Aceh Int. J. Sci. Technol., 12(1) 40-52
April 2023
doi: 10.13170/aijst.12.1.30403

Aceh International Journal of Science and Technology
ISSN: p-2088-9860; e-2503-2398
Journal homepage: http://jurnal.usk.ac.id/aijst

Spatial Autoregressive Modeling on Linear Mixed Models for Dependency Between Regions

Timbang Sirait[1*]

[1]Program Studi D-IV Statistika, Politeknik Statistika STIS
Jalan Otto Iskandardinata No. 64C, Jakarta Timur, DKI Jakarta, 13330, Indonesia
*Corresponding author: timbang@stis.ac.id

Received : January 20, 2023
Received in revised form February 18, 2023
Accepted : February 19, 2023
Online : April 30, 2023

Abstract – This study develops a linear mixed model (LMM) that includes spatial effects between regions with a spatial autoregressive model (SAR model). Between observations (regions) on that LMM are usually assumed to be independent. However, these assumptions are not always fulfilled due to dependency between regions. There are two important parts in spatial modeling: spatial dependence and spatial heterogeneity. In this study, we are concerned with the spatial lag or SAR models because dependency between variables of interest is easier to predict. On the other hand, all observations are real and can be directly seen from the data patterns. In addition, as a challenge for researchers to find all estimators while the values of the spatial dependence, sampling variance, and component variance are all unknown. This study aims to find all parameter estimators using a numerical approach and exact solutions. All exact estimators obtained are consistent estimators.

Keywords: LMM, SAR model, dependency, sampling variance, component variance, consistent

Introduction

Along with the development of statistical modeling, assumptions between independent regions are not always fulfilled due to dependency between regions or, in other words, between regions that have similar properties. This means that one region with another has almost the same characteristics. For example, one high-income region and another is also a high-income region. The similarity of these properties is usually known as the spatial effect between regions. As in the first law of geography Tobler (1979) in Anselin (1988) states, "Everything is related to everything else, but near things are more related than distant things" (Anselin, 1988). Therefore, if the independent assumption on the random effect is not met, the estimators obtained will provide an invalid model because the estimators produced are biased (Salvati, 2004).

Anselin (Anselin, 1988) states that spatial modeling has two important parts: spatial dependence and spatial heterogeneity. Spatial dependence is also known as spatial autocorrelation, while spatial heterogeneity is related to spatial structure. Spatial dependence is part of the modeling that describes the value of one region influenced by the values of regions surrounding its region or location (neighbor). The similarity of properties for locations close to each other indicates this. Meanwhile, spatial heterogeneity describes the structure's instability, which can be in the form of a non-constant error variance or the regression coefficient of a variable that varies spatially, indicated by differences in properties from one location to another.

In addition, Anselin (Anselin, 1988) also stated that spatial dependence modeling is still divided into two types: the spatial lag model and the spatial error model. The spatial lag model, commonly referred to as the spatial autoregressive model (SAR model), is a spatial autoregressive in the dependent variable (the variable of interest). While the spatial error model (SEM) states a spatial autocorrelation in the error variable or dependence on the error. In the SAR model, the dependencies between variables of interest are easier to be predicted because the observations are actual and can be directly seen from the data patterns.

Aceh Int. J. Sci. Technol., 12(1) 40-52
April 2023
doi: 10.13170/aijst.12.1.30403

In real problems, statistical modeling is faced with situations where there is a spatial dependence on variables of interest. It is also faced with conditions between locations (object of research) that are not independent. In addition, researchers are also still troubled by conditions where the value of the spatial dependence is unknown, and even the values of the sampling variance and the component variance are also unknown. Many researchers have developed modeling where they took some set values (fixed values) for both spatial autocorrelation and variances, and a few of them worked by simulations. For example, research conducted by (Quintaes *et al.*, 2011), (Siswantining *et al.*, 2012), (Porter *et al.*, 2013), (Wang *et al.*, 2015), (Zainuddin *et al.*, 2015), (Asfar *et al.*, 2016), (Nusrang *et al.*, 2017), (Watjou *et al.*, 2019), (Watjou *et al.*, 2020), (Amaliana *et al.*, 2021), (Amaliana and Dewi, 2022). All researchers took some fixed values to obtain all estimator values. Quintaes *et al.* (2011) used a regression model with random effects (developing small area estimation). They used real data and concluded that their proposed model was better than the model without random effects. Wang *et al.* (2015) proposed a model with more relaxed assumptions on spatial-temporal random errors. They obtained that the proposed model provided an improvement model. Watjou *et al.* (2019) proposed an approach to estimating the spatial trend of the prevalence of having no health insurance coverage. They obtained that the proposed model (model-based methods) provided the benefits of modeling the weights as flexibly as possible. Watjou *et al.* (2020) used hierarchical spatial smoothing models. They showed the benefits of that model towards obtaining more reliable estimates for areas at the lowest geographical level in case a spatial trend is present in the data. Amaliana *et al.* (2021) showed that the spatial empirical best linear unbiased prediction (SEBLUP) model with radial distance spatial weights matrix provided a better result than models with other spatial weights matrices. Amaliana and Dewi (2022) showed that the empirical best linear unbiased prediction (EBLUP) model is better than the SEBLUP model. Both models used a queen contiguity spatial weighted matrix. Siswantining et al. (2012) used a hierarchical Bayes small area estimation model and obtained the model was more efficient than a direct estimator. Porter *et al.* (2013) used the conditional autoregressive class of spatial models in addition to functional covariates in small area estimation (developing small area estimation model). They concluded that the proposed model could effectively improve the public-use American community survey data estimation. Zainuddin *et al.* (2015) used the SEBLUP using logarithmic transformation and obtained the model provided a better result than the other models. Asfar *et al.* (2016) and Nusrang *et al.* (2017) used the SEBLUP model and obtained that the model provided a better estimation than the EBLUP model. The last five researchers worked with simulations.

Due to the abovementioned problems, the researcher will develop statistical modeling, especially in the SAR model, due to the many dependent conditions that can be predicted through the variables of interest. The researcher is interested to find out the estimating parameters so that the statistical modeling that will be formed can be obtained. A mixed model approached parameter estimation in this study by numerical and exact approaches so that all predicted parameters could be estimated.

Materials and Methods

Linear Mixed Model for Regions Dependence

The linear mixed model, as stated in (McCulloch and Searle, 2001), is as follows:

$$\mathbf{s} = \mathbf{K}\boldsymbol{\alpha} + \mathbf{L}\boldsymbol{\gamma} + \boldsymbol{\delta}, \tag{1}$$

Where, $\mathbf{s} = m \times 1$ observation vector, $\mathbf{K} = m \times (p+1)$ observation matrix, $\boldsymbol{\alpha} = (p+1) \times 1$ coefficient vector of k variables (fixed effects), $\mathbf{L} = m \times q$ observation matrix, $\boldsymbol{\gamma} =$ coefficient vector of l variables (random effects), and $\boldsymbol{\delta} = m \times 1$ sampling variance vector. In a linear mixed model, $\boldsymbol{\delta}$ having a distribution with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{G} = \sigma_\delta^2 \mathbf{I}_m$ and $\boldsymbol{\gamma}$ a distribution with mean vector and covariance matrix are also independent. Simplicity, we can rewrite $\boldsymbol{\delta} \sim N(\mathbf{0}, \sigma_\delta^2 \mathbf{I}_m)$ and $\boldsymbol{\gamma} \sim N(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}_q)$, where σ_δ^2 and σ_γ^2 are sampling variance and component variance, respectively. $\mathbf{I}_m$ Dan $\mathbf{I}_q$ is $m \times 1$ an $q \times 1$ identity matrix, respectively.

