

Research Article

A Bayesian Analysis of Spectral ARMA Model

**Manoel I. Silvestre Bezerra,¹ Fernando Antonio Moala,¹
and Yuzo Iano²**

¹ DEST, FCT, Unesp, Presidente Prudente 19060-900, SP, Brazil

² DECOM, FEEC, Unicamp, Campinas 13083-852, SP, Brazil

Correspondence should be addressed to Manoel I. Silvestre Bezerra, manoel@fct.unesp.br

Received 22 November 2011; Revised 9 April 2012; Accepted 12 April 2012

Academic Editor: Kwok W. Wong

Copyright © 2012 Manoel I. Silvestre Bezerra et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bezerra et al. (2008) proposed a new method, based on Yule-Walker equations, to estimate the ARMA spectral model. In this paper, a Bayesian approach is developed for this model by using the noninformative prior proposed by Jeffreys (1967). The Bayesian computations, simulation via Markov Monte Carlo (MCMC) is carried out and characteristics of marginal posterior distributions such as Bayes estimator and confidence interval for the parameters of the ARMA model are derived. Both methods are also compared with the traditional least squares and maximum likelihood approaches and a numerical illustration with two examples of the ARMA model is presented to evaluate the performance of the procedures.

1. Introduction

The spectral estimation of autoregressive moving average (ARMA) is considered a topic of interest in several applied areas of knowledge, for example, engineering, econometrics, and so forth [1–4]. Different methods have been studied, comprising two focuses, among which, we can point out: (i) the optimal methods, such as the maximum likelihood estimates, which are computational difficult to deal [1, 3, 4]; (ii) the suboptimal methods based on the Yule-Walker equations, which employ linear equations in the parameters estimation [2–8]. Another type of suboptimal method, proposed by [9], makes a simultaneous estimate of both parameters AR and MA using the reformulated Steiglitz-McBride least-squares algorithm.

The estimation methods for the ARMA process have been widely studied, especially for the separate estimation of parameters, where Yule-Walker equations are used to estimate the AR process parameters, and following this, the parameters MA process are estimated through the Durbin method [3, 4]. Among these methods, the best known are the modified

Yule-Walker equations (MYWEs) and the method of least squares Yule-Walker (LSYW). Moreover, these methods provide similar results for ARMA spectral estimation. One of the methods used to make the comparisons is proposed by [8].

Currently, Bayesian inference has been used successfully in practical problems also. In Bayesian inference, the unknown parameter is considered a random variable, thus assuming a probability distribution associated with (prior), which is specified from past data or in the expert opinion.

The main objective of this paper is to show the viability of the Bayesian approach in estimating the parameters of the ARMA process, considering a noninformative prior proposed by Jeffrey [10]. Thus, a comparison between the classical and Bayesian estimation methods is carried out by comparing the estimates of the parameters of the ARMA model and consequently the estimates of the power spectrum. The mean estimates of the parameters and their respective standard deviations are obtained from chains generated by simple Monte Carlo simulations MCMC algorithms.

The paper is organized as follows. In Section 2, the ARMA model and its main features are presented. Section 3 is devoted to the separate estimation methods using Yule-Walker equations, and presents the development of the method proposed, based on the methods described in section. In Section 4, we present a Bayesian analysis considering the Jeffreys and independent normal prior distributions for the parameters. Section 5 deals with applications using Monte Carlo simulations for two numerical examples.

2. The ARMA Spectral Model

Consider x_t an ARMA stationary random process of order (p, q) , defined by the following difference equation:

$$\sum_{i=0}^p \phi_i x_{t-i} = \sum_{j=0}^q \theta_j \varepsilon_{t-j}, \quad \text{assuming } a_0 = 1, \quad 1 \leq t \leq N, \quad (2.1)$$

where ϕ_i , $i = 0, 1, \dots, p$; θ_j , $j = 0, 1, \dots, q$, are the parameters of the process and ε_t is a white noise with mean zero and variance σ^2 . Thus, the vector of parameters to be estimated is defined by

$$\beta = [\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2]. \quad (2.2)$$

We define the power spectral density of ARMA process, in plane- z , from the system function, by

$$S(z) = \frac{\sigma^2 B(z) B^*(1/z^*)}{A(z) A^*(1/z^*)}, \quad (2.3)$$

where $A(z) = \prod_{i=1}^p (1 - \mathbf{p}_i z_i^{-1})$, $A^*(1/z^*) = \prod_{i=1}^p (1 - \mathbf{p}_i^* z_i^*)$, \mathbf{p}_i , $i = 1, \dots, p$, are the poles, and $B(z) = \prod_{i=1}^q (1 - \mathbf{q}_i z_i^{-1})$, $B^*(1/z^*) = \prod_{i=1}^q (1 - \mathbf{q}_i^* z_i^*)$ and \mathbf{q}_i , $i = 1, \dots, q$, are the zeros. As usual, the stars denote complex conjugation.

