한국통계학회논문집 제2권 2호
1995년 12월 pp. 115-125

Bayesian Estimation via the Griddy Gibbs Sampling for the Laplacian Autoregressive Time Series Model

Young Sook Son[1] and Sinsup Cho[2]

Abstract

This paper deals with the Bayesian estimation for the NLAR(1) model with Laplacian marginals. Assuming the independent uniform priors for two parameters of the NLAR(1) model, the griddy Gibbs sampler by Ritter and Tanner(1992) is used to obtain the Bayesian estimates. Random numbers generated from the uniform priors are used as the grids for each parameter. Some simulations are conducted and compared with the maximum likelihood estimation result.

1. Introduction

The New Laplacian AutoRegressive(NLAR) model introduced by Dewald and Lewis(1985) can be applied to marginally double-exponentially distributed time series data with a larger kurtosis or longer tails than Gaussian data. Son and Cho(1988) discussed the properties and the forecasting procedures of the NLAR process. Karlsen and Tjøstheim(1988) obtained the conditional least square(CLS) estimates which are consistent and asymptotically normal for all four parameters of the NLAR(2) model. Also, Son and Cho(1995) discussed the maximum likelihood(ML) estimation for the NLAR(1) and the NLAR(2) models. The results of a simulation study for the ML estimates of the NLAR(1) and the NLAR(2) model show that as Karlsen and Tjøstheim(1988) had pointed out for the CLS estimates, if $|\alpha_i\beta_i|$, $i=1,2,$ in the NLAR(2) and $|\alpha\beta|$ in the NLAR(1) are smaller than 0.1 the ML estimates are also virtually useless because of large bias and mean square error(MSE). But, for moderate sample size in the model with $|\alpha_i\beta_i|$ much larger than 0.1 the ML estimates obtained by using the CLS estimates as initial estimates in the optimization program are better than the CLS estimates in the sense of bias and MSE.

In this paper, we consider the Bayesian estimation of the NLAR(1) model via the Griddy Gibbs sampling technique. Major obstacles to a practical implementation of Bayesian inference

(1) Department of Statistics, Chonnam National University, Kwangju, 500-757.
(2) Department of Computer Science and Statistics, Seoul National University, Seoul, 151-742.

are mainly due to the difficulties in the integration required to obtain posterior densities. In case of Bayesian models with a non-conjugate prior analytical or numerical approximations are often used, but they still involve computational difficulties or require the derivatives of posterior density or likelihood function. Since the likelihood function of the NLAR(1) model contains the absolute terms and is nonlinear it is very difficult to derive the posterior probability density for each parameter. We can avoid these difficulties by using the Gibbs sampling technique, in which the posterior densities can be obtained by a Monte Carlo method without the direct integration.

Since the Gibbs sampler has been introduced in the Bayesian restoration problem of images by Geman and Geman(1984), this Monte Carlo method has been successfully applied to Bayesian inferences and other important theoretical and practical issues in many statistical areas. In the implementation of Gibbs sampling, we need to know the maximum or the upper bound of the posterior density or restrictions on the density, for example, the log concavity. But, the griddy Gibbs sampling technique by Ritter and Tanner(1992) based on a linear approximation of the inverse cumulative distribution function(c.d.f.) requires only a proportionality constant of the posterior density. We can obtain the Bayesian estimates for the NLAR(1) model by applying the griddy Gibbs sampler without knowing the mode or log-concavity of the posterior density.

This paper is constructed as follows. Section 2 describes the NLAR(1) model, and formulates its likelihood function and posterior densities. Section 3 briefly reviews the Gibb sampling and the griddy Gibbs sampling. Secton 4 shows the results of Bayesian estimation via the griddy Gibbs sampling for four models of NLAR(1).

2. The NLAR(1) model and Bayesian formulation

Let $\{X_t\}$ be a stationary sequence of random variables whose marginal distribution is standard Laplacian. Then the NLAR(1) model is constructed as follows : for $t=0, \pm 1, \pm 2, \cdots,$

$$X_t = \left\{ \begin{array}{ll} \beta X_{t-1} & \text{w.p. } \alpha \\ 0 & \text{w.p. } 1-\alpha \end{array} \right\} + \varepsilon_t , \tag{2.1}$$

where

$$\varepsilon_t = \left\{ \begin{array}{ll} L_t & \text{w.p. } 1-p \\ |\beta|\sqrt{1-\alpha}\, L_t & \text{w.p. } p=\alpha\beta^2/\{1-(1-\alpha)\beta^2\} \end{array} \right. , \tag{2.2}$$

$0<\alpha<1$, $0<|\beta|<1$, and $\{L_t\}$ is a sequence of i.i.d. standard Laplacian variables.

