<Research Paper>

Journal of the Korean
Statistical Society
Vol. 14, No. 2, 1985

Bayesian Analysis of GLEM with Half-Normal Prior

Samir K. Bhattacharya*
and Ram Lal**

ABSTRACT

In this paper, Bayesian analysis of the general linear econometric model is carried out by using a multinormal prior for the vector of unknown regression coefficients and a half-normal prior for the standard deviation of the disturbances.

1. Introduction

Consider the general linear econometric model (GLEM) specified by

$$\underset{\sim}{y}=X\underset{\sim}{\beta}+\underset{\sim}{u} \tag{1.1}$$

where $\underset{\sim}{y}'=(y_1,y_2,\dots,y_n)$ is a point in R^n and represents observations on an endogenous variable, $X=[X_{ij}]$, $i=1,2,\dots n$, $j=1,2,\dots,p$, is an $n\times p$ matrix of observations on p exogenous variables, $\underset{\sim}{\beta}'=(\beta_1,\beta_2,\dots\beta_p)$ is a p-dimensional vector of unknown parameters and $\underset{\sim}{u}'=(u_1,u_2,\dots,u_n)$ is an n-dimensional vector of unobservable disturbances or errors. It is assumed that Rank$(X)=p$, and that elements of X are nonstochastic and independent of $\underset{\sim}{u}$, whose elements are iid normal $(0,\sigma^2)$. In this paper, a Bayesian estimator of $\underset{\sim}{\beta}$ is obtained when the prior pdf of the unknown parameter $\underset{\sim}{\beta}$, given σ, is multinormal, and for the prior density of σ, we take a half-normal density $g(\sigma)$ with a parameter "a" as given below:

$$g(\sigma)=\left(\frac{2}{\pi}\right)^{1/2}\frac{1}{a}e^{-\frac{\sigma^2}{2a^2}}\quad (0<\sigma<\infty\ ;\ a>0) \tag{1.2}$$

In obtaining the Bayes estimator, it is assumed that the loss function is squared error.

* Department of Mathematics, Allahabad University, India.
** Allahabad Agricultural Institute, Allahabad, India.

2. Bayes Estimator

In what follows, the prior pdf of $\underset{\sim}{\beta}$, given σ, is taken to be a multinormal density

$$g(\underset{\sim}{\beta}|\sigma)\propto\frac{1}{\sigma^p}\exp\left[-\frac{V}{2\sigma^2}\right], \tag{2.1}$$

where

$$V=(\underset{\sim}{\beta}-\bar{\underset{\sim}{\beta}})'A(\underset{\sim}{\beta}-\bar{\underset{\sim}{\beta}}). \tag{2.2}$$

Here $\bar{\underset{\sim}{\beta}}$ is mean vector of the multinormal prior in (2.1), and A represents the prior covariance matrix of $\underset{\sim}{\beta}$. It is assumed that $\bar{\underset{\sim}{\beta}}$ and A are known on the basis of prior knowledge. The prior $g(\sigma)$ is given in (1.2), and the parameter "a" in (1.2) is also assumed known on the basis of prior knowledge. The joint prior pdf on the parameter space of $(\underset{\sim}{\beta},\sigma)$ is given by

$$g(\underset{\sim}{\beta},\sigma)\propto\frac{1}{\sigma^p}\exp\left[-\frac{V}{2\sigma^2}-\frac{\sigma^2}{2a^2}\right] \tag{2.3}$$

To compute the Bayesian posterior, we need the likelihood function. Under the assumptions stated earlier, the kernel of the likelihood function turns out to be

$$l\propto\frac{1}{\sigma^n}\exp\left[-\frac{W}{2\sigma^2}\right], \tag{2.4}$$

where

$$W=[(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})'X'X(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})+(n-p)S^2], \tag{2.5}$$

$$(n-p)S^2=(\underset{\sim}{y}-X\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y}-X\hat{\underset{\sim}{\beta}}), \tag{2.6}$$

and
$$\hat{\underset{\sim}{\beta}}=(X'X)^{-1}X'\underset{\sim}{y} \tag{2.7}$$

is the OLS estimator of $\underset{\sim}{\beta}$. The transition to Bayesian posterior $g^*(\underset{\sim}{\beta},\sigma)$ is made in the usual manner to obtain

$$g^*(\underset{\sim}{\beta},\sigma)\propto\frac{1}{\sigma^{n+p}}\exp\left[-\frac{(V+W)}{2\sigma^2}-\frac{\sigma^2}{2a^2}\right]. \tag{2.8}$$

