Journal of the Korean Statistical Society 44 (2015) 457–468

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

A piecewise polynomial trend against long range dependence

Changryong Baek *

Department of Statistics, Sungkyunkwan University, 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul 110-745, Republic of Korea

ARTICLE INFO

Article history:
Received 30 October 2013
Accepted 12 February 2015
Available online 8 March 2015

AMS 2000 subject classifications:
primary 62M10
secondary 62G10

Keywords:
Piecewise polynomial trend
Long-range dependence
Local Whittle estimator

ABSTRACT

A sequential testing procedure to distinguish between a piecewise polynomial trend superimposed by short-range dependence and long range dependence is examined. The proposed procedure is based on the local Whittle estimation of long range dependence parameter from the residual series obtained by removing a piecewise polynomial trend. All results are provided with theoretical justifications, and Monte Carlo simulations show that our method achieves good size and provides reasonable power against long range dependence. The proposed method is illustrated to the historical Northern Hemisphere temperature data.

1. Introduction

Many geophysical, environmental, financial and telecommunications time series are known to exhibit long nonperiodic cycles and very strong correlations even for higher lags. In one direction, such characteristics can be captured by considering the so-called long range dependence (LRD) or long memory time series. It is formally defined as the second order stationary time series $X = \{X_n\}_{n\in\mathbb{Z}}$ with slowly decaying autocovariance

$$\gamma_X(h) = \text{Cov}(X_0, X_h) \sim Ch^{2d-1}, \quad C > 0,\ d \in (0, 1/2), \tag{1.1}$$

as $h \to \infty$. The parameter d in (1.1) is called the LRD parameter. See, for example, Beran (1994) and Doukhan, Oppenheim, and Taqqu (2003). Note that autocovariances for LRD series are so slowly decaying that autocovariances are not absolutely summable. If the sum of autocovariances is finite

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| < \infty,$$

then it is referred to as short-range dependence (SRD). Also, it is known that SRD series corresponds to $d = 0$ in the light of (1.1). The popular ARMA(p, q) models are well known examples of SRD series.

It is also very intuitive and natural to model such nonperiodic cycles with simple nonstationary model. For example, one may consider deterministic trend model superimposed by weakly dependent errors, namely,

$$X_t = g\left(\frac{t}{n}\right) + \epsilon_t, \quad t = 1, \dots, n,$$

* Tel.: +82 2 760 0602; fax: +82 502 302 0522.
E-mail address: crbaek@skku.edu.

http://dx.doi.org/10.1016/j.jkss.2015.02.001

where $g(\cdot)$ represents deterministic trend and $\{\epsilon_t\}_{t\in\mathbb{Z}}$ are SRD errors. A historical debate pointed out as early as Klemeš (1974) is that aforementioned two models are quite hard to distinguish from finite samples. This phenomenon is widely observed and well documented in hydrology, teletraffic, economics, finance as well as climatology. See, for example, Boes and Salas (1978), Diebold and Inoue (2001), Granger and Hyung (2004), Mills (2007) and Roughan and Veitch (1999) to list only a few of them.

In practice, however, distinguishing two models is extremely important because implications of each model are dramatically different. For example, consider time series generated by trend plus noise model. Then, it may exhibit LRD-like parabolically decaying sample ACF plots, so one may be tempted to have an impression of *stationary* time series. In fact, however, it is a *nonstationary* model, thus forecasting future values never converges to entire sample average. Therefore, without specifying underlying physical model, two approaches confuse many practitioners and lead to radically opposite scientific conclusions depending on which model to use. One such controversial debating can be found in the Wegman report (Wegman, Said, & Scott, 2006) on the existence of global warming.

This paper considers a statistical testing procedure to distinguish nonstationary polynomial trending shifts against stationary long range dependence. In particular, the focus is on the piecewise trending shifts with unknown change-points rather than smooth polynomial trend for reasons below. First, many underlying physical dynamics in fact imply abrupt changes. For example, water levels in the lake can be dramatically changed due to flooding. Second, a piecewise model is easy to interpret and sometimes gives parsimonious model than smooth changes, so is computationally more attractive. See Chaudhuri, Huang, Loh, and Yao (1994), Friedman (1991), Huh and Park (2004), Lin, Li, and Chen (2008), Sauve (2010) and references therein for the related literature.

