On Bayes Estimation of the First Order Moving Average Model

Mahmoud M. Smadi and Ahmed A. Abu-Taleb
Department of Mathematics and Statistics,   Jordan University of Science and Technology Jordan
P.O.BOX 3030, 22110 Irbid, Jordan

Abstract: In this work, Bayes estimation of the first order moving average model (MA(1)) were studied. Theoretical justification of the Bayes estimates based on the estimated innovations is given. The convergence of Bayes and maximum likelihood estimates are examined via simulation using different parameter values. Also, Bayes estimates were determined when the model is invertible using the estimated innovations.   For long series lengths, it has been noted that the Bayes estimate of θ of invertible MA(1) model assuming uniform prior on θ and inverted gamma prior on σ^2 equals the Bayes estimate of θ for noninvertible MA(1) model. Generally, the simulation results showed that the performance of the Bayes estimates using estimated innovations depends on the values of θ within the invertibility region. As expected, we note that the performance of the maximum likelihood and Bayes estimates are equally likely for long series lengths.

Keywords: Bayes Estimation; Estimated Innovations; Moving Average Model.

INTRODUCTION

Bayesian inferences on autoregressive moving average (ARMA) models are limited due to the complicated form of the likelihood function which makes it difficult to get analytical tractable results. In addition, the Bayesian estimation of ARMA models is based on the causality and inevitability conditions of the model coefficients. The Bayesian inference on time series goes back to Zellner[14], when the posterior and predictive distributions were derived for the first and the second order AR models using vague priors and analysis of regression models with autocorrelated errors. Box and Jenkins[3] introduced Bayesian analysis for ARMA models without any restrictions on causality and inevitability conditions. Monahan[12], used numerical integration technique to implement Bayesian time series. Bromeling and Shaarawy[6] proposed an approximate Bayes estimates based on the estimated innovations for the analysis of MA and ARMA models without any restrictions on the causality and inevitability conditions. The estimated innovations were used later by Chen[9] in Bayesian inference of the bilinear models. Marriot et. al.[11] used Monte Carlo Markov Chain (MCMC) methods to implement the Bayesian inference on the ARMA models. Chen et. al[10], use MCMC methods for exploration of the joint posterior distribution of threshold autoregressive (TAR) models. Smadi[13] used MCMC methods for the Bayesian inference on the threshold autoregression moving average (TARMA) models using the estimated innovations.

In this work, the Bayes estimation of the first order moving average model (MA(1)) has been investigated using the estimated innovations. Asymptotic justification of Bayes estimation of MA(1) model is provided. The convergence of the Bayes and the maximum likelihood estimates are examined via simulation using different parameter values. Moreover, Bayes estimates of the first order invertible moving average model have been considered.

Bayes theorem of estimation is described below[7].

Theorem (1)

$\pi(\theta)$ be a prior distribution of θ. A sample $X_1, X_2,…, X_n$ is then drawn from a population indexed by θ. Let $f(x \mid \theta)$ be the sampling distribution of $X_1, X_2,…,X_n$. The prior distribution is updated by the sample information. The updated prior is called the posterior distribution which is a conditional distribution of θ for a given sample can be described as:

$$\pi(\theta \mid x) = f(x \mid \theta)\pi(\theta) / m(x), \qquad (1)$$

where m(x) is the marginal distribution of X which is given by:

$$m(x) = \int f(x \mid \theta)\pi(\theta)d\theta \qquad (2)$$

The mean of the posterior distribution can be used as a point estimate of θ by means of squared loss function.

The first order moving average model (MA(1)) is given by:

$$X_t = \theta Z_{t-1} + Z_t \qquad t = 0,\pm1,\pm2,... \qquad (3)$$

where $\{Z_t\}$ are *I.I.D* $N(0, \sigma^2)$ random variables.

