TEMA Tend. Mat. Apl. Comput., **9**, No. 2 (2008), 175-184.

Evaluating some Yule-Walker Methods with the Maximum-Likelihood Estimator for the Spectral ARMA Model

M.I.S. BEZERRA[1], Departamento de Matemática, Estatística e Computação, FCT, UNESP, 19060-900, Presidente Prudente, SP, Brasil.

Y. IANO[2], Departamento de Comunicações, FEEC, UNICAMP, 13083-852, Campinas, SP, Brasil.

M.H. TARUMOTO[3], Departamento de Matemática, Estatística e Computação, FCT, UNESP, 19060-900, Presidente Prudente, SP, Brasil.

Abstract. The aim of this work is to compare some ARMA spectral separated estimation methods based on the modified Yule-Walker equation and least squares method with the Maximum-Likelihood estimator, using the convergence curve of the relative mean error (RME), generated by Monte Carlo simulation.

Keywords. ARMA process, parameter estimation, Monte Carlo simulation.

1. Introduction

The autoregressive moving average (ARMA) spectral estimation is considered a topic of interest in several areas of statistics, engineering, econometrics, etc [3], [9], [13]. Different methods have been studied, among which, comprising two focuses, we can point out: 1) the optimal methods, such as the maximum likelihood estimates, which are computational difficult to treat, [3], [9], [13]; and 2) the sub-optimal methods based on the Yule-Walker equations, which employ linear equations in the parameters estimation [9], [10], [13], [5].

The methods of estimation for the autoregressive moving average (ARMA) process have been widely studied, mainly concerning the separate estimation of parameters, where Yule-Walker equations are used to estimate the AR process parameters, and following this, the MA process parameters are estimated through the Durbin method [9], [13].

In this article a new method is proposed of ARMA spectral estimation based on the least squares method of the modified Yule-Walker equations [9], [13], [8]. The

[1]manoel@fct.unesp.br.
[2]yuzo@decom.fee.unicamp.br.
[3]tarumoto@fct.unesp.br.

main idea of this method is to perform an AR filtering in the signal by using the MA estimates obtained through the Durbin method. This filtering generates a new estimate of the signal, which will be used to estimate the parameters.

Initially, in section 2, the ARMA model and its main features have been presented. Section 3 is devoted to the separate estimation methods using Yule-Walker equations. Section 4 shows the development of the method proposed, having as a basis the methods described in section 3. Section 5 shows applications using Monte Carlo simulations for 2 numerical examples.

The precision of the method proposed was analyzed in relation to other separate estimation methods which use the Yule-Walker equations and its convergence through the parametric bootstrap method, is presented in [1]. In order to compare the accuracy of the methods, the relative error in relation to the true values of parameters, and the average variation coefficient of the estimates have been used.

2. The ARMA Spectral Model

Being x_n an ARMA stationary random process of order (p, q), defined by the following equation of differences

$$\begin{aligned} \sum_{i=0}^{p} a_i x_{n-i} &= \sum_{l=0}^{q} b_l \varepsilon_{n-l}, \\ \text{assuming } a_0 &= 1, \;\; 1 \leqq n \leqq N \end{aligned} \tag{2.1}$$

where $a_i, i = 0, 1, ..., p$; $b_l, l = 0, 1, ..., q$, are the parameters of the process and ε_n is a presumed white noise with zero mean and variance σ^2. Thus, the parameters vector to be estimated are defined as:

$$\theta = \left[a_1, ..., a_p, b_0, b_1, ..., b_q, \sigma^2\right]. \tag{2.2}$$

From the process x_n of the equation (1) the autocorrelation function of x_n is defined as:

$$r_k = E\left[x_n x_{n-k}\right]. \tag{2.3}$$

From the autocorrelation function r_k, the power spectral density can be defined

$$S\left(f\right) = \sum_{k=-\infty}^{\infty} r_k \exp\left(-j2\pi k f\right). \tag{2.4}$$

The function $S(f)$ is the transformed of Fourier of r_k, and presumably non-negative and periodical in frequency with period 1 Hz, and assumed to be band limited to ± 0.5 Hz [3],[9]. By using the results described above we can also define more specifically the power spectral density (PSD) of the ARMA process, as [10]

$$S(f) = \frac{\sigma^2 \left| \sum\limits_{k=0}^{q} b_k \exp(-j2\pi f k) \right|^2}{\left| 1 + \sum\limits_{k=1}^{p} a_k \exp(-j2\pi f k) \right|^2}. \tag{2.5}$$

The problem of ARMA spectral estimation consists initially of selecting the suitable model so that we can estimate the parameters vector of the process and then replace the estimates value in the spectral density function. This parametrization of function $S(f)$ is obtained by using the parameter vector θ.