If there is a spatial effect between regions in equation (1), then

$$\boldsymbol{\gamma} = \lambda \mathbf{W}\boldsymbol{\gamma} + \mathbf{v}, \tag{2}$$

where $\lambda =$ spatial autocorrelation coefficient, $\mathbf{W} = q \times q$ spatial weight matrix, and $\mathbf{v} = q \times 1$ random error vector between regions that have a distribution with mean vector and covariance matrix or can be rewritten with $\mathbf{v} \sim N\left(\mathbf{0}, \sigma_v^2 \mathbf{I}_q\right)$. Therefore, $\boldsymbol{\gamma}$ has no more extended constant variance, however, $\boldsymbol{\delta}$ and $\mathbf{v}$ are independent.

We have from the equation (2) as follows:

$$\boldsymbol{\gamma} - \lambda \mathbf{W}\boldsymbol{\gamma} = \mathbf{v} \Leftrightarrow \boldsymbol{\gamma} = \left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{v},$$

as a result,

$$E[\boldsymbol{\gamma}] = E\left[\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{v}\right] = \left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} E[\mathbf{v}] = \mathbf{0}, \text{ because } E[\mathbf{v}] = \mathbf{0}, \text{ and}$$

$$\operatorname{var}[\boldsymbol{\gamma}] = \operatorname{var}\left[\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{v}\right] = \left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{R}\left[\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1}\right]^t = \mathbf{U},$$

Where $\mathbf{U} =$ $q \times q$ symmetric matrix. Therefore, $\boldsymbol{\gamma} \sim N(\mathbf{0}, \mathbf{U})$. However, covariance $\boldsymbol{\gamma}$ becomes no constant because there is a correlation between regions, in the sense that the covariance between regions is not equal to zero.

Spatial Autoregressive

The following model is a spatial autoregressive model (SAR model) that is:

$$\mathbf{s} = \mathbf{K}\boldsymbol{\alpha} + \rho \tilde{\mathbf{W}}\mathbf{s} + \boldsymbol{\delta}, \tag{3}$$

where $\rho =$ spatial autoregressive coefficient and $\tilde{\mathbf{W}} = m \times m$ spatial weight matrix.

If equation (1) is combined with equation (3), then they will produce a new equation as follows:

$$\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s} = \mathbf{K}\boldsymbol{\alpha} + \mathbf{L}\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{v} + \boldsymbol{\delta}. \tag{4}$$

If we let $\boldsymbol{\eta} = \mathbf{L}\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{v} + \boldsymbol{\delta},$ then equation (4) can be rewritten as follows:

$$\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s} = \mathbf{K}\boldsymbol{\alpha} + \boldsymbol{\eta}. \tag{5}$$

The mean and covariance $\boldsymbol{\eta}$ are as follows:

$$E[\boldsymbol{\eta}] = E\left[\mathbf{L}\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{v} + \boldsymbol{\delta}\right] = \mathbf{0}, \text{ because } E[\mathbf{v}] = E[\boldsymbol{\delta}] = \mathbf{0}, \text{ and}$$

$$\operatorname{var}[\boldsymbol{\eta}] = \mathbf{L}\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1} \mathbf{R}\left[\left(\mathbf{I}_q - \lambda \mathbf{W}\right)^{-1}\right]^t \mathbf{L}^t + \mathbf{G} = \mathbf{L}\mathbf{U}\mathbf{L}^t + \mathbf{G} = \mathbf{F},$$

Where $\mathbf{F} =$ $n \times n$ symmetric matrix. Because the covariance of $\boldsymbol{\gamma}$ is not constant, as a result, the covariance of $\boldsymbol{\eta}$ is also not constant, and it can be written with $\boldsymbol{\eta} \sim N(\mathbf{0}, \mathbf{F})$.

Based on equation (5), we have the mean and covariance $\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s}$ as follows:

$$E\left[\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s}\right] = E[\mathbf{K}\boldsymbol{\alpha} + \boldsymbol{\eta}] = \mathbf{K}\boldsymbol{\alpha}, \text{ because } E[\boldsymbol{\eta}] = \mathbf{0},$$

$$\operatorname{var}\left[\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s}\right] = \operatorname{var}[\mathbf{K}\boldsymbol{\alpha} + \boldsymbol{\eta}] = \operatorname{var}[\boldsymbol{\eta}] = \mathbf{F}, \text{ because } \boldsymbol{\alpha} \text{ it is fixed.}$$

Therefore $\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s} \sim N(\mathbf{K}\boldsymbol{\alpha}, \mathbf{F})$, where it is the marginal distribution, $\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s}$. This result shows that the covariance $\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s}$ is also not constant.

Next, if $\boldsymbol{\gamma}$ it is fixed, then equation (5) can be rewritten as follows:

$$\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s} = \mathbf{K}\boldsymbol{\alpha} + \boldsymbol{\eta}, \tag{6}$$

where $\boldsymbol{\eta} = \mathbf{L}\boldsymbol{\gamma} + \boldsymbol{\delta}$.

Based on equation (6), we have the mean and covariance $\boldsymbol{\eta}$ as follows:

$$E[\boldsymbol{\eta}] = E[\mathbf{L}\boldsymbol{\gamma} + \boldsymbol{\delta}] = \mathbf{L}\boldsymbol{\gamma}, \text{ because } E[\boldsymbol{\delta}] = \mathbf{0}, \text{ and } \operatorname{var}[\boldsymbol{\eta}] = \operatorname{var}[\mathbf{L}\boldsymbol{\gamma} + \boldsymbol{\delta}] = \operatorname{var}[\boldsymbol{\delta}] = \mathbf{G}.$$

Next, conditional both the mean and covariance of $\left(\mathbf{I}_m - \rho \tilde{\mathbf{W}}\right)\mathbf{s}$ the given $\boldsymbol{\gamma}$ area

doi: 10.13170/aijst.12.1.30403

$$E\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}|\boldsymbol{\gamma}\right]=E\left[\mathbf{K}\boldsymbol{\alpha}+\boldsymbol{\eta}\right]=\mathbf{K}\boldsymbol{\alpha}+E\left[\boldsymbol{\eta}\right]=\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma},$$

$$\mathrm{var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}|\boldsymbol{\gamma}\right]=\mathrm{var}\left[\mathbf{K}\boldsymbol{\alpha}+\boldsymbol{\eta}\right]=\mathrm{var}\left[\boldsymbol{\eta}\right]=\mathbf{G},$$

and we have $\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}|\boldsymbol{\gamma}\sim N\left(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma},\mathbf{G}\right).$

We recall equation (1) which can be rewritten as follows:

$$\mathbf{s}=\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma}+\boldsymbol{\delta}=\left[\mathbf{K}\ \vdots\ \mathbf{L}\right]\left[\boldsymbol{\alpha}\ \vdots\ \boldsymbol{\gamma}\right]^t+\boldsymbol{\delta}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\delta},$$

