Bootstrap for Maximum Likelihood Estimates of PARMA Coefficients

Anna E. Dudek, Harry Hurd and Wioletta Wójtowicz

Abstract In this chapter we use bootstrap techniques to estimate empirical distributions of parameter estimates for PAR sequences determined by maximum likelihood techniques. The parameters are not the periodic autoregression parameters, but are the coefficients in the Fourier series representing the parameters. We compare two different bootstrap techniques, IID and GSBB, applied to the residuals of the maximum likelihood estimation. The IID method seems a little better, which is not a surprise since the conditions for the GSBB are not completely satisfied. We expect these method to also work satisfactorily for full PARMA estimations, where both PMA and PAR terms are present in the model.

1 Introduction

Let $\{X(t), t \in \mathscr{Z}\}$ be a PARMA (p, q) (periodic autoregressive-moving-average) time series with the known period of the length T i.e.

$$X_t = \sum_{j=1}^{p} \phi_j(t) X_{t-j} + \sum_{k=1}^{q} \theta_k(t) \xi_{t-k} + \sigma(t) \xi_t, \tag{1}$$

where $\phi_j(t) = \phi_j(t+T), \theta_k(t) = \theta_k(t+T), \sigma(t) = \sigma(t+T)$ for all $j = 1, \ldots, p$, $k = 1, \ldots, q$ are periodic coefficients, and ξ_t is mean zero white noise with variance equal to one.

A. E. Dudek (✉) · W. Wójtowicz
AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland
e-mail: aedudek@agh.edu.pl

H. Hurd
Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260, USA

F. Chaari et al. (eds.), *Cyclostationarity: Theory and Methods*,
Lecture Notes in Mechanical Engineering, DOI: 10.1007/978-3-319-04187-2_2,

Examples of PARMA times series can be found e.g. in Hurd and Miamee (2007).

Several methods have been proposed for the estimation of the parameters in the PARMA model (1). The first seems to have been the method of Maximum Likelihood introduced by Vecchia (1985) and more recently a method using the innovations algorithm introduced by Anderson et al. (2007). We have concentrated on the maximum likelihood method applied not to the parameters themselves $\{\phi_j(t),\, j=1,\, ...,\, p;\quad \theta_k(t), k=1,\, ...,\, q;\quad \sigma(t)\}$ for $t=0, 1,\, ...,\, T-1$, but the coefficients in their Fourier transforms. Since, for the maximum likelihood method, the estimates are not expressed directly in terms of the data, computation of estimation error is not straightforward. However, sample distributions of estimates can be computed via bootstrap in a rather straightforward way.

An alternative parametrization that uses Fourier representation was introduced by Jones and Brelsford (1967) to reduce the number of parameters required to represent PARMA model.

$$\phi_j(t) = a_{j,1} + \sum_{m=1}^{\lfloor T/2 \rfloor} a_{j,2m}\cos(2\pi mt/T) + \sum_{m=1}^{\lfloor T/2-1 \rfloor} a_{j,2m+1}\sin(2\pi mt/T), \quad j=1,\, ...,\, p,$$
$$\theta_k(t) = b_{k,1} + \sum_{m=1}^{\lfloor T/2 \rfloor} b_{k,2m}\cos(2\pi mt/T) + \sum_{m=1}^{\lfloor T/2-1 \rfloor} b_{k,2m+1}\sin(2\pi mt/T), \quad k=0,\, ...,\, q,$$

where $\theta_0(t) = \sigma(t)$. The transformation is one-to-one when parameters are unrestricted but provides a simple, and sometimes physically motivated, way to reduce the number of parameters by restricting the the number of frequencies in the Fourier series of $\phi_j(t), \theta_k(t), \sigma(t)$. Maximum likelihood estimates are made of the restricted model parameters (some subset of the unrestricted $\{a_{j,m}, b_{k,m}\}$) and finally the solution can be transformed to $\{\phi_j(t), \theta_k(t)\}$.

2 Bootstrap Methods

In the sequel we propose to use two different bootstrap methods to obtain the confidence intervals for the parameters $\{a_{j,\,m}, b_{k,\,m}\}$. The main reason that bootstrap is of interest here is that estimates (at least the maximum likelihood estimates) of the parameters are made indirectly i.e. by maximizing a likelihood calculation, not by a direct expression of the data. The first considered method is based on the idea of bootstrap of residuals for ARMA time series (for more details see e.g. Lahiri 2003). The latter one is using the Generalized Seasonal Block Bootstrap (GSBB) of Dudek et al. (2014). Below we describe both techniques.