3. Estimation Methods via Yule-Walker Equations

The maximum-likelihood function of the ARMA model is non-linear in the MA parameters and requires a highly interactive computational effort, which hinders the attainment of joint and efficient estimates of AR and MA parameters, and it is many times unfeasible in practice. However, efficient estimates of these parameters can be obtained separately by using linear operations in the AR part, and by applying the Durbin method in the MA part [9], [13].

Methods which use this procedure are known as separate estimation methods. Besides, the methods to be presented here have as their main characteristic the use of Yule-Walker equations [9], [5], [8].

In separate spectral estimation for the ARMA model there are several methods which use Yule-Walker equations, as previously mentioned in introduction. Among these methods the best known are the ones using modified Yule-Walker equations (MYWE) and the method of least squares modified Yule-Walker equations (LSMYW). The latter uses more than p linear equations in the parameters estimation. In both methods errors can be weighted in order to stabilize the variance. Moreover, these methods provide similar results for ARMA spectral estimation.

3.1. Modified Yule-Walker Equations Method

This method is considered to be the simplest for the separate estimation of the ARMA model. Initially we know that the AR parameters of the ARMA process are related to the autocorrelation function $r(k)$, as a set of linear equations, that is,

$$r_k + \sum_{i=1}^{p} a_i r_{k-i} = 0, \ k \geqq q+1. \tag{3.1}$$

The equations above are known as modified Yule-Walker equations (MYWE). The autocorrelation matrix of (3.1) is Toeplitz, but it does not ensure its non-singularity. Through these equations one can find the parameters of AR and MA models separately. For the estimation of the parameters of the MA process the Durbin method can be applied [9].

In many works on ARMA spectral estimation it has been seen that, in general, the MYWE method does not produce good quality estimates for the parameters of the AR process of the ARMA model. In such cases the models have poles far from and zeros near the unit circle (UC), however their angular positions are near each other.

3.2. Least Squares Yule-Walker Method

This method is an attempt to reduce the variance of the MYWE estimator and improve the quality of its estimates. It is known that the autocorrelation function of the ARMA process is defined as in expression (3.1) [8].

From the MYWE we will describe the least squares method of the modified Yule-Walker equations (LSYW) initially for the estimation of the parameters of the AR process, given that the expanded equations of Yule-Walker for the ARMA process can be represented as [9], [10], [13]

$$\mathbf{r} = -\,\mathbf{Ra} \tag{3.2}$$

where $\mathbf{R}$ is the expanded autocorrelation matrix $(M-q)\times p$, which is given by

$$\mathbf{R} = \begin{bmatrix} r_q & r_{q-1} & \cdots & r_{q-p+1} \\ r_{q+1} & r_q & \cdots & r_{q-p+2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{M-1} & r_{M-2} & \cdots & r_{M-p} \end{bmatrix}$$

and $\mathbf{r}$ is the autocorrelations vector $(M-q)\times\ 1$ given by

$$\mathbf{r} = [r_{q+1},\ r_{q+2}, \ldots,\ r_M]^T.$$

The consistent estimates of AR parameters vector $\mathbf{a} = [a_1, ..., a_p]^T$ can be obtained by (3.2). As $M-q>p$, there are more equations than unknowns. In order to estimate the autocorrelations an error $\mathbf{e}$ of size $(M-q)\times 1$ should be introduced so that

$$\widehat{\mathbf{r}} = -\widehat{\mathbf{R}}\mathbf{a} + \mathbf{e}$$

where $\widehat{\mathbf{r}}$ and $\widehat{\mathbf{R}}$ correspond to $\mathbf{r}$ and $\mathbf{R}$ estimators, respectively.

It is convenient to use the unbiased autocorrelations r_k to estimate $\mathbf{r}$ and $\mathbf{R}$, so that the bias of approximation error is null. Thus, the method of weighted least squares can be applied to find the vector which minimizes the sum of squares of weighted errors, that is, [8]

$$\rho = \mathbf{e}^H\mathbf{We} \tag{3.3}$$

where $w(k)$ is a decreasing weight positive sequence and $\mathbf{W}$ is a dimension matrix $(M-q)\times(M-q)$.