The ARMA power spectral density on frequency domain is obtained by substituting $z = \exp(j2\pi f)$ into (2.3) resulting in

$$S(f) = \frac{\sigma^2 |B(f)|^2}{|A(f)|^2} = \frac{\sigma^2 \left| \sum_{k=0}^q \theta_k \exp(-j2\pi f k) \right|^2}{\left| 1 + \sum_{k=1}^p \phi_k \exp(-j2\pi f k) \right|^2}. \quad (2.4)$$

The function $S(f)$ is the Fourier transform of the autocorrelation function $r_k = E[x_t x_{t-k}]$, presumably nonnegative and periodical in frequency with period 1Hz. It is assumed to be band limited to ± 0.5 Hz [1–3].

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

3. Estimation Methods Using Yule-Walker Equations

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

3.1. Least-Squares Yule-Walker Method

This method is an attempt to reduce the variance of the MYWE estimator and improve the quality of their estimates. It is known that the autocorrelation function of the ARMA process is defined as [6]

$$r_k + \sum_{i=1}^p \phi_i r_{k-i} = 0, \quad k \geq q + 1. \quad (3.1)$$

From the MYWE, initially we will describe the least squares method of the modified Yule-Walker equations (LSYW) 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 [2–4, 6]

$$\mathbf{r} = -\mathbf{R}\Phi, \quad (3.2)$$

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

$$\mathbf{R} = \begin{bmatrix} r_q & r_{q-1} & \cdots & r_{q-p+1} \\ r_{q+1} & r_q & \cdots & r_{q-p+2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{M-1} & r_{M-2} & \cdots & r_{M-p} \end{bmatrix}, \quad (3.3)$$

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

$$\mathbf{r} = [r_{q+1}, r_{q+2}, \dots, r_M]^T. \quad (3.4)$$

The consistent estimates of AR vector parameters $\Phi = [\phi_1, \dots, \phi_p]^T$ can be obtained by expression (3.2). As $M - q > p$, there are more equations than unknowns. One error e of size $(M - q) \times 1$ in order to estimate the autocorrelations should be introduced given by

$$\hat{\mathbf{r}} = -\hat{\mathbf{R}}\Phi + \mathbf{e}, \quad (3.5)$$

where $\hat{\mathbf{r}}$ and $\hat{\mathbf{R}}$ 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 least squares can be applied to find the vector which minimizes the sum of squares of errors, that is, [2, 6]

$$S = \sum_{k=q+1}^M |e_k|^2 \quad (3.6)$$

or, in the matrix form:

$$S = \mathbf{e}^T \mathbf{e}. \quad (3.7)$$

Substituting the equation of the errors

$$\mathbf{e} = \hat{\mathbf{r}} + \hat{\mathbf{R}}\Phi \quad (3.8)$$

in (3.7), we obtain

$$S = (\hat{\mathbf{r}} + \hat{\mathbf{R}}\Phi)^T (\hat{\mathbf{r}} + \hat{\mathbf{R}}\Phi). \quad (3.9)$$

The development of the products above in (3.9) one has the following expression results:

$$S = \Phi^T \hat{\mathbf{R}}^T \hat{\mathbf{R}} \Phi + \hat{\mathbf{r}}^T \hat{\mathbf{R}}^T \Phi + \Phi^T \hat{\mathbf{R}}^T \hat{\mathbf{r}} + \hat{\mathbf{r}}^T \hat{\mathbf{r}}. \quad (3.10)$$

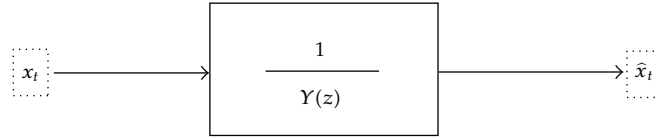


Figure 1: Filtering of the signal through an AR filter so as to obtain a new estimate.

Taking the partial derivatives $\partial S / \partial \Phi = 0$ in expression (3.10),

$$\frac{\partial S}{\partial \Phi} = 2\hat{\mathbf{R}}^T \hat{\mathbf{R}} \Phi + 2\hat{\mathbf{R}}^T \hat{\mathbf{r}} = 0. \quad (3.11)$$

We can obtain the $\hat{\Phi}$ estimator of the AR process, which is given by

$$\hat{\Phi} = -(\hat{\mathbf{R}}^T \hat{\mathbf{R}})^{-1} \hat{\mathbf{R}}^T \hat{\mathbf{r}}. \quad (3.12)$$

The matrix $\hat{\mathbf{R}}^T \hat{\mathbf{R}}$ is usually positive definite, hermitian, and $\hat{\Phi}$ estimator cannot be a minimum-phase estimator [2, 3].

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 $\{\hat{r}_{q-p+1}, \hat{r}_{q-p+2}, \dots, \hat{r}_M\}$ [2, 3]. Moreover, the estimated process parameters MA can be obtained applying the method of Durbin [3].

The performance of the LSYW, $p + q < M < N$, 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 [2, 3, 5].

3.2. The Method Using AR Filter in MA Estimates

The new method for separate estimation consists of making a new AR filtering of the signal x_t by using the estimates of the MA parameters obtained with the Durbin method. Thus determining a new signal \hat{x}_t , proposed by [8], as shown in the Figure 1 and according to the expression:

$$\hat{x}_t = x_t - \hat{\theta}_1 x_{t-1} - \dots - \hat{\theta}_q x_{t-q}. \quad (3.13)$$

In Figure 1, the $H(z) = 1/Y(z)$ is the system transfer function, and $Y(z) = 1 - \hat{\theta}_1 z^{-1} - \dots - \hat{\theta}_q z^{-q}$. This procedure is like repeating the second step of the Durbin method. This is the key idea of the separate estimation method [3, 6].

