MCMC Algorithm for Bayesian Heterogeneous Coefficients of Panel Data Model

Jaya, Mindra; Tantular, Bertho; Ruchjana, Budi Nurani; Zulhanif

Published in:
Journal of Physics: Conference Series

DOI:
10.1088/1742-6596/1320/1/012014

Publication date:
2019

Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

MCMC Algorithm for Bayesian Heterogeneous Coefficients of Panel Data Model

To cite this article: I G N M Jaya *et al* 2019 *J. Phys.: Conf. Ser.* **1320** 012014

IOP Conf. Series: Journal of Physics: Conf. Series **1320** (2019) 012014 doi:10.1088/1742-6596/1320/1/012014

MCMC Algorithm for Bayesian Heterogeneous Coefficients of Panel Data Model

I G N M Jaya[1], B Tantular[2], Zulhanif[3] and B N Ruchjana[4]

[1] Department of Statistics, Universitas Padjadjaran, Indonesia & Spatial Sciences Faculty, University of Groningen, the Netherlands
[2,3] Statistics Department, Universitas Padjadjaran, Indonesia
[4] Mathematics Department, Universitas Padjadjaran, Indonesia

Corresponding author: mindra@unpad.ac.id

Abstract. Panel data models have been applied widely in many subject areas related to economic, social, and epidemiology. In some cases (e.g. epidemiology studies), the phenomena encountered have a complex relationship structured. The risk factors such as house index, healthy behaviour index, rainfall and the other risk factors of particular infectiouse disease may have different effect on the outcome due to the heterogeneity of crossection units. The effect of the covariates on outcome could vary over individual and time units. This condition is called as a non-stationary or instability relationship problem. This problem leads to bias and inefficient of the estimators. It is important to examine the heterogeneous coefficients model for avoiding inefficient estimator. We present in detail a statistical estimation procedure of the heterogeneous coefficients for fixed effect panel data model by means of the hierarchical Bayesian estimation approach. The challenges of the Bayesian approaches are finding the joint posterior distribution and developing the algorithm for estimating the parameters of interest. We find that the joint posterior distribution of the heterogeneous coefficients fixed effect panel data model does not follow any standard known distribution form. Consequently, the analytical solution cannot be applied and simulation approach of Markov Chain Monte Carlo (MCMC) was used. We present the MCMC procedure covering the derivation of the full conditional distribution of the parameters model and present step-by-step the Gibbs sampling algorithm. The idea of this preliminary research can be applied in various fields to overcome the non-stationarity problem.

1. Introduction

Numerous economic, social and epidemiology studies have mainly focused on finding the relationship between the covariates and outcome [1]. The data used are collected from different individuals, times, and their variation. The covariates effect on outcome may vary over individual and time. Also, the standard assumption of the identically distributed in sampling process may not always be satisfied. These conditions may yield heterogeneities problem and possibly inefficient estimators [2], [3], [4]. Hence, examining the heterogeneous coefficients model become important. The Panel Data Model usually uses for presenting the relationship between the covariates and outcome for panel data structure (i.e. combine the individual and time series data). There are three types of Panel Data Models: pooled, fixed, and random effect models. The pooled model assumes that the same regression lines for all individuals as the best model. The fixed effect takes different intercepts for different individuals to accommodate the individual heterogeneity and the random effect model presents the individual heterogeneity as the random variance component with the same lines for all individuals [5]. The last two models only solve the heteroscedasticity problem due to the individual heterogeneity. In some cases (e.g. epidemiology studies), the constant slope may not be appropriate for explaining the relationship of the covariate and outcome. This condition is called as a nonstationary problem [6]. In this paper, we extend the fixed effect panel model by introducing heterogeneous of slope regression coefficients.

To motivate this model, let us consider an effect of the clean and healthy life behavior index on the prevalence rate of diarrhea disease in several districts. The two districts may have equal clean and healthy life behavior index. The standard panel data model allows for two districts with an identical index to have a different expected prevalence rate (i.e. due to different intercepts). Furthermore, it allows for the marginal effect of clean and healthy life behavior index on the prevalence rate of diarrhea disease to vary across districts (i.e. due to different slope coefficients). If the variability of population density is important, then such an interaction between population density and the clean and healthy life behavior index may occur and lead to a non-stationarity problem [1]. The non-stationarity issue has attracted considerable attention in the past two decades. Mou et al (2017) [7] applied use non stationary regression model to estimate the housing price in China and found that the land price has significant effect with different magnitudes. The extension of the constant coefficients to the heterogeneous coefficients enable to reflect the cross section units nature of the data, minimize the bias and increasing the efficiency of the estimator [7].

The standard panel data model with classical estimation approaches (i.e. ordinary least square and maximum likelihood) are not be appropriate used to solve the non-stationarity problem. The standard approaches lead to identifiability problems [8]. Moreover, unbiasedness, consistency, and efficiency become serious problems.

