

Moving Block Bootstrap for a class of bilinear models

Moving Block Bootstrap per una classe di modelli bilineari

Francesco Giordano

Dipartimento di Scienze Economiche e Statistiche

Università degli Studi di Salerno

giordano@unisa.it

Riassunto: Questo lavoro si pone nel contesto della classe (I) dei modelli bilineari introdotti da Grahn (1995) nell'ottica della stima bootstrap della varianza degli stimatori dei minimi quadrati condizionati. L'obiettivo è quello di valutare teoricamente l'applicabilità, in tale contesto, del bootstrap a blocchi (Moving Block Bootstrap) e quindi la possibilità di selezionare la lunghezza del blocco tramite una funzione di Influenza derivata analiticamente.

Keywords: Bilinear model, Moving Block Bootstrap, CLS estimator

1. Introduction

The framework of non-linear time series models offers many statistics of interest whose sample distribution and variance estimator are very difficult to derive. In these cases the use of a general methodology such as Bootstrap is often recommended.

We focus our attention on a class (I) of bilinear models introduced by Grahn (1995). The stochastic process X_t , belonging to the class (I), is defined as:

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{i=1}^q c_i \epsilon_{t-i} + \sum_{i=1}^k \sum_{j=w}^r b_{ij} \epsilon_{t-i} X_{t-j} + \epsilon_t \quad (1)$$

where $w = \max(k, q) + 1$ and ϵ_t is an i.i.d. White Noise with zero mean and variance σ^2 . Without loss of generality we suppose that ϵ_t has a Normal distribution.

Grahn (1995) introduced the Conditional Least Squares (CLS) estimators for the parameters of (1) whereas the variance of the CLS estimators, which is very difficult to derive analytically, is not yet known.

In this paper we propose to use the Moving Block Bootstrap (MBB) (Künsch, 1989) to estimate the variance of the CLS estimators given that, under mild conditions, they are strongly consistent and converge weakly to the Normal distribution (Grahn, 1995). In particular we analyze two problems: the first concerns the functional form of the CLS estimators written in terms of their Influence function; the second problem is related to the selection of the block length according to the procedure of Bühlmann and Künsch (1999) with a proposed analytic Influence function.

The paragraph 2 introduces the notations, definitions and some properties of the MBB procedure and the CLS estimators. A theoretical justification to apply MBB to the CLS estimators is showed in the third paragraph. Finally, a simulation experiment with a theoretical Influence function and some concluding remarks are reported in paragraph 4.

2. Definitions and properties of the Moving Block Bootstrap and the CLS estimators

In this paragraph we introduce briefly the MBB procedure and we report the main results about the CLS estimators of the parameters in model (1).

It is known that X_t , in model (1), is a strictly stationary, causal and ergodic process under some conditions on the coefficients $\{a_i\}$, $i = 1, \dots, p$, and $\{b_{ij}\}$, $i = 1, \dots, k$, $j = w, \dots, r$, independently from the MA coefficients (Grahm, 1995), (Tong, 1990). So we assume that

A0) the process X_t is strictly stationary, causal and ergodic

and $T_n \equiv T_n(X_1, \dots, X_n)$ is a CLS estimator for a parameter of model (1), with n the length of the observed time series. The aim is to estimate the variance of T_n via MBB. Consider the blockwise bootstrap procedure with blocks of length l from the original data. This parameter l is related to the dependence structure in the observed data. For simplicity, we assume that n is a multiple of l and let $d = n/l$.

The bootstrap sample is: $X_{(j-1)l+t}^* = X_{S_j+t}$, $\forall t \in \{1, \dots, l\}$, $\forall j \in \{1, \dots, d\}$ with $S_1, \dots, S_d$, Uniform i.i.d. random variables on $\{0, 1, \dots, n-l\}$. If we draw $S_1 = s_1, \dots, S_d = s_d$ we have a bootstrap sample $(X_1^*, \dots, X_n^*)$, $T_n^* \equiv T_n(X_1^*, \dots, X_n^*)$ is the bootstrap statistic and the bootstrap variance estimator is $\hat{\sigma}_{BOOT}^2(l) = Var^*(T_n^*)$, where Var^* denotes the variance of T_n^* conditional on $X_1, \dots, X_n$ (Künsch, 1989).

If the statistic has a dimension m as in Bühlmann (2002), greater than one, then the MBB procedure should be made using $Y_t = (X_t, \dots, X_{t-m+1})$ (Bühlmann, 2002).