Corresponding Author: A. Abu-Taleb Department of Mathematics and Statistics,   Jordan University of Science and Technology Jordan

The first order moving average model given in (3) is always stationary. According to Brockwell and Davis[4] a model is judged to be invertible if $-1 < \theta < 1$, i.e.

$$Z_t = \sum_{j=0}^{\infty} \theta^j X_{t-j} \qquad (4)$$

For first order moving average model (MA(1)) given in (3), the joint probability density function of $X_1, X_2, ..., X_n$ can be written as[3]:

$$f(\underline{X} \mid \theta, \sigma^2, Z_0) = (2\pi)^{-\frac{n}{2}}(\tau)^{\frac{n}{2}} \exp(-\frac{\tau}{2}\sum_{t=1}^{n}(X_t - \theta Z_{t-1})^2),$$
(5)

where Z_0 is non-observable random variable; assuming Z_0 equals its unconditional expectation zero. The innovations Z_t *(t = 1,2,...,n)* can be calculated recursively.

The conditional likelihood function is given by

$$L(\underline{X}; \theta, \sigma^2 \mid Z_0 = 0) = (2\pi)^{-\frac{n}{2}}(\tau)^{\frac{n}{2}} \exp(-\frac{\tau}{2}\sum_{t=1}^{n}(Z_t - \theta Z_{t-1})^2)$$
(6)

Bromeling[5] derived approximate marginal posterior distributions and posterior means for the parameters of the MA(q) model. Accordingly, the true innovations $\{Z_t\}$ are first estimated as $\{\hat{Z}_t\}$ and calculated recursively using $\hat{Z}_t = X_t - \hat{\theta}\hat{Z}_{t-1}$ $(t = 1,...,n)$, where $\hat{Z}_0 = 0$, and $\hat{\theta}$ is the least squares estimate of θ. Thus, the approximate conditional likelihood function is given by:

$$\hat{L}(\underline{X}; \theta, \sigma^2 \mid \hat{Z}_0 = 0) \propto (\tau)^{\frac{n}{2}} \exp(-\frac{\tau}{2}\sum_{t=1}^{n}(Z_t - \theta\hat{Z}_{t-1})^2)$$
(7)

Bromeling[5] assumed multi-normal-gamma prior density on the model parameters of the MA(q) model. For the MA(1) model, the normal-gamma prior density on θ and τ is:

$$g(\theta, \tau) = \tau^{\alpha-1} e^{-\tau\beta} e^{-\frac{\tau p}{2}(\theta-\mu)^2}, \text{ where } \theta \in R, \tau > 0, \alpha > 0, \beta > 0, \mu \in R$$
(8)

Based on the above, Bromeling[5] concluded the following results:

(i) The marginal posterior density of θ has a *t*-distribution with *n+2α-1* degrees of freedom. The mean and precision are given by

$$E(\theta \mid \underline{X}) = \frac{p\mu - \sum_{t=1}^{n} x_t \hat{z}_{t-1}}{p + \sum_{t=1}^{n} \hat{z}^2_{t-1}} \qquad (9)$$

$$\text{Precision} = \frac{(n+2\alpha)\sum_{t=1}^{n} \hat{z}^2_{t-1}}{\mu^2 p + 2\beta + \sum_{t=1}^{n} x_t^2 - (\sum_{t=1}^{n} x_t \hat{z}_{t-1})^2 / \sum_{t=1}^{n} z^2_{t-1}} \qquad (10)$$

The marginal posterior density of τ is a gamma distribution, where;

$$\alpha' = \frac{n + 2\alpha}{2} \qquad (11)$$

and;

$$2\beta' = \mu^2 p + 2\beta + \sum_{t=1}^{n} x_t^2 - (\sum_{t=1}^{n} x_t \hat{z}_{t-1})^2 / \sum_{t=1}^{n} \hat{z}^2_{t-1}$$
(12)

Then, if $\alpha \to -1/2, \beta \to 0,$ and $p \to 0,$ the normal-gamma prior density given in (8) reduces to Jeffrey's prior described by:

$$g(\theta, \tau) \propto \tau^{-1}, \text{ where } \tau > 0, \mu \in R$$
(13)