Where $\mathbf{X}=\left[\mathbf{K}\ \vdots\ \mathbf{L}\right]$ and $\boldsymbol{\beta}=\left[\boldsymbol{\alpha}\ \vdots\ \boldsymbol{\gamma}\right]^t$. $\mathbf{X}$ is a less-than-full rank matrix, the least squares estimator cannot be evaluated directly, so we need to reparameterize it. Matrix visualization $\mathbf{X},\ \boldsymbol{\beta},\ \boldsymbol{\delta}$ is as follows:

$$\mathbf{X}=\left[\begin{array}{ccccc:cccc} 1 & k_{11} & k_{21} & \cdots & k_{p1} & 1 & 0 & \cdots & 0\\ 1 & k_{12} & k_{22} & \cdots & k_{p2} & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & k_{1m} & k_{2m} & \cdots & k_{pm} & 0 & 0 & \cdots & 1 \end{array}\right],$$

$$\boldsymbol{\beta}=\left[\begin{array}{ccccc:cccc}\mu & \alpha_1 & \alpha_2 & \cdots & \alpha_p & \gamma_1 & \gamma_2 & \cdots & \gamma_q\end{array}\right]^t,\quad \boldsymbol{\delta}=\left[\begin{array}{cccc}\delta_1 & \delta_2 & \cdots & \delta_m\end{array}\right]^t$$

We then reparameterize $\mathbf{X}$ and obtain a full rank matrix as follows:

$$\mathbf{X}_*=\left[\begin{array}{cccc:cccc} k_{11} & k_{21} & \cdots & k_{p1} & 1 & 0 & \cdots & 0\\ k_{12} & k_{22} & \cdots & k_{p2} & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots\\ k_{1m} & k_{2m} & \cdots & k_{pm} & 0 & 0 & \cdots & 1 \end{array}\right],\quad \boldsymbol{\beta}_*=\left[\begin{array}{cccc:cccc}\alpha_1 & \alpha_2 & \cdots & \alpha_p & \mu_1 & \mu_2 & \cdots & \mu_q\end{array}\right]^t,$$

where $\mu_k=\mu+\gamma_k$ for $k=1,2,\cdots,q.$ The least squares estimator of $\boldsymbol{\beta}_*$ is

$$\hat{\boldsymbol{\beta}}_*=\left(\mathbf{X}_*^t\mathbf{X}_*\right)^{-1}\left(\mathbf{X}_*^t\mathbf{s}\right). \tag{7}$$

where $\hat{\mu}=\dfrac{1}{m}\sum_{i=1}^{m}s_i$ and $\hat{\gamma}_k=\hat{\mu}_k-\hat{\mu}.$

Moran's Index

Benerjee *et al.* (2004), Kelejian and Prucha (2004), Gebremariam *et al.* (2006), Zhang and Lin (2007), Gebremariam (2007), Gebremariam *et al.* (2007, 2008), Ngeleza (2011), Krishnapillai and Kinnucan (2012), Huo *et al.* (2012) and Tłuczak (2013) are all in (Sirait *et al.*, 2017b) state that Moran's index is one of many indexes that is used to detect spatial effect between regions. As an illustration (or example), this study will use queen contiguity to find the weight matrix. Detailed formulation and explanation can be seen in (Sirait *et al.*, 2017b).

Results

Parameters Estimation

We recall equation (2) and obtain that $\mathbf{v}=\boldsymbol{\gamma}-\lambda\mathbf{W}\boldsymbol{\gamma},$ where $\mathbf{v}\sim N(\mathbf{0},\mathbf{R}).$ Therefore, the maximum likelihood function of the random vector $\mathbf{v}$ is as follows:

$$L\left(\mathbf{v}\right)=\frac{1}{\left(2\pi\right)^{\frac{q}{2}}\left|\mathbf{R}\right|^{\frac{1}{2}}}e^{-\frac{1}{2}\mathbf{v}^t\mathbf{R}^{-1}\mathbf{v}}. \tag{8}$$

Based on equation (8), the maximum likelihood function $\boldsymbol{\gamma}$ is as follows:

$$L\left(\boldsymbol{\gamma}\right)=\frac{1}{\left(2\pi\right)^{\frac{q}{2}}\left|\mathbf{R}\right|^{\frac{1}{2}}}e^{-\frac{1}{2}\boldsymbol{\gamma}^t\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\boldsymbol{\gamma}}\left|J\right|, \tag{9}$$

where $\left|J\right|=\left\|\mathbf{I}_q-\lambda\mathbf{W}\right\|=$ the absolute value of the determinant of $\mathbf{I}_q-\lambda\mathbf{W}.$

Next, the maximum likelihood function $\left(\mathbf{I}_m - \rho\tilde{\mathbf{W}}\right)\mathbf{s} \mid \boldsymbol{\gamma}$ is as follows:

$$L\left(\left(\mathbf{I}_m - \rho\tilde{\mathbf{W}}\right)\mathbf{s} \mid \boldsymbol{\gamma}\right) = \frac{1}{(2\pi)^{\frac{m}{2}}|\mathbf{G}|^{\frac{1}{2}}} e^{-\frac{1}{2}\left(\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma})\right)^t \mathbf{G}^{-1}\left(\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma})\right)}. \tag{10}$$

Based on both equations (9) and (10), we have the maximum likelihood function as follows:

$$L(\boldsymbol{\alpha},\boldsymbol{\gamma}) = L\left(\left(\mathbf{I}_m - \rho\tilde{\mathbf{W}}\right)\mathbf{s} \mid \boldsymbol{\gamma}\right) L(\boldsymbol{\gamma}) = \left\{\frac{1}{(2\pi)^{\frac{m}{2}}|\mathbf{G}|^{\frac{1}{2}}} e^{-\frac{1}{2}\left(\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma})\right)^t \mathbf{G}^{-1}\left(\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma})\right)}\right\}$$
$$\times\left\{\frac{1}{(2\pi)^{\frac{q}{2}}|\mathbf{R}|^{\frac{1}{2}}} e^{-\frac{1}{2}\boldsymbol{\gamma}^t\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\boldsymbol{\gamma}} \left\|\mathbf{I}_q - \lambda\mathbf{W}\right\|\right\}.$$

We then take natural logarithm and obtain the function as follows

$$\begin{aligned}\log L(\boldsymbol{\alpha},\boldsymbol{\gamma}) = &-\left(\frac{m+q}{2}\right)\log(2\pi) - \frac{1}{2}\log|\mathbf{G}| - \frac{1}{2}\log|\mathbf{R}| \\ &-\frac{1}{2}\left(\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma})\right)^t \mathbf{G}^{-1}\left(\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma})\right) \\ &-\frac{1}{2}\boldsymbol{\gamma}^t\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\boldsymbol{\gamma} + \log\left\|\mathbf{I}_q-\lambda\mathbf{W}\right\|.\end{aligned} \tag{11}$$

Based on equation (11), we take the derivative for $\boldsymbol{\alpha}$ and obtain the estimator as follows:

$$\frac{\partial \log L(\boldsymbol{\alpha},\boldsymbol{\gamma})}{\partial\boldsymbol{\alpha}} = \mathbf{K}^t\mathbf{G}^{-1}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s} - \mathbf{K}^t\mathbf{G}^{-1}\mathbf{K}\boldsymbol{\alpha} - \mathbf{K}^t\mathbf{G}^{-1}\mathbf{L}\boldsymbol{\gamma}.$$

We then take $\left.\frac{\partial \log L(\boldsymbol{\alpha},\boldsymbol{\gamma})}{\partial\boldsymbol{\alpha}}\right|_{\substack{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}\\ \boldsymbol{\gamma}=\hat{\boldsymbol{\gamma}}}} = \mathbf{0},$ and obtain

$$\mathbf{K}^t\mathbf{G}^{-1}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s} - \mathbf{K}^t\mathbf{G}^{-1}\mathbf{K}\hat{\boldsymbol{\alpha}} - \mathbf{K}^t\mathbf{G}^{-1}\mathbf{L}\hat{\gamma} = \mathbf{0},$$