Let $\underline{x} = (x_1, x_2, \cdots, x_n)$ denote the observed data of size n of $\{X_t\}$. Then, from Son and Cho(1995), the likelihood function is given by

$$l(\alpha,\beta)=f(\underline{x}\mid\alpha,\beta)=\frac{1}{2}e^{-|x_1|}\prod_{t=2}^{n}\sum_{i=1}^{2}\sum_{j=1}^{2}\alpha_i p_j d_j^{-1}\exp\{-|x_t-\beta_i x_{t-i}|/d_j\}, \tag{2.3}$$

where $\alpha_1=1-\alpha$, $\alpha_2=\alpha$, $\beta_1=\beta$, $\beta_2=0$, $p_2=p=\alpha\beta^2/\{1-(1-\alpha)\beta^2\}$, $p_1=1-p$, $d_1=1$ and $d_2=|\beta|\sqrt{1-\alpha}$.

Assuming the independent uniform priors, $\alpha\sim U(0,1)$ and $\beta\sim U(-1,1)$, where $\beta\neq 0$, the posterior densities for α and β are

$$f_{\alpha|\beta,\underline{x}}(\alpha\mid\beta,\underline{x})=\frac{1}{2}c_1(\beta,\underline{x})\cdot l(\alpha,\beta) \tag{2.4}$$

and

$$f_{\beta|\alpha,\underline{x}}(\beta\mid\alpha,\underline{x})=\frac{1}{2}c_2(\alpha,\underline{x})\cdot l(\alpha,\beta), \tag{2.5}$$

where $c_1(\beta,\underline{x})$ and $c_2(\alpha,\underline{x})$ are normalizing constants independent of α and β, respectively.

3. The Gibbs sampling and the griddy Gibbs sampling

The Gibbs sampling is a Markovian updating scheme to obtain samples from a joint distribution $f(\theta_1,\theta_2,\cdots,\theta_k)$ via sampling iterated from k available complete(full) conditional distributions,

$$\left\{\begin{array}{c} f(\theta_1\mid\theta_2,\theta_3,\cdots,\theta_k)\\ f(\theta_2\mid\theta_1,\theta_3,\cdots,\theta_k)\\ \vdots\\ f(\theta_k\mid\theta_1,\theta_2,\cdots,\theta_{k-1})\end{array}\right\}. \tag{3.1}$$

Given an arbitrary set of initial values, $(\theta_{2(0)},\cdots,\theta_{k(0)})$, the Gibbs sampling algorithm to obtain a sample $(\theta_1,\theta_2,\cdots,\theta_k)$ from $f(\theta_1,\theta_2,\cdots,\theta_k)$ repeats the following loop by setting $i=0$ in the beginning :

(a) draw $\theta_{1(i+1)}$ from $f(\theta_1\mid\theta_{2(i)},\theta_{3(i)},\theta_{4(i)},\cdots,\theta_{k(i)})$.

(b) draw $\theta_{2(i+1)}$ from $f(\theta_2\mid\theta_{1(i+1)},\theta_{3(i)},\theta_{4(i)},\cdots,\theta_{k(i)})$.

$\vdots$

(k) draw $\theta_{k(i+1)}$ from $f(\theta_k\mid\theta_{1(i+1)},\theta_{2(i+1)},\cdots,\theta_{k-1(i+1)})$.

After repeating the above loop l times, the sample value $(\theta_{1(l)},\theta_{2(l)},\cdots,\theta_{k(l)})$ is obtained.