To obtain the marginal posterior density of $\underset{\sim}{\beta}$, (2.8) has to be integrated over $0<\sigma<\infty$. This integral is evaluated in terms of the modified Bessel function of the third kind $K_\nu(z)$ (Cf. Erdélyi, 1953, p. 5, formula (13)) by using its integral representation (Erdélyi, Op. Cit., p. 82, formula (23)) given by

$$\int_0^\infty v^{-\nu-1}e^{-\frac{z}{2}\left(v+\frac{\mu^2}{v}\right)}\,dv=\frac{2}{\mu^\nu}K_\nu(\mu z) \tag{2.9}$$

where $Re(z)>0$, and $Re(\mu^2 z)>0$. Using this result, the Bayesian posterior density of $\underset{\sim}{\beta}$ is obtained (see Appendix A.1 for derivation) from (2.8) as

$$g^*(\underset{\sim}{\beta}) \propto \frac{K_{\frac{n+p-1}{2}}\left(\sqrt{\frac{V+W}{a^2}}\right)}{(V+W)^{\frac{n+p-1}{4}}} \tag{2.10}$$

which is defined for $\beta_j \in (-\infty, \infty)$, $j=1, 2, \ldots, p$. We now use (2.2) and (2.5) along with a lemma of Box and Tiao (1973, p. 418, lemma 1) to obtain

$$V+W = b_0 + (\underset{\sim}{\beta} - \underset{\sim}{d})'(A+X'X)(\underset{\sim}{\beta} - \underset{\sim}{d}) \tag{2.11}$$

where

$$\underset{\sim}{d} = (A+X'X)^{-1}(A\underset{\sim}{\bar{\beta}} + X'X\underset{\sim}{\hat{\beta}}), \tag{2.12}$$

and

$$b_0 = (\underset{\sim}{\bar{\beta}} - \underset{\sim}{\hat{\beta}})\ A(A+X'X)^{-1}X'X(\underset{\sim}{\bar{\beta}} - \underset{\sim}{\hat{\beta}}) + (\underset{\sim}{y} - X\underset{\sim}{\hat{\beta}})'(\underset{\sim}{y} - X\underset{\sim}{\hat{\beta}}). \tag{2.13}$$

Hence, from (2.10) and (2.11), we have

$$g^*(\underset{\sim}{\beta}) \propto \frac{K_{\frac{n+p-1}{2}}\left(\sqrt{\frac{1}{a^2}[b_0 + (\underset{\sim}{\beta} - \underset{\sim}{d})'(A+X'X)(\underset{\sim}{\beta} - \underset{\sim}{d})]}\right)}{[b_0 + (\underset{\sim}{\beta} - \underset{\sim}{d})'(A+X'X)(\underset{\sim}{\beta} - \underset{\sim}{d})]^{\frac{n+p-1}{4}}} \tag{2.14}$$

which is defined for $\beta_j \in (-\infty, \infty)$, $j=1, 2, \ldots, p$. It is known that under the assumption of squared error loss function, the Bayes estimator of $\underset{\sim}{\beta}$ is simply the posterior expectation of $\underset{\sim}{\beta}$ calculated from $g^*(\underset{\sim}{\beta})$ in (2.14). This evaluation is carried out with the help of a result of Bhattacharya and Saxena (1985) and the details are given in Appendix A.2. The final result is

$$E(\underset{\sim}{\beta}) = \underset{\sim}{d} = (A+X'X)^{-1}(A\underset{\sim}{\bar{\beta}} + X'\underset{\sim}{y}) \tag{2.15}$$

for the Bayes estimator of $\underset{\sim}{\beta}$. It may be pointed out that this estimator (2.15) is mathematically somewhat more general than the Hoerl and Kennard's (1970) ridge estimator

$$\underset{\sim}{\hat{\beta}}_R = (X'X+kI)^{-1}X'y \quad (k>0) \tag{2.16}$$

proposed for situations wherein the matrix $X'X$ is ill-conditioned (cf. Riley 1955, Lawless 1976). However, it should be understood clearly that the resemblance between the Bayes estimator (2.15) and the ridge estimator (2.16) is merely mathematical and the philosophies involved in the two approaches are radically different. While Bayesian inference is relevant for a single observed sample only, the sample space is deemed very important in classical inference.

Appendix A.1

The Bayesian posterior density of $\underset{\sim}{\beta}$ is given by

$$g^*(\underset{\sim}{\beta})=\int_0^\infty g^*(\underset{\sim}{\beta},\sigma)\,d\sigma$$

$$\propto\int_0^\infty \sigma^{-(n+p)}\exp\left[-\frac{(V+W)}{2\sigma^2}-\frac{\sigma^2}{2a^2}\right]d\sigma$$

$$\propto\int_0^\infty v^{-\frac{(n+p+1)}{2}}\exp\left[-\frac{(V+W)}{2v}-\frac{v}{2a^2}\right]dv \tag{A.1.1}$$

This integral is evaluated by writing $\nu=\frac{n+p-1}{2}$, $z=\frac{1}{a^2}$, and $\mu^2=a^2(V+W)$ in (2.9) of the paper, so that we obtain

$$g^*(\underset{\sim}{\beta})\propto\frac{2}{a^{\frac{n+p-1}{2}}[V+W]^{\frac{n+p-1}{2}}}K_{\frac{n+p-1}{2}}\left(\sqrt{\frac{V+W}{a^2}}\right)$$

$$\propto\frac{K_{\frac{n+p-1}{2}}\left(\sqrt{\frac{V+W}{a^2}}\right)}{[V+W]^{\frac{n+p-1}{4}}} \tag{A.1.2}$$

as desired.