It is known that an AR(∞) process (large order) can be an approach for the MA process [1–4]. In this method, a new signal estimate \hat{x}_t will be used to obtain a new AR estimate through the LSYW method in order to subsequently determine a new MA estimate through the Durbin method. The key idea of the method proposed is to consist of obtain a new signal, from the AR filtering, which will provide better estimates when the parameters of the ARMA process are reestimated.

4. Bayesian Estimation

In a Bayesian framework, the inference is based on the posterior distribution of parameters β , denoted by $p(\beta | x)$, which, in turn, is used for inferences and decisions involving β .

The posterior distribution $p(\beta | x)$ is obtained from the combination of the information provided by a prior distribution $f(\beta)$ and the information supplied by the data through the likelihood $L(\beta | x)$. Thus, using Bayes' theorem, the posterior distribution is given by

$$p(\beta | x) \propto f(\beta)L(\beta | x). \quad (4.1)$$

The prior distribution represents the knowledge or uncertainty state about the parameter β before the experiment is run and the posterior distribution describes the updated information about β after the data x is observed.

For an ARMA(p, q) model, we need to estimate the parameters β and σ where $\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$. The likelihood function proposed in [11] for the observations X given the vector parameters (β, σ) can be written as follows:

$$L(\beta, \sigma | X) \propto \frac{1}{\sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^n (x_t - \mu_t)^2 \right\}, \quad (4.2)$$

where $\mu_1 = \sum_{i=1}^p \phi_i x_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$,

$$\begin{aligned} \mu_t &= \sum_{i=1}^p \phi_i x_{t-i} + \sum_{i=1}^{t-1} \theta_i (x_{t-i} - \mu_{t-i}) + \sum_{i=t}^q \theta_i \varepsilon_{t-i}, \quad t = 2, \dots, q, \\ \mu_t &= \sum_{i=1}^p \phi_i x_{t-i} + \sum_{i=1}^{t-1} \theta_i (x_{t-i} - \mu_{t-i}) + \sum_{i=t}^q \theta_i \varepsilon_{t-i}, \quad t = q+1, \dots, n. \end{aligned} \quad (4.3)$$

The likelihood above combined with the prior distribution results in the posterior distribution given by

$$p(\beta, \sigma | x) \propto \frac{1}{\sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^n \varepsilon_t^2 \right\} f(\beta), \quad (4.4)$$

The values of the variables x_{i-p} , $i = 1, \dots, p$ and error terms ε_{i-q} , $i = 1, \dots, q$ are arbitrarily fixed.

To progress with the Bayesian analysis, it is necessary to specify a prior distribution over the parameter space. Different prior distributions can be used in our study according to all currently available information.

If prior information on study parameters is unavailable or does not exist for a device, then initial uncertainty about the parameters can be quantified with a noninformative prior distribution. This is the same to include in the analysis just the information provided by the data.

Therefore, for the ARMA coefficients, we assume that little is known about these parameters such that a uniform distribution can be used as a prior distribution. We also assume a prior that the components of β are independent and hence the overall prior for β

is the product of the priors for its components. Besides, a conventional noninformative prior distribution for the model variance σ^2 can be represented by

$$f(\sigma^2) \propto \frac{1}{\sigma^2}. \quad (4.5)$$

Therefore, the joint prior distribution of β and σ^2 has the form:

$$f(\beta, \sigma^2) \propto \frac{1}{\sigma^2}, \quad (4.6)$$

where $-\infty < \beta_i < \infty$ and $0 < \sigma^2 < \infty$.

Other prior specifications also could be considered, as independent informative normal distributions for the components of β , that is, $\beta_j \sim N(\mu, \sigma_\beta^2)$, $j = 1, \dots, k$, with mean μ and variance σ_β^2 specified a prior and an informative prior as the inverse Gamma distribution for the parameter σ^2 , that is, $\sigma^2 \sim \text{IG}(a_\sigma, b_\sigma)$ with hyperparameters a_σ and b_σ known. Thus, the joint prior for β and σ^2 would be given by

$$f(\beta, \sigma^2) \propto f(\beta)f(\sigma^2). \quad (4.7)$$

To obtain the marginal posterior distribution for each parameter from the model, one needs to solve integrals involving the joint posterior density that are not analytically tractable and which standard integral approximations can perform poorly. In this case, we use of Markov chain Monte Carlo (MCMC) approaches to carry out the Bayesian posterior inference. Specifically, we can run an algorithm for simulating a long chain of draws from the posterior distribution, and base inferences on posterior summaries of the parameters or functional of the parameters calculated from the samples. MCMC is essentially Monte Carlo integration using Markov chains.

Constructing such a Markov chain is not difficult. We first describe the Metropolis-Hastings algorithm. This algorithm is due to [12], which is a generalization of the method first proposed by [13].

Let $g(\gamma)$ be the distribution of interest. Suppose at time t , γ_{t+1} is chosen by first sampling a candidate point η from a proposal distribution $q(\cdot | \gamma_t)$.