Geman and Geman(1984) showed that under mild conditions

$$(\theta_{1(l)}, \theta_{2(l)}, \cdots, \theta_{k(l)}) \xrightarrow{d} (\theta_1, \theta_2, \cdots, \theta_k) \tag{3.2}$$

and hence for each s,

$$\theta_{s(l)} \xrightarrow{d} \theta_s \sim f(\theta_s), \tag{3.3}$$

as $l \to \infty$. Therefore, for large l $\theta_s(l)$ can be regarded as an observation drawn from its marginal distribution, $f(\theta_s)$. After this entire process is performed until the l-th repetition with G independent parallel runs, G simulated samples $(\theta_{1(l)}^{(g)}, \theta_{2(l)}^{(g)}, \cdots, \theta_{k(l)}^{(g)})$, g = 1, 2, ...G, are obtained, which can be used to estimate the marginal density $f(\theta_s)$, for any s. The marginal density is then estimated as follows

$$\hat{f}(\theta_s) = \frac{1}{G} \sum_{g=1}^{G} f(\theta_{s(l)}^{(g)} \mid \theta_{r(l)}^{(g)}, \ r \neq s). \tag{3.4}$$

From the above process for Gibbs sampling we can know that the Gibbs sampling involves generating random variates from all complete conditional densities of (3.1). Thus the implementation of Gibbs sampling is straightforward under the assumption of conjugate priors or the form of standard distributions. In other cases, more sophisticated random variate generating methods such as the rejection method(Ross, 1993) or the generalized ratio of uniforms methods(Wakefield, et al., 1994) can be used to generate random variates from non-normalized densities. But these methods require locating the mode of non-normalized density. The adaptive rejection sampling algorithm(Gilks and Wild, 1992) does not require the mode of the density but the log-concavity of the density. Since the griddy Gibbs sampling by Ritter and Tanner(1992) requires neither the mode of the non-normalized density nor the log-cancavity, this method is applicable if only the non-normalized density can be calculated.

Ritter and Tanner(1992) considered the following griddy Gibbs sampling algorithm :

Step 1. Calculate $p(\theta_i \mid \theta_1, \theta_2, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_k)$ at $\theta_i = \phi_1, \phi_2, \cdots, \phi_m$, to obtain $w_1, w_2, \cdots, w_m$.

Step 2. Use $w_1, w_2, \cdots, w_m$ to obtain an approximation to the inverse c.d.f. of $p(\theta_i \mid \theta_1, \theta_2, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_k)$.

Step 3. Draw an uniform (0, 1) random variate and transform the observation via the approximate inverse c.d.f.

Usual simple approximate c.d.f's are based on a discrete distribution for $\phi_1, \phi_2, \cdots, \phi_m$ with probabilities $p(\phi_i) = w_i / \sum_{j=1}^{m} w_j$ or an uniform distribution on the interval $[a_i, a_{i+1}]$,

$i=1,2,\cdots,m$ with $p(\phi_i)=w_i \Big/ \left\{ \sum_{j=1}^{m} w_j(a_{j+1}-a_j) \right\}$, where $\phi_i \in [\, a_i,\ a_{i+1}\,]$. Though more sophisticated approximations to the c.d.f. can be used besides the above simple approximations, we simply use the latter method in the griddy Gibbs sampling applied to the NLAR(1) model in the next section.

4. Simulation study

In this section, we illustrate how the Bayesian estimates obtained via the griddy Gibbs sampling cope with the NLAR(1) model of (2.1). We consider four NLAR(1) models with (α,β) = (0.1, -0.2), (0.3, 0.4), (0.5, -0.6), and (0.9, 0.8) dealt with in the maximum likelihood estimation of Son and Cho(1995). The samples of size n = 100 are generated for each model. Given the observations $\underline{x}=(x_1,x_2,\cdots,x_{100})$ for each model, our strategy for the griddy Gibbs sampling is as follows : we use the CLS estimate $\hat{\beta}_{cls}$ as an initial value $\beta_{(0)}$ of β. Now, let's begin by letting $l=0$.

● A procedure for drawing α from $f(\alpha \mid \beta, \underline{x})$ of (2.4).

Step 1. Generate m uniform (0, 1) random variates, $u_1,u_2,\cdots,u_m$.

Step 2. Compute $w_j=f_{\alpha|\beta,\underline{x}}(u_j \mid \beta_{(l)}, \underline{x})$, $j=1,2,\cdots,m$.

