

Estimation of an Optimum Spatial Autocorrelation in Linear Mixed Models Containing Spatial Effects

TIMBANG SIRAIT

Program Studi Statistika

Politeknik Statistika STIS

Jalan Otto Iskandardinata No. 64C, Jakarta Timur, DKI Jakarta, 13330
INDONESIA

Abstract: Fay-Herriot model assumes that the random effects between regions (areas) are independent of each other. This allows the regions to be mutually independent so that the estimators obtained are unbiased estimators. In cases where the regions are not mutually independent, it can develop a model in which the assumptions are violated (not fulfilled) or allow the regions to be dependent. The development of this model is known as small area estimation (SAE) with spatial effects. In small area estimation with spatial effects, one of the important parameters is the spatial autocorrelation parameter. In various small area estimation with spatial effects, it is still very rare to generate estimators of the spatial autocorrelation parameter. Most of the parameter values used are known, that is by trying to enter several spatial autocorrelation parameter values to show that the addition of regional aspects can increase the accuracy of the small area estimation. In addition, they have also tried a restricted maximum likelihood approach in estimating the spatial autocorrelation and component variance but they still assume that the sampling variance is known (assigned). Therefore, this research proposes a concentrated log-likelihood function by means of numerical procedure to find an optimum estimate value for spatial autocorrelation coefficient where both a sampling variance and a component variance are unknown. Parameters estimators obtained in the models, fixed and random effects parameters, are proved to be consistent.

Key-Words: SAE, spatial autocorrelation, optimum, fixed and random effects, consistent

Received: May 25, 2021. Revised: October 3, 2021. Accepted: October 14, 2021. Published: October 29, 2021.

1 Introduction

As statistics progress, many policy makers use statistical modeling to make decisions easier. Statistical modeling provides an easier and more detailed description in explaining the actual conditions in the field [1].

Statistical modeling requires supporting data so that the modeling process can run as expected. Supporting data or commonly referred to as sample adequacy is data that can represent actual observations on the specific location [2]. Limited cost and available time as well as wide coverage can be obstacles in fulfilling data adequacy. This allows the data taken to be used for macro decision making [3].

The demand for large data availability is unavoidable, considering that many events or events at the smaller location level have an impact on policy makers having to take policies on a smaller level (micro scale). On the one hand, this demand is very difficult to annul because of the time and cost, which is sometimes impossible to fulfill it [3-4].

Recently, statistical modeling has developed in dealing with situations where supporting data are not available. One of the models used is small area estimation [5-9]. Small area estimation was first developed by Fay-Herriot. The Fay-Herriot model is known as the basic model in the development of small area estimation models [9-10].

In addition to the problem of sample adequacy, events or activities in one certain area can also be caused by events in other areas or surrounding areas. Events like this can add to the polemic in making decisions, so it is necessary to expand statistical modeling that can facilitate the decision-making process by policy makers. One way that is mostly conducted by statisticians is by trying to add a regional aspect in statistical modeling which is known as the spatial aspect [11-12].

In small area estimation, the Fay-Herriot model assumes that the random effects between regions are independent of each other. This allows the regions to be mutually independent so that the estimators obtained are unbiased estimators [10]. However, in most conditions, this assumptions are difficult or

impossible to be fulfilled which results in the decisions taken being not on target because the estimators obtained are biased estimators [13].

In cases where the regions are not mutually independent, it can develop a model in which the assumptions are violated (not fulfilled) or allow the regions to be dependent. The development of this model is known as small area estimation with spatial effects [13-15].

In small area estimation with spatial effects, one of the important parameters is the spatial autocorrelation parameter. In various small area estimation with spatial effects, it is still very rare to generate estimators of the spatial autocorrelation parameter. Most of the parameter values used are known, that is by trying to enter several spatial autocorrelation parameter values to show that the addition of regional aspects can increase the accuracy of the small area estimation [14-15]. In addition, [14-15] have also tried the restricted maximum likelihood (REML) approach in estimating the spatial autocorrelation and component variance but they still assume that the sampling variance is known (assigned).

Therefore, this research proposes a concentrated log-likelihood function by means of numerical procedure to find an optimum estimate value for spatial autocorrelation coefficient where both a random error variance of sampling (sampling variance) and a random error variance between regions (component variance) are unknown.

In this paper, we are motivated to develop linear mixed models especially containing spatial effects in all regions. The objective of this paper is to find the closed-form and the numerical approximation estimators of parameter models and to prove their consistency, especially for closed-form estimators.