Non-parametric, semi-parametric, and the Geographically weighted regression (GWR) approaches are the most popular methods for estimating the individual and time-heterogeneous coefficients [9], [10], [11]. Li Q et al (2002) [9] introduced heterogeneous coefficients model in an individual setting and proposed a semi-parametric estimation to estimate the parameters model. Li D et al (2011) [10] proposed time heterogeneous coefficient fixed effect model and introduced the non-parametric estimation procedure. Cai R et al (2011) [11] used the Geographically Weighted Panel Regressions (GWPR) to model varying relationship between covariates and an outcome.

We propose an alternative approach for individual heterogeneous coefficients of a panel data model using a hierarchical Bayesian estimation approach. The Bayesian approach has several advantages over non/semi-parametric and GWR approaches. Using Bayesian approach, we able to use out of the sample knowledge in the estimating process via a prior distribution. The complex problem can be solved using the hierarchical Bayesian and Bayesian approaches which provides interpretable answers from the credible interval of parameters of interest [12]. However, the Bayesian approach also has some disadvantages. Some difficulties using the Bayesian approach is finding a joint posterior distribution and developing an algorithm for estimation process. Generally, the joint posterior distribution does not follow a standard distribution form and an analytical solution cannot be obtained. The Markov Chain Monte Carlo (MCMC) simulation can be used to solve this problem. However, the simulation approach requires a good computational ability. The Bayesian approach of heterogenous coefficient panel data have been introduced in [4]. However, the derivatives for each full conditional posterior distributions are not explained in detail and the use of prior Gamma prior for error precision makes it difficult to obtain derivatives from the poster distribution.

This paper focuses on the methodological issue on how to develop a Bayesian procedure for estimating parameters of heterogeneous coefficients in the fixed effect panel model and present the detail derivation of the full conditional posterior distributions using inverse Gamma prior for the variance error. The paper structure is organized as follows; in Section 2, the Bayesian approach of heterogeneous coefficients Panel Data Model is described. The MCMC algorithm is accomplished in Section 3. Section 4 contains discussion and conclusion.

2. Heterogeneous coefficients panel data model

2.1. Model

A fixed effect model is developed to accommodate an individual heterogeneity by assuming the individual parameters effect are fixed. Consider a standard fixed effect panel data model [4] given below:

$$\mathbf{y}_i = \alpha_i \mathbf{1}_{\mathrm{T}} + \mathbf{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i \quad ; i = 1, \ldots, n \tag{1}$$

where $\mathbf{y}_i = [y_{i1}, .., y_{iT}]^T$ is a T–column vector of the response variable, $\mathbf{1}_{\mathrm{T}}$ denotes a T -column vector of one's, $\mathbf{X}_i$ is a $T \times K$ design matrix with K explanatory variables, $\boldsymbol{\beta} = [\beta_1 \quad \beta_2 \quad \ldots \quad \beta_K]^{\mathrm{T}}$ is a K-column vector of regression slope coefficients, $\boldsymbol{\varepsilon}_i = [\varepsilon_{i1} \quad \varepsilon_{i2} \quad \cdots \quad \varepsilon_{iT}]^{\mathrm{T}}$ is a T- column vector of random errors. The error term is assumed normally distributed, $\boldsymbol{\varepsilon}_i \sim_{\mathrm{iid}} \mathrm{N}(0, \sigma^2 \mathbf{I}_{\mathrm{T}})$. The least square dummy variable (LSDV) or within estimators are usually used to estimate the parameters of the fixed effect model.

The standard fixed effect model can be extended to non-stationarity model by introducing the means of heterogeneous coefficients. The heterogeneous coefficient fixed effect panel data model [4] can be written below:

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i \quad ; i = 1, \ldots, n \tag{2}$$

where $\boldsymbol{\beta}_i = [\alpha_i \quad \beta_{i1} \quad \ldots \quad \beta_{iK}]^{\mathrm{T}}$. The parameters vector of $\boldsymbol{\beta}_i$ shows the heterogeneous coefficients of fixed effect panel data model. For $T \to \infty$, the $\boldsymbol{\beta}_i$ can be estimated using a standard ordinary least square (OLS) method. However for small T, the OLS estimator leads to inefficient. Non-parametric and semi-parametric approaches are the most popular methods that used to estimate the individual and time-heterogeneous coefficients. The semi-parametric and non-parametric approaches are generally used [9], [10]. Choi et al (2012). [1] use non-parametric model by mean Bayesian to model the spatiotemporally varying coefficients in low birth weight incidence data. Cai R et al (2014) [11] introduced Geographically Weighted Panel Regressions (GWPR) using the weighted least square (WLS) estimators. We propose an alternative approach for individual heterogeneous coefficients panel data model using hierarchical Bayesian estimation approach.