Step 3. Obtain a piecewise linear approximate c.d.f. $F_\alpha(y)$ using $w_1,w_2,\cdots,w_m$, where

$$F_\alpha(y)=\begin{cases} 0 & ,y<a_1 \\ c_i+d_iy & ,a_i \le y < a_{i+1},\ i=1,2,\cdots,m \\ 1 & ,y \ge a_{m+1}. \end{cases} \tag{4.1}$$

Here,
$$c_i=\frac{\sum_{j=1}^{i-1} w_j(a_{j+1}-a_j)-w_ia_i}{\sum_{j=1}^{m} w_j(a_{j+1}-a_j)}, \tag{4.2}$$

$$d_i=\frac{w_i}{\sum_{j=1}^{m} w_j(a_{j+1}-a_j)}, \tag{4.3}$$

and

$$\begin{cases} a_1 = 0 \\ a_j = \frac{1}{2}(u_{j-1} + u_j), \ j = 2, \cdots, m \\ a_{m+1} = 1. \end{cases} \tag{4.4}$$

We note that w_j of c_i and d_i can be replaced by the likelihood function value $l(u_j, \beta_{(l)})$, which implies that we need to know at least the proportionality constant of the postorior density.

Step 4. Draw an uniform (0, 1) random variate u^* and transform a simulated observation $\alpha_{(l+1)}$ via the approximate inverse c.d.f., i.e.,

$$\alpha_{(l+1)} = F_\alpha^{-1}(u^*) = \frac{u^* - c_i}{d_i}, \text{ if } a_i \le u^* < a_{i+1}, \ i = 1,2,\cdots,m. \tag{4.5}$$

● A procedure for drawing β from $f(\beta \mid \alpha, x)$ of (2.5).

Step 5. Generate m uniform random variates $v_1, v_2, \cdots, v_m$, where $v_j \sim U(-1, 1)$, $v_j \neq 0$.

Step 6. Compute $w_j = l(\alpha_{(l+1)}, v_j)$, $j = 1,2,\cdots,m$.

Step 7. Obtain a piecewise linear approximate c.d.f. $F_\beta(y)$ as same as the equations (4.1)-(4.4) of Step 3 after setting $u_j = v_j$, $w_j = l(\alpha_{(l+1)}, v_j)$, $a_1 = -1$, and $a_{m+1} = 1$.

Step 8. Draw an uniform (0, 1) random variate u^* and transform a simulated observation,

$$\beta_{(l)} = F_\beta^{-1}(u^*) = \frac{u^* - c_i}{d_i}, \text{ if } a_i \le u^* < a_{i+1}, \ i = 1,2,\cdots,m.$$

For our analysis we conducted the above entire process through l=100 iterations. At each iteration of the griddy Gibbs sampling algorithm G=2000 independent parallel samples are generated for each model. Thus the resulting data from the each posterior density are G=2000 sampled pairs drawn on the l=100th iteration. The algorithm has been implemented in FORTRAN on a PC/pentium computer. Uniform (0, 1) random variates are generated using PC/IMSL subrutine, RNUN. On convergence, Figure 1 displays the plots of sample quartiles for α and β sampled on each iteration. The dispersions of the posterior densities become samaller as $|\alpha\beta|$ approches 1. Also sample quartiles of the model with $\alpha = 0.1$, $\beta = -0.2$

($|\alpha\beta| = 0.02$) still show cyclical fluctuations even on considerably large iteration. This result can be explained by the fact that the NLAR(1) process with $|\alpha\beta| \approx 0$ shows the behaviors of Laplacian random process. On the other hand, sample quratiles with $\alpha = 0.9$, $\beta = 0.8$ ($|\alpha\beta| = 0.72$) converge on the first iteration. These features are consistent with the results of the CLS and ML esmation. For each model, we calculated the median, mode, mean, standard deviation, the 5th and 95th percentile of simulated samples from each posterior density for α and β. These results are listed in Table 1. The means and standard deviations of the CLS and the ML estimates with 100 replications are also shown in Table 1.

In case of assuming uniform priors, the maximization of the posterior density is equivalent to the maximization of the likelihood function of one parameter with others fixed. Thus, we adopt the modes of posterior densities as the estimates of parameters. The modes of posterior densities are estimated from the probability histograms of each posterior density in Figure 2 which was ploted using PROC GCHART of SAS/GRAPH. From the results of Table 1, we know that the estimates by the griddy Gibbs sampling are well incorporated. Especially, the griddy Gibbs sampling in case of small $|\alpha\beta|$ value gives better estimates than the CLS and the ML method.

Table 1.
Result of a simulation study. Standard deviation(sd.) is in parentheses.

