Journal of Statistical Planning and Inference 177 (2016) 1–27

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

# Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions

Li Pan, Dimitris N. Politis *

*Department of Mathematics, University of California–San Diego, La Jolla, CA 92093-0112, USA*

ARTICLE INFO

*Article history:*
Received 23 March 2014
Accepted 15 October 2014
Available online 7 November 2014

*Keywords:*
Confidence intervals
Forecasting
Time series

ABSTRACT

In order to construct prediction intervals without the cumbersome – and typically unjustifiable – assumption of Gaussianity, some form of resampling is necessary. The regression set-up has been well-studied in the literature but time series prediction faces additional difficulties. The paper at hand focuses on time series that can be modeled as linear, nonlinear or nonparametric autoregressions, and develops a coherent methodology for the construction of bootstrap prediction intervals. Forward and backward bootstrap methods using predictive and fitted residuals are introduced and compared. We present detailed algorithms for these different models and show that the bootstrap intervals manage to capture both sources of variability, namely the innovation error as well as estimation error. In simulations, we compare the prediction intervals associated with different methods in terms of their achieved coverage level and length of interval.



## 1. Introduction

Statistical inference is not considered complete if it is not accompanied by a measure of its inherent accuracy. With point estimators, the accuracy is measured either by a standard error or a confidence interval. With (point) predictors, the accuracy is measured either by the predictor error variance or by a *prediction interval*.

In the setting of an i.i.d. (independent and identically distributed) sample, the problem of prediction is not interesting. However, when the i.i.d. assumption no longer holds, the prediction problem is both important and intriguing; see Geisser (1993) for an introduction. Typical situations where the i.i.d. assumption breaks down include regression and time series.

The literature on predictive intervals in regression is not large; see e.g. Carroll and Ruppert (1988), Patel (1989), Schmoyer (1992) and the references therein. Note that to avoid the cumbersome (and typically unjustifiable) assumption of Gaussianity, some form of resampling is necessary. The residual-based bootstrap in regression is able to capture the predictor variability due to errors in model estimation. Nevertheless, bootstrap prediction intervals in regression are often characterized by finite-sample *undercoverage*. As a remedy, Stine (1985) suggested resampling the studentized residuals but this modification does not fully correct the problem; see the discussion in Olive (2007). Politis (2013) recently proposed the use of *predictive* (as opposed to fitted) residuals to be used in resampling which greatly alleviates the finite-sample undercoverage.

Autoregressive (AR) time series models, be it linear, nonlinear, or nonparametric, have a formal resemblance to the analogous regression models. Indeed, AR models can typically be successfully fitted by the same methods used to estimate a regression, e.g., ordinary Least Square (LS) regression methods for parametric models, and scatterplot smoothing for

* Corresponding author.
*E-mail addresses:* lipan@ucsd.edu (L. Pan), dpolitis@ucsd.edu (D.N. Politis).

http://dx.doi.org/10.1016/j.jspi.2014.10.003
0378-3758/

nonparametric ones. The practitioner has only to be careful regarding the standard errors of the regression estimates but the model-based, i.e., residual-based, bootstrap should in principle be able to capture those.

Therefore, it is not surprising that model-based resampling for regression can be extended to model-based resampling for *auto*-regression. Indeed, standard errors and confidence intervals based on resampling the residuals from a fitted AR model has been one of the first bootstrap approaches for time series; cf. Freedman (1984), Efron and Tibshirani (1986), and Bose (1988).

However, the situation as regards prediction intervals is not as clear; for example, the conditional nature of the predictive inference in time series poses a difficulty. There are several papers on prediction intervals for linear AR models but the literature seems scattered and there are many open questions: (a) how to implement the model-based bootstrap for prediction, i.e., how to generate bootstrap series; (b) how to construct prediction intervals given the availability of many bootstrap series already generated; and lastly (c) how to evaluate asymptotic validity of a prediction interval. In addition, little seems to be known regarding prediction intervals for nonlinear and nonparametric autoregressions.