2 Linear Mixed Models

The general form of the mixed linear model [16] is as follows

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where \mathbf{y} is a $n \times 1$ observation vector, \mathbf{X} is a $n \times (p+1)$ observation matrix (design matrix), $\boldsymbol{\alpha}$ is a $n \times (p+1)$ parameters vector of x variables (fixed effects), \mathbf{Z} is a $n \times q$ design matrix, $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$ are respectively a $q \times 1$ parameters vector of z variables (random effects) and a $n \times 1$ sampling variance vector and they are assumed to be independent. This models assume $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{G})$ in which $\boldsymbol{\Sigma} = \sigma_\varepsilon^2 \mathbf{I}_n$ and $\mathbf{G} = \sigma_\beta^2 \mathbf{I}_q$. σ_ε^2 and σ_β^2 are

respectively sampling variance and component variance. \mathbf{I}_n dan \mathbf{I}_q are identity matrices in which a $n \times 1$ and a $q \times 1$, respectively.

3 Spatial Analysis

In spatial analysis there is a correlation between spaces which is commonly referred to as spatial correlation. In this case, each observation is not stochastically independent [17].

Two-dimensional spatial information is mathematically represented by a matrix. The matrix that describes the spatial dependencies or spatial structure is called the spatial weight matrix. This spatial weight matrix plays an important role in spatial analysis [18].

We refer to [18-20] for the use first-order queen contiguity to find the row-standardized spatial weight matrix.

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

$$\boldsymbol{\beta} = \lambda \mathbf{W}\boldsymbol{\beta} + \mathbf{u}, \quad (2)$$

where λ is a spatial autocorrelation coefficient, \mathbf{W} is a $q \times q$ spatial weight matrix and \mathbf{u} is a $q \times 1$ random error vector among regions. It is assumed that $\mathbf{u} \sim N(\mathbf{0}, \mathbf{H})$ in which $\mathbf{H} = \sigma_u^2 \mathbf{I}_q$. σ_u^2 is a random error variance between regions. $\boldsymbol{\varepsilon}$ and \mathbf{u} are assumed to be independent.

From (2) we have $\boldsymbol{\beta} = (\mathbf{I}_q - \lambda \mathbf{W})^{-1} \mathbf{u}$ and obtain

$$\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{F}), \quad \mathbf{F} = (\mathbf{I}_q - \lambda \mathbf{W})^{-1} \mathbf{H} \left[(\mathbf{I}_q - \lambda \mathbf{W})^{-1} \right]^t,$$

where \mathbf{F} is a $q \times q$ symmetric matrix.

We can rewrite (1) as follows

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}(\mathbf{I}_q - \lambda \mathbf{W})^{-1} \mathbf{u} + \boldsymbol{\varepsilon}. \quad (3)$$

If we consider $\boldsymbol{\theta} = \mathbf{Z}(\mathbf{I}_q - \lambda \mathbf{W})^{-1} \mathbf{u} + \boldsymbol{\varepsilon}$, then (3) can be rewritten as follows

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\theta}. \quad (4)$$

Now, we have $\boldsymbol{\theta} \sim N(\mathbf{0}, \mathbf{L})$, $\mathbf{L} = \mathbf{Z}\mathbf{F}\mathbf{Z}' + \boldsymbol{\Sigma}$, where \mathbf{L} is a $n \times n$ symmetric matrix. Therefore, $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\alpha}, \mathbf{L})$.

If we assume $\boldsymbol{\beta}$ fixed, then $\boldsymbol{\theta} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ and the conditional distribution of \mathbf{y} given $\boldsymbol{\beta}$ is $\mathbf{y} | \boldsymbol{\beta} \sim N(\mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta}, \boldsymbol{\Sigma})$. We can then rewrite (1) as follows $\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{T}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$,

where $\mathbf{T} = [\mathbf{X} \mid \mathbf{Z}]$ and $\boldsymbol{\gamma} = [\boldsymbol{\alpha} \mid \boldsymbol{\beta}]'$.

