Journal of Korean
Data & Information Science Society
2005, Vol. 16, No. 1, pp. 71～78

Reference Prior and Posterior in the AR(1) Model

Yoon Jae Lee[1)]

Abstract

Recently an important issue in Bayesian methodology is determination of noninformative prior distributions, often required when there is no idea of prior information. In this thesis attention is focused on the development of noninformative priors for stationary AR(1) model. The noninformative priors primarily discussed are the Jeffreys prior, and the reference priors. The remarkable points in the result are that the Jeffreys prior coincides with the reference prior for the case that ρ is the parameter of interest.

Keywords : 제프리스 사전분포, 준거 사전분포, 피셔 정보 행렬, AR(1)

1. 머리말

베이지안 분석의 기본 틀은 사전분포와 우도함수를 통해 사후분포를 얻고, 그것으로부터 정보를 추출하거나 의사를 결정하는 것이다. 그러나 사전분포를 정하는데 있어서 사전분포가 주관성을 가질 수 있기에 다소 문젯거리가 되어왔다. 특히, 사전정보가 없을 때는 사전분포를 정하기조차 어려워진다. 따라서 어떤 방법으로 이와 같은 상황에서 합리적이며 객관적인 사전분포를 정하느냐는 베이지안들에게 중요한 문제가 아닐 수 없다. 최근에 이 문제에 대한 해결 방법으로 무정보 사전분포가 베이지안들에게 주목을 받고 있다. 역사적으로 살펴보면 Bayes (1763)는 균일 사전분포 $\pi_U(\theta)=1$을 통해 이를 극복하고자 했다. 그러나, 모수 변환시 불변성이 깨지는 단점이 발생하곤 했다. (Mitchell (1967), Ye and Berger (1991)). Jeffreys (1961)는 $\pi_U(\theta)=1$의 불변성이 깨지는 단점을 보완하여 $\pi_J(\theta) \propto \sqrt{det[I(\theta)]}$를 찾아냈고, Bernardo (1979)는 사전분포와 사후분포의 Kullback-Liebler divergence를 이용해 다음과 같은 정보 값 $I\{\pi(\theta)\}$를 생각하여 준거 사전분포를 만들었다.

---

1) 서울특별시 성동구 행당동 17번지 한양대학교 대학원 수학과 박사과정
E-mail : protossrgx@ hanmail.net

$$I\{\pi(\theta)\} = E_X\left[E_{\theta|X}\left(\log\frac{p(\theta|X)}{\pi(\theta)}\right)\right]$$

여기서 $I\{\pi(\theta)\}$는 주어진 사전분포에 따라 그 사전분포를 사후분포와 비교해 얼마나 정보를 잃는지 평균적으로 그 양을 측정한다. Bernardo는 $I\{\pi(\theta)\}$값을 최대화 시키는 사전분포 $\pi(\theta)$를 가장 적합한 무정보 사전분포라고 생각했다.

본 연구에서 다루게 될 AR(1)시계열 모형의 경우, 고전적인 접근방법으로는 Box and Jenkins (1970)와 Granger and Newbold (1977)에 의한 것들이 있고, 베이지안 접근방법으로는 Broemeling (1985)를 보면 켤레인 정규-감마를 AR(1)에 적용했고, 무정보 사전분포로는 Berger and Yang (1994)이 제한된 조건 하에 준거 사전분포를 찾아냈으며, Lubrano (1995)는 Beta사전분포를 생각해서 그것이 제프리스 사전분포와 일치함을 보였다. 시계열은 경제 모델링에 많이 사용되는데, 경제 상황은 주관에 따라 달리 인식이 되어지고, 또 현 상황에 대해 정확한 정보가 없는 경우가 많기 때문에, 무정보 사전분포가 시계열의 분석에 있어 중요한 위치를 차지한다 할 수 있다. 본 연구에서는 초기치 X_0를 확률변수로 생각하여 좀더 일반적인 AR(1)모형에 대하여 제프리스 사전분포와 준거 사전분포를 구하고자한다. 2장에서는 피셔정보행렬을 구하고 이를 이용하여 제프리스 사전분포와 준거 사전분포를 구한 후, 사후분포를 찾아 그것이 적절성을 만족하는 분포임을 보였다. 그리고, 3장에서는 컴퓨터를 이용한 모의 실험을 통하여 그 유효성을 검증해 보았다.