The candidate η is then accepted with probability:

$$\alpha(\gamma, \eta) = \min\left(1, \frac{g(\eta)q(\gamma_t | \eta)}{g(\gamma)q(\eta | \gamma_t)}\right). \quad (4.8)$$

If the candidate point is accepted, the next state becomes $\gamma_{t+1} = \eta$. If it is rejected, then $\gamma_{t+1} = \gamma_t$ and the chain does not move. The proposal distribution is arbitrary and, provided the chain is irreducible and aperiodic, the equilibrium distribution of the chain will be $g(\gamma)$.

After obtaining a random sample from the MCMC algorithm for each component of β , it is important to investigate issues such as convergence and mixing, to determine whether the sample can reasonably be treated as a set of random realizations from the target posterior distribution. Looking at marginal trace plots is the simplest way to examine the output

Table 1: Estimation classic AR and MA parameters (standard deviation).

| Method | Parameter AR | | | | Parameter MA | |
|-----------|---------------------|--------------------|---------------------|--------------------|---------------------|--------------------|
| | | | | | | |
| $N = 256$ | | | | | | |
| | -0.520 | 1.018 | -0.255 | 0.240 | -0.337 | 0.810 |
| MYWE | -0.8394 (2.1903) | 0.8236 (1.8329) | -0.2752 (0.2346) | 0.2145 (0.2539) | -0.2712 (0.4057) | 0.6068 (0.3770) |
| LSYW | -0.4057 (0.2012) | 0.7724 (0.2699) | -0.2002 (0.1018) | 0.2252 (0.0809) | -0.3132 (0.3061) | 0.6867 (0.2234) |
| LSYWS | -0.5952 (0.4266) | 1.0085 (0.3257) | -0.3210 (0.2659) | 0.2405 (0.1406) | -0.3888 (0.3061) | 0.8040 (0.1698) |
| MLE | -0.3760 (0.4133) | 0.7283 (0.4455) | -0.1764 (0.1545) | 0.2167 (0.1163) | -0.1954 (0.4117) | 0.5418 (0.4204) |

besides formal procedures. This way, all the values of the chain have marginal distribution given by the equilibrium distribution.

For more details of MCMC in a variety of ways to construct these chains, see, for example, [14–16].

We now describe the MCMC implementation used in our time series framework. We generate a posterior Monte Carlo sample by simulating a Markov chain described as follows:

- (i) choose starting values $\beta^o = (\beta_0^o, \beta_1^o, \dots, \beta_k^o)$ and σ_0^2 . At step $i + 1$, we draw a new sample $(\beta_{i+1}, \sigma_{i+1}^2)$ conditional on the current sample (β_i, σ_i^2) using the following proposal distributions (and the usual Metropolis-Hastings acceptance rule);
- (ii) sample a new vector β_{i+1} from the multivariate normal distribution $N_k(\beta_i; V)$ where V is a diagonal matrix;
- (iii) the candidate for β_{i+1} , denoted by β_{prop} , will be accepted with a probability given by the Metropolis ratio:

$$\alpha(\beta_i, \beta_{\text{prop}}) = \min \left\{ 1, \frac{p(\beta_{\text{prop}}, \sigma_i^2 | X)}{p(\beta_i, \sigma_i^2 | X)} \right\}; \quad (4.9)$$

- (iv) sample the new variance σ_{i+1}^2 from inverse Gamma distribution $IG(d, (d-1) * \sigma_i^2)$;
- (v) because this proposal is not symmetric, we find

$$\alpha(\sigma_i^2, \sigma_{\text{prop}}^2) = \min \left\{ 1, \frac{IG(\sigma_i^2) p(\beta_{i+1}, \sigma_{\text{prop}}^2 | X)}{IG(\sigma_{\text{prop}}^2) p(\beta_{i+1}, \sigma_i^2 | X)} \right\}. \quad (4.10)$$

The proposal distribution parameters V and d were chosen to obtain good mixing of the chains.

Table 2: Estimation Bayesian AR and MA parameters (standard deviation).

| $N = 256$ | | Parameter AR | | Parameter MA | | |
|-----------------|---------------------|--------------------|---------------------|--------------------|---------------------|--------------------|
| Method | | | | | | |
| | -0.520 | 1.018 | -0.255 | 0.240 | -0.337 | 0.810 |
| Bayes non | -0.5953 (0.0282) | 1.1073 (0.0291) | -0.2650 (0.0372) | 0.2148 (0.0386) | -0.4052 (0.0440) | 0.8759 (0.0469) |
| Empirical Bayes | -0.5950 (0.0281) | 1.1065 (0.0290) | -0.2645 (0.0372) | 0.2142 (0.0389) | -0.4050 (0.0439) | 0.8754 (0.0425) |

Table 3: Estimation classic AR parameters (standard deviation).

| $N = 256; M = 20$ | | Parameter AR | | | |
|-------------------|--------------------|--------------------|--------------------|--------------------|--|
| Method | 0.1 | 1.66 | 0.093 | 0.8649 | |
| MEYW | 0.1361 (0.1688) | 1.6707 (0.1179) | 0.1254 (0.1179) | 0.8644 (0.0646) | |
| LSYW | 0.1054 (0.0299) | 1.6438 (0.0483) | 0.0978 (0.0323) | 0.8432 (0.0472) | |
| LSYWS | 0.1038 (0.0706) | 1.6537 (0.0528) | 0.0995 (0.0600) | 0.8522 (0.0555) | |
| MLE | 0.1001 (0.0275) | 1.6414 (0.0395) | 0.0942 (0.0363) | 0.8482 (0.0404) | |

5. Numerical Illustration

In this section, we describe the Monte Carlo simulations [17], in order to obtain the estimates of the parameters and the spectral power density corresponding to each of the cited below ARMA processes: (1) ARMA(4,2) [18] and (2) ARMA(4,4) [19]. The simulation and the programs were implemented in the MATLAB and R softwares.