Let $X(1),\, \ldots,\, X(n)$ be a sample from PARMA (p,q) time series. To apply any of bootstrap algorithms first the estimates of $\{a_{j,\,m}, b_{k,\,m}\}$ coefficients need to be calculated. As a result we get the residuals $\widehat{\varepsilon}_i$ for $i=1,\, \ldots,\, N$, where $N = n - (p+q)$.

Since for ARMA time series to generate the valid approximation of the asymptotic distribution the residuals need to be centered, we expect that PARMA model also requires this condition. Thus, we define the centered residuals by

$$\tilde{\varepsilon}_i = \widehat{\varepsilon}_i - \overline{\varepsilon}_N, \tag{2}$$

where $\overline{\varepsilon}_N = (N)^{-1} \sum_{i=1}^{N} \widehat{\varepsilon}_i$.

Method 1-IID bootstrap:

1. For $i = 1, \ldots, n$ let
$$\varepsilon_i^{*(1)} = \tilde{\varepsilon}_{k_i},$$
where k_i is iid from a discrete uniform distribution
$$P(k_i = s) = \frac{1}{N} \quad \text{for} \quad s = 1, \ldots, N.$$
2. Joining selected residuals we get the bootstrap sample $\left(\varepsilon_1^{*(1)}, \ldots, \varepsilon_n^{*(1)}\right)$.

Method 1 assumes the the residuals $\tilde{\varepsilon}_i$ $(i = 1, \ldots, N)$ are at least approximately independent, so selecting each of them separately we do not destroy the dependence structure in the sample. But in practical cases this condition may not hold. For example one may choose to fit the model of lower rank that the true one. Then, the residuals are no longer independent. Moreover, they may reflect the periodic structure of the original data.

GSBB is the new block bootstrap technique for periodic data. It is the generalization of two known block bootstrap methods i.e. the Seasonal Block Bootstrap of Politis (Politis 2001) and the Periodic Block Bootstrap of Chan et al. (2004). Dudek et al. (2014) used it for the overall mean and the seasonal means of the periodically correlated (PC) time series. Moreover, they showed GSBB consistency for triangular row-wice periodically correlated arrays with growing period. The wide spectrum of possible GSBB applications encouraged us to apply it in the considered problem.

To simplify the presentation of GSBB algorithm we assume that the sample size n is an integer multiple of the block length b $(n = bl)$ and also is an integer multiple of of the period length T $(n = wT)$. Each of these conditions can be easily omitted (for more details see Dudek et al. 2014). Moreover, we present the circular version of GSBB i.e. we treat the sample as wrapped on the circle. Whenever the index t of any chosen observation is greater than N we take $t - N$ instead.

Method 2-GSBB:

1. Choose a (positive) integer block size $b(<N)$.
2. For $t = 1, b+1, 2b+1, \ldots, (l-1)b+1$, let

$$\left(\varepsilon_t^{*(2)}, \varepsilon_{t+1}^{*(2)}, \ldots, \varepsilon_{t+b-1}^{*(2)}\right) = \left(\widehat{\varepsilon}_{k_t}, \widehat{\varepsilon}_{k_t+1}, \ldots, \widehat{\varepsilon}_{k_t+b-1}\right),$$

where k_t is iid from a discrete uniform distribution

$$P\left(k_t = t + vT\right) = \frac{1}{w} \quad \text{for} \quad v = 0, 1, \ldots, w - 1.$$

Since we consider the circular version of GSBB, when $t + vd > N$ we take the shifted observations $t + vT - N$.

3. Join the l blocks $(\tilde{\varepsilon}_{k_t}, \tilde{\varepsilon}_{k_t+1}, \ldots, \tilde{\varepsilon}_{k_t+b-1})$ to get the bootstrap sample.

Finally, the bootstrap version $\left(X_1^{*(j)}, \ldots, X_n^{*(j)}\right)$ of the original sample $(X_1, \ldots, X_n)$ is obtained using the estimates of $\{\phi_j(t), \theta_k(t)\}$ and the bootstrap error variables $\left(\varepsilon_1^{*(j)}, \ldots, \varepsilon_n^{*(j)}\right)$. The superscript j denotes the chosen bootstrap method. It is equal to 1 for IID bootstrap and 2 for GSBB.

