Economics Letters 102 (2009) 128–131

Contents lists available at ScienceDirect

Economics Letters

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

An algorithm for robust fitting of autoregressive models

Dimitris N. Politis *

Department of Mathematics, and Economics, University of California at San Diego, La Jolla, CA 92093-0112, USA

ARTICLE INFO

Article history:
Received 20 August 2005
Received in revised form 30 November 2008
Accepted 3 December 2008
Available online 13 December 2008

Keywords:
ARMA models
Linear time series
Outliers

JEL classification:
C1
C5

ABSTRACT

An algorithm for robust fitting of AR models is given, based on a linear regression idea. The new method appears to outperform the Yule–Walker estimator in a setting of data contaminated with outliers.

1. Introduction

Consider data $X_1,\ldots,X_n$ arising as a stretch of a second-order stationary time series $\{X_t, t\in\mathbb{Z}\}$ with autocovariance sequence $\gamma_k=Cov(X_t, X_{t+k})$. We will assume that $EX_t=0$ which, from a practical point of view, means that the data have been de-trended. An autoregressive model of order p, i.e., an AR(p), is defined by the following recursion:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t, \text{ for all } t\in\mathbb{Z}, \quad (1)$$

where Z_t is a second-order stationary white noise with $EZ_t=0$ and $EZ_t^2=\sigma^2$. We will assume the above AR(p) model is *causal*, i.e., that for any $k>0$, Z_{t+k} is uncorrelated to $\{X_{t-s}, s\geq 0\}$; see Brockwell and Davis (1991). Multiplying both sides of (1) by X_{t-k} and taking expectations, we derive:

$$\gamma_k = \phi_1\gamma_{k-1} + \ldots + \phi_p\gamma_{k-p} + \sigma^2\delta_k, \text{ for all } k\geq 0 \quad (2)$$

where $\delta_k=0$ for $k\neq 0$ but $\delta_0=1$.

Plugging the sample autocovariance $\hat{\gamma}_k = n^{-1}\sum_{i=1}^{n-|k|} X_t X_{t+|k|}$ in place of the true γ_k in Eq. (2) for $k=0,1,\ldots,p$, the well-known Yule–Walker (YW) equations are derived:

$$\hat{\gamma}_k = \phi_1\hat{\gamma}_{k-1} + \ldots + \phi_p\hat{\gamma}_{k-p} + \sigma^2\delta_k, \text{ for } k=0,\ldots,p. \quad (3)$$

whose unique solution $\hat{\phi}_1,\ldots,\hat{\phi}_p$ and $\hat{\sigma}^2$ forms the well-known YW estimator that is asymptotically efficient in the context of a Gaussian AR series. Nevertheless, the YW estimator loses its asymptotic efficiency under a non-Gaussian distributional assumption; see e.g. Sengupta and Kay (1989).

In what follows, we describe a simple estimation algorithm for AR model fitting; it is not a fast algorithm but it is promising in improving the finite-sample accuracy of the YW estimators when outliers are present. The new algorithm exemplifies robustness against outliers, and in particular against clusters of (two or more) outliers.

2. Motivation of the new algorithm

In what follows, the focus is on fitting an AR(p) model; for concreteness, the order p is assumed known. Let $\underline{\phi}_p = (\phi_1,\ldots,\phi_p)'$, $\underline{\gamma}_k = (\gamma_1,\ldots,\gamma_k)'$, and $\underline{\hat{\gamma}}_k = \left(\hat{\gamma}_1,\ldots,\hat{\gamma}_k\right)'$. Recall that the YW estimator of $\underline{\phi}_p$ and σ^2 is a (linear) function of $\hat{\gamma}_0, \hat{\gamma}_1,\ldots,\hat{\gamma}_p$; it is asymptotically efficient if the series is Gaussian AR(p) in which case $(\hat{\gamma}_0,\hat{\gamma}_1,\ldots,\hat{\gamma}_p)$ is (approximately) a sufficient statistic for $\underline{\phi}_p$ and σ^2. Therefore, in the Gaussian case, one is justified to just look at functions of $(\hat{\gamma}_0, \hat{\gamma}_1,\ldots,\hat{\gamma}_p)$.

Nevertheless, $(\hat{\gamma}_0, \hat{\gamma}_1,\ldots,\hat{\gamma}_p)$ is not necessarily sufficient without Gaussianity. It seems natural that in the context of second-order stationarity, a general estimator of $\underline{\phi}_p$ and σ^2 would be a function of all the second-order information available, i.e., $(\hat{\gamma}_0, \hat{\gamma}_1,\ldots,\hat{\gamma}_{n-1})$ Since $\hat{\gamma}_k$ is unreliable for large k, i.e., when $n-k$ is small, it makes sense to base our estimator on $\hat{\gamma}_0,\hat{\gamma}_1,\ldots,\hat{\gamma}_{p'}$ where p' is potentially large as compared to p, but small as compared to n; in particular, we require that $n-p'$ is large, i.e., that

$$p\leq p'\leq cn \text{ for some } c\in(0,1). \quad (4)$$

Asymptotically, any value of $c\in(0,1)$ guarantees that $n-cn\to\infty$ when $n\to\infty$ but for practical sample sizes of the order of 100 or 1000 a reasonable choice for c must be small, say in the interval [0.1, 0.2], for

* Tel.: +1 858 534 5861; fax: +1 858 534 5273.
E-mail address: dpolitis@ucsd.edu.