The methods used on the simulations are (1) modified Yule-Walker equations (MYWE); (2) least squares modified Yule-Walker equations (LSYW); (3) least squares modified Yule-Walker equations with AR filtering (LSYWS—proposed method); (4) maximum likelihood estimator (MLE); (5) Bayesian estimator considering the noninformative priors as Jeffreys and independence normal priors for the parameters.

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, $i = 1, \dots, B$, and B is the replications number in each simulation. We carried out this using $N = 256$ observations and $B = 30$ replications. The random numbers are generated from a standard Gaussian distribution (white noise) with mean equal zero and variance equal one, then generated a signal of the process ARMA.

The modified covariance was the criterion in the Durbin method [3], and the large AR process order was chosen using the criterion described by [20] and the number of equations, M , was chosen in accordance with the spectral characteristics to be analyzed. On the other hand, for the spectrum with poles far from the unit circle (UC), as it was given in Example 5.1, a small number was used for M , which has one pole near and other far from the unit cycle (UC).

Table 4: Estimation classic MA parameters (standard deviation).

| $N = 256; M = 20$ | | Parameter MA | | |
|-------------------|---------------------|--------------------|---------------------|--------------------|
| Method | 0.0226 | 0.8175 | 0.0595 | 0.0764 |
| MEYW | -0.0244 (0.2514) | 0.8848 (0.2227) | -0.0020 (0.2106) | 0.1291 (0.2222) |
| LSYW | -0.0649 (0.1701) | 0.8739 (0.2205) | -0.0137 (0.2112) | 0.0982 (0.2342) |
| LSYWS | -0.0681 (0.1914) | 0.8879 (0.2236) | -0.0224 (0.2262) | 0.1070 (0.2403) |
| MLE | 0.0241 (0.0594) | 0.8014 (0.0785) | 0.0653 (0.0721) | 0.0760 (0.1186) |

Table 5: Estimation Bayesian AR parameters (standard deviation).

| $N = 256$ | | Parameter AR | | |
|-----------------|--------------------|--------------------|--------------------|--------------------|
| Method | 0.1 | 1.66 | 0.093 | 0.8649 |
| Bayes non | 0.0967 (0.0151) | 1.6029 (0.0157) | 0.0875 (0.0187) | 0.8102 (0.0189) |
| Empirical Bayes | 0.0824 (0.0154) | 1.5964 (0.0206) | 0.0739 (0.0215) | 0.8121 (0.0231) |

In the first example, the ARMA(4,2) model was used with $M = 10, L = 85$, to $N = 256$, and in the second example, we used ARMA(4,4) model with $M = 20, L = 125$ to $N = 256$. The dashed curve is theoretical and solid curve is the average of estimates.

Example 5.1. The ARMA(4,2) model presented in [18] is given by

$$x_t - 0.52x_{t-1} + 1.018x_{t-2} - 0.255x_{t-3} + 0.24x_{t-4} = \varepsilon_t - 0.337\varepsilon_{t-1} + 0.810\varepsilon_{t-2}. \quad (5.1)$$

Example 5.2. The ARMA(4,4) model presented in [19] is given by

$$\begin{aligned} & x_t + 0.1x_{t-1} + 1.66x_{t-2} + 0.093x_{t-3} + 0.8649x_{t-4} \\ & = \varepsilon_t + 0.0226\varepsilon_{t-1} \\ & \quad + 0.8175\varepsilon_{t-2} + 0.0595\varepsilon_{t-3} + 0.0764\varepsilon_{t-4}. \end{aligned} \quad (5.2)$$

5.1. Discussion

Firstly, we consider a comparative analysis between the classical methods applied to the ARMA(4,2) model fitted by the data of Example 5.1. Thus, Table 1 shows that the LSYWS method provided the best averages for estimate parameters than other methods, but with a slightly larger variability than the LSYW method for the AR model. The LSYW method produces intermediate estimates but with a superior performance than the MYWE and MLE methods. These facts can be observed through the average estimates of the spectra power in

Table 6: Estimation Bayesian MA parameters (standard deviation).

| Method | Parameter MA | | | |
|-----------------|--------------------|--------------------|--------------------|--------------------|
| | 0.0226 | 0.8175 | 0.0595 | 0.0764 |
| Bayes non | 0.0151 (0.0424) | 0.7556 (0.0508) | 0.0518 (0.0478) | 0.0525 (0.0698) |
| Empirical Bayes | 0.0077 (0.0418) | 0.7554 (0.0604) | 0.0517 (0.0616) | 0.0677 (0.0613) |