2. 본론

$X_1, X_2, \cdots, X_n$은 시계열 $AR(1)$모형 $X_t = \rho X_{t-1} + \varepsilon_t$ 을 따른다고 하자. 여기서 초기 치는 $X_0 \sim N(0, 1/(1-\rho^2)r)$ 이고 $\varepsilon_t \overset{iid}{\sim} N(0, 1/r)$ 라고 가정하자. (2.1)

(Berger and Yang (1994)은 초기 치 X_0가 0이고 ε_t의 분산 $1/r$이 1인 경우를 다루었고 여기서는 그보다 확장된 모형을 다루고자한다.)

제프리스 사전분포와 준거 사전분포를 구하기는 위해서는 피셔 정보 행렬이 필요하다. 이를 얻기 위해 우선적으로 주어진 가정으로부터 ρ와 r에 대한 우도함수를 아래와 같이 구한다.

$$l = (1-\rho^2)^{\frac{1}{2}}\left(\frac{r}{2\pi}\right)^{\frac{n+1}{2}} \exp\left\{-\frac{r}{2}\sum_{t=1}^{n}(X_t - \rho X_{t-1})^2 - \frac{(1-\rho^2)r}{2}X_0^2\right\}$$

$$L = \frac{1}{2}\log(1-\rho^2) + \frac{n+1}{2}(\log r - \log 2\pi) - \frac{r}{2}\left[\sum_{t=1}^{n}(X_t - \rho X_{t-1})^2 + (1-\rho^2)X_0^2\right] \qquad (2.2)$$

식(2.2)로부터 피셔 정보 행렬의 구성요소를 아래와 같이 계산하였다.

$$
\begin{aligned}
\frac{\partial^2 L}{\partial \rho^2} &= -\frac{1+\rho^2}{(1-\rho^2)^2} - r\left[\sum_{t=1}^{n}(X_{t-1}^2) - X_0^2\right] \\
\frac{\partial^2 L}{\partial \rho \partial r} &= -\left[\sum_{t=1}^{n}(X_t - \rho X_{t-1})(-X_{t-1}) - \rho X_0^2\right] \\
\frac{\partial^2 L}{\partial r^2} &= -\frac{n+1}{2r^2}
\end{aligned}
\tag{2.3}
$$

피셔 정보 행렬을 좀 더 용이하게 얻기 위해서 다음의 Lemma를 먼저 계산해보도록 한다.

Lemma. $X_1, X_2, \cdots, X_n$이 가정(2.1)을 따르는 시계열 $AR(1)$모형이라고 하면,

$$E\left(r\sum_{t=1}^{n}X_{t-1}^2\right) = \frac{n}{1-\rho^2}$$

이다.

증명) $X_t = \rho^t X_0 + \rho^{t-1}\varepsilon_1 + \rho^{t-2}\varepsilon_2 + \cdots + \rho\varepsilon_{t-1} + \varepsilon_t \ (t \geq 1)$이 된다. 따라서 가정 (2.1)으로부터 $E(X_t) = 0$임을 얻을 수 있다. 2차 적률을 구하기 위해 먼저 분산을 구하면,

$$Var(X_t) = \rho^{2t}\,Var(X_0) + \frac{1}{r}\left(\frac{1-\rho^{2t}}{1-\rho^2}\right)$$

이 된다. 따라서,

$$E(X_t^2) = \rho^{2t}E(X_0^2) + \frac{1}{r}\left(\frac{1-\rho^{2t}}{1-\rho^2}\right)$$

이제 $E\left(r\sum_{t=1}^{n}X_{t-1}^2\right)$을 계산하면

$$E\left(r\sum_{t=1}^{n}X_{t-1}^2\right) = \frac{n}{1-\rho^2} + \frac{1-\rho^{2n}}{1-\rho^2}\left[rE(X_0^2) - \frac{1}{1-\rho^2}\right] = \frac{n}{1-\rho^2}$$

따라서 Lemma의 결과를 얻을 수 있다.

