STATISTICS IN TRANSITION new series, June 2016
Vol. 17, No. 2, pp. 211–219

BAYESIAN INFERENCE FOR STATE SPACE MODEL WITH PANEL DATA

Ranjita Pandey[1], Anoop Chaturvedi[2]

ABSTRACT

The present work explores panel data set-up in a Bayesian state space model. The conditional posterior densities of parameters are utilized to determine the marginal posterior densities using the Gibbs sampler. An efficient one step ahead predictive density mechanism is developed to further the state of art in prediction-based decision making.

Key words: Bayesian analysis, Gibbs sampler, conditional posterior densities, predictive distribution.

1. Introduction

The importance of panel data stems from its ability for addressing questions of economic and social behaviour which cannot be easily answered by using the usual cross section or time series data. The panel data consists of the observations on the same cross section of units under study at different and usually successive time periods. The longitudinal nature of panel data allows for the use of simple techniques to solve otherwise complicated problems and permits cross section and/or time heterogeneity. Tiwari, Yang and Zalkikar (1996) have studied the level of water pollution by recording biochemical oxygen demand (BOD) and dissolved oxygen (DO), at a selected point along the stream at different time points. The present paper adapts and extends their concept to a multiple point scenario. The situation could be visualized in a comprehensive manner, if the BOD and DO measurements are taken at more than one selected point along the length of the stream, at successive time periods. The resulting longitudinal data accounts for the individual effects at the various locations as well as considers impact of other relatively slowly changing left-out variables. Panel data-based studies have been undertaken by Maddala (1971), Mundlak (1978), Hausman (1978), Hausman and Taylor (1981) and Chamberlain (1982) in various fields. Baltagi (2008) advocates the use of panel data for controlling individual

[1] Department of Statistics, University of Delhi. E-mail: ranjitapandey111@gmail.com.
[2] Department of Statistics, University of Allahabad. E-mail: anoopchaturv@gmail.com.

heterogeneity and to incorporate dynamic adjustment in the model while elaborating that the panel data structure gives more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency compared to the classical cross sectional model-based study.

2. Model with panel data and assumptions

The variables involved in the model are denoted by

y_{it} : observed value of the dependent variable for unit i at time point t.

x_{it} : $p\times 1$ vector of observations on p explanatory variable corresponding to the i-th unit at time t, $(i = 1,\ldots,n;\ t = 1,\ldots,T)$

The state space model for panel data is given by

$$y_{it} = x'_{it}\theta_t + \varepsilon_{it} \tag{2.1}$$

The dynamics of the system as it evolves through time is represented by

$$\theta_t = G_t\,\theta_{t-1} + v_t \tag{2.2}$$

G_t is a known $p\times p$ transition matrix, θ_t is a $p\times 1$ unknown parameter vector, θ_0 denotes the initial state of the system, v_t is the systems error, ε_{it} is the observation error. We write ε_{it} as

$$\varepsilon_{it} = \alpha_i + \eta_{it} \tag{2.3}$$

We also assume that α_i, η_{it} and v_t are all independently distributed. Hence, equation (2.1) may be rewritten as

$$y_{it} = x'_{it}\theta_t + \alpha_i + \eta_{it} \tag{2.4}$$

We further assume that

$$\alpha_i \sim N\left(0, \frac{\sigma_\alpha^2}{\lambda}\right) \quad \text{for all } i \tag{2.5}$$

$$\eta_{it} \sim N\left(0, \frac{\sigma_\eta^2}{\lambda}\right) \quad \text{for all } i \text{ and } t \tag{2.6}$$

$$v_t \,|\, \lambda \sim N(0, \lambda^{-1}\Sigma) \quad \text{for all } t \tag{2.7}$$

$$\lambda \sim G\left(\frac{a_0}{2}, \frac{b_0}{2}\right) \tag{2.8}$$

$$\theta_0 \,|\, \lambda \sim N(m_0, \lambda^{-1}\Sigma_0) \tag{2.9}$$

Thus, $\theta_t \,|\, \theta_{t-1}, \lambda \sim N(G_t\theta_{t-1}, \lambda^{-1}\Sigma)$ (2.10) where Σ is a known $p \times p$ positive definite matrix. We can write the model (2.1) as

$$y_t = X_t \theta_t + \varepsilon_t \tag{2.11}$$

and

$$\varepsilon_t = \alpha + \eta_t \tag{2.12}$$

where $y_t = (y_{1t}, \ldots, y_{nt})' \;:\; n \times 1,\; X_t = \begin{bmatrix} x'_{1t} \\ \vdots \\ x'_{nt} \end{bmatrix} \;:\; n \times p$

