Combining bootstrap methods: a Monte Carlo experiment

Luisa Bisaglia[1] and Margherita Gerolimetto[2]

[1] luisa.bisaglia@unipd.it
University of Padua, Via Cesare Battisti 241, 35121 PD, Italy
[2] margherita.gerolimetto@unive.it
Ca' Foscari University Venice, Cannaregio 873, 30121 VE, Italy

Abstract. The idea of this work is to study whether the combination of different bootstrap methods can lead to an improvement in the performance, as it does in the forecasting framework where is widely used. Given that it represents a rather challenging set-up, we will focus on long memory time series and we will explore the combination of three very well-known bootstrap methods for time series with long range dependence. We present a very preliminary Monte Carlo experiment that provides interesting and promising results.

Key words: bootstrap, long memory, combination

1 Introduction

Bootstrap methods are still now among the most popular tools to approximate the distribution of certain statistics that would be difficult to be handled otherwise. Bootstrap methods have been proposed in principle for independent observations and afterwords they have been extended to the case when there is serial dependence, such as in time series data. In this context, long memory time series represent quite a complex setting since the strong dependence characterizing those data often requires specific techniques to form a valid resampling strategy.

Traditional bootstrap strategies for time series can be broadly divided into two categories. The first includes methods that resample the residuals of a parametric estimated model (assuming they are *iid*) and the validity of these methods is based on the correct specification of the model and on the consistency of the estimation methods (e.g [1]). The second category gathers bootstrap techniques that do not strongly rely on parametric restrictions, the most popular in case of long memory time series are the Sieve Bootstrap ([2]), that is considered semiparametric, and the Block Bootstrap ([3]; [4]), that is considered nonparametric. The element in common to all those methods is that the final resampling is implemented in approximately independent quantities (either residuals or blocks) that yet is difficult to guarantee in long memory time series whose serial dependence is often so strong that it persists even when the observations are distant, as emphasized by [5] in a review that presents also an extensive Monte Carlo experiment together with a possible guide for practitioners.

To the best of our knowledge, there is no attempt to consider the possibility of combining bootstrap methods, as it is often done with forecasts. The first contribute to formalize the idea of combining multiple individuals forecasts is by [6]. Still now in the literature it is established that forecast combinations are beneficial. With this in mind, the aim of this work is to investigate whether also for bootstrap methods the idea of combination can lead to a possibly better performance. Given that it represents quite a challenging set-up, we will focus our attention on long memory time series. From the methodological point of view, for now, we will confine ourselves to the Sieve Bootstrap, the Moving Block Bootstrap and the Parametric Bootstrap.

The structure of the paper is as follows. In the second section we present our methodological framework. In the third section we describe the Monte Carlo experiment and some table results.

2 Bootstrap combination

In this section we will present our methodological framework.

2.1 The combination of forecast: state of the art

The literature on forecast combinations has notably flourished (among others see [7]) for a very recent and rich review). Combining multiple forecasts is now widely adopted to improve the performance. There are several reasons why the forecast combinations perform well, the most intuitive one being that the combination is capable of integrating different information (not completely overlapping) provided by the individual forecasts. The advantage from forecast combination does not depend only on the quality of the individual forecast to be combined but also on how the weights assigned to each forecast are determined. There is a big variety of combination schemes that range from simple combinations that avoid weight estimation (e.g. [8]) to very sophisticated methods that assign specific weights to each individual forecast (e.g. [9]). In this large set of options, equally weighted averages, which do not account for past information regarding the precision of the individual forecasts and correlation between forecast errors, work reasonably well compared to potentially more rigorous combination schemes.

In this paper we will borrow the set-up and the literature advancement in the the forecast combination to extend it to bootstrap resampling strategies, as will be presented in the next sections.

2.2 Long memory time series

We will consider a linear process of the form

$$y_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j} \tag{1}$$

where ϵ_t are *iid* variables with zero mean and finite variance and a_j are real-valued constants. When $\sum_{j=0}^{\infty} a_j^2 < \infty$ the process is stationary. Long memory time series are characterized by a significant dependence between distant observation such that the autocovariance function satisfies:

$$\Gamma_y(k) = Cov(y_t, y_{t-k}) \sim Gk^{2d-1} \quad k \to \infty \tag{2}$$

where G is a finite constant and d is the memory parameter that governs the strong persistence of y_t. When $|d| < 1/2$ the series is stationary, but it shows long memory (the autocovariances are not summable) and the spectral density (for C positive constant) behaves as $f_y(\lambda) \sim C\lambda^{-2d} \quad \text{when} \quad \lambda \to 0$. One of the best know form of long memory is fractional integration, denoted by $y_t \sim I(d)$ of which ARFIMA(p,d,q) is undoubtedly the most used.

$$\Phi_p(B)(1-B)^d y_y = \Theta_q(B)\epsilon_t \tag{3}$$

where $\Phi_p(B)$ and $\Theta_q(B)$ are polynomials of order p and q respectively, with all roots outside the unit circle and no roots in common.