Substituting the equation of the errors

$$\mathbf{e} = \widehat{\mathbf{r}} + \widehat{\mathbf{R}}\mathbf{a} \tag{3.4}$$

in the equation (3.3) one has

$$\rho = \left(\widehat{\mathbf{r}}+\widehat{\mathbf{R}}\mathbf{a}\right)^H \mathbf{W} \left(\widehat{\mathbf{r}}+\widehat{\mathbf{R}}\mathbf{a}\right). \tag{3.5}$$

Making the products above in the equation one has the following results

$$\rho = \mathbf{a}^H\widehat{\mathbf{R}}^H\mathbf{W}\widehat{\mathbf{R}}\mathbf{a}+\widehat{\mathbf{r}}^H\mathbf{W}\mathbf{R}^H\mathbf{a} + \mathbf{a}^H\widehat{\mathbf{R}}^H\mathbf{W}\widehat{\mathbf{r}} + \widehat{\mathbf{r}}^H\mathbf{W}\widehat{\mathbf{r}}. \tag{3.6}$$

Taking the partial derivatives $\frac{\partial\rho}{\partial\mathbf{a}} = 0$ in expression (3.6),

$$\frac{\partial\rho}{\partial\mathbf{a}} = \mathbf{2}\widehat{\mathbf{R}}^H\mathbf{W}\widehat{\mathbf{R}}\mathbf{a} + \mathbf{2}\widehat{\mathbf{R}}^H\mathbf{W}\widehat{\mathbf{r}} \tag{3.7}$$

one can obtain the $\widehat{\mathbf{a}}$ estimator of the AR process, which is given by

$$\widehat{\mathbf{a}} = -\left(\widehat{\mathbf{R}}^H\mathbf{W}\widehat{\mathbf{R}}\right)^{-1}\widehat{\mathbf{R}}^H\mathbf{W}\widehat{\mathbf{r}}. \tag{3.8}$$

The matrix $\widehat{\mathbf{R}}^H\mathbf{W}\widehat{\mathbf{R}}$ is usually positive definite, Hermitian and the $\widehat{\mathbf{a}}$ estimator may not be a minimum-phase estimator [9], [10]. The problem is to choose the suitable kind of weights to correct the effect of variance of the autocorrelations estimates. A suggestion could be using windows $(N-k)^3$ and $(N-k)^4$. It has been observed that the weight reduces the compromise when choosing the number of equations, once it automatically reduces the importance of participation of errors e_k as k increases. If matrix $\mathbf{W}$ is equal to the identity (unit weights), the AR estimator will be in the form

$$\widehat{\mathbf{a}}= -\left(\widehat{\mathbf{R}}^H\widehat{\mathbf{R}}\right)^{-1}\widehat{\mathbf{R}}^H\widehat{\mathbf{r}}. \tag{3.9}$$

The estimation of AR parameters of the ARMA model through LSYW method can be interpreted as an application of the covariance method considering the set of values $\{\widehat{r}_{q-p+1}, \widehat{r}_{q-p+2}, ..., \widehat{r}_M\}$ [9], [10]. For the estimation of the parameters of the MA process the Durbin method can be applied again.

The performance of the LSYW is usually better than the modified Yule-Walker equations method (MYWE), where $M = p + q$ (number of equations), using only p equations, but the results will also depend on the kind of spectrum to be estimated [10], [5].

4. Proposed Method using AR Filter in MA Estimates

The novelty of the method proposed for separate estimation consists of making a new AR filtering of the signal x_n by using the estimates of the MA parameters obtained with the Durbin method, thus determining a new signal $\widehat{x}_n$ as shown in Figure 1 and according to the expression below:

$$\widehat{x}_n = \widehat{b}_0 x_n + \widehat{b}_1 x_{n-1} + ... + \widehat{b}_q x_{n-q}. \tag{4.1}$$

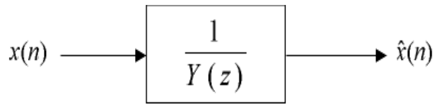

Figure 1: Filtering of the signal through an AR filter so as to obtain a new estimate

In Figure 1 the $Y(z) = \widehat{b}_0 + \widehat{b}_1 z^{-1} + ... + \widehat{b}_q z^{-q}$ is the system transfer function. This procedure is like repeating the second step of the Durbin method. This is the key idea of the separate estimation method [9], [8].