Therefore, the Bayes estimate of the precision parameter τ is given by:

$$\frac{1}{n-1}[\sum_{t=1}^{n} x_t^2 - (\sum_{t=1}^{n} x_t \hat{z}_{t-1})^2 / \sum_{t=1}^{n} \hat{z}^2_{t-1}] \qquad (14)$$

Asymptotic justification of the estimated innovations in Bayes estimation of the MA(1) model is described below using basic convergence results of the probability theory. The following theorem will be used[2].

Theorem (2)

(i) If $Z_n \xrightarrow{P} Z$, then $Z_n \xrightarrow{D} Z$.

(ii) If $Z = Z_0$ (a cnstant). $Z_n \xrightarrow{D} Z \Rightarrow Z_n \xrightarrow{P} Z$.

(iii) If $Z_n \xrightarrow{P} Z_0$ and g is continuous at Z_0, then $g(Z_n) \xrightarrow{P} g(Z_0)$.

(iv) If $Z_n \xrightarrow{D} Z$ and g is continuous, then $g(Z_n) \xrightarrow{D} g(Z)$.

(v) If $Z_n \xrightarrow{D} Z$ and $U_n \xrightarrow{P} u_0$ (a constant), then

(a) $Z_n + U_n \xrightarrow{D} Z + u_0$

(b) $U_n Z_n \xrightarrow{D} u_0 Z$

Theorem (3)

Consider the first order MA (1) model

$$X_t = \theta Z_{t-1} + Z_t (t = 0,...,n)$$
(15)

Let $\{\hat{Z}_t\}$ be the estimated innovations which are calculated recursively, i.e. $\hat{Z}_t = X_t - \hat{\theta}\hat{Z}_{t-1} (t = 1,...,n)$, where $\hat{Z}_0 = 0$ and

$\hat{\theta}$ is the least squares estimate of θ. Then, the approximate conditional joint distribution

$$f(\underline{X} \mid \theta, \sigma^2, \hat{Z}_0) = (2\pi)^{-\frac{n}{2}} (\tau)^{\frac{n}{2}} \exp(-\frac{\tau}{2} \sum_{t=1}^{n} (X_t - \theta \hat{Z}_{t-1})^2)$$

(16)

converges in distribution to

$$f(\underline{X} \mid \theta, \sigma^2, Z_0) = (2\pi)^{-\frac{n}{2}} (\tau)^{\frac{n}{2}} \exp(-\frac{\tau}{2} \sum_{t=1}^{n} (X_t - \theta Z_{t-1})^2)$$

(17)

Proof:

Let $\hat{\theta}$ be the least squares estimate of θ, it has been shown that[3]

$$\hat{\theta} \xrightarrow{P} \theta$$

Then, according to Theorem (2), part (v),

$$\hat{Z}_t = X_t - \hat{\theta}\hat{Z}_{t-1} \xrightarrow{P} Z_t = X_t - \hat{\theta}Z_{t-1}$$

Consequently,

$$\hat{Z}_t \xrightarrow{D} Z_t .$$

Also, according to Theorem (2), part (v),

$$\sum_{t=1}^{n} (X_t - \theta\hat{Z}_{t-1})^2 \xrightarrow{D} \sum_{t=1}^{n} (X_t - \theta Z_{t-1})^2 ,$$

Finally, because of the continuity of the exponential function and Theorem (2), part (iv), the following approximate conditional distribution function

$$f(\underline{X} \mid \theta, \sigma^2, \hat{Z}_0) = (2\pi)^{-\frac{n}{2}} (\tau)^{\frac{n}{2}} \exp(-\frac{\tau}{2} \sum_{t=1}^{n} (X_t - \theta \hat{Z}_{t-1})^2)$$

(18)

converges in distribution to

$$f(\underline{X} \mid \theta, \sigma^2, Z_0) = (2\pi)^{-\frac{n}{2}} (\tau)^{\frac{n}{2}} \exp(-\frac{\tau}{2} \sum_{t=1}^{n} (X_t - \theta Z_{t-1})^2)$$

(19)

Similarly this result can be generalized to the MA(q) model.

