Bayesian Analysis of the Regression Model Generated by a Second Order Autoregressive Process

Lee Se-youhl

2차 자기회귀 과정에 의해 생성된 회귀모형의 베이지안 분석

李 世 烈

I. Introduction

We consider the relalionships

(1.1) $y_t = \beta x_t + u_t$

(1.2) $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \epsilon_t \quad t = 1, 2, \ldots, T.$

These equations define the simple regression model with an error term generated by a second order autoregressive process. Equation (1.1) represents the simple regression model in which β is a regression coefficient, y_t the t-th observation, X_t the t-th fixed element and U_t the t-th error terms. Equation (1.2) denotes the autoregressive scheme generating the error term U_t which involves parameters ρ_1, ρ_2 and error term ϵ_t. We assume that the ϵ_t are normally and independently distributed with zero means and common variance σ^2.

From (1.1) and (1.2), We obtain the relation

(1.3) $y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \beta(x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2}) + \epsilon_t \quad t = 1, 2, \ldots, T.$

We note that y_{-1}, y_0 appear in (1.3). In order to proceed with the analysis, we suppose the forth assumptions about y_{-1} and y_0 are

(1.4) $y_{-1} - \beta x_{-1} = M + \epsilon_{-1}$

(1.5) $y_0 - \beta x_0 = N + \epsilon_0$

Where M,N depend on certain unobservable and unobserved quantities so that they are regarded as parameters. Under these assumptions, y_{-1}, y_0 are normally distributed with means $\beta x_{-1}+M$, βx_0+N respectively and with common variance σ^2.

Under the assumptions associated with (1), the likelihood function for $\beta, \rho_1, \rho_2, \sigma$, M and N is given by;

(2) $$l(\beta, \rho_1, \rho_2, \sigma, M, N \mid y_{-1}, y_0, y_1, \ldots, y_T) \propto \sigma^{-(T+2)} \exp\Big\{ -\frac{1}{2\sigma^2}[y_{-1} - \beta x_{-1} - M]^2 - \frac{1}{2\sigma^2}[y_0 - \beta x_0 - N]^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}[y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2} - \beta(x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2})]^2 \Big\}$$

with $-\infty<\beta<\infty$, $-\infty<\rho_1<\infty$, $-\infty<\rho_2<\infty$, $-\infty<M<\infty$, $-\infty<N<\infty$ and $\sigma>0$, and where the symbol $\propto$ denotes proportionality. Using this likelihood function, we shall derive posterior distributions for the parameters.

II. Derivation of Posterior Distributions

We assume that the prior knowledge about the parameters β, ρ_1, ρ_2, M, N and log σ can be represented by locally uniform and independent distribution; that is,

$$(3)\quad P(\beta) \propto K_1\ ; P(\rho_1) \propto K_2\ ; P(\rho_2) \propto K_3\ ;$$
$$P(M) \propto K_4\ ; P(N) \propto K_5\ ; P(\sigma) \propto \frac{1}{\sigma}.$$

Appling Bayes' Teorem with these prior distributions and the likelihood function in (2), we have the following joint posterior distribution:

$$(4)\quad P(\beta, \rho_1, \rho_2, \sigma, M, N \mid y_{-1}, y_0, y_1, \ldots, y_T)$$
$$= K\sigma^{-1}\, l(\beta, \rho_1, \rho_2, \sigma, M, N \mid y_{-1}, y_0, y_1, \ldots, y_T)$$

where $l(\beta, \rho_1, \rho_2, \sigma, M, N \mid y_{-1}, y_0, y_1, \ldots, y_T)$ is the likelihood function in (2) and K is a normalizing constant.

In order that we are interesting in investigating β, ρ_1 and ρ_2, we eliminate the influence of M and N by integrating (4) over these parameters to yield;

$$(5)\quad P(\beta, \rho_1, \rho_2, \sigma \mid y_{-1}, y_0, y_1, \ldots, y_T)$$
$$= K\sigma^{-(T+1)} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^{T} [y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2} - \beta(x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2})]^2\right\}$$

Upon integrating out the scale parameter σ from (5), we obtain the following Lemma 1.

