Journal of Statistical Planning and Inference 140 (2010) 3820–3840

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Efficiency in estimation of memory

Willa W. Chen

Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA

ARTICLE INFO

Available online 5 May 2010

JEL classification:
C13
C22

Keywords:
Gaussian semiparametric estimation
Tapering
Periodogram

ABSTRACT

We study the efficiency of semiparametric estimates of memory parameter. We propose a class of shift invariant tapers of order (p,q). For a fixed p, the variance inflation factor of the new tapers approaches 1 as q goes to infinity. We show that for $d \in (-1/2, p+1/2)$, the proposed tapered Gaussian semiparametric estimator has the same limiting distribution as the nontapered version for $d \in (-1/2, 1/2)$. The new estimator is mean and polynomial trend invariant, and is computationally advantageous in comparison to the recently proposed exact local Whittle estimator. The simulation study shows that our estimator has comparable or better mean squared error in finite samples for a variety of models.

1. Introduction

The immense interest in long memory modelling has produced numerous important results in memory parameter estimation. Though a full parametric model such as autoregressive fractionally integrated moving average (ARFIMA) model with n observations yields $\sqrt{n}$ consistency for the memory parameter estimates, model misspecification may produce slower convergence rates, nonstandard limiting properties (see Yajima, 1993; Chen and Deo, 2005) or even inconsistent estimates. Semiparametric modelling provides a more robust alternative in estimating the memory parameter by making use of the regularity property of a long memory process at low frequencies of the spectral density without specifying an explicit short memory structure. The spectral density of a covariance stationary long memory process is of the form

$$f(\omega) \sim C|\omega|^{-2d} \quad \text{as } \omega \longrightarrow 0^+, \tag{1}$$

where $d < 0.5$ and C is a positive constant.

Several semiparametric estimations were developed over the years. Among them, the log-periodogram regression estimator (GPH, Geweke and Porter-Hudak, 1983; Robinson, 1995a) and Gaussian semiparametric estimator (GSE, Künsch, 1987; Robinson, 1995b) have drawn the most interest and generated a vast literature empirically and theoretically. The GPH has an explicit expression because it is the slope of a simple linear regression, while the GSE is a local Whittle likelihood estimator which is more efficient. Earlier work on these two estimators, like that of others, assumes that the memory parameter d lines in the interval $(-0.5, 0.5)$, i.e., the series is stationary and invertible (a long memory process is invertible if $d \in (-1, 1/2)$, see Bondon and Palma, 2007; Palma, 2007). To extend the range of application of long memory modelling, the technique of tapering has been suggested, for example, Velasco (1999a) on GPH, Velasco (1999b) and Hurvich and Chen (2000) on GSE. These periodogram-based methods are easy to implement and can be evaluated efficiently via the fast Fourier transform (FFT) algorithm. Moreover, they have the mean and polynomial trend invariant property, which is essential in long memory modelling because of the potential slow convergence rates of estimators of mean μ. However, the drawback of these methods is that tapering inflates the variance of estimates of d by a factor $\Phi > 1$.

E-mail address: wchen@stat.tamu.edu

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2010.04.046

This factor Φ, depending on the data taper, is the same as that of the nonparametric estimates for spectral density (Brillinger, 1981, Section 5.6).

Recently, Shimotsu and Phillips (2005) has proposed the exact local Whittle (ELW) estimator for any fixed value of d and showed that it has the same limiting distribution as the nontapered GSE for the stationary and invertible case, i.e., N(0,1/4), thus is more efficient than the tapered GSE. However, this superiority in efficiency comes with a price. The objective function of the ELW estimator is based on the periodogram of fractionally differenced data with a length-n linear filter depending on the parameter to be estimated, d. Evaluating this objective function bears a much higher computational cost even with an FFT embedded algorithm. Moreover, the ELW estimator is not shift invariant due to the filtering. In Shimotsu and Phillips (2005), the mean is assumed to be known. To relax this assumption, Shimotsu (2010) proposed the two-step feasible ELW estimator for $d \in (-0.5,1.75)$ with the objective function including an estimated mean that depends on d. The first step of this estimator is to obtain an initial value for the optimization procedure using the tapered GSE for d by Velasco (1999b). Clearly, the finite sample properties of the two-step feasible ELW estimator will depend upon the tapered initial estimator.