Künsch (1989) showed that $n\hat{\sigma}_{BOOT}^2(l) \sim 2\pi\hat{h}_{IF}(0)$ where $\hat{h}_{IF}(0)$ is the estimated spectral density function at frequency 0 with respect to the Influence function of the estimator T_n . In particular the block length l is related to the triangular spectral lag window or Bartlett spectral window.

So we have to verify that some "regularity" conditions (A1-A4) in Künsch (1989) are satisfied for the estimator T_n in order to apply the MBB variance estimator. The conditions (A1-A4) mean that, for the estimator T_n ,

- A1) $T_n \xrightarrow{a.s.} T(F^m)$, where F^m is the distribution of Y_t ;
- A2) the Influence function, $IF(y, F^m)$ exists $\forall y \in \mathcal{R}^m$;
- A3) $n^{-1/2} \sum_{t=1}^n IF(Y_t, F^m) \xrightarrow{d} N(0, \sigma_{as}^2)$,
where $\sigma_{as}^2 = \sum_{k=-\infty}^{\infty} E[IF(Y_0, F^m)IF(Y_k, F^m)]$;
- A4) the linearization $T(\hat{F}^m) = T(F^m) + 1/n \sum_{t=1}^n IF(Y_t, F^m) + R_n$,
where $\hat{F}^m$ denotes the empirical distribution function of dimension m and the remainder term is $R_n = o_p(n^{-1/2})$.

These four conditions state the existence of an Influence function for the estimator T_n and its asymptotic variance σ_{as}^2 can be estimated only by its Influence function.

If (A0) holds and $E(X_t^8) < \infty$ the following results are valid for the CLS estimator (Grahm, 1995):

- B1) $T_n \xrightarrow{a.s.} T \equiv \theta$ a parameter in model (1);
- B2) $n^{1/2}(T_n - T) \xrightarrow{d} N(0, V^2)$ with V^2 the asymptotic variance of $n^{1/2}T_n$.

3. Main results

The MBB resampling procedure can be applied to the CLS estimators if the above "regularity" conditions (A1-A4) with respect to T_n are satisfied. So we can state the following theorem

Theorem *If the process X_t is defined as in (1), (A0) holds and $E(X_t^8) < \infty$, then the conditions (A1-A4) are valid for the CLS estimator T_n .*

Proof.

From the (CA) condition in Grahn (1995) the CLS estimator T_n can be written as

$$T_n = f(\hat{M}_m) + o_p(n^{-1/2}) \quad (2)$$

where $\hat{M}_m$ is a vector of the estimated mixed moments up to order $m = 4$ (Grahn (1995), Theorem 3.1); so $m = 4$ is also the dimension of the statistic T_n in (A1); the function $f(\cdot)$ is continuously differentiable in a neighborhood of $\hat{M}_m$, the corresponding parameter of $\hat{M}_m$. Using (B1) it is easy to show that (A1) is valid with m up to order 4 (Grahn (1995), Theorem 3.3). Since $f(\cdot)$ is continuously differentiable and the Influence function exists for the moments $(M_m)_i$, $i = 1, \dots, \nu$, with ν the dimension of vector $\hat{M}_m$, then (A2) holds using (2). So the Influence function of T_n is

$$IF_{T_n}(\cdot, \cdot) = \sum_{i=1}^{\nu} g_i IF_{(\hat{M}_m)_i}(\cdot, \cdot) + o_p(n^{-1/2}) \quad (3)$$

where g_i , $i = 1, \dots, \nu$, are constants which depend only on $f(\cdot)$.

(A3) is true by Theorem 3.1 in Grahn (1995).

To prove (A4) we can rewrite the linearization of T_n , using (2), as

$$T_n = f(\hat{M}_m) + \sum_{i=1}^{\nu} g_i [(\hat{M}_m)_i - (M_m)_i] + R_f(n) + o_p(n^{-1/2})$$

where $R_f(n)$ is the remainder term for the Taylor's expansion of $f(\cdot)$ in the neighborhood of $\hat{M}_m$. Since X_t is geometrically ergodic (Tong, 1990) and it is also exponentially ϕ -mixing, the variance of $(\hat{M}_m)_i$ exists $\forall i$ and it is consistent of order n^{-1} . Finally, by standard results, $R_f(n) = o_p(n^{-1/2})$. $\diamond$

This theorem assures that the asymptotic variance of T_n is completely specified by its Influence function. Besides, the theorem gives its functional form as linear combination of the Influence function of some mixed moments up to order 4.

