Computational Statistics and Data Analysis 100 (2016) 676–693

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Adaptive bandwidth selection in the long run covariance estimator of functional time series[☆]

Lajos Horváth, Gregory Rice[*], Stephen Whipple

Department of Mathematics, University of Utah, Salt Lake City, UT 84112-0090, USA

ARTICLE INFO

Article history:
Received 25 March 2014
Received in revised form 3 June 2014
Accepted 5 June 2014
Available online 12 June 2014

Keywords:
Functional data
Long run covariance
Mean squared error
Optimal bandwidth

ABSTRACT

In the analysis of functional time series an object which has seen increased use is the long run covariance function. It arises in several situations, including inference and dimension reduction techniques for high dimensional data, and new applications are being developed routinely. Given its relationship to the spectral density of finite dimensional time series, the long run covariance is naturally estimated using a kernel based estimator. Infinite order "flat-top" kernels remain a popular choice for such estimators due to their well documented bias reduction properties, however it has been shown that the choice of the bandwidth or smoothing parameter can greatly affect finite sample performance. An adaptive bandwidth selection procedure for flat-top kernel estimators of the long run covariance of functional time series is proposed. This method is extensively investigated using a simulation study which both gives an assessment of the accuracy of kernel based estimators for the long run covariance function and provides a guide to practitioners on bandwidth selection in the context of functional data.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

A common way of obtaining functional data is to break long, continuous records into a sample of shorter segments which may be used to construct curves. For example, tick data measuring the price of an asset obtained over several years, which in principle may contain millions of data points, may be used to construct smaller samples of daily or weekly curves. Over the last decade functional time series analysis has grown steadily due to the prevalence of these types of data; we refer to Hörmann and Kokoszka (2012) and Horváth and Kokoszka (2012) for a review of the subject.

Suppose

$$X_i(t),\ 1 \le i \le n \text{ and } t \in [0, 1] \text{ are observations from a stationary ergodic functional time series with } E\|X_0\|^2 < \infty, \tag{1.1}$$

where $\|\cdot\|$ denotes the standard norm in L^2. An object which arises frequently in this context is the long run covariance function

$$C(t, s) = \sum_{i=-\infty}^{\infty} \mathrm{Cov}(X_0(t), X_i(s)), \quad 0 \le t, s \le 1. \tag{1.2}$$

[☆] Research supported by NSF grant DMS 1305858.

[*] Corresponding author.
E-mail address: rice@math.utah.edu (G. Rice).

http://dx.doi.org/10.1016/j.csda.2014.06.008
0167-9473/© 2014 Elsevier B.V. All rights reserved.

C may be viewed as an extension of the spectral density function evaluated at zero for univariate and multivariate time series, and its usefulness in the analysis of functional time series is similarly motivated. For example, under some regularity conditions $\sqrt{n}\bar{X}(t)$ is asymptotically Gaussian with covariance function C, where

$$\bar{X}(t) = \frac{1}{n}\sum_{i=1}^{n} X_i(t),$$

and hence the distribution of functionals of $\bar{X}$ can be approximated using an approximation of C, see Jirák (2013) and Horváth et al. (2014). Also, the principle components computed as the eigenfunctions of the Hilbert–Schmidt integral operator

$$c(f)(t) = \int_0^1 C(t,s)f(s)ds$$

may be used to give asymptotically optimal finite dimensional representations of dependent functional data, see Ferraty and Vieu (2006). Given its representation as an infinite sum, C is naturally estimated with a kernel estimator of the form

$$\hat{C}_{n,h}(t,s) = \sum_{i=-(n-1)}^{n-1} K\left(\frac{i}{h}\right)\hat{\gamma}_i(t,s), \tag{1.3}$$

where

$$\hat{\gamma}_i(t,s) = \begin{cases} \frac{1}{n}\sum_{j=1}^{n-i}\left(X_j(t)-\bar{X}(t)\right)\left(X_{j+i}(s)-\bar{X}(s)\right), & i \geq 0 \\ \frac{1}{n}\sum_{j=1-i}^{n}\left(X_j(t)-\bar{X}(t)\right)\left(X_{j+i}(s)-\bar{X}(s)\right), & i < 0. \end{cases}$$

We use the standard convention that $\hat{\gamma}_i(t,s) = 0$ when $i \geq n$. It was shown in Horváth et al. (2012) that if

$$K(0) = 1, \quad K(u) = K(-u), \quad K(u) = 0 \quad \text{if } |u| > c \text{ for some } c > 0,\ K \text{ is continuous on } [-c,c], \tag{1.4}$$

and

$$h = h(n) \to \infty, \quad h = o(n), \quad \text{as } n \to \infty, \tag{1.5}$$

then

$$\|\hat{C}_{n,h} - C\| = o_P(1), \tag{1.6}$$

as long as $\{X_i(t)\}_{i=-\infty}^{\infty}$ is a weakly dependent Bernoulli shift (cf. (2.5)–(2.7)).

Although the L^2 consistency of $\hat{C}_{n,h}$ holds under these standard conditions on the kernel K and the bandwidth parameter h, their choice can greatly affect the estimators performance in finite samples. Classically, finite order kernels such as the Bartlett and Parzen kernels were used, see Rosenblatt (1991). More recently though infinite order "flat-top" kernels of the form

$$K_f(t;x) = \begin{cases} 1, & 0 \leq |t| < x \\ (x-1)^{-1}(|t|-1), & x \leq |t| < 1 \\ 0, & |t| \geq 1, \end{cases} \tag{1.7}$$

which are equal to one in a neighborhood of the origin and then decay linearly to zero, were advocated for by Politis and Romano (1996) and Politis and Romano (1999) where it is shown that they give reduced bias and faster rates of convergence when compared to kernels of finite order.

An important consideration though, regardless of the kernel choice, is the selection of the bandwidth parameter h. At present there is no available guidance regarding the choice of the bandwidth parameter for kernel based estimation with functional data.

One popular technique for such problems is cross validation, which has been used with success in scalar spectral density estimation (cf. Beltrao and Bloomfield, 1987). Such methods are difficult to extend to the functional setting however since the already time consuming calculations involved in applying cross validation with scalar data become incalculable with densely observed curves. A separate approach which is more amenable with functional data is the use of plug-in or adaptive bandwidths which aim to minimize the mean squared error using an estimated bandwidth. Among the contributions in this direction are Andrews (1991), Andrews and Monahan (1992), and Bühlmann (1996) who showed that the asymptotically optimal bandwidth for spectral density estimation with scalar ARMA(p, q) data using finite order kernels is of the form $c_d n^{1/r}$, where c_d increases with the strength of dependence of the sequence. Their results are established by comparing the estimators asymptotic bias, which can be computed with standard arguments, to the asymptotic variance for which formulae have been derived in the scalar case, see Priestly (1981). This theory and subsequent simulation studies all indicate that

Download English Version:

https://daneshyari.com/en/article/6869206

Download Persian Version:

https://daneshyari.com/article/6869206

Daneshyari.com