	Characteristics of posterior density						
model	median	mode	mean(sd.)	90% interval		CLS(sd.)	ML(sd.)
α = 0.1	0.293	0.20	0.339(0.217)	[0.063,	0.803]	0.657(0.413)	0.643(0.356)
β = -0.2	0.010	-0.10	0.013(0.395)	[-0.649,	0.629]	-0.073(0.481)	-0.018(0.396)
α = 0.3	0.341	0.25	0.371(0.208)	[0.082,	0.760]	0.603(0.381)	0.508(0.299)
β = 0.4	0.333	0.40	0.286(0.309)	[-0.295,	0.731]	0.294(0.462)	0.334(0.424)
α = 0.5	0.535	0.55	0.531(0.170)	[0.243,	0.814]	0.657(0.292)	0.559(0.175)
β = -0.6	-0.491	-0.56	-0.468(0.192)	[-0.743,	-0.118]	-0.559(0.248)	-0.604(0.179)
α = 0.9	0.830	0.87	0.821(0.082)	[0.674,	0.939]	0.867(0.141)	0.891(0.044)
β = 0.8	0.775	0.78	0.746(0.153)	[0.451,	0.944]	0.825(0.097)	0.806(0.054)

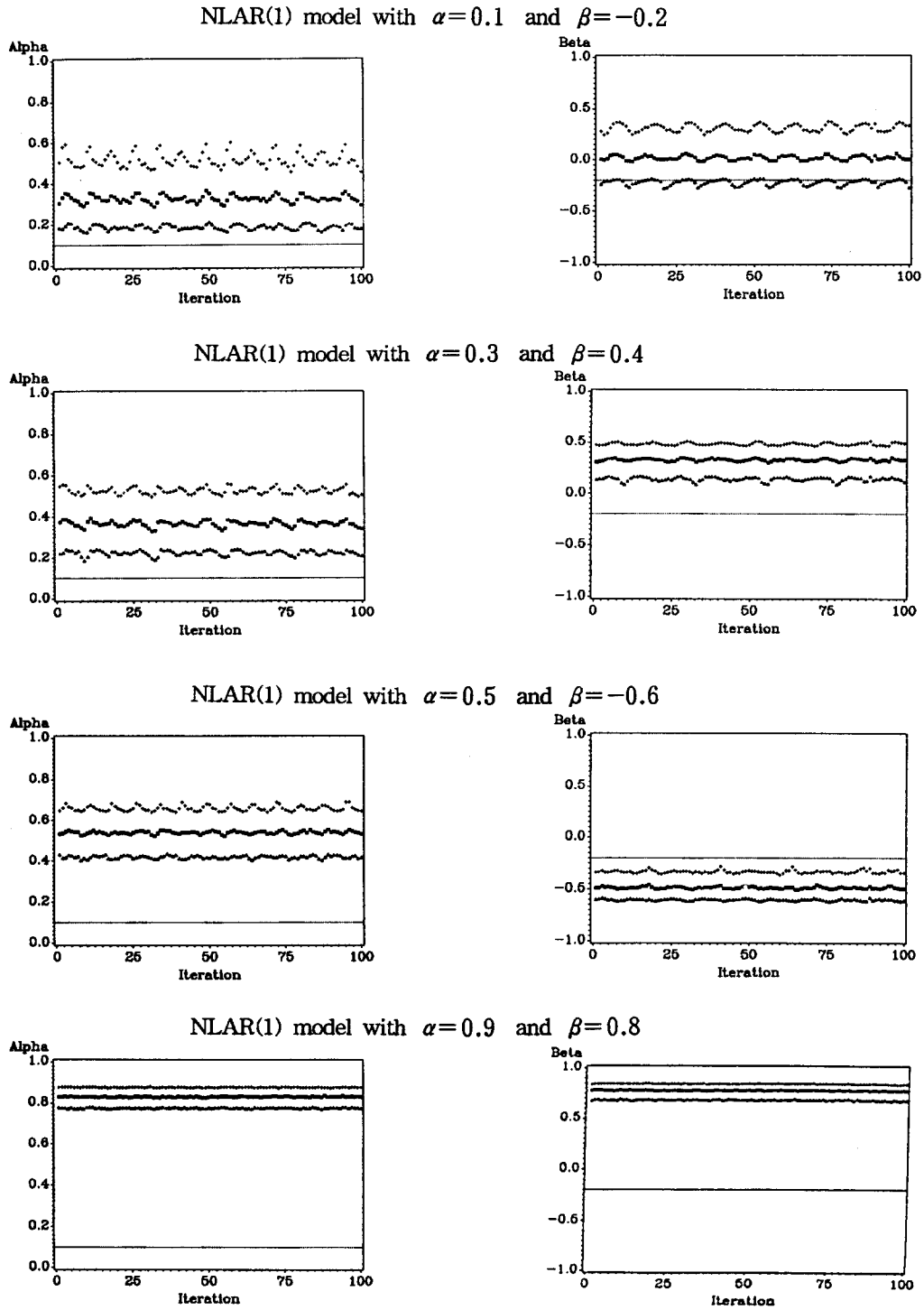
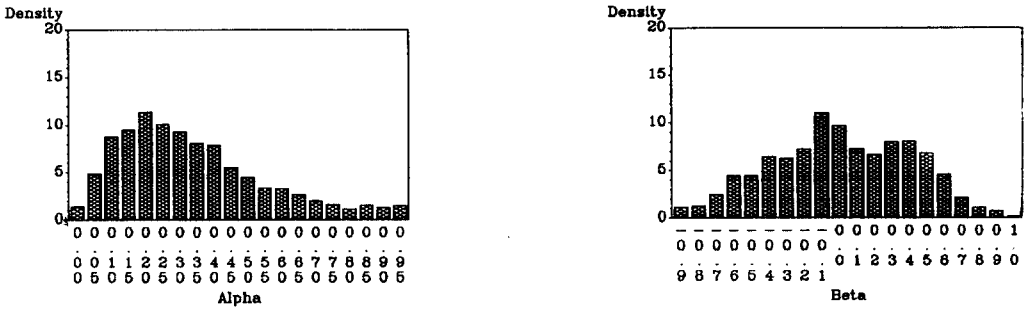
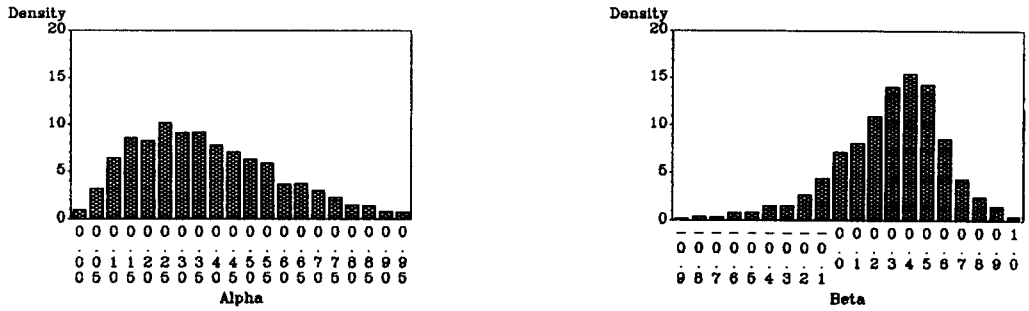


