Journal of the Korean Statistical Society 38 (2009) 131–143

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Kink estimation with correlated noise

Justin Wishart *

School of Mathematics and Statistics, F07, University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 10 June 2008
Accepted 6 August 2008
Available online 5 September 2008

AMS 2000 subject classifications:
primary 62G05
secondary 62G20

Keywords:
Change point
Kink
High-order kernel
Zero-crossing technique
Fractional Gaussian noise
Long-range dependence
Separation rate lemma
Inverse problems
Fractional integration
Climate change

ABSTRACT

In this article we study the estimation of the location of jump points in the first derivative (referred to as kinks) of a regression function f in the presence of noise that exhibits long-range dependence (LRD). The method is based on the zero-crossing technique and makes use of high-order kernels. The effect of LRD is seen to be detrimental to the rate of convergence. Using a fractional integration operator we draw a parallel with certain inverse problems which suggests optimality of our approach. The kink location and estimation technique is demonstrated on some simulated data and the detrimental effect of LRD is shown. We also apply our kink analysis on Australian temperature data.

1. Introduction

The objective of this paper is to detect and estimate the location of a jump in the first derivative (and possibly higher-order derivatives) of a regression function f in the presence of long-range dependent (LRD) noise. For our purposes we suppose that the unknown regression function $f(x)$, $x \in I = [0, 1]$ is observed from the fractional Gaussian noise model,

$$dY(x) = f(x)dx + \varepsilon^{\alpha} dB_H(x) \tag{1}$$

where $B_H(x)$ is a normalised fractional Brownian motion on I with self-similarity parameter $H \in [\frac{1}{2}, 1)$ (see, Mandelbrot and Van Ness (1968)). We assume that ε, the noise level, is small and is of the form $\varepsilon = \varepsilon_n = \sigma n^{-\frac{1}{2}}$ where $\sigma > 0$ is constant. Model (1) was first considered by Wang (1996) to study the effect of long-range dependence on non-parametric methods. For notational convenience throughout this paper $\alpha = 2 - 2H$ will be used as the measure of long-range dependence with $0 < \alpha \le 1$. This model includes the white noise model as a special case when $\alpha = 1$ and the driving process for the noise is a standard Brownian motion.

There has been a large focus in recent times on the estimation of a jump point in the function f for the white noise model and the more generalised LRD model. However, less attention has been paid to detecting and estimating the location of a jump in the first derivative of f (which will be referred to as a kink). To our knowledge, the kink scenario under the assumption of LRD has not been studied and there are many examples of real data that exhibit LRD in fields such as

* Tel.: +61 2 9351 5805.
E-mail address: justinw@maths.usyd.edu.au.

1226-3192/$ – see front matter
doi:10.1016/j.jkss.2008.08.001

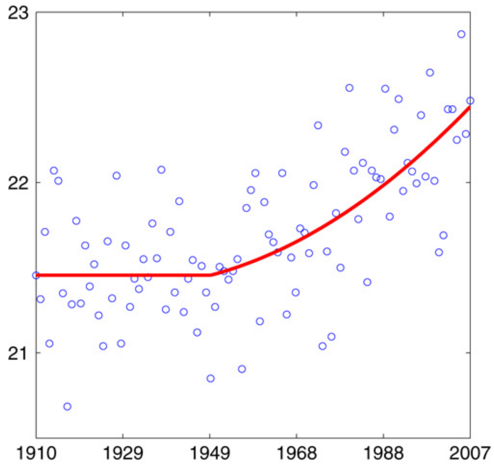

Fig. 1. Scatterplot of average daily air temperature across Australia with fitted curve after kink detection and estimation. Observations are taken annually from 1910–2007. The value of α was estimated to be approximately 0.7.

economics, environmental sciences, geosciences, hydrology, physics and other fields. Knowing the location of a kink is of particular importance. For example if the underlying trend in a model is represented by the regression function f, then a kink would correspond to a change in the trend of the model either from an upward trend to a downward trend or vice versa. In the case of regression smoothing, the knowledge of the location of a kink can be used to improve the visual and numerical performance traditional smoothing methods. It can also be used to locate or explain a qualitative change in behaviour of an observed variable.

A particular dataset we shall study in this paper is yearly average Australian temperature readings. This data is shown in Fig. 1 and after analysis it tests positive for a kink that is located around 1949. This suggests a possible climate change due to the change in trend in air temperature across Australia. Other specific datasets that exhibit kink behaviour are given in Cheng and Raimondo (2008).

1.1. Existing results

For estimating a kink location in f in the case of a white noise model, work has been done by Huh and Carrière (2002) and more recently by Cheng and Raimondo (2008). The kink scenario for the white noise model also appears as a special case in the general theory of Goldenshluger, Tsybakov, and Zeevi (2006) where they determined the optimal rate to be $O(n^{-s/(2s+1)})$, where s is the number of derivatives that exist for the function f away from the kink. In this paper we adapt the zero-crossing technique that was pioneered by Goldenshluger et al. (2006) and followed by Cheng and Raimondo (2008) and we apply it to the LRD setting. In the implementation of the method we use high-order kernels that were introduced by Cheng and Raimondo (2008).

There is a large body of literature for estimation of jump points in f, to name a few; Gijbels and Goderniaux (2004a,b), Gijbels, Hall, and Kneip (1999), Huh and Park (2004), Korostelëv (1987), Korostelëv and Tsybakov (1993), Müller (1992), Raimondo (1998), Raimondo and Tajvidi (2004) and Wang (1995, 1999).

There has been some work on change point estimation specifically for the case of weakly dependent (short-range dependent) errors by Antoch, Hušková, and Prášková (1997) and Tang and MacNeill (1993).

The effect of LRD data on the subject of estimation and detection has been studied by Beran (1992, 1994) and Csörgö and Mielniczuk (1995). The effect of LRD on non-parametric regression has been studied by Cavalier (2004), Johnstone (1999), Johnstone and Silverman (1997) and Wang (1996, 1997). Work on LRD and its effect on change point estimation include Wang (1999). In all of the above cases the inclusion of LRD has a detrimental effect on the convergence rate and estimators converge at a slower rate than in the independent case.

1.2. Result

We show that our method of finding the kink location(s) θ in the case of LRD errors will lead to an accuracy of order,

$$r(n, s, \alpha) \asymp n^{-\alpha s/(2s+\alpha)}. \tag{2}$$

This elucidates the effect of LRD on the convergence rate of kink estimation. The rate deteriorates as α decreases and this is consistent with the literature. This is confirmed by our simulation study and is shown in Table 1 in Section 4.5. Note that for $\alpha = 1$ our rate agrees with the literature in the independent noise scenario, see Cheng and Raimondo (2008).