In the next section we present some simulation study results in which we construct the bootstrap pointwise equal-tailed confidence interval for coefficients $\{a_{j,m}, b_{k,m}\}$. The actual coverage probabilities (ACPs) are calculated to compare the performance of both considered bootstrap algorithms. Although, we do not have any theoretical results confirming the consistency of the proposed bootstrap methods, the preliminary simulation results indicate the validity of our procedures.

3 Simulation Study

Our aim is to check the performance of the proposed bootstrap algorithms in the problem od estimating confidence intervals for PARMA model coefficients. In this section we consider a few examples of PARMA time series and calculate the bootstrap equal-tailed pointwise confidence intervals for the coefficients $\{a_{j,m}, b_{k,m}\}$ for $j = 1, \ldots, p$, $k = 0, \ldots, q$ and $m = 1, \ldots, T$. In our study we use procedures first implemented by Hurd (2007) and now available as R package 'perARMA' (Comprehensive R Archive Network reference Dudek et al. 2013).

To reduce the number of parameters that needs to be estimated and decrease the time of computation we restricted our study only to PAR time series. The following examples are considered:

PAR2: the nonzero coefficient are $a_{1,1} = 0.8$, $a_{1,2} = 0.3$, $a_{2,1} = -0.4$ and $b_{0,1} = 1$;

PAR1: the nonzero coefficient are $a_{1,1} = 0.8$, $a_{1,2} = 0.3$ and $b_{0,1} = b_{0,2} = -0.5$.

Note that **PAR2** model has the constant $\sigma(t)$ function (equal to 1) in contrary to the **PAR1** case, where $\sigma(t)$ is periodic. The names **PAR2** and **PAR1** indicate that these are PAR(1) and PAR(2) time series, respectively. This particular choice was caused by the fact that we wanted to restrict the number of parameters and simultaneously investigate the influence of function $\sigma(t)$ on our results.

Unfortunately, $\sigma(t)$ is not the only important factor. Much bigger impact can have the choice of the model fitted to the data. Each practitioner will decide to take

Table 1 Actual coverage probabilities for simulated **PAR2** series

Method	n	ACP									
		E1							E2		
		$a_{1,1}$ (%)	$a_{1,2}$ (%)	$a_{1,3}$ (%)	$a_{2,1}$ (%)	$a_{2,2}$ (%)	$b_{0,1}$ (%)	$b_{0,2}$ (%)	$a_{1,1}$ (%)	$a_{2,1}$ (%)	$b_{0,1}$ (%)
IID	120	92.2	92.2	95.2	92.0	93.4	92.2	93.2	89.6	93.4	89.4
Bootstrap	240	94.4	94.0	93.6	93.6	94.0	94.4	91.8	87.2	92.0	82.6
GSBB	120	89.0	91.2	89.4	87.0	91.0	86.8	87.0	89.2	91.8	84.8
	240	87.8	89.4	89.6	88.8	89.8	91.2	88.2	89.6	91.6	84.6

Columns 1–7 refer to **E1** case and 8–10 to **E2**. Rows 1–2 and 3–4 contain results for IID bootstrap and GSBB, respectively. For both methods ACPs for two sample sizes $n = 120$ and $n = 240$ are presented

the model of lower order if according to some criteria it is comparable to more complicated one. Having this in mind we decided consider the following cases

E1: **PAR2** estimating 7 coefficients i.e. $a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}$ and $b_{0,1}, b_{0,2}$;
E2: **PAR2** estimating 3 coefficients i.e. $a_{1,1}, a_{1,2}$ and $b_{0,1}$;
E3: **PAR1** estimating 6 coefficients i.e. $a_{1,1}, a_{1,2}, a_{1,3}$ and $b_{0,1}, b_{0,2}, b_{0,3}$;
E4: **PAR1** estimating 2 coefficients i.e. $a_{1,1}$ and $b_{0,1}$.

As a result, in **E1** and **E3** cases we estimate more coefficient than are nonzero in reality, while in **E2** and **E4** we always have one less coefficient of each type.