Figure 1. Sample quartiles(vertical axis) vs. iteration(horizontal axis). Sample size n=100, m=10 grids, l=100 iterations, G =2000 independent samples.

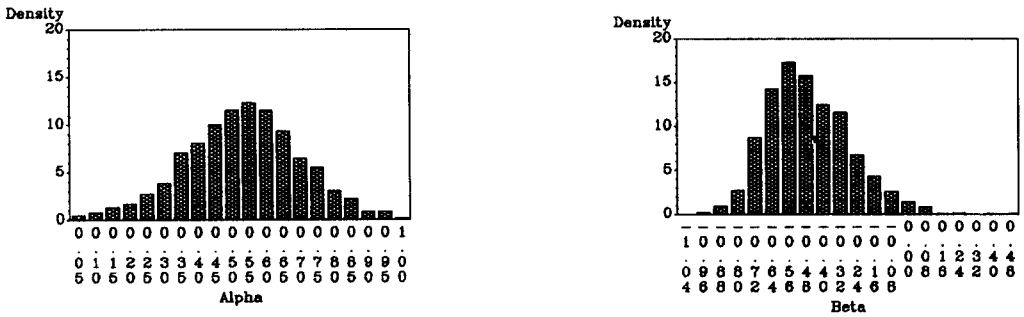
NLAR(1) model with $\alpha=0.1$ and $\beta=-0.2$