\mathbf{T} is a less than full rank matrix and therefore, we need to reparameterization [21]. We reparameterize

γ to be γ_* and change \mathbf{T} to be \mathbf{T}_* . Now, \mathbf{T}_* is a full rank matrix, we obtain the least-square estimator of γ_* is

$$\hat{\gamma}_* = (\mathbf{T}'_*\mathbf{T}_*)^{-1}(\mathbf{T}'_*\mathbf{y}). \quad (5)$$

4 Estimating the Parameters

Maximum likelihood function of (2) is as follows

$$L(\mathbf{u}) = \frac{1}{(2\pi)^{\frac{q}{2}}|\mathbf{H}|^{\frac{1}{2}}} e^{-\frac{1}{2}\mathbf{u}'\mathbf{H}^{-1}\mathbf{u}}. \quad (6)$$

Based on (6), the maximum likelihood function of β is as follows

$$L(\beta) = \frac{1}{(2\pi)^{\frac{q}{2}}|\mathbf{H}|^{\frac{1}{2}}} e^{-\frac{1}{2}\beta'(\mathbf{I}_q - \lambda\mathbf{W})' \mathbf{H}^{-1}(\mathbf{I}_q - \lambda\mathbf{W})\beta} |J|, \quad (7)$$

where J is Jacobian of transformation and $|J| = \|\mathbf{I}_q - \lambda\mathbf{W}\|$ is a absolute value of the determinant of $|\mathbf{I}_q - \lambda\mathbf{W}|$.

We then can find maximum likelihood function of $\mathbf{y} | \beta$ as follows

$$L(\mathbf{y} | \beta) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\alpha - \mathbf{Z}\beta)' \Sigma^{-1}(\mathbf{y} - \mathbf{X}\alpha - \mathbf{Z}\beta)}. \quad (8)$$

By means of (7) and (8), we obtain

$$L(\alpha, \beta) = L(\mathbf{y} | \beta)L(\beta),$$

and take natural logarithm, obtained

$$\begin{aligned} \log L(\alpha, \beta) &= -\left(\frac{n+q}{2}\right)\log(2\pi) - \frac{1}{2}\log(|\Sigma||\mathbf{H}|) \\ &\quad - \frac{1}{2}(\mathbf{y} - \mathbf{X}\alpha - \mathbf{Z}\beta)' \Sigma^{-1}(\mathbf{y} - \mathbf{X}\alpha - \mathbf{Z}\beta) \\ &\quad - \frac{1}{2}\beta'(\mathbf{I}_q - \lambda\mathbf{W})' \mathbf{H}^{-1}(\mathbf{I}_q - \lambda\mathbf{W})\beta \\ &\quad + \log\|\mathbf{I}_q - \lambda\mathbf{W}\|. \end{aligned} \quad (9)$$

If we take derivative respective for α and β , then set this derivative equal to zero so we obtain the estimator of α and β , respectively, as follows

$$\hat{\alpha} = [\mathbf{X}'\Sigma^{-1}\mathbf{X}]^{-1} \mathbf{X}'\Sigma^{-1}(\mathbf{y} - \mathbf{Z}\hat{\beta}) \quad (10)$$

$$\begin{aligned} \hat{\beta} &= \left[\mathbf{Z}'\Sigma^{-1}\mathbf{Z} + (\mathbf{I}_q - \lambda\mathbf{W})' \mathbf{H}^{-1}(\mathbf{I}_q - \lambda\mathbf{W}) \right]^{-1} \\ &\quad \times \mathbf{Z}'\Sigma^{-1}(\mathbf{y} - \mathbf{X}\hat{\alpha}). \end{aligned} \quad (11)$$

Equations (10) and (11) show that there is a dependency or relationship between $\hat{\alpha}$ and $\hat{\beta}$ so that $\hat{\alpha}$ and $\hat{\beta}$ can be combined by modifying the matrix. This technique aims to make it easier to obtain $\hat{\alpha}$ and $\hat{\beta}$ simultaneously. Now, we have

$$\begin{aligned} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} &= \begin{bmatrix} \mathbf{X}'\Sigma^{-1}\mathbf{X} & \mathbf{X}'\Sigma^{-1}\mathbf{Z} \\ \mathbf{Z}'\Sigma^{-1}\mathbf{X} & \mathbf{Z}'\Sigma^{-1}\mathbf{Z} + (\mathbf{I}_q - \lambda\mathbf{W})' \mathbf{H}^{-1}(\mathbf{I}_q - \lambda\mathbf{W}) \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} \mathbf{X}'\Sigma^{-1}\mathbf{y} \\ \mathbf{Z}'\Sigma^{-1}\mathbf{y} \end{bmatrix}. \end{aligned} \quad (12)$$

$$\begin{aligned} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} &= \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \frac{\sigma_\varepsilon^2}{\sigma_u^2}(\mathbf{I}_q - \lambda\mathbf{W})'(\mathbf{I}_q - \lambda\mathbf{W}) \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}. \end{aligned} \quad (13)$$