In the paper at hand we attempt to give answers to the above, and provide a comprehensive approach towards bootstrap prediction intervals for linear, nonlinear, or nonparametric autoregressions. The models we will consider are of the general form:

- **AR model with homoscedastic errors**

$$X_t = m(X_{t-1}, \ldots, X_{t-p}) + \epsilon_t \tag{1.1}$$

- **AR model with heteroscedastic errors**

$$X_t = m(X_{t-1}, \ldots, X_{t-p}) + \sigma(X_{t-1}, \ldots, X_{t-p})\epsilon_t. \tag{1.2}$$

In the above, $m(\cdot)$ and $\sigma(\cdot)$ are unknown; if they can be are assumed to belong to a finite-dimensional, parametric family of functions, then the above describe a linear or nonlinear AR model. If $m(\cdot)$ and $\sigma(\cdot)$ are only assumed to belong to a smoothness class, then the above models describe a nonparametric autoregression. Regarding the errors, the following assumption is made:

$$\epsilon_1, \epsilon_2, \ldots \text{ are i.i.d. } (0, \sigma^2), \text{ and such that } \epsilon_t \text{ is independent from } \{X_s, s < t\} \text{ for all } t; \tag{1.3}$$

in conjuction with model (1.2), we must further assume that $\sigma^2 = 1$ for identifiability. Note, that under either model (1.1) or (1.2), the *causality* assumption (1.3) ensures that $E(X_t|\{X_s, s < t\}) = m(X_{t-1}, \ldots, X_{t-p})$ gives the optimal predictor of $X_t$ given $\{X_s, s < t\}$; here optimality is with respect to Mean Squared Error (MSE) of prediction.

Section 2 describes the foundations of our approach. Pseudo-series can be generated by either a forward or backward bootstrap, using either fitted or predictive residuals—see Section 2.1 for a discussion. Predictive roots are defined in Section 2.2 while Sections 2.3 and 2.4 discuss notions of asymptotic validity. Section 3 goes in depth as regards bootstrap prediction intervals for linear AR models. Section 4 addresses the nonlinear case using two popular nonlinear models as concrete examples. Finally, Section 5 introduces bootstrap prediction intervals for nonparametric autoregressions. A short conclusions section recapitulates the main findings making the point that the forward bootstrap with fitted or predictive residuals serves as the unifying principle across all types of AR models, linear, nonlinear or nonparametric.

## 2. Bootstrap prediction intervals: laying the foundation

### 2.1. Forward and backward bootstrap for prediction

As previously mentioned, an autoregression can be formally viewed as regression. However, in prediction with an AR($p$) model, linear or nonlinear, an additional difficulty is that the one-step-ahead prediction is done conditionally on the last $p$ observed values that are themselves random.

To fix ideas, suppose $X_1, \ldots, X_n$ are data from the linear AR(1) model: $X_t = \phi_1 X_{t-1} + \epsilon_t$ where $|\phi_1| < 1$ and the $\epsilon_t$ are i.i.d. with mean zero. Given the data, the MSE-optimal predictor of $X_{n+1}$ given the data is $\phi_1 X_n$ which is approximated in practice by plugging-in an estimator, say $\hat{\phi}_1$, for $\phi_1$. Generating bootstrap series $X_1^*, X_2^*, \ldots$ from the fitted AR model enables us to capture the variability of $\hat{\phi}_1$ when the latter is re-estimated from bootstrap datasets such as $X_1^*, \ldots, X_n^*$.

For the application to prediction intervals, note that the bootstrap also allows us to generate $X_{n+1}^*$ so that the statistical accuracy of the predictor $\hat{\phi}_1 X_n$ can be gauged. However, none of these bootstrap series will have their last value $X_n^*$ *exactly* equal to the original value $X_n$ as needed for prediction purposes. Herein lies the problem, since the behavior of the predictor $\hat{\phi}_1 X_n$ needs to be captured *conditionally* on the original value $X_n$.

To avoid this difficulty, Thombs and Schucany (1990) proposed to generate the bootstrap data $X_1^*, \ldots, X_n^*$ going backwards from the last value that is fixed at $X_n^* = X_n$. This is the *backward bootstrap* method that was revisited by Breidt et al. (1995) who gave the correct algorithm of finding the backward errors. Note that the generation of $X_{n+1}^*$ is still done in a forward fashion using the fitted AR model conditionally on the value $X_n$.