Journal of the Korean Statistical Society 44 (2015) 646–660

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Revisiting the estimation of the error density in functional autoregressive models

Nadine Hilgert[a], Bruno Portier[b,*]

[a] UMR 729 MISTEA, INRA SUPAGRO, 2 Place Viala, 34060 Montpellier cedex, France

[b] Normandie Université, INSA de Rouen, LMI EA 3226, BP 08, Avenue de l'Université, 76800 Saint-Etienne du Rouvray, France

ARTICLE INFO

Article history:
Received 19 August 2014
Accepted 25 April 2015
Available online 18 May 2015

AMS 2000 subject classifications:
primary 62G07
secondary 62G08

Keywords:
Kernel density estimation
Functional autoregressive models
Multivariate central limit theorem
Goodness-of-fit test

ABSTRACT

In this paper, we study some asymptotic properties of a new estimator of the probability density function of the driven noise in a nonparametric functional autoregressive model. This density estimator, based on the kernel method, does not require the estimation of the residuals of the model (as in the usual plug-in estimator). We prove its pointwise consistency with rate and establish a multivariate central limit theorem, without any assumption on the noise distribution tail. Theoretical results are illustrated with some simulation experiments. We finally propose a goodness-of-fit test of the error distribution, built from a normalized sum of the quadratic deviation between the true density and its estimator evaluated on a finite set of distinct points.

1. Introduction

The model validation is an important issue when modeling time series. In particular, analyzing residuals is a crucial step in this validation. For example, to check the assumption about Gaussian distribution, one may perform a histogram of the residuals or directly estimate their density. Kernel-based methods, as initiated by Parzen and Rosenblatt, are among the most common nonparametric methods used to that purpose. There is a huge literature on kernel density estimation in the context of the independent and identically distributed sample or mixing processes, see for example Devroye and Lugosi (2001) and Silverman (1986). But, when dealing with regression or autoregressive models, the driven noise is not observed. This makes the density estimation and the study of its properties more complex. The probability density function (pdf for short) can only be estimated through the residual errors calculated from the estimation of the unknown component of the model. This one shall thus be estimated with an appropriate convergence rate to induce good properties to the residual error. A common noise density estimator is the Parzen–Rosenblatt kernel estimator, based on this residual error which is then considered as a noise predictor.

Good properties have been proved with this estimation procedure in the framework of parametric models. Chai, Li, and Tian (1991) proved the uniform strong consistency on $\mathbb{R}$ of the noise kernel density estimator (KDE for short) in the linear regression case. In the parametric autoregressive case, Cheng (2005) showed that the asymptotic distribution of the maximum of a suitably normalized deviation of the density estimator from the expectation of the kernel error density (based on the true error) is the same as in the case of the one sample set up, which is given in Bickel and Rosenblatt (1973). These results are extended in Cheng (2010) to the nonlinear case, for which almost sure uniform convergence of the KDE and

* Corresponding author.
E-mail address: bruno.portier@insa-rouen.fr (B. Portier).

http://dx.doi.org/10.1016/j.jkss.2015.04.005

asymptotic normality results are also obtained by Liebscher (1999). Convergence rates are improved by Müller, Schick, and Wefelmeyer (2005) with the use of weighted kernel density estimators. A recent work was also published by Kim, Sin, and Kim (2014) in the same context, with extensions to a goodness-of-fit test for the error density. Conditions on the stationarity of the time-series are given in all these references.

The nonparametric setting is mainly addressed in the literature in the regression framework with independent and identically distributed data. The first works of Ahmad (1992) and Cheng (2004) showed the difficulty of combining model estimation and density estimation based on nonparametric residuals. The consistency was proved under the condition that the estimation error of the nonlinear regression function has to be uniformly weakly consistent. Efromovich (2005) pointed out that the nonparametric framework for error density estimation is "extremely complicated due to its indirect nature". He made developments under the customary assumption that the regression function is differentiable and the error density is twice differentiable. Györfi and Walk (2012) proved the strong L1-consistency of a recursive and a nonrecursive kernel density estimate based on a regression estimate, without assuming an additive noise in the model.