It is known that an AR(∞) process (large order) can be a good approach for the MA process [3], [9], [10], [13]. In this method a new signal estimate $\widehat{x}(n)$ will be used to obtain a new AR estimate through the LSYW method in order to, then, determine a new MA estimate through the Durbin method. The key idea of the method proposed is, from the AR filtering, to obtain a new signal, which will provide better estimates when the parameters of the ARMA process are reestimated. This method has a computational cost a little grater than the other methods because it uses an extra procedure for calculations.

5. Numerical Examples

In this section we will describe the Monte Carlo simulations [7], [11], [14], in order to obtain the relative average error of the estimates, and observe of the convergence the methods with relation each of the ARMA processes, with different characteristics. The models are: 1) ARMA(4,2) [12] and the 2) ARMA(5,5) [6]. The simulation and the programs were implemented in MATLAB. The methods used are:

1. Modified Yule-Walker Equations (**MYWE**);
2. Least Squares Modified Yule-Walker Equations (**LSYW**) (**W** is equal to the identity);
3. Least Squares Modified Yule-Walker Equations with AR filtering (**LSYWp** - Proposed Method) (**W** is equal to the identity);
4. Maximum-Likelihood Estimator (**MLE**).

In order to evaluate the performance of the methods the random data have been generated from the same seed and that change it with the sequences, $i = 1, ..., B$, [2] and B is the replications number in each simulations. We conduced this using two size $N = 256$ and $N = 1024$. In the first, we use $B = 1000, ..., 20000$, and in the second, this numbers was $B = 1000, ..., 10000$. The random numbers, was generated from a Gaussian distribution, that is white noise, with mean equal zero and variance equal one, then generated a signal of the process ARMA.

The modified covariance was criterion in the Durbin method [9], and the large AR process order was choose using the criterion describes by [4]. For the number of Yule-Walker equations, we considered two cases in each example. In the first example was used ARMA(4,2) model with $M = 10, L = 13$ ($N = 256$) and $M = 10,\ L = 506$ ($N = 1024$) and in the second example we used ARMA(5,5) model $M = 20,\ L = 90$ ($N = 256$) and $M = 20,\ L = 325$ ($N = 1024$).

Comparisons among the methods have been made by using the relative error (RE) of the estimates mean in relation to the true values of the parameters of the ARMA theoretical models. The statistics relative error (RE) and relative mean error (RME) are defined as follows:

$$RE_j = \sqrt{\frac{\sum_{i=1}^{m}\left(\widehat{\theta}_i^{(j)} - \theta_i\right)^2}{\sum_{i=1}^{n}\theta_i^2}} \text{ and } RME = \frac{1}{B}\sum_{j=1}^{B} RE_j. \tag{5.1}$$

Where θ_i is the i-th value of the true parameter; $\widehat{\theta}_i^{(j)}$ is the i-th estimates of the j-th estimate of the B bootstrap repetitions, adopting $m = p$ for the AR estimates and $m = q$ for the MA estimates. The results of the Monte Carlo simulations are presented in graphics which show the RME of the estimates of the parameters. The simulations and programs have been implemented in MATLAB.

The ARMA(4,2) model presented in [12] is given by

$$x_n - 1.6402x_{n-1} + 2.2044x_{n-2} - 1.4808x_{n-3} + 0.8145x_{n-4} =$$
$$= \varepsilon_n + 1.5857\varepsilon_{n-1} + 0.9604\varepsilon_{n-2}.$$

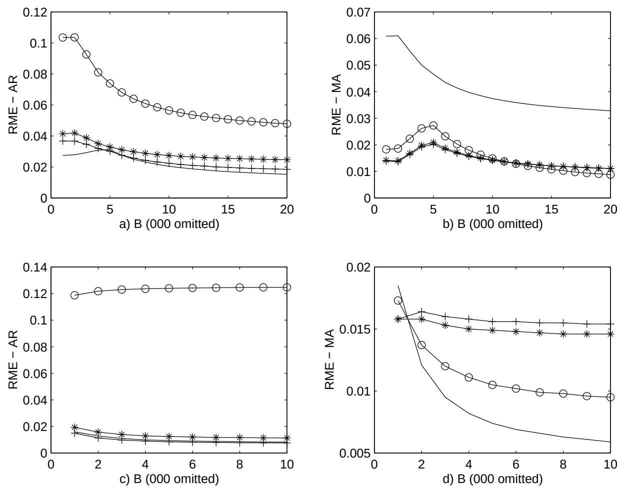

Figure 2: Relative Mean Error of the Model ARMA(4,2) - MYWE o-o; LSYW *-*; LSYWp +-+; MLE –; a) and b) N=256; c) and d) N=1024.

