BAYESIAN ESTIMATION FOR ARMA MODELS

By Ch. Krisnandari Ekowati
ekowatichristine@gmail.com

Teacher Training and Educational Science Faculty, Nuca Cendana University - Indonesia

ABSTRACT: The concept of estimating a parameter is needed to help estimate a situation or observational data before making a decision. There are estimation methods that have been developed, namely the moment method, which is the oldest method, the maximum likelihood method (MLE), and the Bayes method, which is the latest method in determining the estimator of a parameter. Furthermore, the concept of forecasting is also one of the important ways to make a decision. Time series analysis technique (time series) is one of the forecasting methods that are often used, were specifically selected ARMA models. In the Bayesian approach, the parameters in the ARMA model are seen as quantities whose variance is represented by a probability distribution called a prior distribution. Within the framework of Bayes decision theory, estimator selection can be thought of as a problem of decision theory in uncertain circumstances. By using a multivariate Wishart normal prior distribution, the Bayesian estimator for $\psi = z + u^*$ and the Bayesian estimator for $\tau: \hat{\tau} = (\alpha + 1)Ls + L$ with $L = (1, 0, \ldots, 0)' \in \mathbb{R}^{r-p}$ and $S = S + \frac{u}{u+1}(u-y)(u-y)'$. Using the Gamma multivariate prior normal distribution, the Bayesian estimator for $\psi^\text{Bayes} = Z + u_o$ and the Bayesian estimator for $\tau$ is $\hat{\tau} = \frac{(\alpha + 1)(u-y)}{u} + u_o = (s + R y)\beta + \frac{1}{2} \left((u - u) - \sqrt{u(u - u)}\right)$. Forecasting one step ahead, namely $\hat{y}_n(I) = \sum_{i=1}^{n} y_{t-i} + \sum_{j=1}^{r} \theta_j e_{t-j}$.

Keywords: Bayesian estimation, forecasting, time series, ARMA model.

INTRODUCTION

Forecasting is one of the essential elements in decision making because the effectiveness of a decision depends on several factors that cannot see when the decision taken. The role of forecasting explores into many fields such as economics, finance, state administration, geophysics and population (Abraham, 1983).

The time series analysis technique which is one of the forecasting methods is used in many fields of scientific work, especially in making operational forecasting (Box & Jenkins, 1970).

The ARMA (AutoRegressive Moving Average) model has been used widely as a statistical model. The statistical model is a mathematical model that contains heretical terms to accommodate the inevitable differences. This heresy interpreted as a random variable that has a probability distribution with a zero mean. The problem is how to estimate the parameters of the ARMA model and use the model to calculate the forecast value for the future. In this case, there are two approaches, namely the classical approach and the Bayesian approach (Box & Tiad, 1973).

In the classical approach, the parameter in the ARMA model is an unknown fixed amount. Random samples are taken from the population and based on observed prices of the samples; knowledge of parameters in the ARMA model obtained (Abraham Bovas, 1983).

In the Bayesian approach, the parameters in the ARMA model seen as quantities whose variance is represented by a probability distribution called a prior distribution. It is a subjective distribution, based on one's beliefs and formulated before data is taken. Then the sample is drawn from the population, and the prior distribution adjusted to the information of this sample, the adjusted priors are called the posterior distribution. This adjustment is done using Bayes rules (Box & Tiad, 1973).

The two approaches above are informal because they do not produce the best single estimator. The selection of estimators, according to the classical method, is a matter of whether it cannot be more important than efficiency and so forth. The same problem is also faced by Bayesians to choose estimators such as posterior mean, mode or median. If what desired is a general idea of price parameters, then an informal approach is sufficient. If the loss function can be determined then the point estimator can be done according to the decision theory far obtained (Lehman, 1983).

Based on the background of the problem outlined above, the problem that will discuss in this paper is how to estimate parameters in the ARMA time series model through the Bayesian approach.