In this paper, we consider the problem of estimating the probability density function of the driven noise in a functional autoregressive model of order 1. More precisely, we assume that we observe a time series $X_n \in \mathbb{R}^d$ whose dynamics is described by the following model, for all $n \geq 0$,

$$X_{n+1} = f(X_n) + \varepsilon_{n+1}. \tag{1}$$

The function f of $\mathbb{R}^d$ in $\mathbb{R}^d$ is unknown and $\varepsilon = (\varepsilon_n)_{n\geq 0}$ is a sequence of independent and identically distributed (*i.i.d.* for short) random variables with zero mean, covariance matrix Γ and unknown probability density function p. The initial state X_0 is given and is independent of ε. The objective of this work is then to propose an estimator of p with good convergence properties, even though f is unknown and the (ε_n) are not observable.

In a recent paper, Hilgert and Portier (2012) studied the asymptotic properties of a plug-in estimator, the usual kernel density estimator built from the nonparametric residuals $\widehat{\varepsilon}_n$ of model (1) defined by $\widehat{\varepsilon}_n = X_n - \widehat{f}_{n-1}(X_{n-1})$, where $\widehat{f}_n$ is a nonparametric estimator of function f. They showed the strong consistency over $\mathbb{R}^d$ of the estimator of p and obtained a multivariate central limit theorem (CLT for short). These convergence results involve the convergence towards 0 of the quadratic mean of the prediction errors $(f(X_n) - \widehat{f}_n(X_n))$, which requires the uniform strong consistency of $\widehat{f}_n$. Convergence rates are slow, and specifying them requires the knowledge of the shape of the distribution tail of p, which is unknown. This last requirement is restrictive and incompatible with the use of such a plug-in estimator for the construction of a goodness-of-fit test for the noise density.

To overcome this problem, it is necessary to consider another estimator which does not involve the nonparametric residuals $(\widehat{\varepsilon}_n)$. In this paper, we introduce a kernel-based estimator $\widehat{p}$ that only requires the estimation of function f at a given point x_0, instead of requiring it along the whole trajectory $X_0, X_1, \ldots, X_{n-1}$, as does the plug-in estimator. The advantage of this approach is that the asymptotic properties of $\widehat{p}$ only depend on the pointwise convergence of $\widehat{f}_n$, which is easy to establish and leads to satisfying rates of convergence, free of any assumption on the shape of p. To that aim, we introduce a bidimensional kernel, which prevents from obtaining the usual convergence rate in $\sqrt{nh_n^{2d}}$ where h_n denotes the bandwidth of the kernel density estimator. Nevertheless, it is sufficient to derive a goodness-of-fit test for the density p.

Our paper is organized as follows. Model assumptions are specified in Section 2. In Section 3, we introduce the nonparametric estimators. Section 4 is devoted to the asymptotic results: almost sure convergence and multivariate CLT are given. A numerical evaluation is presented in Section 5 to highlight the contribution of the paper. Section 6 illustrates the use of the multivariate CLT in a goodness-of-fit test for the density p. All technical proofs are postponed in appendices. Appendix A is concerned with two technical lemmas and Appendix B with the proof of the main theorem of the paper.

2. Model assumptions and consequences

This section is devoted to the assumptions made on the model and their resulting properties. We give conditions under which $(X_n)_{n\geq 0}$ is asymptotically stationary and possesses an invariant distribution.

To ensure the stability of model (1) and the existence of an asymptotic invariant distribution, we set the following usual assumptions (see for example Duflo, 1997).

Assumption A1. The function f is continuous and there are two positive constants $r_f < 1$ and C_f such that for any $x \in \mathbb{R}^d$,

$$\|f(x)\| \leq r_f\|x\| + C_f. \tag{2}$$

Assumption A2. The initial state X_0 and $\varepsilon = (\varepsilon_n)_{n\geq 0}$ have a finite moment of order $m > 2$. The probability density function p of ε is positive ($p > 0$).

The main consequence of Assumptions A1 and A2 is that the process $X = (X_n)_{n\geq 0}$ is asymptotically stationary and possesses an invariant distribution μ, which has a finite moment of order m and a probability density function denoted h,