If we recall $\Sigma = \sigma_\varepsilon^2\mathbf{I}_n$ and $\mathbf{H} = \sigma_u^2\mathbf{I}_q$ then $\Sigma^{-1} = (\sigma_\varepsilon^2)^{-1}\mathbf{I}_n$ and $\mathbf{H}^{-1} = (\sigma_u^2)^{-1}\mathbf{I}_q$. Therefore, we can rewrite (12) and obtain the result in (13).

We recall again $\Sigma = \sigma_\varepsilon^2\mathbf{I}_n$ and $\mathbf{H} = \sigma_u^2\mathbf{I}_q$ and obtain the determinant of Σ and \mathbf{H} ($|\Sigma|$ and $|\mathbf{H}|$) are $(\sigma_\varepsilon^2)^n$ and $(\sigma_u^2)^q$, respectively, and therefore (9) can be rewritten as follows

$$\begin{aligned} \log L(\alpha, \beta) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma_\varepsilon^2) \\ &\quad - \frac{1}{2\sigma_\varepsilon^2}(\mathbf{y} - \mathbf{X}\alpha - \mathbf{Z}\beta)'(\mathbf{y} - \mathbf{X}\alpha - \mathbf{Z}\beta) \\ &\quad - \frac{q}{2}\log(2\pi) - \frac{q}{2}\log(\sigma_u^2) \\ &\quad - \frac{1}{2\sigma_u^2}\beta'(\mathbf{I}_q - \lambda\mathbf{W})'(\mathbf{I}_q - \lambda\mathbf{W})\beta \\ &\quad + \log\|\mathbf{I}_q - \lambda\mathbf{W}\|. \end{aligned} \quad (14)$$

The values of σ_ε^2 and σ_u^2 are unknown and its values can be estimated from sample. If we take derivative respective for σ_ε^2 and σ_u^2 , then set this derivative equal to zero so we obtain the estimator of σ_ε^2 and σ_u^2 , respectively, as follows

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n}(\mathbf{y} - \mathbf{X}\hat{\alpha} - \mathbf{Z}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\alpha} - \mathbf{Z}\hat{\beta}). \quad (15)$$

and

$$\hat{\sigma}_u^2 = \frac{1}{q}\hat{\beta}'(\mathbf{I}_q - \lambda\mathbf{W})'(\mathbf{I}_q - \lambda\mathbf{W})\hat{\beta}. \quad (16)$$

$$\begin{aligned} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} &= \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \frac{\hat{\sigma}_\varepsilon^2}{\hat{\sigma}_u^2}(\mathbf{I}_q - \lambda\mathbf{W})'(\mathbf{I}_q - \lambda\mathbf{W}) \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}. \end{aligned} \quad (17)$$

By means of (15) and (16), so (13) can be rewritten as obtained in (17).

5 Concentrated Log-likelihood

If we change respective σ_ε^2 and σ_u^2 with their estimators in (14) then we obtain the result as follows

$$\log L^{con}(\hat{\alpha}, \hat{\beta}) = C - \frac{n}{2} \log \left(\frac{1}{n} (\mathbf{y} - \mathbf{X}\hat{\alpha} - \mathbf{Z}\hat{\beta})' (\mathbf{y} - \mathbf{X}\hat{\alpha} - \mathbf{Z}\hat{\beta}) \right) - \frac{q}{2} \log \left(\frac{1}{q} \hat{\beta}' (\mathbf{I}_q - \lambda \mathbf{W})' (\mathbf{I}_q - \lambda \mathbf{W}) \hat{\beta} \right) + \log \|\mathbf{I}_q - \lambda \mathbf{W}\|. \tag{18}$$

where $C = -\left(\frac{n+q}{2}\right)(\log(2\pi) + 1)$. We refer (18) to as a concentrated log-likelihood function.