Appendix A.2

The posterior expectation of $\underset{\sim}{\beta}$ can be computed from the Bayesian posterior density $g^*(\underset{\sim}{\beta})$ in (A.1.2) by using the results of Bhattacharya and Saxena (1985). The details are too long to be presented here in full, but we shall sketch the main derivation. Bhattacharya and Saxena (op. cit.) define a multivariate modified Bessel distribution with parameters $(\nu,\underset{\sim}{a},A,\lambda,b)$ by the joint density function of $(X_1,X_2,\dots,X_p)$ as

$$f(\underset{\sim}{x}')=C\cdot\frac{K_\nu(\sqrt{2\lambda[b+Q]})}{[b+Q]^{\nu/2}} \tag{A.2.1}$$

where

$$C=\frac{|A|^{1/2}}{\pi^{p/2}}\left(\frac{\lambda}{2}\right)^{p/2}\cdot\frac{b^{1/2(\nu-\frac{p}{2})}}{K_{\nu-\frac{p}{2}}(\sqrt{2b\lambda})} \tag{A.2.2}$$

and $\underset{\sim}{x}'=(x_1,x_2,\dots,x_p)\in R^p$, $\underset{\sim}{a}'=(a_1,a_2,\dots a_p)\in R^p$, b and λ are real positive scalars, and A is a symmetric square matrix of order p such that

$$Q=(\underset{\sim}{x}-\underset{\sim}{a})'A(\underset{\sim}{x}-\underset{\sim}{a}) \tag{A.2.3}$$

is a positive definite quadratic form over the field of real numbers. Then, they obtain

$$E(X_j-a_j)=\int_{R^p}\!\!\cdots\!\int(x_j-a_j)f(\underset{\sim}{x}')\,d\underset{\sim}{x}'$$

$$=C\int_{R^p}\!\!\cdots\!\int y_j\frac{K_\nu(\sqrt{2[b+\underset{\sim}{y}'A\underset{\sim}{y}]})}{[b+\underset{\sim}{y}'A\underset{\sim}{y}]^{\nu/2}}\,d\underset{\sim}{y}' \tag{A.2.4}$$

Since A is symmetric positive definite, there exists a real nonsingular matrix B of order p such that $A=B'B$. Clearly $|A|=|B|^2$. We then apply the transformation $\underset{\sim}{z}'=\underset{\sim}{y}B'$, so that $\underset{\sim}{y}'A\underset{\sim}{y}=\underset{\sim}{y}'B'B\underset{\sim}{y}=\underset{\sim}{z}'\underset{\sim}{z}$. The Jacobian of this transformation is

$$J=\frac{\partial(z_1, z_2, \ldots, z_p)}{\partial(y_1, y_2, \ldots, y_p)}=|B|=|A|^{1/2} \tag{A.2.5}$$

Since y_j is linear in z_i, that is $y_j=\sum_{i=1}^{p} l_i z_i$, equation (A.2.4) yields

$$E(X_j-a_j)=\frac{C}{|A|^{1/2}}\int\cdots\int_{R^p}\sum_{i=1}^{p} l_i z_i \frac{K_\nu\left(\sqrt{2\lambda\left[b+\sum_{j=1}^{p} z_j^2\right]}\right)}{\left[b+\sum_{j=1}^{p} z_j^2\right]^{\nu/2}}\, d\underset{\sim}{z}'=0 \tag{A.2.6}$$

so that $E(X_j)=a_j$, for $j=1, 2, \ldots, p$, that is, we have

$$E(\underset{\sim}{X}')=\underset{\sim}{a}' \tag{A.2.7}$$

We apply this result to the posterior density in (2.14) of the paper to obtain

$$\begin{aligned} E(\underset{\sim}{\beta})=\underset{\sim}{d}&=(A+X'X)^{-1}(A\bar{\underset{\sim}{\beta}}+X'X\hat{\underset{\sim}{\beta}}) \\ &=(A+X'X)^{-1}(A\bar{\underset{\sim}{\beta}}+X'\underset{\sim}{y}), \end{aligned}$$

proving the result in (2.15) of the paper.

References

(1) Bhattacharya, S.K. and Saxena, A.K. (1985). A Modified Bessel Integral with Statistical Applications, Communicated.

(2) Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis*, Addison-Wesley, London.

(3) Erdélyi, A. (1985). *Higher Transcendental Functions*, Vol. Ⅱ, McGraw-Hill, N.Y.

(4) Hoerl, A.E. and Kennard, R.W. (1970). Ridge Regression: Biased Estimation of Nonorthogonal Problems, *Technometrics*, Vol. 12, 55~67.

(5) Lawless, J.F. and Wang, P. (1976). A Simulation Study of Ridge and Other Regression Estimators, *Comm. Stat.*, Vol. 1, 307~323.

(6) Riley, J.D. (1955). Solving Systems of Linear Equations with a Positive Definite Symmetric but Possibly Ill-Conditioned Matrix, *Mathematics of Computation*, Vol. 9, 96~101