As a result, the estimator $\boldsymbol{\alpha}$ is as follows:

$$\hat{\boldsymbol{\alpha}} = \left(\mathbf{K}^t\mathbf{G}^{-1}\mathbf{K}\right)^{-1}\mathbf{K}^t\mathbf{G}^{-1}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s} - \mathbf{L}\hat{\gamma}\right]. \tag{12}$$

Next, we take derivative for $\boldsymbol{\gamma}$ and obtain

$$\mathbf{L}^t\mathbf{G}^{-1}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s} - \mathbf{L}^t\mathbf{G}^{-1}\mathbf{K}\hat{\boldsymbol{\alpha}} - \mathbf{L}^t\mathbf{G}^{-1}\mathbf{L}\hat{\boldsymbol{\gamma}} - \left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\hat{\boldsymbol{\gamma}} = \mathbf{0},$$

Therefore, the estimator $\boldsymbol{\gamma}$ is as follows:

$$\hat{\boldsymbol{\gamma}} = \left[\mathbf{L}^t\mathbf{G}^{-1}\mathbf{L} + \left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\right]^{-1}\mathbf{L}^t\mathbf{G}^{-1}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s} - \mathbf{K}\hat{\boldsymbol{\alpha}}\right]. \tag{13}$$

Equations (12) and (13) show a dependency or relation between $\hat{\alpha}$ dan $\hat{\beta}$, so both estimators can be combined by modifying the matrix. The aims are to make it easier to obtain these estimators simultaneously; simultaneously, the displayed matrix becomes simpler. Those matrices are as follows:

$$\begin{bmatrix}\mathbf{K}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{K}^t\mathbf{G}^{-1}\mathbf{L} \\ \mathbf{L}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{L}^t\mathbf{G}^{-1}\mathbf{L} + \left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\gamma}}\end{bmatrix} = \begin{bmatrix}\mathbf{K}^t\mathbf{G}^{-1}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s} \\ \mathbf{L}^t\mathbf{G}^{-1}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix},$$

as a result,

$$\begin{bmatrix}\hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\gamma}}\end{bmatrix} = \begin{bmatrix}\mathbf{K}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{K}^t\mathbf{G}^{-1}\mathbf{L} \\ \mathbf{L}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{L}^t\mathbf{G}^{-1}\mathbf{L} + \left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{K}^t\mathbf{G}^{-1}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s} \\ \mathbf{L}^t\mathbf{G}^{-1}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}. \tag{14}$$

As equations (1) and (2) have been stated that $\mathbf{G}=\sigma_{\gamma}^{2}\mathbf{I}_{m}$ and $\mathbf{R}=\sigma_{v}^{2}\mathbf{I}_{q}$ so that $\mathbf{G}^{-1}=\left(\sigma_{\delta}^{2}\mathbf{I}_{m}\right)^{-1}=\left(\sigma_{\delta}^{2}\right)^{-1}\mathbf{I}_{m}$ and $\mathbf{R}^{-1}=\left(\sigma_{v}^{2}\mathbf{I}_{q}\right)^{-1}=\left(\sigma_{v}^{2}\right)^{-1}\mathbf{I}_{q}$. Therefore, equation (14) can be rewritten as follows:

$$\begin{bmatrix}\hat{\boldsymbol{\alpha}}\\ \hat{\boldsymbol{\gamma}}\end{bmatrix}=\begin{bmatrix}\mathbf{K}^{t}\mathbf{K} & \mathbf{K}^{t}\mathbf{L}\\ \mathbf{L}^{t}\mathbf{K} & \mathbf{L}^{t}\mathbf{L}+\dfrac{\sigma_{\delta}^{2}}{\sigma_{v}^{2}}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{K}^{t}\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \mathbf{L}^{t}\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}. \tag{15}$$

It was mentioned above that $\mathbf{G}=\sigma_{\delta}^{2}\mathbf{I}_{m}$ and $\mathbf{R}=\sigma_{v}^{2}\mathbf{I}_{q}$ so that $|\mathbf{G}|=\left(\sigma_{\delta}^{2}\right)^{m}$, $|\mathbf{R}|=\left(\sigma_{v}^{2}\right)^{q}$, $\mathbf{G}^{-1}=\left(\sigma_{\delta}^{2}\right)^{-1}\mathbf{I}_{m}$ and $\mathbf{H}^{-1}=\left(\sigma_{v}^{2}\right)^{-1}\mathbf{I}_{q}$. Therefore, equation (11) can be rewritten as follows:

$$\begin{aligned}\log L(\boldsymbol{\alpha},\boldsymbol{\gamma})=&-\left(\frac{m+q}{2}\right)\log\left(2\pi\right)-\frac{m}{2}\log\left(\sigma_{\delta}^{2}\right)-\frac{q}{2}\log\left(\sigma_{v}^{2}\right)\\ &-\frac{1}{2\sigma_{\delta}^{2}}\left(\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-\left(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma}\right)\right)^{t}\left(\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-\left(\mathbf{K}\boldsymbol{\alpha}+\mathbf{L}\boldsymbol{\gamma}\right)\right)\\ &-\frac{1}{2\sigma_{v}^{2}}\boldsymbol{\gamma}^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)\boldsymbol{\gamma}+\log\left\|\mathbf{I}_{q}-\lambda\mathbf{W}\right\|,\end{aligned} \tag{16}$$

Where both of σ_{δ}^{2} and σ_{v}^{2} are unknown variances, whose values can be estimated using a sample of observations. We then take the derivatives of equation (16) in each σ_{δ}^{2} and σ_{v}^{2} set it to zero. Each estimator is as follows:

$$\hat{\sigma}_{\delta}^{2}=\frac{1}{m}\left(\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-\left(\mathbf{K}\hat{\boldsymbol{\alpha}}+\mathbf{L}\hat{\boldsymbol{\gamma}}\right)\right)^{t}\left(\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-\left(\mathbf{K}\hat{\boldsymbol{\alpha}}+\mathbf{L}\hat{\boldsymbol{\gamma}}\right)\right). \tag{17}$$

$$\hat{\sigma}_{v}^{2}=\frac{1}{q}\hat{\boldsymbol{\gamma}}^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)\hat{\boldsymbol{\gamma}}. \tag{18}$$

We will use the estimators of variance from equations (17) and (18), and therefore, equation (15) can write are as follows:

$$\begin{bmatrix}\hat{\boldsymbol{\alpha}}\\ \hat{\boldsymbol{\gamma}}\end{bmatrix}=\begin{bmatrix}\mathbf{K}^{t}\mathbf{K} & \mathbf{K}^{t}\mathbf{L}\\ \mathbf{L}^{t}\mathbf{K} & \mathbf{L}^{t}\mathbf{L}+\dfrac{\hat{\sigma}_{\delta}^{2}}{\hat{\sigma}_{v}^{2}}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{K}^{t}\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \mathbf{L}^{t}\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}. \tag{19}$$

