Resampling techniques for cyclostationary time series.
Long memory, weak dependence and heavy tails perspective.

Elżbieta Gajecka
PWSZ Nowy Sącz, Poland, egajecka@gmail.com

Jacek Leśkow
Institute of Mathematics, Cracow Technical University, Poland , jleskow@pk.edu.pl

Abstract

The main goal of this article is to review recent results regarding resampling techniques for nonstationary time series and stochastic processes. The type of nonstationarity under study is focused on periodic and almost periodic behavior of first and second moment characteristics of time series or stochastic processes. Such models are widely applied in the processing of telecommunication signals, diagnostics of mechanical components, climatology, finance and many others and are frequently called *cyclostationary processes*. For an exhaustive review of this research topic the reader is refereed to the paper of Gardner et al (2006) where more than 1500 research papers on the topic are classified according to their research area. Our focus will be on such time series that also have long memory but fulfill mixing or weak dependence conditions. We will show how long memory, weak dependence and heavy tails play a role in the consistency of subsampling. As a conclusion we will get that the subsampling is the most versatile procedure in the presence of heavy tails and weak dependence.

This work is supported by the grant of the National Center of Science of Poland, grant number UMO-201310/M/ST1/00096.

Keywords: periodically correlated time series; almost periodically correlated time series; weak dependence; resampling.

1. Introduction

In the analysis of nonstationary time series, one very frequently encounters the problem of having to use the asymptotic distribution for the considered estimation procedure. Usually, the asymptotic distribution, even if normal, can have very difficult parameters to estimate and this makes the practical application problematic. Therefore, a popular remedy is to apply resampling procedures like: the moving block bootstrap (MBB - see Synowiecki (2007)), the seasonal block bootstrap (SBB - see Leśkow and Synowiecki (2010)), the periodic block bootstrap (PBB), the generalized seasonal block bootstrap (GSBB - see Dudek et al (2014)), the circular block bootstrap (CBB - see Dudek (2014)) and the subsampling (see Leśkow and Synowiecki (2010) , Lenart, Leśkow and Synowiecki (2008) or Dehay, Dudek and Leśkow (2014). In this article we would like to shortly review the fundamental results obtained by the above mentioned autors and also point out significant improvements of those results in the case when the time series under study are either long memory or heavy tailed or both.

In this paper, our starting point is the time series $\{X_t, t \in \mathbb{Z}\}$ with the mean function $m(t) = EX_t$ and the autocovariance function $B(t,\tau) = Cov(X_t, X_{t+\tau})$. In the sequel, it will be assumed that $m(t)$ is periodic or almost periodic in t and $B(t,\tau)$ is periodic or almost periodic in t for each τ. Such time series will be called PC (in the periodic case) or APC (in the almost periodic case).

A time series $\{X_t\}_{t\in\mathbb{Z}}$ is called (strictly) periodically stationary (PS) with period T if, for every n, any collection of times $t_1, ..., t_n \in \mathbb{Z}$, and Borel sets $A_1, ..., A_n \subset \mathbb{R}$,

$$P_{t_1+T,...,t_n+T}(A_1, ..., A_n) = P_{t_1,...,t_n}(A_1, ..., A_n),$$

and there are no smaller values of $T > 0$ for which above equation holds.

In the statistical inference for PC, APC or PS time series it is essential to estimate the mean function $m(t)$ and the autocovariance function $B(t, \tau)$. The classical results, based on asymptotic theorems and asymptotic confidence intervals are presented, for example, in Dehay and Leśkow (1996). However, the asymptotic distribution is very difficult to apply since it depends on parameters that are hard to estimate. The usual way to circumvent this problem is to apply various resampling techniques and this is exactly the aim of this paper. In order to be able to apply resampling techniques such as MBB, CBB, PBB, SBB, GSBB and subsampling one has to show the consistency of such methods. In the context of resampling techniques, consistency means that the quantiles obtained by resampling are asymptotically identical with the quantiles from the asymptotic distributions. So, given the consistency, resampling techniques give a clear advantage: using the sample and resampling we can derive valid confidence intervals. For more details, the reader is refereed e.g. to the paper of Leśkow and Synowiecki (2010).