Abadir et al. (2007) have also proposed a version of GSE called the fully extended local Whittle (FELW) estimator for $d \in (-1.5,\infty)$. The objective function of FELW is based on the classical periodogram with corrections that depend on the region of d. Though the optimization can be computed efficiently, as pointed out by Shimotsu (2010), the limiting distribution is valid on an interval with holes at $-0.5,0.5,1.5,\ldots$ due to discontinuities of the objective function at these points. Moreover, the estimator does not accommodate processes with deterministic components such as a linear trend.

As high frequency data become available, computational efficiency becomes imperative in practice. It has been found that many statistical procedures are often not feasible when applied to a sizeable time series due to the computational difficulties, see for example, Deo et al. (2006). The data tapering approach, which upholds the typical advantages of a frequency domain method, is clearly more favorable in terms of implementation. The question remains whether the tapered estimates can achieve the same efficiency as the nontapered one. This problem was studied in different settings by Dahlhaus (1988) who considered the Whittle estimates of an ARMA model with strongly peaky spectral density and showed that the tapered estimates have the same optimal efficiency as MLE if the taper function has a variance inflation factor $\Phi \to 1$ as $n \to \infty$. Two such functions are Tukey's split cosine bell and the polynomial taper. This suggests that the same taper functions can potentially yield fully efficient tapered GSE or GPH estimates. However, these two tapers do not have the shift invariance property which is necessary for obtaining mean and trend invariant estimates. Furthermore, the class of models considered in Dahlhaus (1988) assumes that the peaks and troughs of the spectral density are known functions, thus excludes the long memory parameter estimation.

In this paper, we propose a new class of shift invariant tapers defined with an order of (p,q), where p measures the strength of tapers. the new tapered DFT is a linear combination of $p+q$ nontapered DFTs with coefficients obtained by minimizing the sum of squared autocorrelations of tapered DFTs of a white noise process at $p+q$ consecutive Fourier frequencies. We adapt the differencing-followed-by-tapering routine of Hurvich and Chen (2000) by applying a (p,q)th order taper on the pth differences of a series with $d \in (-0.5,p+0.5)$ to obtain pth order polynomial trend invariant and bias reduced periodogram. With the new tapers, the tapered GSE have a variance inflation factor $\Phi_{p,q}$ that is a strictly monotone decreasing function of q. These tapered estimates are more efficient than the tapered GSE implemented by Hurvich and Chen (2000) because their tapers are of order $(p,1)$ under the new definition. By letting q increase with n at a certain rate, we show that the new tapered GSE has a limiting distribution of N(0,1/4) in part because $\Phi_{p,q} \to 1$ as $q \to \infty$ for any fixed p. Thus the new tapered GSE achieves the same efficiency as the nontapered version for $d \in (-0.5,0.5)$ without compromising the computational efficiency and also enjoys the shift/trend invariance property.

In the next section, we begin with a brief discussion on tapering followed by the formulation of the new tapers. Section 4 presents the asymptotic results for the new tapered GSE. Appendix A examines finite-sample performance of the new estimates and compares to that of two-step feasible ELW via simulation. Proof details are left to Appendices B and C.

2. The maximal efficient tapers

The idea of tapering in improving a Fourier approximation has a long history dating back to 1900 (see Brillinger, 1981, Chapter 3). Tukey (1967) introduced this technique to time series for reducing periodogram bias due to strong peaks and troughs in spectral density. A data taper is a sequence of constants $h_t=h(t/n)$, $h(x)=0$ for $x \notin (0,1]$. Given a series of n observations, x_t, $t=1,\ldots,n$, the tapered DFT and periodogram are defined as

$$J_x(\omega) = \left(2\pi \sum |h_t|^2\right)^{-1/2} \sum_{t=1}^{n} h_t x_t e^{i\omega t}, \quad I_x(\omega) = |J_x(\omega)|^2. \tag{2}$$

The expectation of periodogram is

$$\mathbb{E}(I_x(\omega)) = \int_{-\pi}^{\pi} K_n(\lambda-\omega) f(\lambda)\, d\lambda, \quad K_n(\omega) = \left(2\pi \sum |h_t|^2\right)^{-1} \left| \sum_{t=1}^{n} h_t e^{i\omega t} \right|^2. \tag{3}$$

The case of no tapering corresponds to the boxcar window $h_t \equiv 1$, $t=1,\ldots,n$ and $K_n(\cdot)$ being the Fejer kernel. A suitable chosen taper function other than the boxcar gives K_n smaller sidelobes so that the spectral density away from the