Concentrated Log-Likelihood Function

If the estimators of σ_{δ}^{2} and σ_{v}^{2} are substituted in the function of $\log L(\boldsymbol{\alpha},\boldsymbol{\gamma})$ equation (16), then it is obtained the concentrated log-likelihood function $\log L^{con}(\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\gamma}})$, as follows:

$$\begin{aligned}\log L^{con}(\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\gamma}})=&-\left(\frac{m+q}{2}\right)\left(\log\left(2\pi\right)+1\right)\\ &-\frac{m}{2}\log\left(\frac{1}{m}\left(\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-\left(\mathbf{K}\hat{\boldsymbol{\alpha}}+\mathbf{L}\hat{\boldsymbol{\gamma}}\right)\right)^{t}\left(\left(\mathbf{I}_{m}-\rho\tilde{\mathbf{W}}\right)\mathbf{s}-\left(\mathbf{K}\hat{\boldsymbol{\alpha}}+\mathbf{L}\hat{\boldsymbol{\gamma}}\right)\right)\right)\\ &-\frac{q}{2}\log\left(\frac{1}{q}\hat{\boldsymbol{\gamma}}^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)^{t}\left(\mathbf{I}_{q}-\lambda\mathbf{W}\right)\hat{\boldsymbol{\gamma}}\right)+\log\left\|\mathbf{I}_{q}-\lambda\mathbf{W}\right\|\end{aligned} \tag{20}$$

Now, let $\mathbf{W}$ have eigenvalues $\tau_{1},\tau_{2},\cdots,\tau_{q}$, (Anselin *et al.*, 1996) stated that the acceptable spatial autocorrelation coefficients are as follows:

$$\frac{1}{\tau_{\text{minimum}}}<\rho<\frac{1}{\tau_{\text{maksimum}}}\ \text{dan}\ \frac{1}{\tau_{\text{minimum}}}<\lambda<\frac{1}{\tau_{\text{maksimum}}},$$

doi: 10.13170/aijst.12.1.30403

as well as $\mathbf{W}$ always having the largest eigenvalue equal to one and $\frac{1}{\tau_{\text{minimum}}} \leq -1,$, as a result, $\frac{1}{\tau_{\text{minimum}}} < \rho < 1$ and $\frac{1}{\tau_{\text{minimum}}} < \lambda < 1.$

It is difficult for us to derive the estimators of ρ and λ analytically which maximize $\log L^{con}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}).$ Therefore, this study uses a numerical method approach, that is, sequences forming method of ρ and λ (see Sirait *et al.*, 2017a), in which from such sequences will be obtained a value of ρ and λ respectively that give the largest $\log L^{con}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}).$ The procedures are as follows:

a. We recall equation (7) to find the estimator $\hat{\boldsymbol{\beta}}_*$ and then use it to find the estimator $\hat{\boldsymbol{\beta}}.$ As we know, that $\hat{\boldsymbol{\beta}}$ consists of $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}.$
b. We make the values of sequences of ρ and $\lambda,$ respectively, where $\rho = \text{seq(start value, end value, increasing)}$ and $\lambda = \text{seq(start value, end value, increasing)}.$ In this study, the increased value is 0.01.
c. Based on the result of (a) and (b), we substitute them for equation (20).
d. We check the values of ρ , and λ that produce the largest $\log L^{con}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}).$ Those values will be the optimum estimates.

Next, we find the sampling and component variance using equations (17) and (18), respectively. We first substitute ρ , and λ with its optimum estimates, that is:

$$\hat{\sigma}_\delta^2 = \frac{1}{m}\left(\left(\mathbf{I}_m - \hat{\rho}\tilde{\mathbf{W}}\right)\mathbf{s} - \left(\mathbf{K}\hat{\boldsymbol{\alpha}} + \mathbf{L}\hat{\boldsymbol{\gamma}}\right)\right)^t \left(\left(\mathbf{I}_m - \hat{\rho}\tilde{\mathbf{W}}\right)\mathbf{s} - \left(\mathbf{K}\hat{\boldsymbol{\alpha}} + \mathbf{L}\hat{\boldsymbol{\gamma}}\right)\right). \quad (21)$$

$$\hat{\sigma}_v^2 = \frac{1}{q}\hat{\boldsymbol{\gamma}}^t \left(\mathbf{I}_q - \hat{\lambda}\mathbf{W}\right)^t \left(\mathbf{I}_q - \hat{\lambda}\mathbf{W}\right)\hat{\boldsymbol{\gamma}}. \quad (22)$$

Based on all the estimates, the estimates vector $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$ equation (19) can be rewritten as follows:

$$\begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\gamma}} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^t\mathbf{K} & \mathbf{K}^t\mathbf{L} \\ \mathbf{L}^t\mathbf{K} & \mathbf{L}^t\mathbf{L} + \frac{\hat{\sigma}_\delta^2}{\hat{\sigma}_v^2}\left(\mathbf{I}_q - \hat{\lambda}\mathbf{W}\right)^t\left(\mathbf{I}_q - \hat{\lambda}\mathbf{W}\right) \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{K}^t\left(\mathbf{I}_m - \hat{\rho}\tilde{\mathbf{W}}\right)\mathbf{s} \\ \mathbf{L}^t\left(\mathbf{I}_m - \hat{\rho}\tilde{\mathbf{W}}\right)\mathbf{s} \end{bmatrix}. \quad (23)$$

In this study, the above estimators are called the estimators of parameters of spatial autoregressive modeling on linear mixed models for dependency between regions.

Properties of Estimators

One of the properties of the feasible estimators is consistent estimators (A. M. Mood, F. A. Graybill and D.C. Boes, 1974). In this study, we will check consistency for the estimators in equation (14), that is

$$\begin{bmatrix} \mathbf{K}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{K}^t\mathbf{G}^{-1}\mathbf{L} \\ \mathbf{L}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{L}^t\mathbf{G}^{-1}\mathbf{L} + \left(\mathbf{I}_q - \lambda\mathbf{W}\right)^t \mathbf{R}^{-1}\left(\mathbf{I}_q - \lambda\mathbf{W}\right) \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\gamma}} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^t\mathbf{G}^{-1}\left(\mathbf{I}_m - \rho\tilde{\mathbf{W}}\right)\mathbf{s} \\ \mathbf{L}^t\mathbf{G}^{-1}\left(\mathbf{I}_m - \rho\tilde{\mathbf{W}}\right)\mathbf{s} \end{bmatrix}$$

$$\mathbf{A}\hat{\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{K}^t\mathbf{G}^{-1} & \mathbf{0}_* \\ \mathbf{0} & \mathbf{L}^t\mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \left(\mathbf{I}_m - \rho\tilde{\mathbf{W}}\right)\mathbf{s} \\ \left(\mathbf{I}_m - \rho\tilde{\mathbf{W}}\right)\mathbf{s} \end{bmatrix}$$

$$\mathbf{A}\hat{\boldsymbol{\theta}} = \mathbf{B}\mathbf{d}$$

$$\hat{\boldsymbol{\theta}} = \mathbf{A}^{-1}\mathbf{B}\mathbf{d},$$

where

$$\mathbf{A}=\begin{bmatrix}\mathbf{K}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{K}^t\mathbf{G}^{-1}\mathbf{L}\\ \mathbf{L}^t\mathbf{G}^{-1}\mathbf{K} & \mathbf{L}^t\mathbf{G}^{-1}\mathbf{L}+\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^t\mathbf{R}^{-1}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)\end{bmatrix},\ \hat{\boldsymbol{\theta}}=\begin{bmatrix}\hat{\boldsymbol{\alpha}}\\ \hat{\boldsymbol{\gamma}}\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}\mathbf{K}^t\mathbf{G}^{-1} & \mathbf{0}_*\\ \mathbf{0} & \mathbf{L}^t\mathbf{G}^{-1}\end{bmatrix},$$