2.2. Bayesian Estimation

A Bayesian approach is widely used for economic, social and epidemiology in some recent decades due to the advantages of the Bayesian rather than classical approach. Jaya et al (2016) [13] used Bayesian varying coefficients model to estimate the effect of risk factors on dengue disease incidence in Bandung. Wheeler et al (2014) [14] estimate the hedonic price analysis by mean Bayesian approach and Law and Haining (2014) [15] used approach to modeling binary data to model the case of high-intensity crime areas. The Bayesian approach is essentially using a Bayes theorem idea by considering. Three components of a likelihood function, $p(\mathbf{y}|\boldsymbol{\beta}, \sigma)$, the a prior distribution, $p(\boldsymbol{\beta}, \sigma)$, and the a joint posterior distribution, $p(\boldsymbol{\beta}, \sigma|\mathbf{y})$. The three components can be presented in simple formula as given below [4], [16], [17]:

$$p(\boldsymbol{\beta}, \sigma|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\beta}, \sigma) p(\boldsymbol{\beta}, \sigma)}{p(\mathbf{y})} \tag{3}$$

where $p(\mathbf{y})$ is a normalizing constant, the posterior distribution, of $p(\boldsymbol{\beta}, \sigma|\mathbf{y})$is proportional to ($\propto$) the likelihood function multiplied by the prior distribution.

$$p(\boldsymbol{\beta}, \sigma|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\beta}, \sigma) p(\boldsymbol{\beta}, \sigma) \tag{4}$$

A main focus in the Bayesian estimation is obtaining the joint posterior distribution. After that, the joint posterior distribution is used to obtain the point and interval estimate of the parameter interest [10]. However, the joint posterior distribution is usually not in a closed form solution and an analytical solution cannot be obtained. As an alternative, a simulation using Markov Chain Monte Carlo

(MCMC) can be developed for performing Bayesian inferences [17], [18], [19]. The MCMC algorithm has a good performance for either a simple or a complex model.

2.2.1. Likelihood Function

First step for applying the Bayesian method is, defining the likelihood function of the observed variable $\mathbf{y}$ is given by:

$$p(\mathbf{y}|\boldsymbol{\beta},\sigma)=\prod_{i=1}^{n}\frac{1}{(2\pi\sigma^2)^{\frac{T}{2}}}\left\{\exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)\right)\right\} \quad (5)$$

2.2.2. Prior Distributions

A convenient hierarchical prior is assuming that the $\boldsymbol{\beta}_i$ for $i=1\ldots,n$ are-independently drawn from a Normal distribution [3,11]:

$$\boldsymbol{\beta}_i \sim \mathrm{N}(\boldsymbol{\mu}_\beta,\mathbf{V}_\beta) \quad (6)$$

where $\boldsymbol{\beta}_i=[\alpha_i \quad \beta_{i1} \quad \ldots \quad \beta_{iK}]^T$ denotes the vector regression parameter of panel data model which varies by cross section unit.

The second stage of the hierarchical prior is given by

$$\boldsymbol{\mu}_\beta \sim \mathrm{N}(\boldsymbol{\mu}_\beta,\boldsymbol{\Sigma}_\beta) \quad (7)$$

and

$$\boldsymbol{V}_\beta^{-1} \sim W(\boldsymbol{v}_\beta,\boldsymbol{V}_\beta^{-1}) \quad (8)$$

where *W*(.) denotes the Wishart distribution. The Wishart distribution can be parameterized so that $\mathbf{E}(\boldsymbol{V}_\beta^{-1})=\boldsymbol{v}_\beta\boldsymbol{V}_\beta^{-1}$. For the error variance, it is assumed to follow an Inverse Gamma prior distribution [4] as follows,

$$\sigma^2 \sim IG\left(\underline{\sigma}^2,\underline{\gamma}\right) \quad (9)$$

where σ^2 denotes the variance of error components.

The structure of hierarchical Bayesian for the heterogeneous coefficients fixed effect panel data model can be written as:

$$\begin{cases} \mathbf{y}_i|\boldsymbol{\beta}_i,\sigma^2 \sim \mathrm{N}(\mathbf{X}_i\boldsymbol{\beta}_i,\sigma^2\mathbf{I}_T) \\ \boldsymbol{\beta}_i \sim \mathrm{N}_p(\boldsymbol{\mu}_\beta,\mathbf{V}_\beta) \\ \sigma^2 \sim \mathrm{IG}\left(\underline{\sigma}^2,\underline{\gamma}\right) \\ \boldsymbol{\mu}_\beta \sim \mathrm{N}\left(\underline{\boldsymbol{\mu}}_\beta,\underline{\boldsymbol{\Sigma}}_\beta\right) \\ \mathbf{V}_\beta^{-1} \sim \mathrm{W}\left(\underline{\boldsymbol{v}}_\beta,\underline{\mathbf{V}}_\beta^{-1}\right) \end{cases} \quad (10)$$

2.2.3. The Joint Posterior Distribution

Here we present the deviation of the joint posterior distribution by multiplying the likelihood function over to the prior distribution.