Lemma 1. Under the assumptions associated with (1), the joint posterior distribution of β, ρ_1 and ρ_2, $P(\beta, \rho_1, \rho_2 \mid y)$ is obtained from (5) and is given by ;

$$(6)\quad P(\beta, \rho_1, \rho_2 \mid y) = K\left\{\sum_{t=1}^{T} [y_t - \beta x_t - \rho_1(y_{t-1} - \beta x_{t-1}) - \rho_2(y_{t-2} - \beta x_{t-2})]^2\right\}^{-T/2}$$
$$= K\left\{\sum_{t=1}^{T} [Y_t - \beta X_t]^2\right\}^{-T/2}$$

where

$$X_t = x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2}$$
$$Y_t = y_t - \rho_{t-1} - \rho_2 y_{t-2}$$

and K is a normalizing constant. This distribution summarizes all the information about β, ρ_1 and ρ_2. Further the marginal distributions of β and (ρ_1, p_2) may be obtained from (6).

Integrating out the parameters ρ_1, ρ_2 from (6), the marginal distribution of β, $P(\beta \mid y)$ is obtained as follows.

Theorem 2. Under the assumptions associated with (1), the marginal distribution of β, $P(\beta|y)$ is obtained from (6) and is given by ;

$$(7)\quad P(\beta|y) = K\left\{\Sigma U_{t-2}^2\, \Sigma U_{t-1}^2 - (\Sigma U_{t-2} U_{t-1})^2\right\}^{-1/2} \cdot \left[\Sigma_t^2 - \frac{(\Sigma U_{t-2} U_t)^2}{\Sigma U_{t-2}^2} - \frac{(\Sigma U_{t-2}^2 \cdot \Sigma U_{t-1} U_t - \Sigma U_{t-2} U_{t-1}\, \Sigma U_{t-2} U_t)^2}{\Sigma_{t-2}^2\left\{\Sigma U_{t-2}^2 \Sigma U_{t-1}^2 - (\Sigma U_{t-2} U_{t-1})^2\right\}}\right]^{-\frac{T}{2}+1}$$

where $U_t = y_t - Bx_t$ and K is a normalizing constant. Integrating out the parameter β from (6), we may obtain the joint marginal distribution of (ρ_1, ρ_2) as follows.

Theorem 3. Under the assumptions associated with (1), the joint marginal distribution of (ρ_1, ρ_2), $P(\rho_1, \rho_2 \mid y)$ is obtained from (6) and is given by;

$$(8)\quad P(\rho_1, \rho_2 \mid y) = K(\Sigma X_t^2)^{-1/2}\left[\Sigma Y_t^2 - \frac{(\Sigma X_t Y_t)^2}{\Sigma X_t^2}\right]^{-\frac{T-1}{2}}.$$

in which

$$X_t = x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2}$$

$$Y_t = y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2}$$

and K is a normalizing constant. Finally, we obtain the following theorem for the conditional distribution of β, $P(\beta|\rho_1, \rho_2, y)$.

Theorem 4. Under the assumptions associated with (1) the conditional distribution of β for fixed values of ρ_1 and ρ_2, $P(\beta|\rho_1, \rho_2, y)$ is obtained from (6) and is given by.

$$(9.1)\quad P(\beta|\rho_1, \rho_1, \rho_2, y) = \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{T-1}{2})\ \sqrt{\pi(T-1)}}$$

$$\left\{S^2(\rho)\right\}^{-1/2}\left\{1+\frac{[\beta-\hat{\beta}(\rho)]^2}{(T-1)S^2(\rho)}\right\}^{-T/2}$$

where

$$\hat{\beta}(\rho) = \frac{\Sigma X_t Y_t}{\Sigma X_t^2}$$

$$S^2(\rho) = \frac{\Sigma[Y_t - \hat{\beta}(\rho)X_t]^2}{(T-1)\Sigma X_t^2} \qquad \rho = (\rho_1, \rho_2)$$

$$X_t = x_t - \rho_1 x_{t-1} - \rho_2 x_{t-2}$$

$$Y_t = y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2}.$$

Also, the distribution of $\dfrac{\beta-\hat{\beta}(\rho)}{S(\rho)}$ is given by;

$$(9.2)\quad P\left(\frac{\beta-\hat{\beta}(\beta)}{S(\rho)}|y\right) = P(t_{T-1})$$

where t_{T-1} is a Student-t variable with $(T-1)$ degree of freedom.