위의 Lemma를 이용하여(2.3)을 계산하면 피셔 정보 행렬을 아래와 같이 완성할 수 다.

$$
\begin{pmatrix}
\frac{1+\rho^2}{(1-\rho^2)^2} + \frac{n-1}{1-\rho^2} & -\frac{\rho}{(1-\rho^2)\,r} \\
-\frac{\rho}{(1-\rho^2)\,r} & \frac{n+1}{2r^2}
\end{pmatrix}
\tag{2.4}
$$

Theorem 1. $X_1, X_2, \cdots, X_n$이 가정(2.1)을 따르는 시계열 AR(1)모형에서 관심모수를 ρ라 할 때, 제프리스 사전분포는

$$\pi_J(\rho, r) \quad \propto \quad \frac{1}{(1-\rho^2)r}\sqrt{(n+1)-(n-1)\rho^2}$$

이다.

이제 Berger and Bernardo (1989)가 제안한 알고리즘을 이용하여 준거 사전정보를 구하여보자.

Theorem 2. $X_1, X_2, \cdots, X_n$이 가정(2.1)을 따르는 시계열AR(1)모형에서 관심모수를 ρ라 할 때, ρ와 r의 결합 준거 사전분포는 다음을 따른다. (제프리스 사전분포와 준거 사전분포가 동일한 것을 알 수 있다.)

$$\pi_R(\rho, r) \quad \propto \quad \frac{1}{(1-\rho^2)r}\sqrt{(n+1)-(n-1)\rho^2}$$

증명) (2.4)의 피셔 정보행렬을 이용하여 다음 4단계를 풀어보면,

STEP 1.

$$\pi(r|\rho) = \sqrt{\frac{n+1}{2r^2}} = \frac{1}{r}\sqrt{\frac{n+1}{2}}$$

STEP 2. Compact 집합 $k_i = [-1+1/i, 1-1/i]\times[1/i, i]$로 정하면 $k_1 \subset k_2 \subset \cdots$ 이 되고 $\bigcup_i k_i = \Omega$이다.

$$\Omega_{i,\rho} = \{r\,|\,(\rho, r) \subset k_i\} = [1/i, i] \quad , \text{모든} \ \rho$$

$$C_i(\rho) = \left(\int_{1/i}^{i} \frac{1}{r}\sqrt{\frac{n+1}{2}}\, dr\right)^{-1} = \left(2\sqrt{\frac{n+1}{2}}\log i\right)^{-1}$$

$$P_i(r|\rho) = \frac{1}{r}\sqrt{\frac{n+1}{2}}\left(2\sqrt{\frac{n+1}{2}}\log i\right)^{-1} = \frac{1}{2r\log i}$$

STEP 3. 주변 준거 사전분포는 다음과 같이 구할 수 있다.

$$\pi_i(\rho) = \exp\left\{\frac{1}{2}\int_{1/i}^{i} \frac{1}{2r\log i}\log\left[\frac{|I(\theta)|}{|(n+1)/2r^2|}\right]dr\right\}$$

$$= \exp\left\{\frac{1}{2}\log\left|\frac{n^2+n-(n^2-n)\rho^2}{(n+1)(1-\rho^2)^2}\right|\int_{1/i}^{i}\frac{1}{2r\log i}\, dr\right\}$$

$$= \left| \frac{n^2+n-(n^2-n)\rho^2}{(n+1)(1-\rho^2)^2} \right|^{\frac{1}{2}}$$

STEP 4.

$$\pi_R(\rho, r) = \lim_{i\to\infty} \left| \frac{\frac{n^2+n-(n^2-n)\rho^2}{(n+1)(1-\rho^2)^2}}{\frac{n^2+n-(n^2-n)\rho_0^2}{(n+1)(1-\rho_0^2)^2}} \right|^{\frac{1}{2}} \frac{1}{r}\sqrt{\frac{n+1}{2}}$$

$$= \lim_{i\to\infty} \left| \frac{(1-\rho_0^2)^2[(n^2+n)-(n^2-n)\rho^2]}{(1-\rho^2)^2[(n^2+n)-(n^2-n)\rho_0^2]} \right|^{\frac{1}{2}} \frac{1}{r}\sqrt{\frac{n+1}{2}}$$

$$\propto \frac{1}{(1-\rho^2)r}\sqrt{n+1-(n-1)\rho^2}$$

Theorem 3. Theorem2에서 구한 ρ와 r의 결합 준거 사후분포는 다음과 같고, 적절성을 만족한다.