When the piecewise polynomial trend is zero order, that is the constant mean changes, there has been some progress in this direction. For example, Berkes, Horváth, Kokoszka, and Shao (2006) suggested CUSUM-based methods to devise tests, and Bai and Perron (1998) can be similarly adapted based on the sup -F statistic calculated from the least squares method. However, it is noteworthy that aforementioned methods are not obvious and even may not be possible to extend to local polynomial trending shifts. For instance, as it is stated in Bai and Perron (1998) on p. 52, the validity of sup -F test for multiple breaks with polynomial trending shifts still remains open.

In this work, the testing procedure to distinguish between higher order piecewise polynomial shifts against LRD is proposed based on the local Whittle (LW, in short) LRD parameter estimation from the residual series after estimating a piecewise polynomial trend. It generalizes the idea in Baek and Pipiras (2012, 2014) to a higher order piecewise polynomial trend. Detailed description of the testing procedure and theoretical evidence is elaborated in Section 2. The method is examined through Monte Carlo simulations in Section 3, and applications to historical Northern Hemisphere temperature data are illustrated in Section 4. We conclude with Section 5.

2. Proposed testing procedure

In this section, we discuss the testing procedure for a piecewise trend model against LRD based on the LW estimation. We first consider when the number of shifts is known, then provide sequential testing procedure to jointly estimate the locations of shifts and the number of breaks.

2.1. Tests when the number of polynomial shifts is known

Consider the following piecewise polynomial model with R breaks (PT-R model, in short)

$$X_t = \sum_{r=0}^{R}\left\{\beta_{r0} + \beta_{r1}\left(\frac{t}{n}\right) + \cdots + \beta_{rp}\left(\frac{t}{n}\right)^p\right\} 1_{\{k_r<t\le k_{r+1}\}} + u_t \quad t = 1,\ldots,n, \tag{2.1}$$

where $\{k_r, i = 1,\ldots,R\}$ are the *unknown* locations of discontinuous points, i.e. break points, with the convention that $k_0 = 0$ and $k_{R+1} = n$. Assume that $\{u_t\}$ is SRD series. For the clarity of presentation, we will first assume that the number of breaks R is known. A binary segmentation method to estimate unknown number of breaks R will be discussed later on.

In a matrix form, it is written as

$$\mathbf{X} = \mathbf{Z}^R\boldsymbol{\beta} + \mathbf{U}$$

where $\mathbf{X} = (X_1,\ldots,X_n)'$, $\mathbf{U} = (u_1,\ldots,u_n)'$, $\mathbf{Z}^R = \mathrm{diag}(\mathbf{Z}_0,\ldots,\mathbf{Z}_R)$ is the diagonally partitioned block matrix, namely

$$\mathbf{Z}^R = \begin{pmatrix} \mathbf{Z}_0 & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{O} & \mathbf{Z}_1 & \cdots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{O} & \cdots & \cdots & \mathbf{Z}_R \end{pmatrix}, \qquad \mathbf{Z}_r = \begin{pmatrix} 1 & (k_r+1)/n & \cdots & (k_r+1)^p/n^p \\ 1 & (k_r+2)/n & \cdots & (k_r+2)^p/n^p \\ \vdots & \vdots & \vdots & \vdots \\ 1 & k_{r+1}/n & \cdots & k_{r+1}^p/n^p \end{pmatrix}_{(k_{r+1}-k_r)\times(p+1)},$$

and coefficient vector $\boldsymbol{\beta} = (\boldsymbol{\beta}_0,\ldots,\boldsymbol{\beta}_R)'$ with $\boldsymbol{\beta}_r = \left(\beta_{r0},\ldots,\beta_{rp}\right)'$, $r = 0,\ldots,R$.