The purpose of writing this scientific work is:

1. To understand how the ARMA time series model;
2. To know that to estimate the parameters of the ARMA model in addition to the likelihood function approach can also use the Bayesian approach.

The benefits of writing this scientific work are:
1. By knowing the problems in the ARMA time series model, is expected to provide an alternative to solving the estimation problem in the ARMA time series model.

2. Can provide input to readers, especially those who are relevant to this paper.

DISCUSSION

The writing method of this scientific paper is a study of literature.

1. **Determine Bayesian Estimator**

Within the framework of Bayes decision theory, estimator selection can think of as a problem of decision theory in uncertain circumstances. Here the decision corresponds with the possible estimators. The loss function is determined to provide losses for errors in estimating estimated parameters. If it is possible to determine the loss function, then the best action is the action that minimizes the expected loss function (Lehman E.L, 1983).

**Theorem 1.1**

For the quadratic loss function, if $E(\theta)$ there is a Bayes estimator for $\theta$ is $E(\theta)$ (Box & Tiad, 1973).

**Proof:**

Because $k > 0$, to minimize $R(\theta - \delta)$ it is sufficient to minimize $E(\theta - \delta)^2$.

Because $-\theta = \delta - E(\theta) + E(\theta) - \theta$, then:

$$E(\theta - \delta)^2 = E(\theta - E(\theta) + E(\theta) - \delta)^2 = E[\delta - E(\theta)]^2 + 2E[(\delta - E(\theta))(E(\theta) - \theta)] + E[E(\theta) - \theta]^2$$

$$= [\delta - E(\theta)]^2 + E[E(\theta) - \theta]^2$$

Since the last term does not meet $\delta$, the minimum price depends only on the term $[\delta - E(\theta)]^2$. The term will be zero only if $\delta = E(\theta)$. So this price is the one that minimizes the risk function for each $\theta$.

Furthermore, if the parameter is a vector, that is $\theta = (\theta_1, \theta_2, ..., \theta_k)^T$ with $k \geq z$. In this case, the loss function $d \in R^k$ and the random variable $X_1, ..., X_k \in R^2$. Bayes estimator for $\theta$ will be found in the following theorem.

**Theorem 1.2**

For the quadratic loss function, if $E(\theta)$ and $Kov(\theta)$ exist, the Bayes estimator $\theta$ is $E(\theta)$ (Box & Tiad, 1973).

2. **ARMA Time Series Model**

Consider the time series model (Abraham, 1983):

$$y_t = \sum_{i=1}^{p} \theta_i y_{t-i} + \sum_{i=1}^{q} \theta_i f_{t-j} + e_t$$

Which $e_t$ is the normal random i.i.d random variable with $e_t \sim N(0, \tau^{-1})$ and unknown. We write the model (2.1) with ARMA ($p, q$).

Suppose there are $n$ observations $(y_1, y_2, ..., y_n)^T$, then the residue is written as:

$$e = y_t - \sum_{i=1}^{p} \theta_i y_{t-i} + \sum_{i=1}^{q} \theta_i f_{t-j} + e_t$$

(2.2)

With the condition that the first $p$ observation and suppose $e_p = e_{p-1} = ... = e_r = 0$, with $r = \min (0, p+1 - q)$.

Then, the complete model can be written in the form of a matrix as follows:

$$e = y - z\psi$$

(2.3)

Where:

$$e = (e_{p+1}, e_{p+2}, ..., e_n)^T$$

$$y = (y_{p+1}, y_{p+2}, ..., y_n)^T$$

$$\psi = (\theta_1, \theta_2, ..., \theta_p, \theta_1, \theta_2, ..., \theta_q)^T$$

With residues assumed to have the following properties:

1. $E(e_t) = 0$
2. $Var (e_t) = \tau^{-1}I$
3. $Kov (e_t, e_k) = 0, t \neq k$

3. **Prior and Posterior Analysis**

In Bayesian analysis, prior distribution decisions usually play an essential role. Using the Bayes theorem, initial information expressed by prior distributions and sample information expressed by the likelihood function are combined to form a posterior distribution. One way to avoid this difficulty is to limit the distribution of priors in certain distribution families that depend on the form of the likelihood function. The Bayesian statement has developed the concept of prior peer distribution. The prior normal Wishart or normal gamma distribution is used to obtain Bayesian estimators (Dudewicz & Mishra, 1988).