$$\pi_J(\rho, r \mid X, X_0) \propto r^{\frac{n+1}{2}-1}\exp\left\{-\frac{r}{2}\left[\sum_{t=1}^{n}(X_t-\rho X_{t-1})^2+(1-\rho^2)X_0^2\right]\right\} \times\sqrt{\frac{n+1-(n-1)\rho^2}{1-\rho^2}}$$

증명)

$$\pi_J(\rho, r \mid X, X_0)$$

$$\propto (1-\rho^2)^{\frac{1}{2}}\left(\frac{r}{2\pi}\right)^{\frac{n+1}{2}}\exp\left\{-\frac{r}{2}\left[\sum_{t=1}^{n}(X_t-\rho X_{t-1})^2+(1-\rho^2)X_0^2\right]\right\} \times\frac{1}{(1-\rho^2)r}\sqrt{n+1-(n-1)\rho^2}$$

$$\propto (1-\rho^2)^{-\frac{1}{2}} r^{\frac{n+1}{2}-1}\exp\left\{-\frac{r}{2}\left[\sum_{t=1}^{n}(X_t-\rho X_{t-1})^2+(1-\rho^2)X_0^2\right]\right\} \times\sqrt{n+1-(n-1)\rho^2}$$

$$\propto r^{\frac{n+1}{2}-1}\exp\left\{-\frac{r}{2}\left[\sum_{t=1}^{n}(X_t-\rho X_{t-1})^2+(1-\rho^2)X_0^2\right]\right\} \times\sqrt{\frac{n+1-(n-1)\rho^2}{1-\rho^2}}$$

여기서 $A=\sum_{t=1}^{n}(X_t-\rho X_{t-1})^2+(1-\rho^2)X_0^2$ 라고 하면, 모든 t에대하여 $X_t<\infty$이고 $|\rho|<1<\infty$때문에, 항상 $A<\infty$이다.

$$\int_{-1}^{1}\int_0^\infty r^{\frac{n+1}{2}-1}e^{-\frac{A}{2}r}\sqrt{\frac{n+1-(n-1)\rho^2}{1-\rho^2}}\,dr\,d\rho \tag{2.5}$$

$n=1$인 경우, (2.5)는 아래와 같다.

$$\begin{aligned}&\int_{-1}^{1}\sqrt{\frac{2}{1-\rho^2}}\left[\int_0^\infty e^{-\frac{A}{2}r}dr\right]d\rho\\ &=\int_{-1}^{1}\sqrt{\frac{2}{1-\rho^2}}\cdot\frac{2}{A}\,d\rho\\ &<\int_{-1}^{1}\sqrt{\frac{2}{1-\rho^2}}\cdot M_1\,d\rho\\ &=\int_0^{\pi/2}\sqrt{\frac{2}{1-\sin^2\theta}}\cdot M_1\cos\theta\,d\theta\\ &=\sqrt{2}\pi M_1\end{aligned}$$

$n\geq 2$인 경우, (2.5)는 아래와 같다.

$$\begin{aligned}&\int_{-1}^{1}\sqrt{\frac{(n+1)-(n-1)\rho^2}{1-\rho^2}}\left[\int_0^\infty r^{\frac{n+1}{2}-1}e^{-\frac{A}{2}r}dr\right]d\rho\\ &=\int_{-1}^{1}\sqrt{\frac{(n+1)-(n-1)\rho^2}{1-\rho^2}}\left[\Gamma\left(\frac{n+1}{2}\right)\left(\frac{A}{2}\right)^{-\frac{n+1}{2}}\right]d\rho\end{aligned} \tag{2.6}$$

위식 (2.6)에 대하여 아래와 같은 부등식이 성립한다.

$$\begin{aligned}&\int_{-1}^{1}\sqrt{\frac{\frac{n+1}{n-1}-\rho^2}{1-\rho^2}}\cdot M_2 d\rho\\ &<\ 2\int_0^1\sqrt{\frac{3}{1-\rho^2}}\cdot M_3\ d\rho\\ &=\ 2\int_0^{\pi/2}\sqrt{\frac{3}{1-\sin^2\theta}}\cdot M_3\ \cos\theta\ d\theta\ =\sqrt{3}\,\pi M_3\end{aligned}$$