$\alpha = (\alpha_1, \ldots, \alpha_n)' \;:\; n \times 1,\; \eta_t = (\eta_{1t}, \ldots, \eta_{nt})' \;:\; n \times 1,\; \varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})' \;:\; n \times 1$

The distributions of α and η_t are given by

$$\alpha \sim N_n \left(0, \frac{\sigma_\alpha^2}{\lambda} I_n \right) \tag{2.13}$$

$$\eta_t \sim N_n \left(0, \frac{\sigma_\eta^2}{\lambda} I_n \right) \tag{2.14}$$

3. Conditional Posterior densities

In the posterior analysis of the model we treat α as an unknown parameter and derive its conditional posterior density also along with the conditional posterior densities of other parameters, and utilize these conditional posterior densities in employing Gibbs sampler. Under model specifications and the underlying assumptions, we have

$$f^*(y_t | \{\theta_t\}_{t=0}^T, \alpha, \lambda) = \frac{\lambda^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}} \sigma_\eta^n} \exp\left[-\frac{\lambda}{2\sigma_\eta^2} (y_t - X_t\theta_t - \alpha)'(y_t - X_t\theta_t - \alpha) \right] \tag{3.1}$$

so that $$f^*(\underset{\sim}{y} | \{\theta\}_{t=0}^T, \alpha, \lambda) = \frac{\lambda^{\frac{nT}{2}}}{(2\pi)^{\frac{nT}{2}} \sigma_\eta^{nT}} \exp\left[-\frac{\lambda}{2\sigma_\eta^2} \sum_{t=1}^T (y_t - X_t\theta_t - \alpha)'(y_t - X_t\theta_t - \alpha) \right] \tag{3.2}$$

The joint density of $(\underset{\sim}{y}, \alpha, \lambda, \theta_0, \{\theta_t\}_{t=1}^T)$, obtained by combining expressions (2.8), (2.9), (2.10), (2.13) and (3.2), is given as follows

$$f(\underset{\sim}{y},\alpha,\lambda,\theta_0,\{\theta_t\}_{t=1}^{T} = \frac{\lambda^{\frac{nT}{2}}}{(2\pi)^{\frac{nT}{2}}\sigma_\eta^{nT}}\exp\left[-\frac{\lambda}{2\sigma_\eta^2}\sum_{t=1}^{T}(y_t - X_t\theta_t - \alpha)'(y_t - X_t\theta_t - \alpha)\right]\times$$

$$\frac{\lambda^{\frac{pT}{2}}}{(2\pi)^{\frac{pT}{2}}\Sigma^{\frac{T}{2}}}\exp\left[-\frac{\lambda}{2}\sum_{t=1}^{T}(\theta_t - G_t\theta_{t-1})'\Sigma^{-1}(\theta_t - G_t\theta_{t-1})\right]\times$$

$$\frac{\lambda^{\frac{p}{2}}}{(2\pi)^{\frac{p}{2}}\Sigma^{\frac{1}{2}}}\exp\left[-\frac{\lambda}{2}(\theta_0 - m_0)'\Sigma_0^{-1}(\theta_0 - m_0)\right]\times$$

$$\left(\frac{b_0}{2}\right)^{\frac{a_0}{2}}\times\frac{1}{\Gamma\left(\frac{a_0}{2}\right)}\times\lambda^{\frac{a_0}{2}-1}\exp\left[-\frac{b_0}{2}\lambda\times\right]\frac{\lambda^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}\sigma_\alpha^n}\exp\left[-\frac{\lambda}{2}\left(\frac{\alpha'\alpha}{\sigma_\alpha^2}\right)\right]$$

(3.3)

We define

$$B_t^{-1} = \begin{cases} \Sigma_0^{-1} + G_1'\Sigma^{-1}G_1 & \text{, for } t = 0 \\ \Sigma^{-1} + \frac{1}{\sigma_\eta^2}X_t'X_t + G_{t+1}'\Sigma^{-1}G_{t+1} & \text{, for } t = 1,2,\ldots,T-1 \\ \Sigma^{-1} + \frac{1}{\sigma_\eta^2}X_T'X_T & \text{, for } t = T \end{cases} \quad (3.4)$$

$$b_t = \begin{cases} G_1'\Sigma^{-1}\theta_1 + \Sigma_0^{-1}m_0 & \text{, for } t = 0 \\ \Sigma^{-1}G_t\theta_{t-1} + G_{t+1}'\Sigma^{-1}\theta_{t+1} + \frac{1}{\sigma_\eta^2}\left(X_t'y_t - X_t'\alpha\right) & \text{, for } t = 1,2,\ldots,T-1 \\ \Sigma^{-1}G_T\theta_{T-1} + \frac{1}{\sigma_\eta^2}(X_T'y_T - X_T'\alpha) & \text{, for } t = T \end{cases} \quad (3.5)$$

$$\theta^{(s)} = \{\theta_t\}_{t=0}^{T} \quad \text{for } t \neq s \quad (3.6)$$

Theorem 1: The conditional posterior density of θ_S given ($\theta^{(s)}$, α, λ) is normal with mean vector $B_S\, b_S$ and covariance matrix $\lambda^{-1}\, B_S$.