The ARMA(5,5) model presented in [6] is given by

$$x_n - 0.8713x_{n-1} - 1.539x_{n-2} + 1.371x_{n-3} + 0.6451x_{n-4} - 0.5827x_{n-5} =$$
$$= \varepsilon_n - 1.051\varepsilon_{n-1} + 0.0718\varepsilon_{n-2} + 0.05164\varepsilon_{n-3} + 0.5322\varepsilon_{n-4} - 0.5735\varepsilon_{n-5}.$$

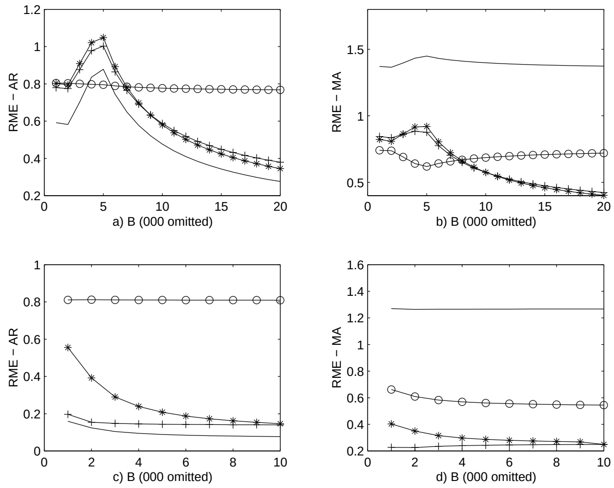

Figure 3: Relative Mean Error of the Model ARMA(5,5) - MYWE o-o; LSYW *-*; LSYWp +-+; MLE –; a) and b) N=256; c) and d) N=1024.

Table 1: Medians and p-values (probability error) - ARMA(4,2)

	N=256				N=1024			
Method	AR(4)		MA(2)		AR(4)		MA(2)	
	p-value	median	p-value	median	p-value	median	p-value	median
MYWE	$\leq$0.001	0.0359	0.153	0.0014	0.006	0.1151	0.008	-0.0049
LSYW	$\leq$0.001	0.0051	0.004	0.0002	0.006	0.0040	0.009	-7.5E-4
MLE	$\leq$0.001	-0.0024	$\leq$0.001	0.0238	0.006	0.0009	0.008	-0.0079

The figure 2 present the RME of the example 1 with respect the AR and MA parts. The figure 3 present the RME of the example 2 with respect the AR and MA parts. In this figures we observe that RME estimates are different among some methods, but we do not know if they are significant. Therefore the statistic tests are necessary.

The RME estimates are analyzed through the Wilcoxon statistical test in order to know if there is a significative difference among the RME curves. The comparison were based in the hypotheses: H_0:The RME of the method A is equal of the RME of the proposed method, the null hypotheses (difference =0) versus H_a: The RME of the method A is different of the proposed method (difference $\neq$0). This results we can observed in Table 1 (ARMA(4,2)) and in Table 2 (ARMA(5,5)).

In the table 1, in order to model ARMA(4,2), it can be observed that: the proposed method had curves of convergence better than the methods MYWE and LSYW (p-value $\leq$ 0.001), in the estimates AR, while in the estimates MA, was shown better than the methods LSYW (p-value = 0.004) and MLE (p-value $\leq$ 0.001), and equal MYWE, when N = 256. To N = 1024, the proposed method introduced curves of convergence better regarding the all of the three studied meth-

Table 2: Medians and p-values (probability error) - ARMA(5,5)

	N=256				N=1024			
Method	AR(5)		MA(5)		AR(5)		MA(5)	
	p-value	median	p-value	median	p-value	median	p-value	median
MYWE	0.004	0.1650	0.076	0.0758	0.006	0.6657	0.006	0.3224
LSYW	0.681	-0.0025	0.467	-0.0043	0.006	0.0772	0.009	0.0498
MLE	0.185	0,0684	≤0.001	0.7800	0.006	-0.0531	0.006	1.024

ods (p-value = 0.006), for estimates AR. Already for the estimates MA, the other methods were shown better.