As for the estimation of the long memory parameter d, the literature is rich of proposals, we will focus here on the most traditional ones, implemented in several packages (for example in R). They are the Log-periodogram estimator (LPE) proposed by [10] and further developed by [11] and the Local Wittle estimator (LWE), proposed by [12] and also further developed by [11]. Due to lack of space, we refer to the original literature for the technical details of the estimators.

2.3 The combination of bootstraps

Following the intuition behind the combination of forecasts briefly recalled before, here we present our (very preliminary) proposal of a Combined Bootstrap focussing on the specific case of stationary long memory time series.

It is given a (stationary long memory time series) y_t, $t = 1, ..., T$ and a statistic S (in our case LPE or LWE) whose distribution should be approximated via bootstrap. The idea is to obtain a combined bootstrap sample y_t^{**} by linearly combining (with a simple combination scheme, for now) the bootstrap samples outcoming from each of three methods suitable for long memory, specifically Parametric Bootstrap, Sieve Bootstrap, Moving Block Bootstrap. The Parametric Bootstrap ([1]) moves from the assumption that the time series y_t, $t = 1, ..., T$ is the realization of an $ARFIMA(p, d, q)$ process (3). In very simple terms, once the model is estimated, y_t^* will be computed from the estimated model using simulated normal random variables (with mean and variance obtained from the parametric model) in place of the residuals and the unknown parameters replaced by the estimated values. The Sieve Bootstrap ([2]) is a semiparametric technique and it is based on the approximation of the data generating process with an $AR(h)$ model. The bootstrap samples are drawn from the studentized residuals obtained from the autoregressive approximation. The Block Bootstrap is an entirely nonparametric strategy that is based on resampling block of observations that can be overlapping ([4]) or not ([3]). The approximate independence of the block is necessary for this bootstrap being successful and in case of long memory this is very difficult to find ([13]). [14] recommend in this case the use of overlapping blocks (so called Moving Block Bootstrap) and provide some rules of thumb for the blocks length.

The statistic S is evaluated on the Combined Bootstrap sample. The algorithm for $b = 1, ..., B$ bootstrap samples is the following:

Step 1: Resample y_t with the Parametric Bootstrap and obtain $y^{*PB}_{b,t}$

Step 2: Resample y_t with the Sieve Bootstrap and obtain $y^{*SB}_{b,t}$

Step 3: Resample y_t with Moving Block Bootstrap and $y^{*MBB}_{b,t}$

Step 4: Combine (with a simple linear scheme) the three bootstrap samples and obtain $y^{**}_{b,t}$:

$$y^{**}_{b,t} = \frac{y^{*PB}_{b,t} + y^{*SB}_{b,t} + y^{*MBB}_{b,t}}{3} \tag{4}$$

Step 5: Repeat steps 1-4 to obtain B bootstrap samples and calculate B values of the desired statistic S^{**}_b

Step 6: Approximate the distribution of the statistic S with the bootstrap distribution (where I is the indicator function):

$$FS^{**}(s) = \frac{1}{B}\sum_{b=1}^{B} I(S^{**}_b \leq s)$$

As mentioned, this bootstrap algorithm is at the moment very preliminary. Following the hints coming from the literature on combination of forecasts, we begin here with a simple linear combination scheme that combines the most suitable time domain tools for bootstrapping long memory time series, with the intention to put forward this idea via possibly more complex and sophisticated weighting scheme and through the consideration of other bootstrapping methods.

3 Monte Carlo experiment

In order to investigate the performance of the Combined Bootstrap with respect to the other bootstrap methods we intend to carry out a Monte Carlo experiment. Here we will present some preliminary results that will be the guidance for a future more extensive simulation design.

Along the lines of the Monte Carlo experiment in [5], for now, we consider $S = 1000$ Monte Carlo simulations from an $ARFIMA(0, d, 0)$ model, for $T =$ 250, 500 sample sizes and for $d = 0.2$, 0.4 values of the long memory parameter. The statistics under examinations are the estimators LWE and LPE of the long memory parameter where $m_{LWE} = 1+T^{0.65}$ and $m_{LPE} = 1+T^{0.8}$ (following the related literature and [5]). For the implementation of the Parametric Bootstrap, an $ARFIMA(0,d,0)$ model is estimated via maximum likelihood. As for the Sieve Bootstrap, the autoregression is estimated via maximum likelihood and the order of the model is selected via AIC minimization. Finally, for the Moving Block Bootstrap the block length is $k = [T^{1/2}]$, as in [14]. The simulations are run with the software R (R Core Team, 2023).