Proof: . Utilizing (3.3), the conditional posterior density of θ_0 is obtained as

$$f^*(\theta_0|\underset{\sim}{y},\alpha,\lambda) \propto \exp\left[-\frac{\lambda}{2}\left\{\theta_0'\left(\Sigma_0^{-1}+G_1'\Sigma^{-1}G_1\right)\theta_0-2\theta_0'\left(\Sigma_0^{-1}m_0+G_1'\Sigma^{-1}\theta_1\right)\right\}\right]$$

$$\propto \exp\left[-\frac{\lambda}{2}\left(\theta_0-b_0B_0\right)' B_0^{-1}\left(\theta_0-b_0B_0\right)\right] \tag{3.7}$$

We now write expression (3.7) as

$$f^*(\theta_0|\underset{\sim}{y},\alpha,\lambda)=C_0\exp\left[-\frac{\lambda}{2}\left(\theta_0-B_0b_0)'B_0^{-1}(\theta_0-B_0b_0\right)\right]$$

where C_0 is the normalizing constant. Its value is obtained as

$$C_0^{-1}=\int_{R^p}\exp\left[-\frac{\lambda}{2}(\theta_0-b_0B_0)'B_0^{-1}(\theta_0-b_0B_0)\right]d\theta_0=\frac{(2\pi)^{\frac{p}{2}}|B_0|^{\frac{1}{2}}}{\lambda^{\frac{p}{2}}}$$

For $s = 1,\ldots, T-1$, We have

$$f^*(\theta_s|\underset{\sim}{y},\alpha,\lambda,\theta^{(s)}) \propto \exp\Bigg[-\frac{\lambda}{2}\Bigg\{\theta_s'\left(\Sigma^{-1}+\frac{1}{\sigma_\eta^2}X_s'X_s+G_{s+1}'\Sigma^{-1}G_{s+1}\right)\theta_s$$

$$-2\theta_s'\left(\frac{1}{\sigma_\eta^2}(X_s'y_s-X_s'\alpha)+G_{s+1}'\Sigma^{-1}\theta_{s+1}+\Sigma^{-1}G_s\theta_{s-1}\right)\Bigg\}\Bigg]$$

$$\propto \exp\left[-\frac{\lambda}{2}(\theta_s-B_sb_s)'B_s^{-1}(\theta_s-B_sb_s)\right] \tag{3.8}$$

We now write expression (3.8) as

$$f^*(\theta_s|\underset{\sim}{y},\alpha,\lambda,\theta^{(s)})=C_s\exp\left[-\frac{\lambda}{2}(\theta_s-B_sb_s)'B_s^{-1}(\theta_s-B_sb_s)\right]$$

where C_s is evaluated as

$$C_s^{-1}=\int_{R^p}\exp\left[-\frac{\lambda}{2}(\theta_s-B_sb_s)'B_s^{-1}(\theta_s-B_sb_s)\right]d\theta_s=\frac{(2\pi)^{\frac{p}{2}}|B_s|^{\frac{1}{2}}}{\lambda^{\frac{p}{2}}}$$

Finally, for $s= T$ we have $f^*\left(\theta_T|\underset{\sim}{y},\alpha,\lambda,\{\theta_t\}_{t=0}^{T-1}\right)$

$$\propto \exp\left[-\frac{\lambda}{2}\left\{\theta_T'\left(\Sigma^{-1}+\frac{1}{\sigma_\eta^2}X_T'X_T\right)\theta_T-2\theta_T'\left(\frac{1}{\sigma_\eta^2}\left(X_T'y_T-X_T'\alpha\right)+\Sigma^{-1}G_T\theta_{T-1}\right)\right\}\right]$$

$$\propto \exp\left[-\frac{\lambda}{2}(\theta_T - B_T b_T)' B_T^{-1}\left(\theta_T - B_T b_T\right) - b_T' B_T b_T\right]$$
$$\propto \exp\left[-\frac{\lambda}{2}(\theta_T - B_T b_T)' B_T^{-1}\left(\theta_T - B_T b_T\right)\right] \tag{3.9}$$