The convergence of the approximate Bayes estimates using the estimated innovations and the maximum likelihood estimates will be examined via simulation using different parameters and series lengths.

Bayes Estimation of the Invertible MA(1) Model

The first order invertible moving average model, where $-1 < \theta < 1$, will be considered. Abu-Salih and Abd-alla[1] obtained in a closed form the Bayes estimates of the parameters of the stationary AR(1) model using different informative and non-informative priors. The AR(1) model is given by:

$$X_t = \varphi X_{t-1} + Z_t \qquad (t = 0,...,n)$$

(20)

where $\{Z_t\}$ ~IID $N(0, \sigma^2)$, $-1 < \varphi < 1$.

The following joint prior on φ and σ^2 is used in the Bayesian estimation:

$$g(\varphi, \sigma^2) = \frac{\alpha^d e^{\frac{\alpha}{\sigma^2}}}{(\sigma^2)^{d+1} \Gamma(d)} \frac{1}{2}, \sigma^2 > 0, -1 < \varphi < 1,$$

(21)

where φ and σ^2 are independent, with inverted gamma prior on σ^2 with hyper parameters d and α and uniform prior on φ.

Assuming squared loss function, Abu-Salih and Abd-alla[1] derived in a closed form the marginal posteriors of φ and σ^2 for the stationary AR(1) model. The derivation for the above invertible MA(1) model is similar to the stationary AR(1) model; where X_t is regressed on $\hat{Z}_{t-1}$ rather than X_{t-1}. The final expressions of the marginal posterior means will be followed, details of the derivations and integration results can be found in Abu-Salih and Abd-alla[1].

The Bayes estimates of θ is given by

$$\hat{\theta} = r + \frac{(m!)^2 (4a)^m [\frac{1}{((a+(1+r)^2)^m} - \frac{1}{((a+(1-r)^2)^m}]}{\frac{(2m)!}{(2m)!(2m)} (\frac{1}{2a} \sum_{j=1}^{m} b(j) + k)}$$

(22)

where

$$k = \frac{1}{\sqrt{a}} (\tan^{-1} \frac{1-r}{\sqrt{a}} + \tan^{-1} \frac{1+r}{\sqrt{a}}),$$

$$b(j) = \frac{j!(j-1)!(4a)^j}{(2j)!} (\frac{1-r}{\{(1-r)^2 + a\}^j} + \frac{1+r}{\{(1-r)^2 - a\}^j}),$$

$$r = \frac{\sum_{i=2}^{n} x_i \hat{z}_{i-1}}{\sum_{i=2}^{n} \hat{z}_{i-1}^2}, \quad a = \{\frac{(2\alpha + \sum_{i=2}^{n} x_i^2)}{\sum_{i=2}^{n} x_{i-1}^2}\} - r^2,$$

$$m = \frac{n+2d}{2} - 1$$

for large n,

$$\hat{\theta} \approx r = \frac{\sum_{i=1}^{n} x_i \hat{z}_{i-1}}{\sum_{i=2}^{n} \hat{z}_{i-1}^2},$$

(23)

The Bayes estimator of σ^2 is given by:

$$\hat{\sigma}^2 = \frac{a \sum_{i=2}^{n} \hat{z}_{i-1}^2}{2m-1} (1 - \frac{b(m)}{\sum_{j=1}^{m} b(j) + 2\sqrt{ak}}),$$

(24)

For large n,

$$\hat{\sigma}^2 \approx \frac{a \sum_{i=2}^{n} \hat{z}_{i-1}^2}{2m-1}$$

(25)