3. 모의 실험

X_t를 MA과정으로 표현하면, $X_t = \rho^t X_0 + \rho^{t-1}\varepsilon_1 + \rho^{t-2}\varepsilon_2 + \cdots + \rho\varepsilon_{t-1} + \varepsilon_t$로 표현된다. 이를 보면 X_t는 X_0의 영향을 받는다는 것을 알 수 있고, $|\rho| \to 1$로 가거나 t가 작게되면 X_0의 영향력은 훨씬 커진다. 본 모의 실험은 Berger and Bernardo (1989)의 제안에 따라, 사후분포에 대한 신용구간들의 프리퀜티스트 포함확률을 ρ값과 n값에 따라 비교하고자한다. 이를 위해 필요한 사후분포는 아래(3.1)과 같고 $F(\rho | X, X_0)$는 (3.1)의 CDF이다.

$$\pi(\rho | X, X_0) \propto \int_{-1}^{\rho} \sqrt{\frac{(n+1)-(n-1)\rho^2}{1-\rho^2}} \cdot \left[\sum_{t=1}^{n} (X_t - \rho X_{t-1})^2 + (1-\rho^2) X_0^2 \right]^{-\frac{n+1}{2}} d\rho \quad (3.1)$$

알고리즘. $r=1$로 고정시키고, 아래와 같이 $n=100, 150, 200$으로 $\rho = 0.25, 0.5, 0.75$으로 설정한 후, 다음의 2단계에 따라 0.05(0.95)분위수를 계산한다.

1) 2000회의 가정(2.1)을 따르는 (X, X_0)를 추출을 한다

2) 2000번의 관측치중 $F(\rho | X, X_0) \le 0.05(0.95)$를 만족하는 개수를 세어 0.05(0.95)분위수를 구한다.

결과는 아래 표1에 나타냈다. 괄호의 왼쪽이 0.05분위수이고 오른쪽이 0.95분위수가 된다. 표1을 살펴보면, n값이 크고 ρ값이 작을수록 좋은 결과가 나옴을 알 수 있다.

<표 1> 신용구간들의 프리퀜티스트 포함 확률

ρ \ n	100	150	200
0.25	(0.041, 0.937)	(0.042 , 0.949)	(0.049, 0.950)
0.5	(0.031 , 0.940)	(0.037 , 0.954)	(0.045 , 0.951)
0.75	(0.529 , 0.998)	(0.243 , 0.942)	(0.068 , 0.950)

참고문헌

1. Bayes, T.R. (1763). Essay towards solving a problem in doctrine of chances. Reprinted in *Biometrika* (1958), 243-315.
2. Bernardo, J.M. (1979). Reference posterior distributions for Bayes

inference. *Journal of Royal Statistical Society* B 41, 113-147.
3. Berger, J.O. and Bernardo. (1989). Estimating a product of means : Bayesian analysis with reference priors. *Journal of the American Statistical Association* 84 ,200-207.
4. Berger, J.O. and Yang. (1994). Noninformative Priors and Bayesian Testing for the AR(1) Model. *Econometric Theory.* 10, 461-482.
5. Box, G.E.P. and Gwilym, M.J. (1970). *Time Series Analysis*, Forecasting and Control, Holden-Day, San Francisco.
6. Broemeling, L.D. (1985). *Bayesian Analysis of Linear Models*, New York :Marcel Dekker.
7. Granger, C.W.J. and Paul, N. (1977). *Forecasting Economic Time Series*, Academic Press, New York.
8. Jeffreys, H. (1961). *Theory of probability*. London : Oxford University Press.
9. Lubrano (1995). Testing for unit root in a Bayesian framework. *Journal of Econometircs A*, 69, 81-109.
10. Mitchell, A. (1967). Discussion of paper by I.J.Good. *Journal of Royal Statistical Society B* 29, 423-424.
11. Ye, K.Y. and J. Berger. (1991) Noninformative priors for inference in exponential regression models. *Biometrika* 78, 645-656.

[2004년 10월 접수, 2004년 12월 채택]