To simulate **E1-E4** 'makeparma' procedure provided by 'perARMA' package was used. This function enables to construct a PARMA type sequence of required length for inputed matrices of coefficients. Two different sample lengths n were taken 120 and 240. As presented approach is based on Fourier representation of model coefficients, we use also 'ab2phth' and 'phth2ab' procedures that enable to transform matrices of coefficients to their Fourier representation and conversely. For each simulated series we fit PAR model using 'parmaf' procedure. The function returns estimates of parameters $\{a_{j,m}, b_{k,m}\}$ as well as series of residuals of fitted model. Next for residuals we apply one of two proposed bootstrap method: IID bootstrap or GSBB. The number of generated bootstrap samples B was 300. In a case of GSBB method we also need to comment the choice of block length b. Since so far, there is no method of optimal block length choice we decided to take $b = \lfloor \sqrt[3]{n} \rfloor$ and $b = T$. The period length T is equal to 12. Taking $b = T$ we wanted to check how the performance of GSBB changes when the longer block is taken. Moreover, $b = T$ is a case when GSBB reduce to SBB. Since the results in both cases were comparable in the sequel we only discuss $b = \lfloor \sqrt[3]{n} \rfloor$ case. Finally, to calculate the bootstrap equal-tailed pointwise confidence intervals bootstrap version of coefficients $\{a_{j,m}, b_{k,m}\}$ were calculated (using 'makeparma' function). The 95 % confidence level was taken. The whole procedure was repeated 500 times and the ACPs were calculated. The results for **E1-E2** are presented in Tables 1 and 2 and for **E3-E4** in Tables 3 and 4.

Table 2 Average lengths of confidence intervals for **PAR2** model

Method	n	Average length of CI									
		E1							E2		
		$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	$a_{2,1}$	$a_{2,2}$	$b_{0,1}$	$b_{0,2}$	$a_{1,1}$	$a_{2,1}$	$b_{0,1}$
IID bootstrap	120	0.3323	0.4693	0.3862	0.3247	0.4796	0.2403	0.3552	0.3332	0.3282	0.2531
	240	0.2323	0.3280	0.2667	0.2314	0.3330	0.1734	0.2498	0.2349	0.2328	0.1828
GSBB	120	0.3127	0.4472	0.3656	0.3179	0.4627	0.2289	0.3403	0.3214	0.3106	0.2515
	240	0.2218	0.3187	0.2566	0.2265	0.3277	0.1658	0.2479	0.2389	0.2501	0.1863

Columns of values 1–7 refer to **E1** case and 8–10 to **E2**. Rows 1–2 and 3–4 contain results for IID bootstrap and GSBB, respectively. For both methods ACPs for two sample sizes $n = 120$ and $n = 240$ are presented

Table 3 Actual coverage probabilities for simulated **PAR2** series

Method	n	ACP							
		E1						E2	
		$a_{1,1}$ (%)	$a_{1,2}$ (%)	$a_{1,3}$ (%)	$b_{0,1}$ (%)	$b_{0,2}$ (%)	$b_{0,3}$ (%)	$a_{1,1}$ (%)	$b_{0,1}$ (%)
IID bootstrap	120	88.6	90.2	92.8	59.4	70.4	94.8	59.0	24.0
	240	90.0	90.4	90.6	78.6	70.4	99.0	36.2	5.8
GSBB	120	87.0	88.0	88.0	72.4	80.8	84.4	58.6	21.0
	240	86.9	87.3	88.3	72.7	82.6	88.1	31.2	4.4

Columns 1–6 refer to **E3** case and 7–8 to **E4**. Rows 1–2 and 3–4 contain results for IID bootstrap and GSBB, respectively. For both methods ACPs for two sample sizes $n = 120$ and $n = 240$ are presented

Table 4 Average lengths of confidence intervals for **PAR2** model

Method	n	ACP							
		E3						E4	
		$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	$b_{0,1}$	$b_{0,2}$	$b_{0,3}$	$a_{1,1}$	$b_{0,1}$
IID bootstrap	120	0.3343	0.3308	0.4965	0.3186	0.4782	0.3407	0.1973	0.2278
	240	0.2076	0.2067	0.3181	0.2010	0.2901	0.1863	0.1334	0.1648
GSBB	120	0.3276	0.3358	0.4772	0.2747	0.4141	0.2369	0.1852	0.2056
	240	0.3248	0.3287	0.4680	0.2735	0.4116	0.2411	0.1210	0.1589

Columns of values 1–6 refer to **E3** case and 7–8 to **E4**. Rows 1–2 and 3–4 contain results for IID bootstrap and GSBB, respectively. For both methods ACPs for two sample sizes $n = 120$ and $n = 240$ are presented

When the number of estimated parameters is big enough i.e. in **E1** case IDD bootstrap definitely outperforms GSBB. In fact it is exactly what one may expect as in these examples the residuals are approximately independent and IID method is the most appropriate. The ACPs for all coefficients are about 2–3 % lower than the nominal confidence level independently on the sample size n. For GSBB the corresponding values are even 6 % lower. Similar conclusions can be taken in **E3** case for a type parameters, although all ACPs are definitely lower than for **E1**. Surprisingly, IID bootstrap is not working well for b type coefficients.