Moreover, $\mathbf{d}=\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}$. $\mathbf{0}_*$ $\mathbf{0}$ are the $(p+1)\times m$ and $q\times m$ matrices, respectively.

$$\begin{aligned}\operatorname{asy.var}\{\mathbf{d}\}&=\operatorname{asy.var}\left\{\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}\right\}\\
&=E\left\{\left(\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}-E\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}\right)\left(\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}-E\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}\right)^t\right\}\\
&=E\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}\begin{bmatrix}\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\\ \left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\end{bmatrix}^t-\begin{bmatrix}E\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\\ E\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\end{bmatrix}\begin{bmatrix}E\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\\ E\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\end{bmatrix}^t\\
&=\begin{bmatrix}\operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right] & \operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\\ \operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right] & \operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\end{bmatrix},\end{aligned}$$

where

$E\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]$ is a scalar vector and

$$\operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]=\mathbf{F}=\mathbf{LUL}^t+\mathbf{G}=\mathbf{L}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^{-1}\mathbf{R}\left[\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^{-1}\right]^t\mathbf{L}^t+\mathbf{G}.$$

therefore,

$$\operatorname{asy.var}\{\hat{\boldsymbol{\theta}}\}=\operatorname{asy.var}\{\mathbf{A}^{-1}\mathbf{Bd}\}=\mathbf{A}^{-1}\mathbf{B}\operatorname{asy.var}\{\mathbf{d}\}\mathbf{B}^t\{\mathbf{A}^{-1}\}^t$$

$$\begin{aligned}\lim_{m\to\infty}\operatorname{asy.var}\{\hat{\boldsymbol{\theta}}\}&=\lim_{m\to\infty}\left[\left(\frac{1}{m}\mathbf{A}\right)^{-1}\mathbf{B}\frac{1}{m}\operatorname{asy.var}\{\mathbf{d}\}\mathbf{B}^t\{\mathbf{A}^{-1}\}^t\right]\\
&=\left[\lim_{m\to\infty}\left(\frac{1}{m}\mathbf{A}\right)^{-1}\right]\mathbf{B}\left[\lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\{\mathbf{d}\}\right]\mathbf{B}^t\{\mathbf{A}^{-1}\}^t\\
&=\left[\lim_{m\to\infty}\left(\bar{\mathbf{A}}\right)^{-1}\right]\mathbf{B}\left[\lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\{\mathbf{d}\}\right]\mathbf{B}^t\{\mathbf{A}^{-1}\}^t\\
&=\bar{\mathbf{A}}^{-1}\mathbf{B}\left[\lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\{\mathbf{d}\}\right]\mathbf{B}^t\{\mathbf{A}^{-1}\}^t,\end{aligned}$$

where $\mathbf{A}$, $\bar{\mathbf{A}}$, and $\mathbf{B}$ are all constant matrices and nonsingular.

$$\lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\{\mathbf{d}\}=\begin{bmatrix}\lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right] & \lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\\ \lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right] & \lim_{m\to\infty}\frac{1}{m}\operatorname{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]\end{bmatrix},$$

where

Aceh Int. J. Sci. Technol., 12(1) 40-52
April 2023
doi: 10.13170/aijst.12.1.30403

$$\begin{aligned}\lim_{m\to\infty}\frac{1}{m}\text{asy.var}\left[\left(\mathbf{I}_m-\rho\tilde{\mathbf{W}}\right)\mathbf{s}\right]&=\lim_{m\to\infty}\frac{1}{m}\left\{\mathbf{L}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^{-1}\mathbf{R}\left[\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^{-1}\right]^t\mathbf{L}^t+\mathbf{G}\right\}\\&=\mathbf{L}\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^{-1}\left\{\lim_{m\to\infty}\frac{1}{m}\mathbf{R}\right\}\left[\left(\mathbf{I}_q-\lambda\mathbf{W}\right)^{-1}\right]^t\mathbf{L}^t+\lim_{m\to\infty}\frac{1}{m}\mathbf{G}\\&=\mathbf{0},\end{aligned}$$

as a result, $\lim_{m\to\infty}\frac{1}{m}\text{asy.var}\{\mathbf{d}\}=\begin{bmatrix}\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}\end{bmatrix}=\mathbf{0}$. Therefore,

$$\lim_{m\to\infty}\text{asy.var}\left\{\hat{\boldsymbol{\theta}}\right\}=\bar{\mathbf{A}}^{-1}\mathbf{B}\left[\lim_{m\to\infty}\frac{1}{m}\text{asy.var}\{\mathbf{d}\}\right]\mathbf{B}^t\left\{\mathbf{A}^{-1}\right\}^t=\bar{\mathbf{A}}^{-1}\mathbf{B}[\mathbf{0}]\mathbf{B}^t\left\{\mathbf{A}^{-1}\right\}^t=\mathbf{0}.$$

These results show that $\hat{\boldsymbol{\theta}}$ it is a consistent estimator. As a result, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$ are also consistent estimators.

Discussion

Illustration

If we are interested in building a new model, we first determine the dependent and independent variables according to the theory. For simplicity, this study uses arbitrary dependent and independent variables. For example, suppose there is a dependent variable or variable of interest, say , and three independent variables, say which are observed for two periods in which each consists of 10 observations or regions or locations. We assume spatial effects in a dependent variable (spatial autoregressive/spatial lag) and between regions in random effects. Now, we have the models as follows:

$$\mathbf{s}=\mathbf{K}\boldsymbol{\alpha}+\rho\tilde{\mathbf{W}}\mathbf{s}+\mathbf{L}\boldsymbol{\gamma}+\boldsymbol{\delta}, \tag{24}$$

$$\boldsymbol{\gamma}=\lambda\mathbf{W}\boldsymbol{\gamma}+\mathbf{v}, \tag{25}$$

where

$\mathbf{s}=\begin{bmatrix}\mathbf{s}_1&\mathbf{s}_2\end{bmatrix}^t$, where $\mathbf{s}_i=\begin{bmatrix}s_{i1}&s_{i2}&\cdots&s_{i10}\end{bmatrix}^t$ for $i=1,2$,

$\boldsymbol{\alpha}=\begin{bmatrix}\mu&\alpha_1&\alpha_2&\alpha_3\end{bmatrix}^t$, $\mathbf{v}=\begin{bmatrix}v_1&v_2&\cdots&v_{10}\end{bmatrix}^t$, $\boldsymbol{\gamma}=\begin{bmatrix}\gamma_1&\gamma_2&\cdots&\gamma_{10}\end{bmatrix}^t$

$\underset{20\times4}{\mathbf{K}}=\begin{bmatrix}\underset{10\times4}{\mathbf{K}_1}\\\underset{10\times4}{\mathbf{K}_2}\end{bmatrix}$, where $\underset{10\times4}{\mathbf{K}_i}=\begin{bmatrix}1&k_{i11}&k_{i21}&k_{i31}\\1&k_{i12}&k_{i22}&k_{i32}\\\vdots&\vdots&\vdots&\vdots\\1&k_{i110}&k_{i210}&k_{i310}\end{bmatrix}$, $\underset{20\times10}{\mathbf{L}}=\begin{bmatrix}\underset{10\times10}{\mathbf{L}_1}\\\underset{10\times10}{\mathbf{L}_2}\end{bmatrix}$, where $\mathbf{L}_i=\begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&\ddots&0\\0&0&0&1\end{bmatrix}$,

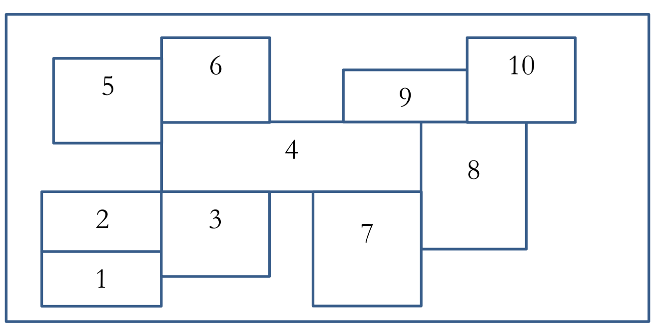

Figure 1. Illustration of the 10 neighboring regions.