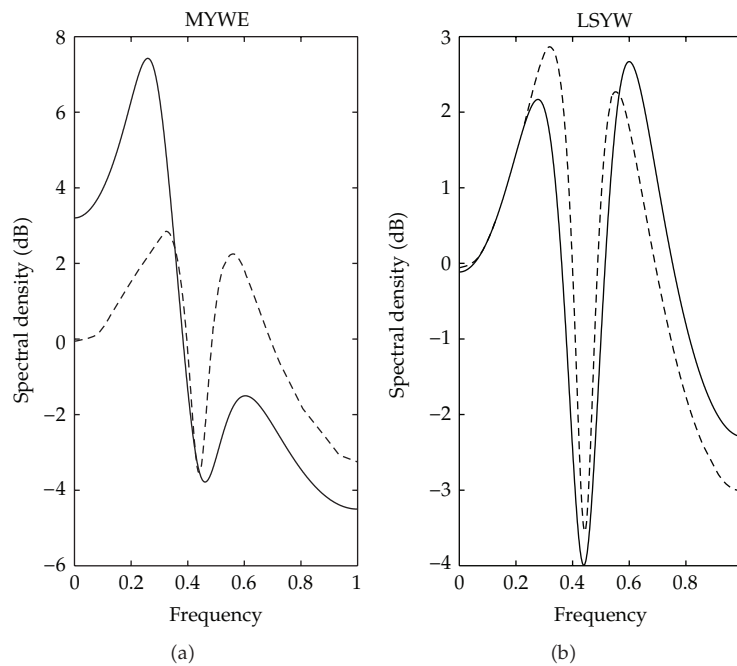


Figure 2: Density Spectral—ARMA(4,2)—Methods Classics.

Figures 2 and 3, which shows that the best estimate is really provided by LSYWS method proposed by [8].

Comparing the classical methods and the Bayesian methods studied in this paper, Table 2 shows that the estimates of parameters from both Bayesian methods are similar to the LSYWS method, but with a smaller variability (see Figure 4).

By using the noninformative Jeffreys prior, we obtained a slightly better estimate than the empirical prior, most likely due to the mean prior given by LME does not produce so good results. Tables 3 and 4 present the results of estimation methods MEYW, LSYW, LSYWS, and MLE applied to the ARMA(4,4) model from Example 5.2 (see Figures 5 and 6). In this case, the average estimates of the parameters were approximately equal for all considered methods. Just the standard deviations of the mean estimates showed a difference with the Bayesian standard deviation, see Tables 5 and 6. The standard deviation, with Jeffreys prior, had lower value than the other methods, but with a value slightly smaller than the standard deviation obtained by the MLE method.

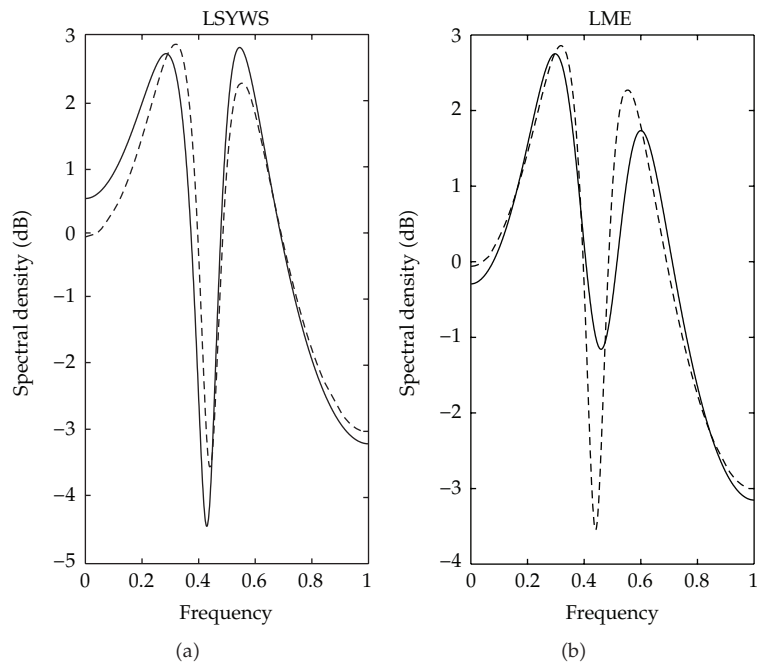


Figure 3: Density Spectral—ARMA(4,2)—Methods Classics.