We now write equation (3.9) as

$$f^*\left(\theta_T | \underset{\sim}{y}, \alpha, \lambda, \{\theta_t\}_{t=0}^{T-1}\right) = C_T \exp\left[-\frac{\lambda}{2}(\theta_T - B_T b_T)' B_T^{-1}(\theta_T - B_T b_T)\right]$$

where C_T is the normalizing constant, which is obtained as

$$C_T^{-1} = \int_{R^p} \exp\left[-\frac{\lambda}{2}(\theta_T - B_T b_T)' B_T^{-1}(\theta_T - B_T b_T)\right] d\theta_T = \frac{(2\pi)^{\frac{p}{2}} |B_T|^{\frac{1}{2}}}{\lambda^{\frac{p}{2}}}$$

Thus, the theorem is proved.

Let us write

$$Q = \frac{T}{\sigma_\eta^2} + \frac{1}{\sigma_\alpha^2}, \; q = \frac{1}{\sigma_\eta^2}[y_t - X_t\theta_t] \tag{3.10}$$

and $\quad a^* = nT + pT + p + a_0 + n$

$$b^*(\alpha) = \frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}(y_t - X_t\theta_t - \alpha)'(y_t - X_t\theta_t - \alpha)' + \sum_{t=1}^{T}(\theta_t - G_t\theta_{t-1})'\Sigma^{-1}(\theta_t - G_t\theta_{t-1})'$$
$$+(\theta_0 - m_0)'\Sigma_0^{-1}(\theta_0 - m_0) + b_0 + \frac{\alpha'\alpha}{\sigma_\alpha^2} \tag{3.11}$$

Theorem 2: The conditional posterior of α is given by normal distribution with mean $Q^{-1}q$ and variance covariance matrix $\lambda^{-1}Q^{-1}$.

Proof : From equation (3.3) we obtain

$$f^*\left(\alpha | \underset{\sim}{y}, \lambda, \theta_0, \{\theta_T\}_{t=1}^{T}\right) \propto \exp\left[-\frac{\lambda}{2}\left\{\alpha'\left(\frac{T}{\sigma_\eta^2} + \frac{1}{\sigma_\alpha^2}\right)\alpha - \frac{2\alpha'}{\sigma_\eta^2}(y_t - X_t\theta_t)\right\}\right]$$
$$\propto \exp\left[-\frac{\lambda}{2}\left\{(\alpha - Q^{-1}q)' Q(\alpha - Q^{-1}q)\right\}\right] \tag{3.12}$$

We rewrite expression (3.12) as follows

$$f^*\left(\alpha | \underset{\sim}{y}, \lambda, \theta_0, \{\theta_T\}_{t=1}^{T}\right) = K_\alpha \exp\left[-\frac{\lambda}{2}\left\{\alpha - Q^{-1}q)' Q\left(\alpha - Q^{-1}q\right)\right\}\right]$$

where the normalizing constant K_α is obtained as

$$K_\alpha^{-1} = \int_{R^n} \exp\left[-\frac{\lambda}{2}\left\{(\alpha - Q^{-1}q)' Q(\alpha - Q^{-1}q)\right\}\right] d\alpha = \frac{(2\pi)^{\frac{n}{2}}}{\lambda\left[\frac{T}{\sigma_\eta^2} + \frac{1}{\sigma_\alpha^2}\right]}$$

which establishes the theorem.

Theorem 3: The conditional posterior of λ is given by

$$f^*(\lambda | \underset{\sim}{y}, \alpha, \{\theta_t\}_{t=0}^{T}) = K_\lambda \lambda^{\frac{a^*}{2}-1} \exp\left[-\frac{\lambda}{2} b^*\right] \tag{3.13}$$

Proof: We obtain the following from expression (5.3.3)

$$f^*(\lambda | \underset{\sim}{y}, \alpha, \{\theta_t\}_{t=0}^{T}) \propto \lambda^{\frac{1}{2}(nT + pT + p + a_0 + n) - 1} \exp\left[-\frac{\lambda}{2}\left\{\frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}(y_t - x_t\theta_t - \alpha)'(y_t - x_t\theta_t - \alpha)\right.\right.$$

$$\left.\left. + \sum_{t=1}^{T}\left\{(\theta_t - G_t\theta_{t-1})'\Sigma^{-1}(\theta_t - G_t\theta_{t-1})\right\} + (\theta_0 - m_0)'\Sigma_0^{-1}(\theta_0 - m_0) + b_0 + \frac{\alpha'\alpha}{\sigma_\alpha^2}\right\}\right]$$

$$\propto \lambda^{\frac{a^*}{2}-1} e^{-\frac{b^*}{2}\lambda}$$

or

$$f^*(\lambda | \underset{\sim}{y}, \alpha, \{\theta_t\}_{t=0}^{T}) = \frac{\left(\frac{b^*}{2}\right)^{\frac{a^*}{2}}}{\Gamma\left(\frac{a^*}{2}\right)} \lambda^{\frac{a^*}{2}-1} e^{\frac{b^*}{2}\lambda}$$

Thus, the theorem is proved.