In the table 2, regarding the model ARMA(5,5), it can be notice that the convergence curve of the proposed method was better than MYWE (p-value = 0.004), and equal to the other two methods, in the estimates AR. In the estimates MA, it was better than MLE and equal to the methods MYWE and LSYW (p-value = 0.006), when $N = 256$. To $N = 1024$, the proposed method, also showed through the tests, that yours curves of convergence is better than the methods: MYWE and LSYW, for the estimates AR; and also better than all of the three methods for the estimates MA (p-value = $0.006, 0.009$ and 0.006, respectively).

6. Conclusion

In this article it have been presented the convergence of some methods of separate estimation and the method of Maximum-Likelihood, for the model spectral it ARMA, among them a proposed method of separate estimation. Through the Wilcoxon statistical test, it was possible to verify that the curves of convergence of RME of the estimates of the parameters had some significant results, showing that, in some cases the proposed method better than the other methods. Therefore, depending on the model in study, the estimation methods for the model ARMA, they can have similar results, or one of them can be shown more satisfactory than the others, beside these methods of separate estimate used the equations of Yule-Walker.

Resumo. Neste trabalho comparamos alguns métodos de estimação separada espectral ARMA, baseados nas equações de Yule-Walker e no método de mínimos quadrado, com o estimador de máxima verossimilhança, utilizando curvas de convergência do erro médio relativo, via simulações de Monte Carlo.

References

[1] M.I.S. Bezerra, Precision of Yule-Walker Methods for the ARMA Spectral Model, in "Proceedings of the IASTED Conference on Circuits, Signals, and Systems", pp. 54-59, IASTED, Clearwater Beach, 2004.

[2] M.I.S. Bezerra, G. Stanghenhaus, Uma Comparação entre L-Estimadores de Regressão, in "Anais do X Simpósio Nacional de Probabilidade e Estatística", p. 76, ABE, Caxambú, 1996.

[3] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, "Time Series Analysis: Forecasting and Control", Third Edition, Prentice-Hall, Englewood Cliffs, 1994.

[4] P.M.T. Broersen, Autoregressive Model Orders for Durbin's MA and ARMA Estimators, *IEEE Trans. on Signal Processing*, **48**, No. 8 (2000), 2454-2457.

[5] J. A. Cadzow, High performance spectral estimation - a new ARMA method, *IEEE Trans. on Acoustics, Speech and Signal Processing*, **ASSP-28**, No. 5 (1980), 524-529.

[6] P.A. Enqvist, A Convex Optimization Approach to ARMA(m,n) Model Design from Covariance and Cepstral Data, *SIAM J. Control Optimization*, **43**, No. 3 (2004), 1011-1036.

[7] B. Efron, R.J. Tibshirani, "An Introduction to the Bootstrap", Chapman & Hall, New York, 1993.

[8] B. Friedlander, B. Porat, The modified Yule-Walker method of ARMA spectral estimation, *IEEE Trans. on Aerospace and Electronic Systems*, **AES-20**, No. 2 (1984), 158-172.

[9] S.M. Kay, "Modern Spectral Estimation Theory and Applications", Prentice-Hall, Englewood Cliffs, 1988.

[10] S.L.Jr. Marple, "Digital Spectral Analysis", Prentice-Hall, Englewood Cliffs, 1987.

[11] D.N. Politis, Computer Intensive Methods in Statistical Analysis, *IEEE Signal Processing Magazine*, **15** (1998), 39-55.

[12] P. Stoica, N. Sandgren, Smoothed Nonparanetric Spectral Estimation Via Cepstrum Thresholding, *IEEE Signal Processing Magazine*, **23**, No. 6 (2006), 34-45.

[13] C.W. Therrien, "Discrete Random Signals and Statistical Signal Processing", Prentice-Hall, Englewood Cliffs, 1992.

[14] A.M. Zoubir, B. Boashash, The Bootstrap and its Application in Signal Processing, *IEEE Signal Processing Magazine*, **15** (1998), 56-76.