**Theorem 3.1**

If $V > 0$, $m \in R^k$, $u \in R^k$ and $R$ matrix type $k \times k$ then:

$$V (m - u)' R (m - u) + (m - \mu)' R (m - \mu) = (v+1) (m - u^*)' R (m - u^*) + \frac{v}{v+1} [u - X (m - u)] (u - X (m - u))^T$$

Where $u^* = (V u + \mu) / (v+1)$.

**Proof:**

With $u^* = (V u + \mu) / (v+1)$.

**Theorem 3.2**

$x$ a random sample of a multivariate normal distribution population with a mean vector $m$ value is unknown and the precision matrix $R$ is unknown. The prior distribution with $m$ and $R$ is $\varepsilon (m, R) = \varepsilon_1 (m | R) \varepsilon_2 (R)$. With $\varepsilon_1 \sim N(u, vR) \exists u \in R^k$ and $u > 0$ and
\( \varepsilon_2 \sim \mathcal{W}(\mathbf{c}, \mathbf{S}) \) \( \exists \infty > k - 1 \) of the positive definite symmetry matrix.

Then the posterior distribution with \( m \) and \( R \) is;
\[
\Pi(m, R) = \Pi_1(m \mid R) \Pi_2(R).
\]
dan \( \Pi \sim \mathcal{W}(\infty, 1, S^*) \) where:
\[
u^* = (u + x)/(v+1)
\]
\[
S^* = S + \frac{v}{v+1} (u - x)(u - x)'
\]

**Theorem 3.3 (Box & Tiad, 1983):**
Given the likelihood function \( f(x|m, R) \)
\[
\varepsilon(m, R) \sim \mathcal{W}(\mathcal{W}(m - u), (S + R))
\]
and the posterior density function
\[
\varepsilon(m, R) \approx \varepsilon_1(m \mid w) \varepsilon_2(w)
\]
then the marginal posterior distribution \( R \) is the Wishart distribution with degrees of freedom \( \infty + 1 \) and a precision matrix
\[
S^* = S + \frac{v}{v+1} (u - x)(u - x)'
\]

**Proof:**
\[
\Pi(R) \propto \int_{x \in \mathbb{R}^k} [\det(R)]^{-\frac{k}{2}} \exp \left\{-\frac{1}{2} (x - m)' R (x - m) \right\} \Pi(R) \]

\[
\Pi(R) \propto \det(R) \frac{-k}{2} \exp \left\{-\frac{1}{2} \text{tr}(S \cdot R) \right\}
\]

Relation (2.4) is a Wishart distribution with degrees of freedom \( \infty + 1 \) and a precision matrix \( S^* \)

As a result of 2.6:
Given the likelihood function:
\[
L(\psi, \tau \mid \alpha) \propto \tau^{-\frac{k}{2}} \exp \left\{-\frac{1}{2} \text{tr}(Y - Z\psi)'(Y - Z\psi) \right\}
\]

and the density function prior
\[
\varepsilon(m, R) \mathcal{W}(\mathcal{W}(m - u), (S + R))
\]
then the Bayes estimator for \( \tau \) is
\[
\hat{\tau} = (\alpha + 1) L' S + L
\]

**Proof:**
Use theorem 2.5 by taking \( R=\tau I \), \( m=Z\psi \) dan \( k=\text{n-p} \), then the posterior marginal distribution \( \tau \) is:
\[
\Pi(\tau \mid I) \propto \det(\tau)^{-\frac{k}{2}} \exp \left\{-\frac{1}{2} \text{tr}(S \cdot \tau I) \right\}
\]

