Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Statistical Planning and Inference 136 (2006) 3415–3434
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Bayesian wavelet analysis of autoregressive fractionally integrated moving-average processes

Kyungduk Ko[a], Marina Vannucci[b,*]

[a]Department of Mathematics, Boise State University, Boise, ID 83725-1555, USA
[b]Department of Statistics, Texas A&M University, College Station, TX, 77843-3143, USA

Received 30 October 2003; accepted 27 January 2005
Available online 22 March 2005

Abstract

Long memory processes are widely used in many scientific fields, such as economics, physics and engineering. In this paper we describe a wavelet-based Bayesian estimation procedure to estimate the parameters of a general Gaussian ARFIMA (p, d, q), autoregressive fractionally integrated moving average model with unknown autoregressive and moving average parameters. We employ the decorrelation properties of the wavelet transforms to write a relatively simple Bayes model in the wavelet domain. We use an efficient recursive algorithm to compute the variances of the wavelet coefficients. These depend on the unknown characteristic parameters of the model. We use Markov chain Monte Carlo methods and direct numerical integration for inference. Performances are evaluated on simulated data and on real data sets.

MSC: 62C10; 62F15; 42C40; 65T60; 60G18

Keywords: ARFIMA processes; Bayesian inference; Wavelets

1. Introduction

Long memory processes are widely used in many fields and applications range from financial data to data from biology and hydrology, to mention a few. Fractional ARIMA (p, d, q), first introduced by Hosking (1981) and Granger and Joyeux (1980), are well

* Corresponding author. Tel.: +1 979 845 0805; fax: +1 979 845 3144.
E-mail address: mvannucci@stat.tamu.edu (M. Vannucci).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2005.01.010

known examples of long memory processes. Classical methods for modelling and inference for these processes involve the calculation of the exact likelihood and its maximization with respect to the parameters, see for example Beran (1994). The complexity of such inferential procedures is mainly due to the dense long memory covariance structure that makes the exact likelihood of the data difficult to handle. Exact maximum likelihood estimators, as well as the calculations for posterior distributions in suitable form for inference, are therefore usually impractical for large data sets. Some improvements were achieved by Pai and Ravishanker (1996, 1998) and by Koop et al. (1997), who proposed Bayesian approaches based on the algorithm of Sowell (1992a) to compute the exact likelihood and used importance sampling and Markov chain Monte Carlo methods for a posteriori inference.

Wavelets have proven to be a powerful tool for the analysis and synthesis of data from long memory processes. Wavelets are strongly connected to such processes in that the same shapes repeat at different orders of magnitude, Wornell (1996). The ability of the wavelets to simultaneously localise a process in time and scale domain results in representing many dense matrices in a sparse form. When transforming measurements from a long memory process, wavelet coefficients are approximately uncorrelated, in contrast with the dense long memory covariance structure of the data, see Tewfik and Kim (1992), among others.

Here we propose a Bayesian approach to the wavelet analysis of fractional ARIMA (p, d, q) processes. We first transform the data into wavelet coefficients and use an efficient recursive algorithm from Vannucci and Corradi (1999a) to compute the exact variances and covariances of the wavelet coefficients. We exploit the de-correlation properties of the wavelets to write a simple model in the wavelet domain. The exact variances of the wavelet coefficients are embedded in the model and depend on the parameters of the ARFIMA process. We carry out posterior inference by Markov chain Monte Carlo methods and, in the simpler case of integrated processes, i.e. ARFIMA$(0, d, 0)$, by direct numerical integration. We perform extensive simulation studies and also test our method on benchmark data sets. In all examples we provide comparisons with other methods.

Other work that uses wavelets in the analysis of discrete-time long memory processes can be found in McCoy and Walden (1996), who proposed an approximate wavelet coefficients-based maximum likelihood iterative estimation procedure for ARFIMA$(0, d, 0)$ only. Jensen (1999, 2000) constructed wavelet-based MLE estimators of the long memory parameter for ARFIMA(p, d, q). His simulations studies showed better overall performances of the wavelet-based estimates with respect to the approximate MLE. Our approach is novel in many respects. We combine Bayesian methods with wavelet-based modelling of long memory processes, for both ARFIMA$(0, d, 0)$ and ARFIMA(p, d, q) models. For ARFIMA(p, d, q), unlike previous approaches, we also produce estimates of the unknown autoregressive and moving average parameters. Our approach is fairly general and may be applicable to other classes of long memory processes.

The paper is organized as follows: In Section 2, we introduce the necessary mathematical concepts on ARFIMA models and on wavelet methods. In Section 3, we describe the Bayesian model and the posterior inference. We report results from simulations in Section 4 and from applications to GNP data and to the well known Nile river data set in Section 5. Some concluding remarks are given in Section 6.