-1/2-\tau}) \\
&= 2T\bar{\epsilon}_i^2(k^*)/s_i^2(k^*) + O_P(T^{-\tau}(\log\log T)^{1/2}),
\end{aligned} \tag{6.6}$$

and

$$\begin{aligned}
\sum_{t=1}^{T} \zeta_{i,t}^2 &= T\bar{\epsilon}_i^2(k^*)/s_i^2(k^*) + T\bar{\epsilon}_i(k^*)o(T^{-1/2-\tau}) \\
&= T\bar{\epsilon}_i^2(k^*)/s_i^2(k^*) + O_P(T^{-\tau}(\log\log T)^{1/2}).
\end{aligned} \tag{6.7}$$

Combining (6.4)–(6.7) yields that for any

$$0 < \tau_1 < \min\{\delta/(4+2\delta), \tau\},$$

$$\max_{k_0 \le k^* \le T-k_0} | - 2\mathfrak{R}(i,k^*) - T\frac{\bar{\epsilon}_i^2(k^*)}{s_i^2(k^*)}| = O_P(T^{-\tau_1}). \tag{6.8}$$

Now applying the Taylor expansion

$$(a+x)^{1/2} = a^{1/2} + x/(2a^{1/2}) + o(x/(a^{1/2})),$$

we have for any $0 < \tau_2 < \tau_1$,

$$\begin{aligned}
M'_{N,T} &= \{\max_{1 \le i \le N, d \le k^* \le T-d} [-\log \mathfrak{R}(i,k^*)]\}^{1/2} \\
&= \max_{1 \le i \le N, k_0 \le k^* \le T-k_0} T^{1/2}\{|\bar{\epsilon}_i(k^*)/s_i^2(k^*)|\} + O_P(T^{-\tau_2}).
\end{aligned} \tag{6.9}$$

Using the same arguments as the proof of Theorem 3.1.2 of Csörgő and Horváth (1997), we have

$$\lim_{N,T\to\infty} Pr\left\{A(\log u(T))(M'_{N,T})^{\frac{1}{2}} \le x + D_r(\log u(T))\right\} = \exp(-e^{-x}),$$

because $A(\log u(T))O_P(T^{-\tau_2}) = o(1)$, and it follows from (6.9) that

$$A(\log u(T))(M'_{N,T}) - D_r(\log u(T)) = A(\log u(T))\{\max_{1 \le i \le N, k_0 \le k^* \le T-k_0} T^{1/2}\{|\bar{\epsilon}_i(k^*)/s_i^2(k^*)|\}\} + O_P(1).$$

Proof of Theorem 2. Under H_1, we can obtain

$$\frac{1}{2T}Z_{i,k^*} = \frac{1}{T}\sup\{\sum_{t=1}^{T}\log(1+\hat{\lambda}_i\tilde{\epsilon_{i,t}}(k^*))\}$$

$$\xrightarrow{a.s}\sup\begin{cases}\frac{1}{T}\sup\{\sum_{t=1}^{k^*}\log(1+\hat{\lambda}_i(Y_{i,t}-[\hat{\beta}_0 1+G_iY_{i,t-1}])\}, & 1\le t\le k^*,\\ \frac{1}{T}\sup\{\sum_{t=k^*+1}^{T}\log(1+\hat{\lambda}_i(Y_{i,t}-[\hat{\beta}_0^* 1+H^iY_{i,t-1}])\}, & k^*+1\le t\le T,\end{cases}$$

$$\xrightarrow{a.s}\sup\begin{cases}\sup cE\{\log(1+\hat{\lambda}_i(Y_{i,t}-[\hat{\beta}_0 1+G_iY_{i,t-1}]))\}, & 1\le t\le k^*,\\ \sup cE\{\log(1+\hat{\lambda}_i(Y_{i,t}-[\hat{\beta}_0^* 1+H_iY_{i,t-1}]))\}, & k^*+1\le t\le T,\end{cases}$$

where $G_i=\sum_{k=1}^{p}\beta_k w_i^k+\beta_{p+1}1$ and $H_i=\sum_{k=1}^{p}\beta_k^* w_i^k+\beta_{p+1}^*1$.

By Jensen's inequality,

$$E\{\log(1+\hat{\lambda}_i(Y_{i,t}-[\hat{\beta}_0 1+G_iY_{i,t-1}]))\}\le\log E(1+\hat{\lambda}_i(Y_{i,t}-[\hat{\beta}_0 1+G_iY_{i,t-1}]))=0,$$

$$\frac{1}{2T}Z_{i,k^*}\xrightarrow{a.s}\sup(1-c)E\{\log(1+\hat{\lambda}_i(Y_{i,t}-[\hat{\beta}_0^* 1+H_iY_{i,t-1}]))\}\le(1-c)c_0.$$

Hence, $Z_{i,k^*}\to\infty$. The proof is complete.