$$p(\boldsymbol{\beta},\sigma^2|\mathbf{y},) \propto p(\mathbf{y}|\boldsymbol{\beta},\sigma^2)p(\boldsymbol{\beta})p(\sigma^2)$$

$$=\frac{1}{(2\pi\sigma^2)^{\frac{nT}{2}}}\times\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)\right)$$

$$\times \frac{1}{(2\pi)^{\frac{p}{2}}|\mathbf{V}_\beta|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)^T\mathbf{V}_\beta^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)\right)$$
$$\times \frac{\underline{\gamma}^{\underline{\sigma}^2}}{\Gamma(\underline{\sigma}^2)}(\sigma^2)^{-(\underline{\sigma}^2+1)}\times\exp\left\{-\frac{\underline{\gamma}}{\sigma^2}\right\}$$

$$\propto (\sigma^2)^{-\left[\frac{nT+2\underline{\sigma}^2}{2}\right]-1}\times\exp\left(-\frac{1}{2}\left[\frac{1}{\sigma^2}\left(2\underline{\gamma}+\sum_{i=1}^{n}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)\right)\right.\right.$$
$$\left.\left.+\sum_{i=1}^{n}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)^T\mathbf{V}_\beta^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)\right]\right) \quad (11)$$

The posterior distribution (Eq.11) of heterogeneous coefficients of fixed effect panel data does not follow a standard known distribution. Hence, we used the MCMC approach with the Gibbs sampling algorithm for estimating the parameters of interest.

3. MCMC Derivation

The analytical solution is hard to obtain from the joint posterior distribution of the heterogeneous coefficients of fixed effect panel data. We introduce the MCMC procedure with Gibbs sampling algorithm to estimate the parameters of interest. The Gibbs sampling algorithm works with the full conditional distribution of the parameter model.

3.1. Full conditional posterior distribution of $\boldsymbol{\beta}_i|\mathbf{y}_i,\sigma^2,\boldsymbol{\mu}_\beta,\mathbf{V}_\beta$

The full conditional posterior distribution of $\boldsymbol{\beta}_i|\mathbf{y}_i,\sigma^2,\boldsymbol{\mu}_\beta,\mathbf{V}_\beta$ can be defined as:

$$p(\boldsymbol{\beta}_i|\mathbf{y}_i,\sigma^2,\boldsymbol{\mu}_\beta,\mathbf{V}_\beta)\propto p(\mathbf{y}_i|\boldsymbol{\beta}_i,\sigma^2)p(\boldsymbol{\beta}_i)$$
$$=\frac{1}{(2\pi\sigma^2)^{\frac{T}{2}}}\times\exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)\right)$$
$$\times\frac{1}{(2\pi)^{\frac{p}{2}}|\mathbf{V}_\beta|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)^T\mathbf{V}_\beta^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)\right)$$
$$\propto\exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_i)-\frac{1}{2}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)^T\mathbf{V}_\beta^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta)\right)$$
$$=\exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}_i^T\mathbf{y}_i-\boldsymbol{\beta}_i^T\mathbf{X}_i^T\mathbf{y}_i-\mathbf{y}_i^T\mathbf{X}_i\boldsymbol{\beta}_i+\boldsymbol{\beta}_i^T\mathbf{X}_i^T\mathbf{X}_i\boldsymbol{\beta}_i)\right.$$
$$\left.-\frac{1}{2}(\boldsymbol{\beta}_i^T\mathbf{V}_\beta^{-1}\boldsymbol{\beta}_i-\boldsymbol{\mu}_\beta^T\mathbf{V}_\beta^{-1}\boldsymbol{\beta}_i-\boldsymbol{\beta}_i^T\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta+\boldsymbol{\mu}_\beta^T\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta)\right)$$
$$\propto\exp\left(-\frac{1}{2}(\sigma^{-2}\boldsymbol{\beta}_i^T\mathbf{X}_i^T\mathbf{X}_i\boldsymbol{\beta}_i+\boldsymbol{\beta}_i^T\mathbf{V}_\beta^{-1}\boldsymbol{\beta}_i-2\boldsymbol{\beta}_i^T\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta-2\sigma^{-2}\boldsymbol{\beta}_i^T\mathbf{X}_i^T\mathbf{y}_i+\boldsymbol{\mu}_\beta^T\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta)\right)$$
$$=\exp\left(-\frac{1}{2}(\boldsymbol{\beta}_i^T(\sigma^{-2}\mathbf{X}_i^T\mathbf{X}_i+\mathbf{V}_\beta^{-1})\boldsymbol{\beta}_i-2\boldsymbol{\beta}_i^T(\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta+\sigma^{-2}\mathbf{X}_i^T\mathbf{y}_i)+\boldsymbol{\mu}_\beta^T\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta)\right) \quad (12)$$