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1,10} \\ w_{21} & w_{22} & \cdots & w_{2,10} \\ \vdots & \vdots & \ddots & \vdots \\ w_{10,1} & w_{10,2} & \cdots & w_{10,10} \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1^t \\ \mathbf{w}_2^t \\ \vdots \\ \mathbf{w}_{10}^t \end{bmatrix} = \left[\mathbf{w}_j^t \right],\ j = 1, 2, \cdots, 10,$$

$\tilde{\mathbf{W}} = \begin{bmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} \end{bmatrix} = \begin{bmatrix} \mathbf{w}_{1j}^t & \mathbf{0}^t \\ \mathbf{0}^t & \mathbf{w}_{2j}^t \end{bmatrix} = \left[\mathbf{w}_{ij}^t \right]$, for $i = 1, 2,\ j = 1, 2, \cdots, 10,$ and $\boldsymbol{\delta} = \left[\boldsymbol{\delta}_1 \quad \boldsymbol{\delta}_2 \right]^t$, where

$\boldsymbol{\delta}_i = \left[\delta_{i1} \quad \delta_{i2} \quad \cdots \quad \delta_{i10} \right]^t$, for $i = 1, 2.$

Table 1. Data for dependent and independent variables

Time	Location/ Region	Variables			
		Dependent	Independent		
		S	***k*₁**	***k*₂**	***k*₃**
1	1	79	46	59	52
	2	86	46	59	41
	3	69	55	62	38
	4	76	52	54	37
	5	71	49	64	47
	6	73	49	54	39
	7	80	48	63	37
	8	73	49	52	40
	9	72	55	52	51
	10	86	43	50	40
2	1	90	87	85	75
	2	85	73	79	77
	3	84	78	92	74
	4	85	81	86	60
	5	72	73	89	68
	6	79	77	84	67
	7	80	69	84	65
	8	92	77	88	72
	9	87	87	87	57
	10	84	75	84	65

Note: data illustration

Based on figure 1, the row-standardized spatial weight matrix is as follows:

doi: 10.13170/aijst.12.1.30403

$$\mathbf{W}=\begin{bmatrix}0&\frac{1}{2}&\frac{1}{2}&0&0&0&0&0&0&0\\ \frac{1}{3}&0&\frac{1}{3}&\frac{1}{3}&0&0&0&0&0&0\\ \frac{1}{3}&\frac{1}{3}&0&\frac{1}{3}&0&0&0&0&0&0\\ 0&\frac{1}{7}&\frac{1}{7}&0&\frac{1}{7}&\frac{1}{7}&\frac{1}{7}&\frac{1}{7}&\frac{1}{7}&0\\ 0&0&0&\frac{1}{2}&0&\frac{1}{2}&0&0&0&0\\ 0&0&0&\frac{1}{2}&\frac{1}{2}&0&0&0&0&0\\ 0&0&0&\frac{1}{2}&0&0&0&\frac{1}{2}&0&0\\ 0&0&0&\frac{1}{4}&0&0&\frac{1}{4}&0&\frac{1}{4}&\frac{1}{4}\\ 0&0&0&\frac{1}{3}&0&0&0&\frac{1}{3}&0&\frac{1}{3}\\ 0&0&0&0&0&0&0&\frac{1}{2}&\frac{1}{2}&0\end{bmatrix},\quad \tilde{\mathbf{W}}=\begin{bmatrix}\mathbf{W}&\mathbf{0}\\ \mathbf{0}&\mathbf{W}\end{bmatrix}.$$

We then check spatial effect using Moran's index as follows:

$$I=\frac{\sum_{j=1}^{20}\sum_{j^*=1}^{20}\tilde{w}_{jj^*}\left(y_j-\bar{y}\right)\left(y_{j^*}-\bar{y}\right)}{\sum_{j=1}^{20}\left(y_j-\bar{y}\right)^2}=\frac{\mathbf{y}^{*t}\tilde{\mathbf{W}}\mathbf{y}^*}{\mathbf{y}^{*t}\mathbf{y}^*},$$

where $\bar{y}=\frac{1}{20}\sum_{i=1}^{20}y_i$, $\mathbf{y}^*=\mathbf{y}-\bar{y}\mathbf{1}$, and $\underset{20\times1}{\mathbf{1}}=\begin{bmatrix}1&1&\cdots&1\end{bmatrix}^t$.

We decide that there is a spatial effect if Moran's index is greater than its expected value. Now, we have a Moran's index and its expected values are equal to 0.3331 and -0.0526 respectively. Therefore, we conclude that there is a spatial effect for the equation model.

We then estimate spatial autoregressive and spatial autocorrelation using an equation (20). We now have the acceptable spatial autoregressive and the spatial autocorrelation are $-1.6242<\rho<1$ and $-1.6242<\lambda<1$, respectively. In this study, we take the increased values both for spatial autoregressive and spatial autocorrelation to be 0.01. We then run it and obtain the values of spatial autoregressive and spatial autocorrelation as follows:

a. Based on equation (7), we obtain $\hat{\boldsymbol{\alpha}}=\begin{bmatrix}80.1500 & -0.1814 & 0.5588 & -0.1330\end{bmatrix}^t$ and
$\hat{\boldsymbol{\gamma}}=[-15.3790\quad -14.5706\quad -27.1705\quad -20.2562\quad -32.6891\quad -24.2337\quad -23.8305$
$-17.8936\quad -19.4292\quad -14.9080]^t$.

b. $\rho=\text{seq}(-1.6142, 0.99, 0.01)$ and $\lambda=\text{seq}(-1.6142, 0.99, 0.01)$.

c. For every value both of ρ and λ (from point b) and of vectors $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$ (from point a), we substitute all values in equation (20).

d. We obtain both the estimates of $\hat{\rho}=-0.0042$ and $\hat{\lambda}=0.8958$ that give the largest $\log L^{con}\left(\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\gamma}}\right)=-85.9451$, and therefore they become the optimal estimates respectively.