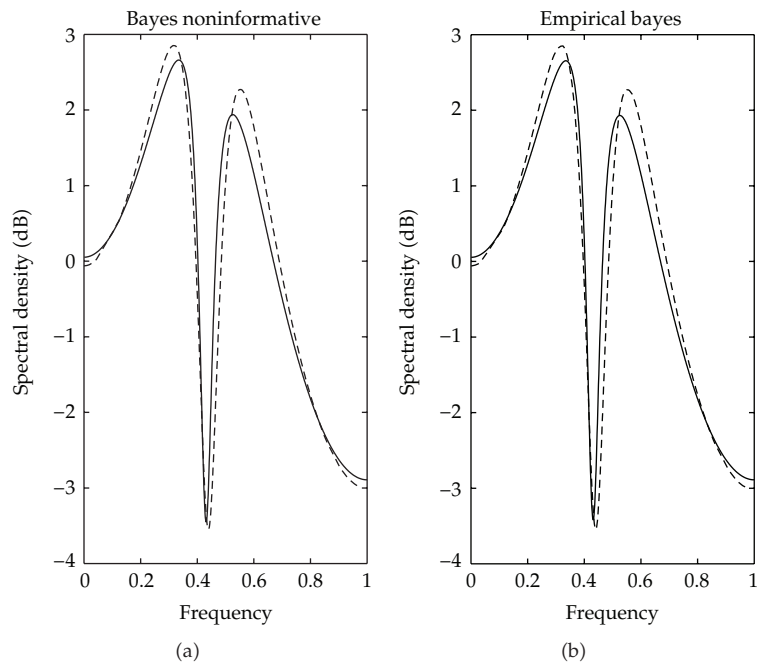


Figure 4: Density Spectral—ARMA(4,2)—Methods Bayesian.

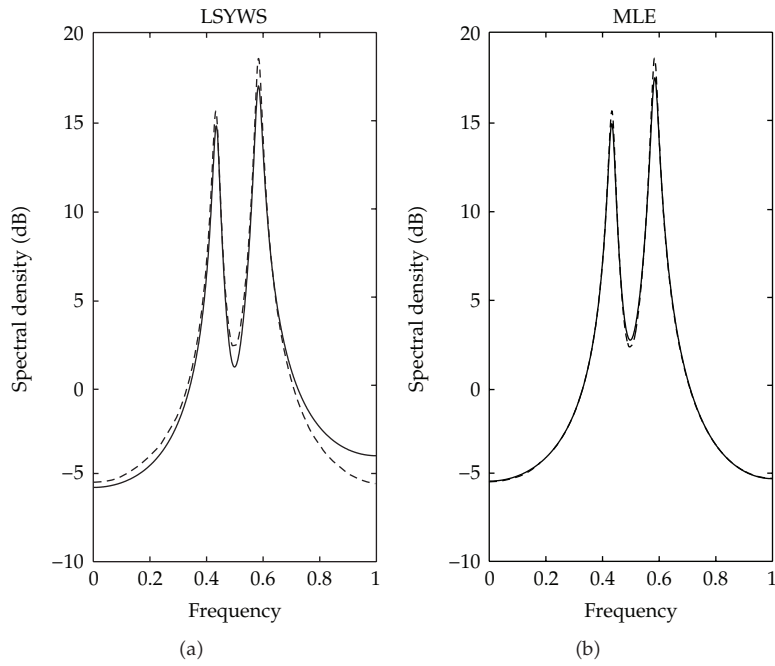


Figure 5: Density Spectral—ARMA(4,4)—Methods Classics.

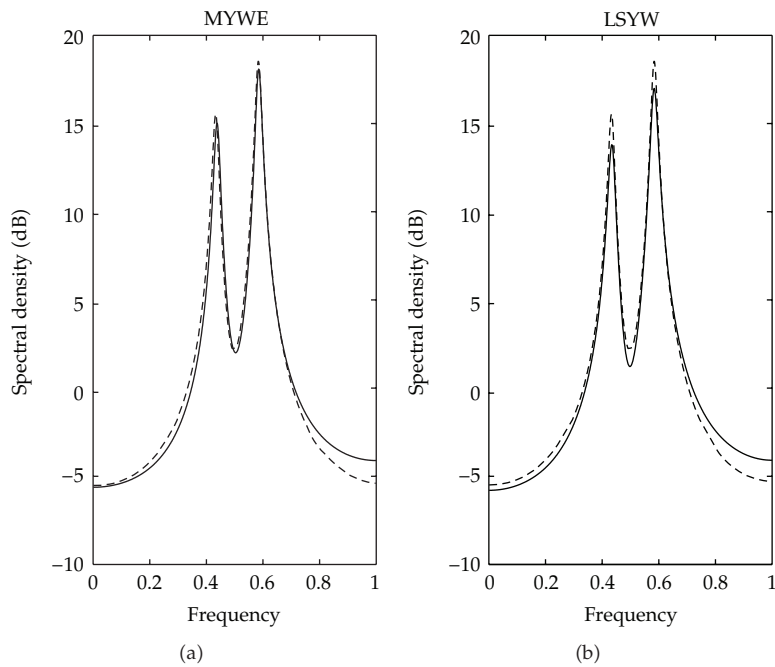


Figure 6: Density Spectral—ARMA(4,4)—Methods Classics.