NLAR(1) model with $\alpha=0.3$ and $\beta=0.4$



NLAR(1) model with $\alpha=0.5$ and $\beta=-0.6$



NLAR(1) model with $\alpha=0.9$ and $\beta=0.8$

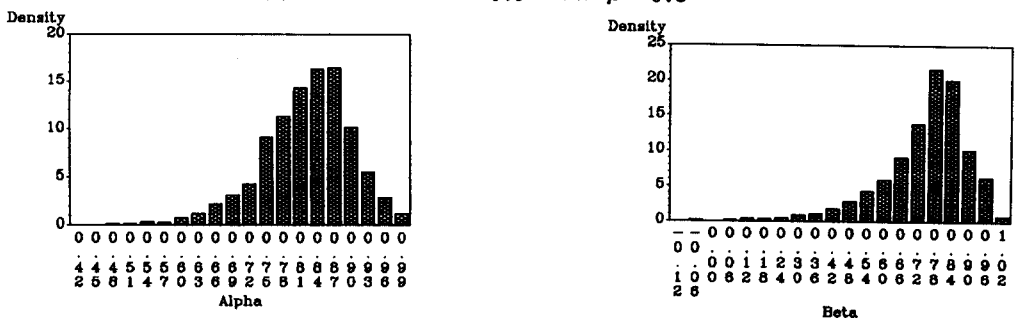


Figure 2. Posteiror densities of α and β obtained using the griddy Gibbs sampling. Sample size n=100, m=10 grids, l=100 iterations, G =2000 independent samples.

5. Discussion and futher sudy

We have used the griddy Gibbs sampling to obtain Bayesian estimates for the NLAR(1) model. Under the uniform priors assumption, it does not require any oprimization theory. It suffices to have the likelihood function. But, we can not but point out that this griddy Gibbs sampling is very computationally intensive and takes a lot of time. For example, it took 1.8 seconds on a PC/pentium computer to obtain the ML estimates for four NLAR(1) models, but on the other hand 13 hours to obtain Bayesian estimates by the griddy Gibbs sampling with $m=10$, $l=100$, $G=2000$. There are little or no difficulties in applying the griddy Gibbs sampling if the computer can intensively be used and enough time can be given.

The Bayesian estimation via the griddy Gibbs sampling considered in this paper can be applied to the New Exponential AutoRegressive(NEAR) model with exponential marginals which has a difficulty of discontinuity in its likelihood function. Also, we expect that the Bayesian estimation via the griddy Gibbs sampling be well performed irrespective of the troublesome ploblems entailed in the Bayesian estimation of random coefficient autoregressive time series model where the random coefficients have a truncated normal distribution(Son, 1994) or a rescaled beta distribution(Liu and Tiao(1980)). Research on these subjects is under study and will be reported in the near future.

References

[1] Dewald, L.S. and Lewis, P.A.W. (1985). A new Laplace second-order autoregressive time series model-NLAR(2), *IEEE Transactions Information Theory*, IT-31, No. 5, 645-651.

[2] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs dsitributions, and the Bayesian restoration of images, *IEEE transaction on pattern analysis and machine intelligence,* Vol. PAMI-6, No. 6, 721-741.

[3] Gilks, W.R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling, *Applied Statistics*, Vol. 41, No. 2, 337-348.

[4] Karlsen, H. and Tjøstheim, D. (1988). Cinsistent estimates for the NEAR(2) and NLAR(2) time series models, *Journal of Royal Statistical Society*, Series B, Vol. 50, No. 2, 313-320.

[5] Liu, L. M. and Tiao, G. C. (1980). Random coefficient first-order autoregressive models, *Journal of Econometrics*, Vol. 13, 305-325.

[6] Ritter, C. and Tanner, M.A. (1992). Facilitating the Gibbs sampler : The Gibbs stopper and the griddy-Gibbs sampler, *Journal of the American Statistical Association,* Vol. 87, No. 419, 861-868.

[7] Ross, S.M. (1993). A course in simulation, Macmillian publishing Company, New York.

[8] Son, Y.S. (1994). Estimation of random coefficient AR(1) model for panel data, *Proceedings of Aurumn Conference, Korean Statistical Society,* 71-80.

[9] Son, Y.S. and Cho, S.S. (1988). A Laplacian autoregressive time series model, *Journal of the Korean Statistical Society,* Vol. 17, No. 2, 101-120.

[10] Son, Y.S. and Cho, S.S. (1995). Maximum likelihood estimation for the Laplacian autoregressive time model, unpublished manuscript.

[11] Wakefield, J.C. Smith, A.F.M. Racine-Poon, A. and Gelfand, A.E. (1994). Bayesian analysis of linear and non-linear population models by using the Gibbs sampler, *Applied statistics,* Vol. 43, No. 1, 201-221.