4. Empirical results

Taking into account two representations of model (1) as

$$\text{M1)} \quad X_t = aX_{t-1} + 0.3\epsilon_{t-1}X_{t-2} + \epsilon_t; \quad a = 0.5$$

$$\text{M2)} \quad X_t = b\epsilon_{t-1}X_{t-2} + \epsilon_t; \quad b = 0.3$$

with $\epsilon_t \sim N(0, \sigma^2)$, i.i.d., $\sigma^2 = 1$, which satisfy (A0) and $E[X_t^8] < \infty$, we consider $\hat{a}$ as the CLS estimator for a in (M1) and $b\hat{\sigma}^2$ as the CLS estimator of $b\sigma^2$ in (M2).

We want to estimate their variance via MBB. In this context we compute the block length l using a jackknife estimate of the Influence function (l_{jack}) (Bühlmann and Künsch, 1999)

and the theoretical Influence function defined in (3)(l_{th}). By direct calculations we get the theoretical Influence functions without scale factors:

$$IF_{\hat{a}}(X_t, F) = X_t X_{t-1} - \frac{\hat{a}}{2} X_t^2; \quad IF_{b\hat{\sigma}^2}(X_t, F) = X_t X_{t-1} X_{t-2} - \frac{b\hat{\sigma}^2}{3} X_{t-2}^2.$$

The experiment of simulation is drawn using 200 Monte-Carlo runs, 500 bootstrap replicates, two values $\{200, 500\}$ as length of the observed time series and two block lengths l_{jack} and l_{th} . Define $\sigma_n^2 = nVar(T_n)$, it is computed by 10000 Monte-Carlo runs. Let $n\hat{\sigma}_{BOOT}^2 = nVar^*(T_n^*)$, $SD[n\hat{\sigma}_{BOOT}^2] = \sqrt{Var(n\hat{\sigma}_{BOOT}^2)}$ and $RMSE = E[(n\hat{\sigma}_{BOOT}^2 - \sigma_n^2)^2]/\sigma_n^2$. The table 1 shows that the MBB variance estimates can be considered equiv-

Table 1: Comparison of the bootstrap variance estimation using jackknife Influence function (l_{jack}) and theoretical Influence function (l_{th}) with the CLS estimators $\hat{a}$ and $b\hat{\sigma}^2$ in models (M1) and (M2), respectively. RMSE is the relative mean squared error.

CLS Estimator	n	σ_n^2	l	$E[n\hat{\sigma}_{BOOT}^2] - \sigma_n^2$	$SD[n\hat{\sigma}_{BOOT}^2]$	$RMSE$
$b\hat{\sigma}^2$	200	2.1840	l_{jack}	-0.5956	1.3063	0.9399
$b\hat{\sigma}^2$	200	2.1840	l_{th}	-0.7857	1.0299	0.7660
$b\hat{\sigma}^2$	500	2.3637	l_{jack}	-0.4546	1.4208	0.9372
$b\hat{\sigma}^2$	500	2.3637	l_{th}	-0.6433	1.3452	0.9368
$\hat{a}$	200	1.4538	l_{jack}	-0.3086	0.4837	0.2256
$\hat{a}$	200	1.4538	l_{th}	-0.3005	0.4997	0.2330
$\hat{a}$	500	1.5756	l_{jack}	-0.3382	0.4785	0.2182
$\hat{a}$	500	1.5756	l_{th}	-0.3277	0.4553	0.1991

alent between (l_{jack}) and (l_{th}). Nevertheless the first two rows in table 1 point out a difference in term of RMSE because the estimator $b\hat{\sigma}^2$ is more complex than $\hat{a}$ and the estimation of l_{jack} is nonparametric with respect to the Influence function (Bühlmann and Künsch, 1999). It is important to note the bias is a well known effect due to the Bartlett spectral window estimate. The use of the theoretical Influence function has some advantages when the statistic is more complex. Besides, from a computational point of view it gives a faster method to apply the algorithm of Bühlmann and Künsch (1999) to estimate the block length l . Nevertheless, there is the problem to find which moments are defined in (3) and how to estimate the coefficients g_i . A possible field of future research is to derive a data-driven method to identify the moments in (3) and estimate the same coefficients g_i consistently.

References

- Bühlmann P. (2002) Bootstraps for time series, *Statistical Science*, 17, 52–72.
 Bühlmann P. and Künsch H.R. (1999) Block length selection in the bootstrap for time series, *Computational Statistics & Data Analysis*, 31, 295–310.
 Grahn T. (1995) A conditional least squares approach to bilinear time series estimation, *Journal of Time Series Analysis*, 57, 509–529.
 Künsch H.R. (1989) The jackknife and the bootstrap for general stationary observations, *The Annals of Statistics*, 17, 1217–1241.
 Tong H. (1990) *Non-linear Time Series. A dynamical System Approach*, Clarendon Press.