4. Implementation of Gibbs Sampler

Let the generated Gibbs sample be denoted by $\left(\left\{\theta_{tj}^{(k)}\right\}_{t=0}^{T}, \lambda_j^{(k)}, \alpha_j^{(k)}\right)$; *j = 1, 2,…,N* , where N is the total number of replications. Then, by employing the Gibbs sampler, posterior density of θ_S given $\underset{\sim}{y}$ can be estimated as

$$\hat{f}(\theta_s | \underset{\sim}{y}) = \frac{1}{N}\sum_{j=1}^{N} f^*(\theta_s | \theta_{s-1,j}^{(k)}, \theta_{s+1,j}^{(k)}, \lambda_j^{(k)}, \underset{\sim}{y}, \alpha_j^{(k)}) \tag{4.1}$$

for $s = 0, 1, \ldots\ldots, T$. Notice that the estimated posterior density of θ_S in (4.1) depends on the two values adjacent to θ_s. Hence, an estimate of $\theta_{s/T}$ is the mean of the estimated density (4.1) which is obtained as

$$\bar{\theta}_{s/T} = B_s\left(\frac{1}{N}\sum_{j=1}^{N} b_{sj}^{(k)}\right) \tag{4.2}$$

where k is the number of iterations during implementation of the Gibbs sampler. $b_{sj}^{(k)}$ is the value of b_S based on $\left(\left\{\theta_{tj}^{(k)}\right\}_{j=1}^{N},\ t = s-1,\ s+1\right)$. Then the fitted value of y_t for our model is

$$\tilde{y}_t = X_t \bar{\theta}_{t|T} \tag{4.3}$$

Similarly, by employing the Gibbs sampler posterior density of α given $\underset{\sim}{y}$ can be estimated as

$$\hat{f}(\alpha|\underset{\sim}{y}) = \frac{1}{N}\sum_{j=1}^{N} f^*\left(\alpha|\left\{\theta_{tj}^{(k)}\right\}_{j=1}^{N},\ \lambda_j^{(k)},\ \underset{\sim}{y}\right) \tag{4.4}$$

Further, an estimator of α is $\bar{\alpha} = \frac{1}{N}\sum_{j=1}^{N} \bar{\alpha}_j^{(k)}$

where $\bar{\alpha}_j^{(k)} = \left(Q_j^{(k)}\right)^{-1} q_j^{(k)},\ Q_j^{(k)} = \left(\frac{T}{\sigma_\eta^2} + \frac{1}{\sigma_\alpha^2}\right),\ q_j^{(k)} = \left(y_t - X_{tj}^{(k)}\theta_{tj}^{(k)}\right)$

5. Conclusion

The state space model is utilized to obtain Bayesian estimators for the parameters which can improve panel data-based prediction wherein the observations are available on the behaviour of a ‘panel’ of decision units at multiple successive time epochs. The use of panel data has become increasingly popular in econometrics in recent years. This analysis provides an elaborate theoretical framework and is therefore expected to contribute effectively to improved and more precise panel data-based prediction for applied researchers and practitioners.

Acknowledgements

Authors are thankful to the reviewers for their valuable comments and useful suggestions which have improved the quality of the original manuscript. The first author acknowledges with thanks financial assistance from DU-DST PURSE Grant and R& D Grant from the University of Delhi.

REFERENCES

BALTAGI, B. H., (2008). Econometric analysis of panel data. Wiley.

CHAMBERLAIN, G., (1982). Multivariate regression models for panel data, Journal of Econometrics 18, No. 1.

HAUSMAN, J., (1978). Specification tests in econometrics. Econometrica 46, No. 6.

HAUSMAN, J., TAYLOR, W., (1981). Panel data and unobservable individual effects. Econometrica 49, No. 6.

MADDALA, G. S., (1971). The use of variance components models in pooling cross-section and time series data. Econometrica 39, No. 2.

MUNDLAK, Y., (1978). On the pooling of time series and cross section data. Econometrica 46, No. 1.

TIWARI, R. C., YANG, Y., ZALKIKAR, J. N., (1996). Time series analysis of BOD data using the Gibbs sampler. Enviormetrics 7: 567−78.