The consistency of resampling usually requires some asymptotic independence assumption for the time series. The most common and frequently used assumption is the α-mixing.

The Section 2 of this article will be devoted to presentation the fundamental results related to MBB and GSBB algorithms. The consistency of those techniques requires α-mixing. We will show the advantages and some disadvantages of such an assumption in the analysis of nonstationary time series. In Section 3, however, we will present a new type of asymptotic independence assumption, introduced in Doukhan and Louhichi (1999). Such an assumption provides a new perspective, especially in the analysis of nonstationary time series that are long memory and heavy tails. We will discuss possibilities of obtaining consistency of various resampling techniques under this new assumption. Last Section of this paper contains references.

Section 2. Resampling techniques for nonstationary time series.

The Moving Block Bootstrap

In this subsection the moving block bootstrap (MBB) will be described. This procedure was introduced by Liu and Singh (1992) for the strictly stationary time series. As it has turned out later (see Leśkow and Synowiecki (2010)) this procedure can be also adapted to nonstationary time series with periodic structure.

Let $(X_1, ..., X_n)$ be the observed sample from the time series and $B(j, b) = (X_j, ..., X_{j+b-1})$ be $b-$block of the data. The length of the $b-$block is $b = b_n$. Assume, without a loss of generality, that $k = n/b \in \mathbb{N}$.

The MBB Algorithm

- Let the i.i.d. random variables $i_1, i_2, ..., i_k$ come from the distribution

$$P(i_j = t) = \frac{1}{n-b+1} \; for \; t = 1, ..., n-b+1.$$

- To obtain the MBB resample

$$(X_1^*, X_2^*, ..., X_n^*)$$

  the blocks $(B(i_1, b), B(i_2, b), ..., B(i_k, b))$ are concatenated.

Notice that the MBB procedure does not require knowledge about the length of the period.

The consistency of the MBB for the PC and APC time series was obtained by Synowiecki (2007) under the moment conditions of the order of $2 + \delta$ and for α-mixing. Notice that Synowiecki in his research did not put any restriction regarding the order of convergence of the $\alpha-$mixing sequence. Such assumptions allow for the heavy tails distribution, like t Student distribution with the numbers of degrees of freedom of the order of $2+\delta$ or for the GED distribution. The long memory here is modelled via α-mixing assumption. However, the α-stable distributions (like Cauchy) can not be included in such considerations due to the moment condition.

The Generalized Seasonal Block Bootstrap

Let $(X_1, ..., X_n)$ be a sample from the periodic time series with period T. Let $b = b_n$ be the block length and $n = \omega T$, where $\omega \in N$.

The GSBB Algorithm

- We choose an positive integer block size $b < n$ such that $n = lb$, $l \in \mathbb{N}$.
- For $t = 1, b+1, ..., (l-1)b+1$ we define B_t^* as follows

$$B_t^* = (X_t^*, ..., X_{t+b-1}^*) = (X_{\tau_t}, ..., X_{\tau_{t+b-1}}),$$

  where τ_t is a discrete uniform random variable taking values in the set

$$\{t - TR_{1,n}, t - T(R_{1,n} - 1), ..., t - 2T, t - T, t, t + T, t + 2T, ..., t + T(R_{2,n} - 1), t + TR_{2,n}\}.$$

  Here $R_{1,n} = [(t-1)/T]$ and $R_{2,n} = [(n-b-t)/T]$. Random variables $\tau_1, \tau_2, ..., \tau_l$ are i.i.d. from the distribution

$$P(\tau_\omega = 1 + (\omega - 1)b + tT) = \frac{1}{\omega}, \qquad t = 0, ..., l-1.$$

  Here τ_t is the beginning of the block B_t^*, and it is restricted to be randomly chosen from a set containing only periodic shifts of t.