Suppose that $\tau_1, \tau_2, \dots, \tau_n$ are the eigenvalues of \mathbf{W} , [23] shows that the acceptable spatial autocorrelation coefficients is $\frac{1}{\tau_{\text{minimum}}} < \lambda < 1$. In this research, we use numerical method in (18) to find an optimum estimate of $\hat{\lambda}$ by means of R program version 3.0.3. The numerical method is a forming of sequence of λ [22, 24-27]. The procedures are as follows:

1. Recall (5) to find $\hat{\gamma}_*$ and use it to find $\hat{\gamma}$. Now, we have $\hat{\alpha}$ and $\hat{\beta}$ because $\hat{\gamma}$ consist of them.
2. Make a sequences of λ and the sequences are $\lambda = \text{seq}(\text{start value, end value, increasing})$. In this research, the increasing is 0.01.
3. Based on the results obtained in steps 1 and 2, use the values in (18).
4. Finding a value of λ that gives a largest value of $\log L_n^{con}$ and it will be an optimum estimate of $\hat{\lambda}$.

We can also find the estimates of $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_u^2$ based on the results obtained in the first and fourth steps of the procedures. Now, we have all the estimate values, that is $\hat{\lambda}$, $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_u^2$.

Finally, we can rewrite (17) as follows

$$\begin{bmatrix} \hat{\alpha}^* \\ \hat{\beta}^* \end{bmatrix} = \left[\begin{array}{cc} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \frac{\hat{\sigma}_\varepsilon^2}{\hat{\sigma}_u^2} (\mathbf{I}_q - \hat{\lambda}\mathbf{W})' (\mathbf{I}_q - \hat{\lambda}\mathbf{W}) \end{array} \right]^{-1} \times \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}. \tag{19}$$

The parameters estimates obtained in equation (19) are called linear mixed model parameters estimators with spatial influence among regions based on an optimum estimate of autocorrelation coefficient.

6 Properties of Estimators

A good estimator has a minimum requirement for consistency. An estimator is said to be consistent if a sample size taken is very large or a result is always correct or at least very close to the true value [28].

Theorem (Consistency). *Estimators of α and β as obtained in (12) are consistent estimators.*

Proof. Recall (12), and it can be rewritten as follows

$$\mathbf{V}\hat{\delta} = \mathbf{S}\mathbf{g}$$

$$\hat{\delta} = \mathbf{V}^{-1}\mathbf{S}\mathbf{g},$$

where

$$\mathbf{V} = \begin{bmatrix} \mathbf{X}'\Sigma^{-1}\mathbf{X} & \mathbf{X}'\Sigma^{-1}\mathbf{Z} \\ \mathbf{Z}'\Sigma^{-1}\mathbf{X} & \mathbf{Z}'\Sigma^{-1}\mathbf{Z} + (\mathbf{I}_q - \lambda\mathbf{W})' \mathbf{H}^{-1} (\mathbf{I}_q - \lambda\mathbf{W}) \end{bmatrix},$$

$$\hat{\delta} = \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix}, \mathbf{S} = \begin{bmatrix} \mathbf{X}'\Sigma^{-1} & \mathbf{0}_* \\ \mathbf{0} & \mathbf{Z}'\Sigma^{-1} \end{bmatrix}, \text{ and } \mathbf{g} = \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}.$$

$\mathbf{0}_*$ is a $p \times n$ vector and $\mathbf{0}$ is a $q \times n$ vector.

Now, we have an $\text{asy. var}\{\mathbf{g}\}$ and $\text{asy. var}(\mathbf{y})$ as follows

$$\begin{aligned} \text{asy. var}\{\mathbf{g}\} &= \text{asy. var} \left\{ \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \right\} \\ &= E \left\{ \left(\begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} - E \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} - E \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \right)' \right\} \\ &= E \left\{ \left(\begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} E(\mathbf{y}) \\ E(\mathbf{y}) \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} E(\mathbf{y}) \\ E(\mathbf{y}) \end{bmatrix} \right)' \right\} \\ &= E \begin{bmatrix} \mathbf{y} & \mathbf{y}' \\ \mathbf{y} & \mathbf{y} \end{bmatrix} - \begin{bmatrix} E(\mathbf{y}) & E(\mathbf{y})' \\ E(\mathbf{y}) & E(\mathbf{y}) \end{bmatrix} \\ &= E \begin{bmatrix} \mathbf{y} & \mathbf{y}' \\ \mathbf{y} & \mathbf{y} \end{bmatrix} - \begin{bmatrix} E(\mathbf{y}) & \\ & E(\mathbf{y}) \end{bmatrix} \\ &\quad \times \begin{bmatrix} E(\mathbf{y}') & E(\mathbf{y}') \end{bmatrix} \\ &= E \begin{bmatrix} \mathbf{y}\mathbf{y}' & \mathbf{y}\mathbf{y}' \\ \mathbf{y}\mathbf{y}' & \mathbf{y}\mathbf{y}' \end{bmatrix} \\ &\quad - \begin{bmatrix} E(\mathbf{y})E(\mathbf{y}') & E(\mathbf{y})E(\mathbf{y}') \\ E(\mathbf{y})E(\mathbf{y}') & E(\mathbf{y})E(\mathbf{y}') \end{bmatrix} \end{aligned}$$

