Empirical properties of forecasts with the functional autoregressive model

Devin Didericksen · Piotr Kokoszka · Xi Zhang

Received: 10 December 2010 / Accepted: 20 April 2011 / Published online: 3 May 2011
© Springer-Verlag 2011

Abstract We study the finite sample performance of predictors in the functional (Hilbertian) autoregressive model $X_{n+1} = \Psi(X_n) + \varepsilon_n$. Our extensive empirical study based on simulated and real data reveals that predictors of the form $\hat{\Psi}(X_n)$ are practically optimal in a sense that their prediction errors are comparable with those of the infeasible perfect predictor $\Psi(X_n)$. The predictions $\hat{\Psi}(X_n)$ cannot be improved by an improved estimation of $\Psi$, nor by a more refined prediction approach which uses predictive factors rather than the functional principal components. We also discuss the practical limits of predictions that are feasible using the functional autoregressive model. These findings have not been established by theoretical work currently available, and may serve as a practical reference to the properties of predictors of functional data.

Keywords Autoregressive process · Functional data · Prediction

1 Introduction

Over the last two decades, functional data analysis (FDA) has grown into a substantial field of statistical research, with new methodology, numerous useful applications and interesting novel theoretical developments. In this brief note, we cannot even outline
the central ideas, as the field has become very broad, so we merely mention comprehensive introductory expositions of Ramsay and Silverman (2002), Ramsay and Silverman (2005), and Ramsay et al. (2009), and more theoretical works by Bosq (2000), Ferraty and Vieu (2006), Bosq and Blanke (2007) and Ferraty and Romain (2011).

The research summarized in this paper pertains to the functional autoregressive (FAR) process studied theoretically by Bosq (2000), and extensively used in both practical and theoretical studies since then, see Besse and Cardot (1996), Damon and Guillas (2002), Antoniadis and Sapatinas (2003), Horváth et al. (2010), Hörmann and Kokoszka (2010), Gabrys et al (2010), among numerous other contributions. The FAR(1) model is given by the equation

\[ X_{n+1} = \Psi(X_n) + \epsilon_{n+1}, \]  

(1)

in which the errors \( \epsilon_n \) and the observations \( X_n \) are curves, and \( \Psi \) is a linear operator transforming a curve into another curve. Precise definitions and assumptions are stated in Sects 2. Model (1) has been introduced to predict curve–valued time series. In addition to Bosq (2000), an informative introduction and review of several prediction methods is given by Besse et al. (2000).

Recently Kargin and Onatski (2008) proposed a sophisticated method of one step ahead prediction in model (1) based on predictive factors, and developed an advanced theory that justifies the optimality of their method, we provide a description in Sect. 2. The initial question that motivated this research was whether the method of Kargin and Onatski (2008) is superior in finite samples to the standard method described in Bosq (2000), which estimates the operator \( \Psi \) and forecasts \( X_{n+1} \) by \( \hat{\Psi}(X_n) \). We found that the predictive factors method never dominates the standard method, and in some cases it performs poorly. We also found that the standard method is almost perfect in a sense that its average prediction errors are typically, within a standard error, the same as if we had perfect knowledge of the operator \( \Psi \). Thus it cannot be hoped that this method can be substantially improved. Surprisingly, this is the case even though the estimates \( \hat{\Psi} \) of the operator \( \Psi \) are typically very poor. We found that it is possible to improve these estimates, we developed a simple algorithm to do it, but this improvement does not affect the quality of prediction. Finally, we realized some natural limits of predictions that can be expected from model (1); predictions with \( \hat{\Psi}(X_n) \) are often not better than those with the mean function. We describe in this paper how we arrived at all these conclusions. It is hoped that this contribution will provide informative and useful insights into finite sample properties of estimators and predictors in the FAR(1) model, whose theoretical properties have already been studied in depth.

The paper is organized as follows. In Sect. 2, we describe the two prediction methods and state the assumptions for their validity. Before comparing them, we address in Sect. 4 the question of the estimation of \( \Psi \), and show how better estimates can be constructed. Sects. 3 and 5 describe, respectively, the design of the simulation study and its outcomes. We conclude with Sect. 6 which discusses general properties of predictors derived from model (1).