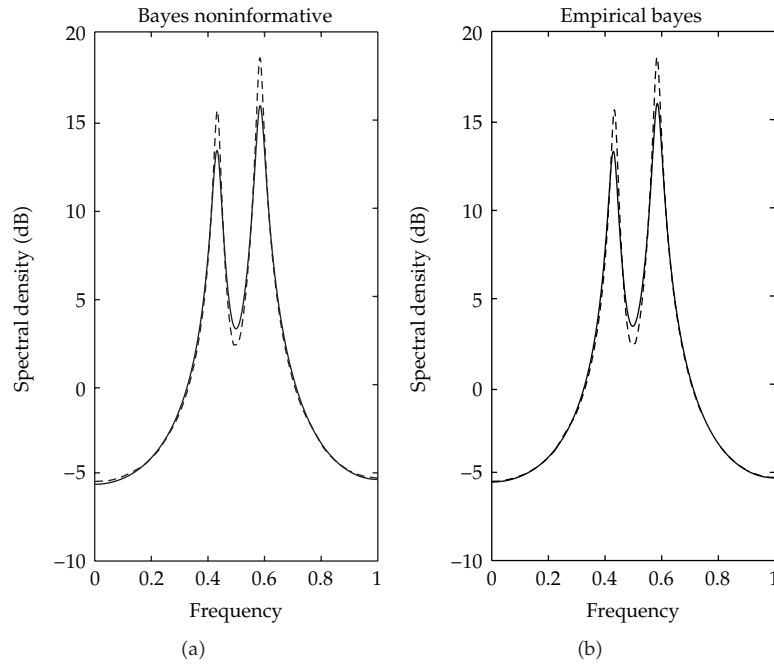


Figure 7: Density Spectral—ARMA(4,4)—Methods Bayesian.

This analysis can also be confirmed by plots of the power spectra estimates 2 (see Figures 6 and 7). In this case, this similarity can be due to the ARMA(4,4) has a stable power spectrum and with more parameters to be estimated than case previous.

6. Conclusions

In this article, we describe the Bayesian approach by using the noninformative prior distribution proposed by Jeffreys [10] to estimate the density spectral of the ARMA model. This Bayesian approach is compared with non-Bayesian approaches LSYWS method and LSYW, both based on Yule-Walker equations and least-squares method. Other methods proposed in the literature as MLE and MYWE (Mod-fitted Yule-Walker Equations) are also compared.

Two ARMA models with different orders were introduced to illustrate the proposed methodologies and to examine their performances.

The study showed that the Bayesian approaches produced more accurate estimates than the other approaches by considering ARMA(4,2) model, but similar results are found for ARMA(4,4). This way, we conclude that Bayesian approach for less stable ARMA models is justified, and more stable power spectrums any one procedures could be chosen. The Bayesian approach provides best fitting of spectral applicable for any order of the ARMA model than the other approaches. Finally, the choice of the approaches becomes irrelevant for a moderate large size. In addition, generally speaking, it was seen that the results depend on the spectrum to be estimated.

Acknowledgments

This work is supported by Capes, RH-TVD, CNPq, Faepex/Unicamp, and FUNDUNESP.

References

- [1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis*, Prentice Hall, Englewood Cliffs, NJ, USA, 3rd edition, 1994, Forecasting and Control.
- [2] S. L. Marple, Jr., *Digital Spectral Analysis with Applications*, Prentice Hall Signal Processing Series, Prentice Hall, Englewood Cliffs, NJ, USA, 1987.
- [3] S. M. Kay, *Modern Spectral Estimation Theory and Applications*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1988.
- [4] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1992.
- [5] J. A. Cadzow, "High performance spectral estimation—a new ARMA method," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 5, pp. 524–529, 1980.
- [6] B. Friedlander and B. Porat, "The modified Yule-Walker method of ARMA spectral estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, no. 2, pp. 158–173, 1984.
- [7] 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, Clearwater Beach, Fla, USA, 2004.
- [8] M. I. S. Bezerra, Y. Iano, and M. H. Tarumoto, "Evaluating some Yule-Walker methods with the maximum-likelihood estimator for the spectral ARMA model," *Tendências em Matemática Aplicada e Computacional*, vol. 9, no. 2, pp. 175–184, 2008.
- [9] B. L. Jackson, "Frequency-domain Steiglitz-McBride method for least-squares IIR filter design, ARMA modeling, and periodogram smoothing," *IEEE Signal Processing Letters*, vol. 15, pp. 49–52, 2008.
- [10] H. Jeffreys, *Theory of Probability*, Oxford University Press, London, UK, 3rd edition, 1967.
- [11] J. M. Marriot, N. M. Spencer, and A. N. Pettit, "Bayesian approach to selecting covariates for prediction," Tech. Rep., Trent University, 1996.
- [12] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.
- [13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. Teller, and H. Teller, "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, pp. 1087–1091, 1953.
- [14] A. F. M. Smith and G. O. Roberts, "Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society Series B*, vol. 55, no. 1, pp. 3–23, 1993.
- [15] A. E. Gelfand and A. F. M. Smith, "Sampling-based approaches to calculating marginal densities," *Journal of the American Statistical Association*, vol. 85, no. 410, pp. 398–409, 1990.
- [16] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London, UK, 1996.
- [17] D. N. Politis, "Computer intensive methods in statistical analysis," *IEEE Signal Processing Magazine*, vol. 15, pp. 39–55, 1998.
- [18] A. S. W. Batista, *Métodos de estimação dos parâmetros dos modelos ARMA para análise espectral [Mastership in Electrical Engineering]*, FEE, UNICAMP, 1992.
- [19] R. L. Moses, V. Simonyte, P. Stoica, and T. Söderström, "An efficient linear method for ARMA spectral estimation," *International Journal of Control*, vol. 59, no. 2, pp. 337–356, 1994.
- [20] P. M. T. Broersen, "Finite sample criteria for autoregressive order selection," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3550–3558, 2000.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