The performance is expressed with two measures ([5]). The first is the Root Mean Square Deviance (RMSD) of the bootstrap probability density function to the true probability function (obtained via Monte Carlo):

$$RMSD(boot) = \sqrt{\frac{1}{S}\sum_{j=1}^{S}\left(p_{boot}(\hat{d}_j - d_0) - p_{mc}(\hat{d}_j - d_0)\right)^2} \qquad (5)$$

where *boot* is one of the bootstrap methods, d_0 is the true value, $p_{mc}(\hat{d}_j - d_0)$ is the ordinate of a kernel density based on the S estimates $\hat{d}_j - d_0$, $j = 1, ..., S$ (evaluated on $\hat{d}_j - d_0$), $p_{boot}(\hat{d}_j - d_0)$ is average over the S Monte Carlo simulations of the kernel density based on $\hat{d}_j^b - \hat{d}_j$ the estimates from the $b = 1, .., B$ bootstrap samples (evaluated on $\hat{d}_j - d_0$). The closer is the RMSD to zero, the better the performance of the bootstrap strategy. The second measure of performance is the coverage frequency of confidence intervals. The percentiles coming from any of the bootstrap strategy are constructed in general as follows: $CI^*_{1-\alpha}(\hat{d}) = \left(\hat{d}_{\alpha/2}; \hat{d}_{1-\alpha/2}\right)$ where $\hat{d}_\alpha$ is the $B\alpha$ ordered value of the bootstrap estimate of d with, in turn LWE and LPE. The closer the coverage frequency to $1-\alpha$ (for us 95%), the better the performance of the bootstrap.

In table (1) are provided some of the results of the preliminary experiment. We propose the results for $d = 0.4$ and $T = 500$, the others are available upon request, but their meaning is substantially analogous.

Table 1. RMDS and Coverage Frequency of bootstrap distributions for LWE and LPE

LWE			LPE		
Bootstrap	RMSD	Cov Freq	Bootstrap	RMSD	Cov Freq
Comb Boot	**0.825**	**0.901**	Comb Boot	**1.802**	**0.989**
SB Boot	1.058	0.831	SB Boot	1.759	0.972
MBB Boot	1.317	0.740	MBB Boot	2.311	0.944
PB Boot	1.468	0.991	PB Boot	2.150	0.994

To describe these preliminary results we need first to clarify that the data generating process considered in this experiment is the same as the model estimated in the Parametric Bootstrap algorithm, meaning that the Parametric Bootstrap is implemented in the best scenario and is expected to perform at best. We did this on purpose, so to have some kind of upper bound of the ideal performance with which the other bootstraps can be compared. As we can see, the Combined Bootstrap has a performance comparable with the parametric bootstrap in terms of coverage frequency and even better than the Parametric Bootstrap in terms of RMSD. The intuition we get form these set of preliminary results is that the Combined Bootstrap, if developed with possibly more sophisticated tools (such as a different weighting scheme for the linear combination or a more wide and informed selection of the component bootstraps) could significantly represent a valid alternative to existing bootstrap strategy.

References

1. Andrews, D.W.K., Lieberman, O., Marmer, V.: Higher-order improvements of the parametric bootstrap for long-memory Gaussian processes. Journal of Econometrics 133: 673-702. (2024)
2. Buhlmann, P: Sieve bootstrap for time series. Bernoulli 3: 123 - 148. (1997)
3. Carlstein, E.: The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence. Annals of Statistics 14: 1171-1179. (1986)
4. Kunsch, H.R: The Jacknife and bootstrap for general stationary observations. Annals of Statistics 17: 1217-1261. (1989)
5. Arteche, J.: Bootstrapping long memory time series: Application in low frequency estimators. Econometrics and Statistics 29: 1-15. (2024)
6. Bates, J.M., Granger, C.W.J.: The Combination of Forecasts. OR 2: 451-68. (1969)
7. Wang, X., Hyndman, R., Li, F., Kang, Y.: Forecast combinations: An over 50-year review. International Journal of Forecasting 39: 1518-1547. (2023)
8. Petropoulos F., Svetunkov, I.: A simple combination of univarie models. International Journal of Forecasting 36: 110-115. (2020)
9. Wang, X., Kang, Y., Petropoulos, F., Li, F.: The uncertainty estimation of feature-based forecast combinations. Journal of the Operational Research Society 73: 979-993. (2022)
10. Geweke, J. Porter-Hudack, S.: The Estimation and Application of Long-Memory Time Series Models. Journal of Time Series Analysis 4: 221-237. (1983)
11. Robinson, P.M.: Log-periodogram regression of time series with long-range dependence. The Annals of Statistics 23: 1048-1072. (1995)
12. Fox, R, Taqqu, M.S.: Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. The Annals of Statistics 14: 517 -532. (1986)
13. Lahiri, S.N.: On the moving block bootstrap under long range dependence. Statistics and Probability Letters 18: 405-413. (1993)
14. Kim, Y.M., Nordman, D.J.: Properties of a Block Bootstrap under Long-range Dependence. Sankhya: The indian Journal of Statistics 73-A: 79-109 (2011)
15. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. (2023)