According to the theorem of the Wishart distribution, namely: if \( V \sim \mathcal{W}_k(T, \alpha) \) and \( L \) are a vector, then \( L' V L \sim \mathcal{W}_k(T' L A^2 \mathcal{W}_k \) with \( L = (1, 0, \ldots, 0)' \).
Using theorem 2.9 and taking $R=I$, $w = \tau$, $Z=\psi$ dan $k = n-p$, then the marginal distribution of posterior $z \psi$ is a multivariate t distribution with degrees of freedom $2(\alpha + (n-p)/2)$, location vector $u_0$ and precision matrix
\[
[(\alpha + (n-p)/2)(s + R)]
\]
According to theorem 2.2, the Bayes estimator for $z \psi = E(z \psi | y)$ or $z \psi = u_0$. So the Bayes approach for $\psi$ adalah $z = z u_0$.

4. ARMA Process Forecasting
Estimation models that have assessed for time series data they have, will be used for time series inference in the future based on their past time behaviour. Based on a model, we want to reduce the conditional distribution of future observations if known past observations. This step, which is the last step in forming a time series model, is known as forecasting. If the best forecast is intended as a forecast that minimizes the forecast of the heresy squared mean, then the conditional distribution of observations that will come with the mean distribution is a good forecast (Box & Jenkins, 1970).

We write $\hat{y}_n(I)$ the expected value of $y_{n+1}$, if it is known that the previous time series has reached time $n$. the forecast is one step ahead, namely:
\[
\hat{y}_n(I) = \frac{E(y_{n+1})}{s_n} + \text{Var}[e_{n+1}](I) = e_{n+1}
\]
and the forecast error variance is
\[
\text{Var}[e_{n+1}](I) = \tau^+.
\]

As a result, the size of disperse depends only on the parameters $\psi$ and $\tau$. Therefore by using the estimated values, the confidence interval for future observations can be calculated. Because $\epsilon_i$s is normally distributed, we can describe the entire distribution of future heresies and thus future observations. In this case;
\[
[e_{n+1}](S) = y_{n+1} - \hat{y}_n(I) = e_{n+1}
\]

Therefore, a 100 $(1-2d)$% confidence interval forecast can be written as:
\[
[y_{n+1}](S) = [\hat{y}_n(I) - Z \times \text{SD}[e_{n+1}](S), \hat{y}_n(I) + Z \times \text{SD}[e_{n+1}](S)]
\]
Where fn $(z_d) = 1- d$ and fn are the standard normal random variable cumulative distribution functions.

CONCLUSIONS
Based on the discussion of the problem in Chapter II, it can be concluded that:
1. Using the Wish normal multivariate Wishart distribution, the Bayesian estimator for $\psi$ is $z + u^*$ and the Bayesian estimator for $\tau$ is $\hat{\tau} = (z + u^*) U S^* L$, dengan $L = (1,0,…,0)' \in R^{n-p}$ and
\[
S_n = S + \frac{(u - y)(u - y)'}{1+u}
\]
2. Using the Gamma multivariate prior normal distribution, the Bayesian estimator for $\psi$ is $z + u$, and the Bayesian estimator for $\tau$ is $\hat{\tau} = (z + u) (z + u)^2/2$.

with $u_0 = (s u + R y)$.

\[
\beta = \beta + \frac{1}{2}(u_0 - u)'S(u_0 - u)
\]

3. Forecasting one step ahead, namely:
\[
\hat{y}_n(I) = \sum_{i=1}^{S} o_i y_{n-j} + \sum_{i=1}^{q} \theta e_{n-j}
\]

SUGGESTIONS
The Bayesian approach can use to determine the predictive distribution for $y_{n+1}$. A conditional distribution is required given past observations and also a posterior distribution for the parameters of the ARMA model (p, q). This is an interesting theme for readers, especially fans of statistics, to continue this paper.

REFERENCES
Wald Abraham, (1966), Sequential Analysis, John Wiley & Sons, Inc.