Proof. Since

$$\Sigma[Y_t - \hat{\beta}(\rho)X_t]^2 = \Sigma Y_t^2 - \hat{\beta}(\rho)\Sigma X_t^2$$

$$= \Sigma Y_t^2 - \frac{(\Sigma X_t Y_t)^2}{\Sigma X_t^2}$$

and

$$\Sigma[Y_t - \beta X_t]^2 = \Sigma Y_t^2 - \hat{\beta}^2(\rho)\Sigma X_t^2$$

$$+ [\beta - \hat{\beta}(\rho)]^2 \Sigma X_t^2,$$

the conditional distribution of β, $P(\beta|\rho_1, \rho_2, y)$ is obtained from (6) and (8):

$$(9.3)\quad P(\beta|\rho_1, \rho_2, y) = \frac{P(\beta, \rho_1, \rho_2, |y)}{P(\rho_1, \rho_2 |y)}$$

$$\propto \left\{(T-1)\,S^2(\rho)\right\}^{-1/2} \left\{1+\frac{(\beta-\hat{\beta}(\rho))^2}{(T-1)S^2(\rho)}\right\}^{-1/2}$$

Hence, we have the equation:

$$(9.4)\quad P(\beta|\rho_1, \rho_2, y)$$

$$= K\left\{S^2(\rho)\right\}^{-1/2}\left\{1+\frac{(\beta-\hat{\beta}(\rho))^2}{(T-1)S^2(\rho)}\right\}^{-1/2}.$$

Put

$$(9.5)\quad t_{T-1} = \frac{\beta - \hat{\beta}(\rho)}{S(\rho)}$$

which we deal with as function of β. Then (9.4) and (9.5) reduce the equation

$$P(t_{T-1}) = P(\beta|\rho_1, \rho_2, y)\left|\frac{\alpha\beta}{dt_{T-1}}\right|$$

to

$$P(t_{T-1}) = K\left(1+\frac{t_{T-1}}{T-1}\right)^{-1/2}$$

Therefore,

$$(9.6)\quad K = \frac{\Gamma(\frac{T}{2})}{\Gamma(\frac{T-1}{2})\sqrt{\pi(T-1)}}$$

and

$$(9.7)\quad P(t_{T-1}) = \frac{\Gamma(\frac{T}{2})}{\Gamma(\frac{T-1}{2})\sqrt{\pi(T-1)}}\left(1+\frac{t_{T-1}^2}{T-1}\right)^{-1/2}$$

are derived from the fact that $P(t_{T-1})$ is a distribution. This equation means that t_{T-1} is a

Student-t variable with (T-1) degree of freedom. And the equation (9.1) is derived from (9.4) and (9.6).

Theorem (3) and (4) allow to write the marginal distribution of β as:

$$(10) \quad P(\beta|y)=\int\int P(\beta|\rho_1, \rho_2, y)\, P(\rho_1, \rho_2, |y) \alpha\rho_1 \alpha\rho_2 \cdot$$

Theorem (4) leads to the conditional distribution of β, $P(\beta|\rho_1, \rho_2, y)$ yields the same Student-t distribution as the case of the first order autoregresive process (cf, [1]).

Literature cited

[1] Arnold Zellner & George C. Tiao 1964, *Baysian Analysis of the Regression Model with Autocorrelated Errors*, J. Amer Statist Assoc 59, p.763-770.

[2] Box, G.E.P & Tiao, G.C. 1962, *A Further Look at Robustness via Bayes, Theorem,* Bimetrika 49, p.419-433.

[3] George E.P Box & George C. Tiao 1973, *Bayesian Inference in Statistical Analysis,* Addison-Wesley.

[4] Tiao, G.C & Zellner, A 1964, *Bayes' Theorem and* the Use of Prior Knowledge in Regression Analysis, Biometrika 51, p.219-230.

[5] Wayne A. Fuller 1976, *Introduction to Statistical Time Series,* New York, John Wiley & Sons.

國 文 抄 錄

본 논문에서는 2차 자기회귀과정에 의해 생성된 오차항을 갖는 회귀모형을 베이지안의 견지에서 분석하였다. 이 모형에 들어 있는 모수에 대하여 국소 일양 사전분포를 도입하여 회귀계수의 사후 분포와 자기회귀과정에 있는 모수의 사후분포를 산출하고 이들 분포의 성질 몇 가지를 간단히 논하였다.