It is noted that for long series lengths, the approximate Bayes estimate (Eq. 23) of θ in the invertible MA(1) model, assuming uniform prior on θ and inverted gamma prior on σ, equals the approximate Bayes estimate of θ (Eq. 9) in the noninvertible MA(1) model, assuming Jeffrey's prior. Thus, it can be concluded that both estimates of invertible and non-invertible models are the same. In addition, there is no effect of the inverted gamma prior on σ on the Bayes estimate of θ for long record lengths. On the other hand, as one expects, there is an effect of the inverted gamma prior on σ when the Bayes estimate of σ when comparing the mean of the marginal posterior density of τ derived by Bromeling (Eq. 14) and the asymptotic Bayes estimate of the invertible model (Eq. 25)

RESULTS AND DISCUSSION

A simulation was carried out to examine the convergence of the Bayes estimates based on the estimated innovations. The results were compared with the maximum likelihood estimates. The means and root mean square errors (RMSE) of the maximum likelihood estimates $\hat{\theta}$ and $\hat{\sigma}^2$ and the Bayes estimates $\hat{\theta}_B$ and $\hat{\sigma}_B^2$ of θ and σ^2 were calculated using different combinations of parameter values (θ, σ^2); namely, (θ=-0.3, σ^2 =1), (θ=0.3, σ^2 =1), (θ=-0.6, σ^2 =1), (θ=0.6, σ^2 =1) and (σ^2 =1, θ=-0.3). The results are displayed in Tables 1-5. The simulation results were based on 1000 replicates for series lengths of 30, 50, 100, 200, 400, 700 and 1000.

Table 1: MLE and Bayes estimates with RMSE of θ (θ = 0.3, σ^2 =1)

n	$\hat{\theta}$	RMSE($\hat{\theta}$)	$\hat{\theta}_B$	RMSE($\hat{\theta}_B$)
30	0.3978	0.3192	0.2248	0.2849
50	0.3490	0.1938	0.2934	0.2162
100	0.3148	0.1092	0.2943	0.1487
200	0.3089	0.0714	0.2993	0.1051
400	0.3056	0.0497	0.3006	0.0510
700	0.3027	0.0362	0.2996	0.0371
1000	0.3019	0.0313	0.3001	0.0322

Table 2: MLE and Bayes estimates with RMSE of θ (θ = 0.6, σ^2 =1)

n	$\hat{\theta}$	RMSE($\hat{\theta}$)	$\hat{\theta}_B$	RMSE($\hat{\theta}_B$)
30	0.7902	0.3254	0.5696	0.2223
50	0.7202	0.2362	0.5875	0.1552
100	0.6737	0.1700	0.5956	0.1035
200	0.6373	0.1210	0.5971	0.0734
400	0.6176	0.0807	0.5977	0.0505
700	0.6070	0.0478	0.5988	0.0373
1000	0.6026	0.0311	0.5983	0.0322

Table 3: MLE and Bayes estimates with RMSE of θ (θ = -0.3, σ^2 =1)

n	$\hat{\theta}$	RMSE($\hat{\theta}$)	$\hat{\theta}_B$	RMSE($\hat{\theta}_B$)
30	-0.2832	0.2512	-0.2903	0.2249
50	-0.2906	0.1639	-0.2927	0.1599
100	-0.2976	0.1078	-0.2993	0.1079
200	-0.2970	0.0725	-0.2975	0.0730
400	-0.2974	0.0481	-0.2979	0.0511
700	-0.2995	0.0355	-0.2995	0.0371
1000	-0.2991	0.0305	-0.2993	0.0323

Table 4: MLE and Bayes estimates with RMSE of θ (θ = -0.6, σ^2 =1)

N	$\hat{\theta}$	RMSE($\hat{\theta}$)	$\hat{\theta}_B$	RMSE($\hat{\theta}_B$)
30	-0.6089	0.2216	-0.5679	0.2152
50	-0.6220	0.1820	-0.5787	0.1610
100	-0.6221	0.1340	-0.5890	0.1045
200	-0.6217	0.1150	-0.5940	0.0735
400	-0.6110	0.0800	-0.5970	0.0504
700	-0.6036	0.0509	-0.5977	0.0379
1000	-0.6005	0.0287	-0.5985	0.0322