Using Gaussian manipulations, we obtain:

$$p(\boldsymbol{\beta}_i|\mathbf{y}_i,\sigma^2,\boldsymbol{\mu}_\beta,\mathbf{V}_\beta)\propto\exp\left(-\frac{1}{2}(\boldsymbol{\beta}_i-\overline{\boldsymbol{\beta}}_i)^T\overline{\mathbf{V}}_i^{-1}(\boldsymbol{\beta}_i-\overline{\boldsymbol{\beta}}_i)\right) \quad (13)$$

where

$$\overline{\boldsymbol{\beta}}_i = \overline{\mathbf{V}}_i\left(\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta + \sigma^{-2}\mathbf{X}_i^T\mathbf{y}_i\right) \text{and } \overline{\mathbf{V}}_i = \left(\mathbf{V}_\beta^{-1} + \sigma^{-2}\mathbf{X}_i^T\mathbf{X}_i\right)^{-1} \tag{14}$$

(Eq.14) is proportional to Normal distribution:

$$\boldsymbol{\beta}_i|\mathbf{y}_i, \sigma^2 \sim \text{Normal}\left(\overline{\boldsymbol{\beta}}_i, \overline{\mathbf{V}}_i\right) \tag{15}$$

3.2. Full conditional posterior distribution of $\sigma^2|\mathbf{y}, \boldsymbol{\beta}, \underline{\sigma}^2, \underline{\gamma}$

Full conditional posterior distribution of $\sigma^2|\mathbf{y}, \boldsymbol{\beta}, \underline{\sigma}^2, \underline{\gamma}$ can be defined as:

$$\begin{aligned}
p\left(\sigma^2|\mathbf{y}, \boldsymbol{\beta}, \underline{\sigma}^2, \underline{\gamma}\right) &\propto p(\mathbf{y}|\boldsymbol{\beta}, \sigma^2)p(\sigma^2) \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{nT}{2}}} \times \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)\right) \\
&\quad \times \frac{\underline{\gamma}^{\underline{\sigma}^2}}{\Gamma(\underline{\sigma}^2)}(\sigma^2)^{-(\underline{\sigma}^2+1)} \times \exp\left\{-\frac{\underline{\gamma}}{\sigma^2}\right\} \\
&\propto (\sigma^2)^{-(nT/2)}(\sigma^2)^{-(\underline{\sigma}^2+1)} \exp\left(-\frac{1}{\sigma^2}\left\{2^{-1}\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i) + \underline{\gamma}\right\}\right) \\
&= (\sigma^2)^{-\left(\underline{\sigma}^2+\frac{nT}{2}+1\right)} \exp\left(-\frac{1}{\sigma^2}\left\{\underline{\gamma} + 2^{-1}\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)\right\}\right) \\
&\propto (\sigma^2)^{-(\bar{\sigma}^2+1)} \exp\left(-\frac{\bar{\gamma}}{\sigma^2}\right)
\end{aligned} \tag{16}$$

3.3. Full conditional posterior distribution of $\boldsymbol{\mu}_\beta|\mathbf{y}, \boldsymbol{\beta}, \sigma^2\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{\Sigma}}_\beta$

Full conditional posterior distribution of $\boldsymbol{\mu}_\beta|\mathbf{y}, \boldsymbol{\beta}, \sigma^2\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{\Sigma}}_\beta$ can be defined as:

$$\begin{aligned}
p\left(\boldsymbol{\mu}_\beta|\mathbf{y}, \boldsymbol{\beta}, \sigma^2\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{\Sigma}}_\beta\right) &\propto p(\mathbf{y}|\boldsymbol{\beta}, \sigma^2)p(\boldsymbol{\beta})p(\boldsymbol{\mu}_\beta) \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{nT}{2}}} \times \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)\right) \\
&\quad \times \frac{1}{(2\pi)^{\frac{p}{2}}|\mathbf{V}_\beta|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta\right)^T\mathbf{V}_\beta^{-1}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta\right)\right) \\
&\quad \times \frac{1}{(2\pi)^{\frac{p}{2}}|\underline{\boldsymbol{\Sigma}}_\beta|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left(\boldsymbol{\mu}_\beta - \underline{\boldsymbol{\mu}}_\beta\right)^T\underline{\boldsymbol{\Sigma}}_\beta^{-1}\left(\boldsymbol{\mu}_\beta - \underline{\boldsymbol{\mu}}_\beta\right)\right) \\
&\propto \exp\left(-\frac{1}{2}\left(\sum_{i=1}^{n}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta\right)^T\mathbf{V}_\beta^{-1}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta\right) + \left(\boldsymbol{\mu}_\beta - \underline{\boldsymbol{\mu}}_\beta\right)^T\underline{\boldsymbol{\Sigma}}_\beta^{-1}\left(\boldsymbol{\mu}_\beta - \underline{\boldsymbol{\mu}}_\beta\right)\right)\right) \\
&= \exp\left(-\frac{1}{2}\left(\boldsymbol{\mu}_\beta^T\left(n\mathbf{V}_\beta^{-1} + \underline{\boldsymbol{\Sigma}}_\beta^{-1}\right)\boldsymbol{\mu}_\beta - \mathbf{2}\boldsymbol{\mu}_\beta^T\left(\mathbf{V}_\beta^{-1}\sum_{i=1}^{n}\boldsymbol{\beta}_i + \underline{\boldsymbol{\Sigma}}_\beta^{-1}\underline{\boldsymbol{\mu}}_\beta\right) + \sum_{i=1}^{n}\boldsymbol{\beta}_i^T\mathbf{V}_\beta^{-1}\boldsymbol{\beta}_i\right)\right)
\end{aligned} \tag{17}$$