- Joining $l+1$ blocks $(X_{\tau_t}, X_{\tau_t+1}, ..., X_{\tau_t+b-1})$ we get the bootstrap sample $(X_1^*, X_2^*, ..., X_{(l+1)b}^*)$. The first n points $X_1^*, X_2^*, ..., X_n^*$ are retained and this implies that the bootstrap series has the same length as the original one. If n is an integer multiple of b, then the whole last block is superfluous.

Let us take $\delta > 0$, $\sup_t E|X_t|^{4+\delta} < \infty$ and $\sum_{\tau=1}^{\infty} \tau \alpha_X^{\delta/(4+\delta)}(\tau) < \infty$, where $\alpha_Y(\tau)$ is the α- mixing coefficient for the series $\{X_t\}$. If $b \to \infty$ as $n \to \infty$, but with $b = o(n)$, then the GSBB-based resampling estimator $\hat{m}^*(i)$ of $[\hat{m}(i)$ is consistent. Moreover, the GSBB-based resampling estimator $\hat{\mu}^*$ is also consistent. For the details, the reader is refereed to Dudek, Leśkow, Paparoditis and Politis (2014).

The assumptions in Dudek, Leśkow, Paparoditis and Politis (2014) allow for the heavy tails distribution, like t Student distribution with the numbers of degrees of freedom of the order of $4 + \delta$ or for the GED distribution. The long memory condition in this research requires α-mixing assumption. However, the α-stable distributions (like Cauchy) can not be included in such considerations due to the moment condition. Similar considerations can be also made for three other resampling techniques available for PC, APC and PS time series, that is for PBB, SBB and CBB. The details can be found in papers of Dudek (2014), Chan et al (2004) and Lahiri (1999).

Section 3. Long memory and weak dependence for cyclostationary time series

In the previous Section we were considering six different resampling techniques for PC, APC and PS time series. For asymptotic validity of these procedures it was necessary to assume the α-mixing condition. As it was mentioned earlier, this condition is practically unverifiable for nonstationary and non-Gaussian time series. The applied research of Cioch, Knapik and Leśkow (2013) or Wyłomańska et al (2014) provides ample evidence for nongaussianity and nonstationarity with heavy tails of considered signals.

A PC or PS time series $\{X_t\}_{t \in \mathbb{Z}}$ has a long memory if its autocovariance function $\gamma(s, h)$ for each $s \in \{1, \ldots, T\}$ satisfies the following formula

$$\sum_{0<|h|<n} \gamma(s, h) \sim C(s) n^\beta$$

where $\beta \in [0, 1)$.

Let us assume that the notation for the long memory with parameter $\beta \in [0, 1)$ will be $LM(\beta)$.

In our research, the concept of heavy tails is quite general. The models considered by our research group include : α-stable time series, GED distributions or time series that are modelled with the help of the multivariate t distribution with the number of degrees of freedom $\nu \in [2, 3]$. For details, see the recent paper of Drake, Leśkow and Medina (2015). For the sake of brevity, we will not give specific definitions that, on the other hand, are quite well known. We will only mention here that a heavy-tailed distribution is any distribution with the tails heavier than normal distribution.

In this section, we will use the concept of weak dependence, introduced by Doukhan and Louichi (1999). It is easy to see that this condition is easier to verify for non-gaussian nonstationary time series. For such time series we have the following new result.

Theorem. Assume that the nonstationary time series $\{X_t; t \in Z\}$ is PC or PS. Assume also that for each t the random variable X_t is heavy tailed, that is it is either α-stable or it has a t Student distribution with the number of degrees of freedom $\nu \in [1, 3]$ or it is a GED type random variable. Let us assume also that this time series is λ-weakly dependent with long memory parameter $\beta \in [0, 1)$. Then the subsampling estimator of the mean function $m(t)$ is consistent.