0165-1765/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2008.12.004

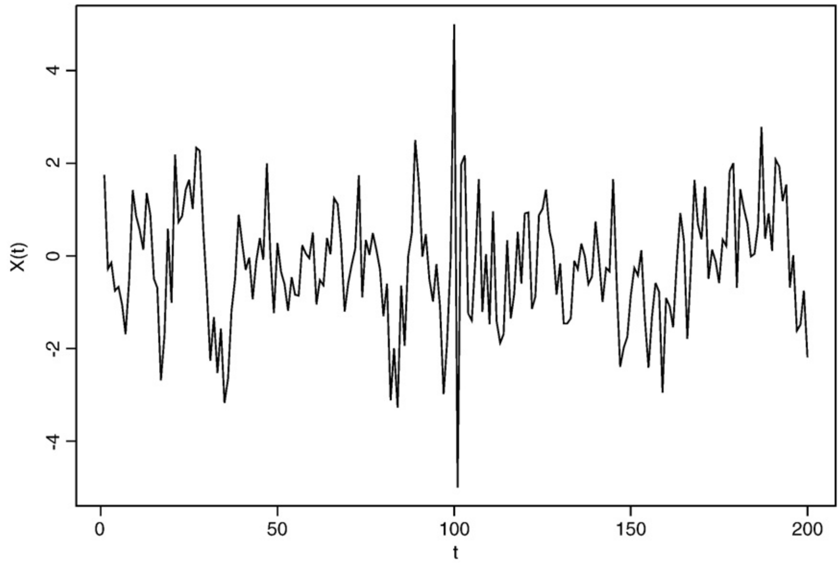

Fig. 1. Plot of Gaussian AR(1) series $X_1,\ldots, X_{200}$.

$n-cn$ to be large. An easy way to satisfy (4) is to let $p'=\max(p, \lfloor cn \rfloor)$ with c as above.

As Eq. (2) implies the YW Eq. (3) by plugging in $\hat{\gamma}_k$ for γ_k, the same Eq. (2) also implies the 'Extended' Yule–Walker (EYW) equations:

$$\hat{\gamma}_k=\phi_1\hat{\gamma}_{k-1}+\ldots+\phi_p\hat{\gamma}_{k-p}+\sigma^2\delta_k, \text{for } k=0,\ldots,p'. \tag{5}$$

3. The notion of robustness

It is desirable to have estimators that are robust to 'outliers', i.e., data points whose value is extreme compared to the bulk of the data. Outliers are generally due either to contaminated/corrupted data, or to heavy-tailed error distributions. For discussion on robustness see Franke et al. (1984), and Hampel et al. (1986).

To give an example, consider the Gaussian AR(1) model $X_t=\phi_1X_{t-1}+Z_t$ with Z_t i.i.d. N(0,1). Suppose $n=200$, and that the 100th observation has been corrupted by an outlier resulting into $X_{100}=B$. All $\hat{\gamma}_k$ are subsequently corrupted; instead of the usual $\hat{\gamma}_k=\gamma_k+O_P(1/\sqrt{n})$ we now have $\hat{\gamma}_k=\gamma_k+O_P(1/\sqrt{n})+O(B/n)$. The situation is further aggravated if a second outlier is found close to the first one, say at the 101st observation. So assume X_{101} is of the order of magnitude of B (positive or negative); consequently, estimation of the lag-0 and lag-1 autocovariances is adversely affected more than the others since $\hat{\gamma}_k=\gamma_k+O_P(1/\sqrt{n})+O(B/n)$ for $k>1$, but $\hat{\gamma}_k=\gamma_k+O_P(1/\sqrt{n})+O(B^2/n)$ for $k=0$ or 1. Hence, the YW estimator $\hat{\phi}_1=\hat{\gamma}_1/\hat{\gamma}_0=\phi_1+O_P(1/\sqrt{n})+O(B^2/n)$.

Fig. 1 shows a plot of a Gaussian AR(1) series with $\phi_1=1/2$, and $n=200$. The corrupted values are $X_{100}=5$ and $X_{101}=-5$. The estimated

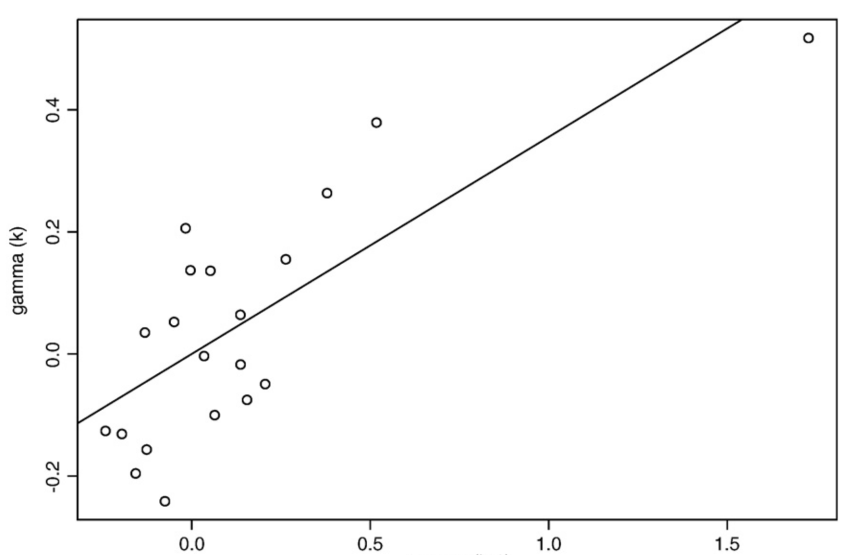

Fig. 2. Scatterplot of $\hat{\gamma}_k$ vs. $\hat{\gamma}_{k-1}$ for $k=1,\ldots, p'$ with $p'=20$ and LS line superimposed.