Adopting the Gaussian manipulations in (Eq. 17) we obtain:

$$p\left(\boldsymbol{\mu}_\beta|\mathbf{y}, \boldsymbol{\beta}, \sigma^2\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{\Sigma}}_\beta\right) \propto \exp\left(-\frac{1}{2}\left(\boldsymbol{\mu}_\beta - \overline{\boldsymbol{\mu}}_\beta\right)^T\overline{\boldsymbol{\Sigma}}_\beta^{-1}\left(\boldsymbol{\mu}_\beta - \overline{\boldsymbol{\mu}}_\beta\right)\right) \tag{18}$$

where

$$\bar{\boldsymbol{\Sigma}}_\beta = n\mathbf{V}_\beta^{-1} + \underline{\boldsymbol{\Sigma}}_\beta^{-1} \text{ and } \bar{\boldsymbol{\mu}}_\beta = \bar{\boldsymbol{\Sigma}}_\beta\left(\mathbf{V}_\beta^{-1}\sum_{i=1}^{n}\boldsymbol{\beta}_i + \underline{\boldsymbol{\Sigma}}_\beta^{-1}\underline{\boldsymbol{\mu}}_\beta\right) \quad (19)$$

(E.q 16) is proportional to ~~a~~ Normal distribution:

$$\boldsymbol{\mu}_\beta|\mathbf{y},\boldsymbol{\beta},\sigma^2\underline{\boldsymbol{\mu}}_\beta,\underline{\boldsymbol{\Sigma}}_\beta \sim \text{Normal}(\bar{\boldsymbol{\mu}}_\beta,\bar{\boldsymbol{\Sigma}}_\beta) \quad (20)$$

3.4. The full conditional posterior distribution of $\mathbf{V}_\beta^{-1}|\mathbf{y},\boldsymbol{\beta},\sigma^2,\underline{\upsilon}_\beta,\underline{\mathbf{V}}_\beta^{-1}$

The full conditional posterior distribution of $\mathbf{V}_\beta^{-1}|\mathbf{y},\boldsymbol{\beta},\sigma^2,\underline{\upsilon}_\beta,\underline{\mathbf{V}}_\beta^{-1}$ can be defined as:

$$p(\mathbf{V}_\beta^{-1}|\mathbf{y},\boldsymbol{\beta},\sigma^2,\underline{\upsilon}_\beta,\underline{\mathbf{V}}_\beta^{-1}) \propto p(\mathbf{y}|\boldsymbol{\beta},\sigma^2)p(\boldsymbol{\beta})p(\mathbf{V}_\beta^{-1})$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{nT}{2}}} \times \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)\right)$$

$$\times \frac{1}{(2\pi)^{\frac{np}{2}}|\mathbf{V}_\beta|^{\frac{n}{2}}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)^T\mathbf{V}_\beta^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)\right)$$

$$\times \frac{1}{(2)^{\frac{\underline{\upsilon}_\beta p}{2}}\Gamma_p\left(\frac{\underline{\upsilon}_\beta}{2}\right)|\underline{\mathbf{V}}_\beta^{-1}|^{\frac{\underline{\upsilon}_\beta}{2}}}|\mathbf{V}_\beta^{-1}|^{(\underline{\upsilon}_\beta-p-1)/2}\exp\left(-\frac{1}{2}\text{trace}(\underline{\mathbf{V}}_\beta\mathbf{V}_\beta^{-1})\right) \quad (21)$$

The exponential terms of (Eq. 21) can be manipulated as below:

$$\sum_{i=1}^{n}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)^T\mathbf{V}_\beta^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta) = \sum_{a=1}^{p}\sum_{b=1}^{p}\mathbf{V}_{\beta(ab)}^{-1}\mathbf{S}_{\boldsymbol{\mu}_\beta(ab)}$$

$$= \text{trace}\left(\mathbf{V}_\beta^{-1}\mathbf{S}_{\boldsymbol{\mu}_\beta}^{\mathbf{T}}\right) = \text{trace}\left(\mathbf{V}_\beta^{-1}\mathbf{S}_{\boldsymbol{\mu}_\beta}\right)$$

$$\mathbf{S}_{\boldsymbol{\mu}_{\beta(ab)}} = \sum_{i=1}^{n}(\boldsymbol{\beta}_{i(a)} - \boldsymbol{\mu}_{\beta(a)})(\boldsymbol{\beta}_{i(b)} - \boldsymbol{\mu}_{\beta(b)}) \text{ and}$$

$$\mathbf{S}_{\boldsymbol{\mu}_\beta} = \sum_{i=1}^{n}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)^T \text{ and}$$

$$\sum_{a=1}^{p}\sum_{b=1}^{p}Y_{(ab)}X_{(ab)} = \text{Trace}(\mathbf{YX}^T) \quad (22)$$