Proof. The first part of the proof of this theorem is based on the Central Limit Theorem for weakly dependent random variables provided by the research of Doukhan and Louhichi (1999). The second part of the proof is using the general subsampling consistency theorem developed by Politis et al (1999).

Section 4. Applications

The cyclostationary time series and signals models are extremely widely applied to variety of practical situatiion. Here we will provide only two of many recent applications our research group was involved with. The first example is related to observing humans and their walking patterns. The details can be found in Maiz et al (2013). Another application is related to the wearing process of the wheel bearing. The source of the data was the acceleration. The detailed description of the experiment and the method can be found in Cioch, Knapik and Leśkow (2013).

1 References

Cioch, W., Knapik, O. and Leśkow, J. (2013), *Finding a frequency signature for a cyclostationary signal with applications to wheel bearing diagnostics*, Mechanical Systems and Signal Processing, **38**, pp. 55 - 64.

Chan V., Lahiri S. N., Meeker W. (2004), *Block bootstrap estimation of the distribution of cumulative outdoor degradation*, Technometrics, 46:215-224

Dehay D., Dudek A.E., Leśkow J. (2014), *Subsampling for continuous-time almost periodically correlated processes*, Journal of Statistical Planning and Inference 150:142-158

Dehay D., Leśkow J. (1996), *Functional Limit Theory for the Spectral Covariance Estimator*, Journal of Applied Probability, Vol. 33, No. 4, p. 1077-1092.

Drake, C., Leśkow, J. and Medina, A. (2015), *Imputation of missing observations for heavy tailed cyclostationary time series*, to appear in : Cyclostationarity: Theory and Methods II. Fakher Chaari editor. Springer Series in Applied Condition Monitoring. Doukhan P., Louhichi S. (1999), *A new weak dependence condition and applications to moment inequalities*, Stoch. Proc. Appl. 84, p. 313-342

Dudek A.E. (2014), *Circular block bootstrap for coefficients of autocovariance function of almost periodically correlated time series*, Metrika, Published online: 09 August 2014

Dudek A.E., Leśkow J., Politis D., Paparoditis E. (2014), *A generalized block bootstrap for seasonal time series*, J. Time Ser. Anal., 35, p. 89-114

Gardner W,. Napolitano A., Paura L. (2006), *Cyclostationarity: half a century of research*, Signal Process., 86:639-697

Lahiri S. (1999), *Theoretical comparison of block bootstrap methods*, Ann. Statist., 27:386-404

Lenart, Ł., Leśkow, J. and Synowiecki, R. (2008), *Susampling in testing autocovariance for periodically correlated time series*, Journal of Time Series Analysis, Vol 29, Issue 6, pp. 995 - 1018.

Leśkow J., Synowiecki R. (2010), *On bootstrapping periodic random arrays with increasing period*, Metrika, 71:253-279

Liu R., Singh K. (1992), *Moving blocks jacknife and bootstrap capture weak dependence*, In LePage R. and Billard L., editors, Exploring the Limits of Bootstrap, p. 225-248, Wiley, New York.

Maiz, S., El Badaoui, M., Bonnardot, F., Dudek, A., Leśkow, J. (2013), *Deterministic/cyclostationary signal separation using bootstrap*, 11th IFAC International Workshop on Adaptation and Learning in Control and Signal Processing, Caen, France, July 3 - 5, 2013.

Politis D. N., Romano J. P., Wolf M. (1999), *Subsampling*, Springer-Verlag, New York

Synowiecki R. (2007), *Consistency and application of moving block bootstrap for non-stationary time series with periodic and almost periodic structure*, Bernoulli, 13:1151-1178

Wyłomańska, A., Obuchowski, J., Zimroz, R. and Hurd, H. (2014), *Periodic autoregressive modeling of vibration time series from planetary gearbox used in bucket wheel excavator*, in: *Cyclostationarity: Theory and Methods. Lecture Notes in Applied Condition Monitoring*, Fakher Chaari and Jacek Leśkow, editors, Springer Verlag.