For $b_{0,1}$ and $b_{0,2}$ the ACPs are about 35 % lower than the nominal coverage probability for $n = 120$ and about 15–25 % for $n = 240$. Let us recall that $b_{0,1}$ and $b_{0,2}$ were the nonzero coefficients. Additionally, for $b_{0,3}$ which in reality is equal to zero IDD bootstrap seems to produce too wide confidence intervals. For $n = 120$ the ACP is almost perfectly equal to 95 % but for $n = 240$ it is close to 1. On the other hand, GSBB provides constantly too low ACPs independently on n, but they are higher compering to IID bootstrap and seem to converge slowly to the nominal coverage probability.

Finally, **E2** and **E4** provide the evidence how destructive influence of the too small set of estimated parameters can be. In fact the performance of the both bootstrap techniques is good for **E2** and the differences between those methods are small. **E2** is a case, where the shocks are constant. The ACPs are similar comparing to **E1** for coefficient of a type and decrease about 5 % for $b_{0,1}$. The noticeable problems appear in **E4** example. For both methods and $n = 120$ the ACPs are very low to became extremely small for $n = 240$. This may indicate that bootstrap is inconsistent in this problem. Let us recall that estimating only $b_{0,1}$ we treat the rest of bs as zeros, which means that $\sigma(t)$ is a constat function. As a result the residuals are definitely dependent.

Although we are aware that we did not provide any theoretical confirmation of validity of the bootstrap methods in the considered problem, the simulation study results seem to be very encouraging. They indicate the consistency of bootstrap. Moreover, probably the practitioner will not be able to use the universal method independently on the PARMA series structure. IDD bootstrap seems to be the best choice when the shocks are constant, while block bootstrap is more appropriate in the opposite case. Additionally, one needs to be extremely careful choosing the size of parameters set that need to be estimated. Despite the longer time of computation, the larger set should be taken to avoid the bootstrap inconsistency.

Acknowledgments Research of Anna Dudek was partially supported by the Polish Ministry of Science and Higher Education and AGH local grant.

References

Anderson PL, Tesfaye YG, Meerschaert MM (2007) Fourier-PARMA models and their application to river flows. J Hydrol Eng 12(5):462–472

Chan V, Lahiri SN, Meeker WQ (2004) Block bootstrap estimation of the distribution of cumulative outdoor degradation. Technometrics 46:215–224

Dudek AE, Hurd H, Wójtowicz W (2013) PerARMA: pacakge for periodic time series analysis, R package version 1.5. http://cran.r-project.org/web/packages/perARMA/

Dudek AE, Leśkow J, Politis DN, Paparoditis E (2014) A generalized block bootstrap for seasonal time series, article first published online: 27 NOV 2013, doi:10.1002/jtsa.12053

Hurd HL, Miamee AG (2007) Periodically correlated random sequences: spectral. Theory and practice. Wiley, New York

Jones R, Brelsford W (1967) Time series with periodic structure. Biometrika 54:403–408

Lahiri SN (2003) Resampling methods for dependent data. Springer, New York

Politis DN (2001) Resampling time series with seasonal components, in Frontiers in data mining and bioinformatics. In: Proceedings of the 33rd symposium on the interface of computing science and statistics. Orange County, California, 13–17 June, pp 619–621

Vecchia AV (1985) Maximum likelihood estimation for periodic autoregressive moving average models. Technometrics 27:375–384

Springer

http://www.springer.com/978-3-319-04186-5

Cyclostationarity: Theory and Methods
Chaari, F.; Leskow, J.; Napolitano, A.; Sanchez-Ramirez, A. (Eds.)
2014, VIII, 186 p. 66 illus., 28 illus. in color., Softcover
ISBN: 978-3-319-04186-5