Substitute the exponential term in (Eq.21) using (Eq.22) so that we obtain:

$$p(\mathbf{V}_\beta^{-1}|\mathbf{y},\boldsymbol{\beta},\sigma^2,\underline{\upsilon}_\beta,\underline{\mathbf{V}}_\beta^{-1}) = \exp\left(-\frac{1}{2}\text{trace}\left(\mathbf{S}_{\boldsymbol{\mu}_\beta}\mathbf{V}_\beta^{-1}\right)\right)|\mathbf{V}_\beta^{-1}|^{\frac{((\underline{\upsilon}_\beta+n)-p-1)}{2}}$$

$$\times \exp\left(-\frac{1}{2}\text{trace}(\underline{\mathbf{V}}_\beta\mathbf{V}_\beta^{-1})\right)$$

$$= |\mathbf{V}_\beta^{-1}|^{((\underline{\upsilon}_\beta+n)-p-1)/2}\exp\left(-\frac{1}{2}\text{trace}\left(\left(\mathbf{S}_{\boldsymbol{\mu}_\beta} + \underline{\mathbf{V}}_\beta\right)\mathbf{V}_\beta^{-1}\right)\right)$$

$$= |\mathbf{V}_\beta^{-1}|^{(\bar{\upsilon}_\beta-p-1)/2}\exp\left(-\frac{1}{2}\text{trace}(\bar{\mathbf{V}}_\beta\mathbf{V}_\beta^{-1})\right) \quad (23)$$

(Eq. 23) is proportional to ~~a~~ the Wishart distribution:

$$\mathbf{V}_\beta^{-1}|\mathbf{y},\boldsymbol{\beta},\sigma^2,\underline{\upsilon}_\beta,\underline{\mathbf{V}}_\beta^{-1} \propto W(\bar{\upsilon}_\beta,\bar{\mathbf{V}}_\beta^{-1}) \quad (24)$$

where

$$\bar{\upsilon}_\beta = \underline{\upsilon}_\beta + n \text{ and } \bar{\mathbf{V}}_\beta = \sum_{i=1}^{n} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)^T + \underline{\mathbf{V}}_\beta \quad (25)$$

The MCMC with Gibbs Sampling for the Bayesian heterogeneous coefficients fixed effect panel data model can be written as:

1. Set $\boldsymbol{\beta}_{i(0)} = \boldsymbol{\mu}_{\beta(0)} = \widehat{\boldsymbol{\beta}}$, $\widehat{\boldsymbol{\beta}}$ is the OLS estimator of $\boldsymbol{\beta}$
2. Do *i*=1 to *n*
   2.1. Do *m = 1 to M,*
      a. Given $\boldsymbol{\beta}_i = \boldsymbol{\beta}_{i(\mathbf{m-1})}$ and $\boldsymbol{\mu}_\beta = \boldsymbol{\mu}_{\beta(m-1)}$ generate

$$\sigma^2_{(m)} \sim \text{IG}\left(\underline{\sigma}^2 + \frac{nT}{2}, \underline{\gamma} + 2^{-1} \sum_{i=1}^{n} (\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)^T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i)\right)$$

      and

$$\mathbf{V}^{-1}_{\beta(m)} \sim \text{W}\left((\underline{\upsilon}_\beta + n), \left[\sum_{i=1}^{n} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)^T + \underline{\mathbf{V}}_\beta\right]^{-1}\right)$$

      b. Given $\mathbf{V}_\beta^{-1} = \mathbf{V}_\beta^{-1}{}_{(m)}$, generate

$$\boldsymbol{\mu}_{\beta(m)} \sim \text{N}\left((n\mathbf{V}_\beta^{-1} + \underline{\boldsymbol{\Sigma}}_\beta^{-1})\left(\mathbf{V}_\beta^{-1}\sum_{i=1}^{n} \boldsymbol{\beta}_i + \underline{\boldsymbol{\Sigma}}_\beta^{-1}\underline{\boldsymbol{\mu}}_\beta\right), (n\mathbf{V}_\beta^{-1} + \underline{\boldsymbol{\Sigma}}_\beta^{-1})\right)$$

      c. Given $(\sigma^2, \boldsymbol{\mu}_\beta, \mathbf{V}_\beta^{-1})_{(m)}$, generate

$$\boldsymbol{\beta}_i \sim \text{N}\left((\mathbf{V}_\beta^{-1} + \sigma^{-2}\mathbf{X}_i^T\mathbf{X}_i)^{-1}(\mathbf{V}_\beta^{-1}\boldsymbol{\mu}_\beta + \sigma^{-2}\mathbf{X}_i^T\mathbf{y}_i), (\mathbf{V}_\beta^{-1} + \sigma^{-2}\mathbf{X}_i^T\mathbf{X}_i)^{-1}\right)$$