The curve for all values ρ and λ can be seen in Figure 2. Both the values of $\rho=-0.0042$ and $\lambda=0.8958$ give the largest $\log L^{con}\left(\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\gamma}}\right)=-85.9451$. Therefore, both $\rho=-0.0042$ and $\lambda=0.8958$ are to be the optimal estimates, those are $\hat{\rho}$ and $\hat{\lambda}$.

doi: 10.13170/aijst.12.1.30403

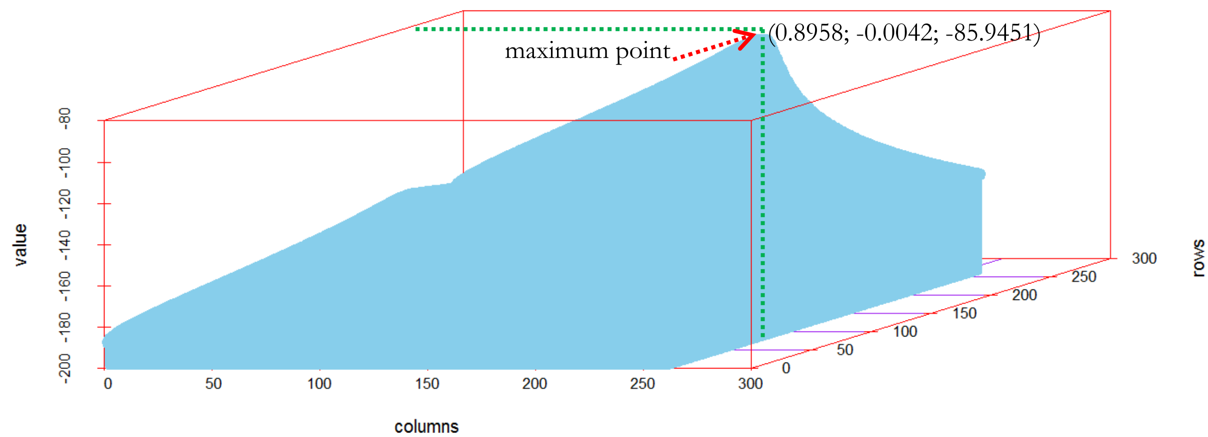

Figure 2. Graph of the concentrated log-likelihood function for both values of lambda and rho

Based on equations (21) and (22), we have the estimate $\hat{\sigma}_{\delta}^2 = 9.4442$ and $\hat{\sigma}_v^2 = 40.1673$ respectively. We then continue to equation (23) and finally obtain (after partition matrix $[\hat{\boldsymbol{\alpha}} \quad \hat{\boldsymbol{\gamma}}]^t$) as follows:

$\hat{\boldsymbol{\alpha}} = \begin{bmatrix} 58.3789 & \text{-}0.0687 & 0.4138 & \text{-}0.0981 \end{bmatrix}^t$ and

$\hat{\boldsymbol{\gamma}} = \begin{bmatrix} 6.9206 & 8.0647 & \text{-}2.0433 & 2.1317 & \text{-}7.181 & \text{-}1.5694 & 0.9114 & 5.1534 & 3.3125 & 7.9522 \end{bmatrix}^t$

Moreover, the estimators of the equation model are:

$$\hat{s}_{ij} = 58.3789 - 0.0687 k_{i1j} + 0.4138 k_{i2j} - 0.0981 k_{i3j} - 0.0042 \mathbf{w}_{ij}^t \mathbf{s} + 0.8958 \mathbf{w}_j^t \hat{\boldsymbol{\gamma}}, \quad \text{for } i = 1, 2,\ j = 1, 2, \cdots, 10.$$

For $i = 1$ and $j = 1$, we obtain the estimated value of $\hat{s}_{11} = 76.9056$. It means that the value of a dependent variable at region 1 in the first time period is 76.9056. The surrounding regions influence this value, that is, region 2 and Region 3. In the same way, we can continue to estimate the values of $\hat{s}_{ij}$ for all i and j.

Conclusion

Firstly, the concentrated log-likelihood function produces optimal spatial autoregressive and spatial autocorrelation estimates simultaneously. Secondly, all estimators can be obtained even though both sampling and component variance are unknown. Furthermore the last one, all estimators are proven to be consistent.Modeling for this study is still restricted to the region level. For further research, it can be lowered to the subregions level.

References

Amaliana, L. and Dewi, M.L.S. 2022. EBLUP-small area estimation method for per capita expenditures in Bali. E-Jurnal Matematika, 11(1), pp. 58-63.

Amaliana, L., Astuti, A.B. and Rahmi, N.S. 2021. Per capita expenditure modeling using spatial EBLUP approach – SAE. Mathematics and Statistics, 9(3), pp. 319-325.

Anselin, L. 1988. Spatial econometrics: methods and models. USA: Springer.

Anselin, L., Bera, A.K., Florax, R. and Yoon, M.J. 1996. Simple diagnostic test for spatial dependence. Regional Science and Urban Economics, 26(1), pp. 77-104.

Asfar, Kurnia, A. and Sadik, K. 2016. Optimum Spatial Weighted in Small Area Estimation. Global Journal of Pure and Applied Mathematics, 12(5), pp. 3977-3989.

doi: 10.13170/aijst.12.1.30403

McCulloch, C.E. and Searle, S.R. 2001. Generalized, linear, and mixed models. Canada: John Wiley and Sons, Inc..

Mood, A.M., Graybill, F.A. and Boes, D.C. 1974. Introduction to the Theory of Statistics, 3rd ed.. Auckland: McGraw-Hill, Inc..

Nusrang, M., Annas, S., Asfar, Hastuty and Jajang. 2017. Spatial EBLUP dalam Pendugaan Area Kecil. Jurnal Sainsmat, VI(1), pp. 59-66.

Porter, A., Holan, S., Wikle, C. and Cressie, N. 2013. Spatial Fay-Herriot models for small area estimation, Australia: National Institute for Applied Statistics Research Australia University of Wollongong.

Quintaes, Viviane, Hansen, Nícia, Silva, Pessoa, D. and Djalma. 2011. A Fay–Herriot model for estimating the proportion of households in poverty in Brazilian Municipalities. Dublin, s.n.

Salvati, N. 2004. Small area estimation by spatial models: the spatial empirical best linear unbiased prediction (spatial EBLUP), Italy: Working Paper 2004/03, "G. Parenti" Department of Statistics, University of Florence.

Siswantining, T., Saefuddin, A., Notodiputro, K.A., Nuryartono, N. and Mangku, I.W. 2012. Adaptation of hierarchical bayes SAE to spatial satscan. IOSR Journal of Mathematics (IOSRJM), 2(4), pp. 01-08.

Sirait, T., Sumertajaya, I.M., Mangku, I.W., Asra, A. and Siregar, H. 2017a. Multivariate spatial error three-stage least squares fixed effect panel: simultaneous models and estimation of their parameters. Far East Journal of Mathematical Sciences (FJMS), 102(12), pp. 2941-2970.

Sirait, T., Sumertajaya, I.M., Mangku, I.W., Asra, A. and Siregar, H., 2017b. Simultaneous equation models for spatial panel data with application to Klein's model. Bogor: Dissertation, Bogor Agricultural University.

Wang, X., Chen, M.H., Kuo, R.C. and Dey, D.K. 2015. Bayesian spatial-temporal modeling of ecological zero-inflated count data. Stat Sin., 25(1), pp. 189-204.

Watjou, K., Faes, C. and Vandendijck, Y. 2020. Spatial modelling to inform public health based on health surveys: impact of unsampled areas at lower geographical scale. Int. J. Environ. Res. Public Health, 17(3), pp. 1-19.

Watjou, K., Faes, C., Kirby, R.S., Aregay, M., Carroll, R. and Vandendij, Y. 2019. Spatial smoothing models to deal with the complex sampling design and nonresponse in the Florida BRFSS survey. Spat Spatiotemporal Epidemiol., Volume 29, p. 59–70.

Zainuddin, H.A., Notodiputro, K.A. and Sadik, K. 2015. A simulation study of logarithmic transformation model in spatial empirical best linear unbiased prediction (SEBLUP) method of small area estimation. Forum Statistika dan Komputasi: Indonesian Journal of Statistics, 20(2), pp. 78-84.