Table 5: MLE and Bayes estimates with RMSE of σ^2 (σ^2 =1, θ = 0.3)

n	$\hat{\sigma}^2$	RMSE($\hat{\sigma}^2$)	$\hat{\sigma}_B^2$	RMSE($\hat{\sigma}_B^2$)
30	0.9145	0.2835	0.8420	0.2849
50	0.9600	0.2041	0.9103	0.2162
100	0.9821	0.1092	0.9574	0.1487
200	0.9866	0.1022	0.9750	0.1051
400	0.9948	0.0720	0.9891	0.0738
700	0.9978	0.0552	0.9947	0.0556
1000	0.9990	0.0463	0.9966	0.0464

Generally, the simulation results showed that the performance of the Bayes estimates using the estimated innovations depends on the values of θ within the invertibility region. As we expect, we note that the performance of the maximum likelihood and the Bayes estimates are equally likely for long series lengths.

CONCLUSION

It can be concluded that for long series lengths, the approximate Bayes estimate of θ of invertible MA(1) model assuming uniform prior on θ and inverted gamma prior on σ, equals the Bayes estimate of θ of noninvertible MA(1) model, assuming Jeffrey's prior. The simulation results showed that the performance of the Bayes estimates using the estimated innovations depends on the values of θ within the invertibility region. For long series lengths, we note that both maximum likelihood and Bayes estimates performs equally likely.

REFERENCES

1. Abu-Saleh M. S. and Abd-alla A. A., 1985. Bayes Estimation of first order autoregressive processes, in Statistical Aspects of Water Quality Monitoring, Editors: A. H. El-Shaarawy and R. E. Kwaitkowski, Elsevier Science Publishers B. V.
2. Bickel P. J. and Doksum K. D. 1977. Mathematical Statistics, Pp 460. Holden-Day, Inc.
3. Box G. E. P. and Jenkins G. M. 1976. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.
4. Brockwell P.J., Davis R. A. 1991. Time Series: Theory and Methods, Springer, New York.
5. Bromeling L. D. 1985. Bayesian analysis of linear models. Dekker, New York.
6. Bromeling L. D. and Shaarawy S. 1988. Time Series: A Bayesian analysis in time domain. In Bayesian Analysis of Time Series and Dynamic Models. (J. C. Spall, ed.). Dekker, New York.
7. Casella G. and Berger R. L. 2002. Statistical Inference. Duxbury.
8. Chen C. W. S. 1992 Bayesian analysis of bilinear time series models: A Gibbs sampling approach. Communication in Statistics. Theory and Methods. 21(12), 3407-3425.
9. Chen C. W. S. 1992. Bayesian inference and forecasting in bilinear time series models. Communication in Statistics. Theory-Methods 21(6), 1725-1743.
10. Chen C. W. S. and Lee J. C. 1995. Bayesian inference of threshold models, Journal of time series analysis. 16, No. 5 483-491.
11. Marriot J., Ravishanker N. Gelfand A., Pai J. 1995 Bayesian Analysis of ARMA Processes: Complete sampling-based inference under exact likelihoods, Bayesian statistics and econometrics: Essays in honor of Arnold Zellner, D. Bery, K. Calloner and J Geweke, eds., Wiley, 241-254.
12. Monahan J. F. 1983 Fully Bayesian analysis of ARMA time series models, Journal of econometrics: 21, 307-331.
13. Smadi M. M. 1997 Bayesian inference of threshold autoregressive moving average models using sample-based methods. PH. D. Dissertation, Colorado State University, Fort Collins, U.S.A.
14. Zellner A. 1971. An Introduction to Bayesian Inference in Econometrics. New York: Wiley