   2.2. End do;
3. End do;
4. Burn-In: Throw away the first N_0 of observations of $(\boldsymbol{\beta}_i, \sigma^2, \boldsymbol{\mu}_\beta, \mathbf{V}_\beta^{-1})_{(m)}$. Usually, the number of burn-in (N_0) is 10% of number of iterations. In the computation, the prior hyperparameters ($\underline{\boldsymbol{\mu}}_\beta$, $\underline{\Sigma}_\beta$, $\underline{\upsilon}_\beta$, $\underline{\mathbf{V}}_\beta^{-1}$, $\underline{\sigma}^2$, $\underline{\gamma}$) are specified by the analyst. The values of the hyperparameters are specified based on the type of prior that will be used whether informative or uninformative priors. Usually the uninformative prior is used. The uninformative prior is achieved by setting $\underline{\boldsymbol{\mu}}_\beta$, $\underline{\upsilon}_\beta$, $\underline{\sigma}^2$, $\underline{\gamma}$ are closed to zero and $\underline{\Sigma}_\beta$ takes large value e.g. 10^5.

Hypothesis testing

The common advantage of Bayesian approach is for hypothesis testing. We can apply the credible interval for testing the significant parameters of interest.

$$H_0: \beta_i = 0 \ vs \ H_1: \beta_i \neq 0$$

Using central limit theorem distribution, for the probability error type one is 0.05, the null hypothesis is rejected if the 95% credible interval does not include 0, then one may conclude that the coefficient is significantly different from 0, and the predictor is important.

4. Discussion and Conclusion

The heterogeneous coefficients panel data model is the extension of standard panel data model which can be used to solve the non-stationarity problem. The non/semi-parametric and GWPR models have been applied to estimate the non-stationarity models. Here we present the alternative approach to cover the non-stationarity problem for panel data modeling by introducing the heterogeneous coefficients model. We use hierarchical Bayesian approach to estimate the parameters of the model. The detail derivation and MCMC procedures are presented. We find that the joint posterior distribution of heterogeneous coefficients panel data model does not follow standard distribution form. MCMC with Gibbs sampling algorithm is used to estimate the conditional posterior distribution and obtain the parameters estimate. The Bayesian approaches give an advantage in hypothesis testing. Using credible interval which is obtained from 2.5% and 97.5% quantiles of iteration samples. The null hypothesis testing is accepted if the credible interval includes zero value.

Acknowledgments
This paper is funded by RFU Unpad contract: 1732 d/UN6.RKT/LT/2018. The authors thank Rector Universitas Padjadjaran. The important and constructive recommendation to improve the quality of this paper is also gratefully acknowledged.

References

[1] Choi J, Lawson AB, Cai B, Hossai MM, Kirby RS, Liu J 2012 *Stat Methods Med Res.* **21** 445–56
[2] Ernst AF, Albers CJ 2017 *Peer J* . **5** 1-26
[3] Klotz S 2004 *Cross Section Dependence in Spatial Econometrics Models With an Application to German Start up Activity Data* (North America: Transaction Publisher) p 20
[4] Koop G 2003 *Bayesian Econometrics* (California: Willey) p 44
[5] Baltagi BH 2005 *Econometrics Analysis of Panel Data* (London: John Wiley & Sons Ltd) p 14
[6] Blangiardo M, Cameletti M, Baio G, Rue H 2013 *Spatial and Spatio-temporal Epidemiology*. 4 33–49
[7] Mou Y, He, Zhou 2017 *Sustainability* **9** 1-17
[8] Rao P 1992 *Identifiability in Stochastic Models Characterization of probability distribution* (London: Academic Press, Inc)
[9] Li Q, Huang CJ, Li D, Fu TT 2002 *J Bus Econ Stat.* **20** 412-22
[10] Li D, Chen J, Gao J 2011 *J Econom*. **14** 387-408
[11] Cai R, Yu D, Oppenheimer M 2014 *Int J Food Agric Econ.* **39** 230-52
[12] Kennedy P 2008 *A Guide to Econometrics 6E.* (London: Blackwell Publishing) p 216
[13] Jaya IGNM, Ruchjana BN, Abdulah AS and Hermawan E 2016 *Int J Appl Math Stat* **54** 94-103
[14] Wheeler DC, Páez A, Spinney J, Waller LA 2014 *Papers in Regional Science* **93** 663-683
[15] Law J, Haining R 2004 *Geographical Analysis* **36** 197-216
[16] Jaya IGNM, Folmer H, Ruchjana BN, Kristiani F, Andriyana Y 2017 *Modeling of Infectious Diseases: A Core Research Topic for the Next Hundred Years* (United States: Springer) p 243
[17] LeSage J, Pace RK 2009 *Introduction to Spatial Econometrics* (Boca Raton: Chapman & Hall) p 124
[18] Hsio C 2007 *Test* **16** 1-22
[19] LeSage JP, Vance C, Chih YY 2017 *Reg. Sci. Urban Econ.* **62** 46-55