$$\text{asy. var } \{\mathbf{g}\} = \begin{bmatrix} E(\mathbf{y}\mathbf{y}') - E(\mathbf{y})E(\mathbf{y}') \\ E(\mathbf{y}\mathbf{y}') - E(\mathbf{y})E(\mathbf{y}') \\ E(\mathbf{y}\mathbf{y}') - E(\mathbf{y})E(\mathbf{y}') \\ E(\mathbf{y}\mathbf{y}') - E(\mathbf{y})E(\mathbf{y}') \end{bmatrix} \\ = \begin{bmatrix} \text{asy. var}(\mathbf{y}) & \text{asy. var}(\mathbf{y}) \\ \text{asy. var}(\mathbf{y}) & \text{asy. var}(\mathbf{y}) \end{bmatrix},$$

and

$$\text{asy. var}(\mathbf{y}) = \mathbf{L} = \mathbf{Z}\mathbf{F}\mathbf{Z}' + \mathbf{\Sigma} \\ = \mathbf{Z}(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \mathbf{H} \left[(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \right]^t \mathbf{Z}' + \mathbf{\Sigma}.$$

Therefore,

$$\text{asy. var } \{\hat{\boldsymbol{\delta}}\} = \text{asy. var } \{\mathbf{V}^{-1}\mathbf{S}\mathbf{g}\} \\ = \mathbf{V}^{-1}\mathbf{S} \text{asy. var } \{\mathbf{g}\} \{\mathbf{V}^{-1}\mathbf{S}\}' \\ = \mathbf{V}^{-1}\mathbf{S} \text{asy. var } \{\mathbf{g}\} \mathbf{S}' \{\mathbf{V}^{-1}\}'$$

and

$$\lim_{n \rightarrow \infty} \text{asy. var } \{\hat{\boldsymbol{\delta}}\} = \lim_{n \rightarrow \infty} \left[\left(\frac{1}{n} \mathbf{V} \right)^{-1} \mathbf{S} \frac{1}{n} \text{asy. var } \{\mathbf{g}\} \right] \\ \times \mathbf{S}' \{\mathbf{V}^{-1}\}' \\ = \left[\lim_{n \rightarrow \infty} \left(\frac{1}{n} \mathbf{V} \right)^{-1} \right] \mathbf{S} \left[\lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var } \{\mathbf{g}\} \right] \\ \times \mathbf{S}' \{\mathbf{V}^{-1}\}' \\ = \left[\lim_{n \rightarrow \infty} (\bar{\mathbf{V}})^{-1} \right] \mathbf{S} \left[\lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var } \{\mathbf{g}\} \right] \\ \times \mathbf{S}' \{\mathbf{V}^{-1}\}' \\ = \bar{\mathbf{V}}^{-1} \mathbf{S} \left[\lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var } \{\mathbf{g}\} \right] \mathbf{S}' \{\mathbf{V}^{-1}\}' ,$$

where \mathbf{V} , $\bar{\mathbf{V}}$ and \mathbf{S} are nonsingular constant matrices.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var } \{\mathbf{g}\} = \lim_{n \rightarrow \infty} \frac{1}{n} \begin{bmatrix} \text{asy. var}(\mathbf{y}) & \text{asy. var}(\mathbf{y}) \\ \text{asy. var}(\mathbf{y}) & \text{asy. var}(\mathbf{y}) \end{bmatrix} \\ = \begin{bmatrix} \lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var}(\mathbf{y}) \\ \lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var}(\mathbf{y}) \\ \lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var}(\mathbf{y}) \\ \lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var}(\mathbf{y}) \end{bmatrix},$$

where

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var}(\mathbf{y}) = \lim_{n \rightarrow \infty} \frac{1}{n} \left\{ \mathbf{Z}(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \mathbf{H} \right. \\ \left. \times \left[(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \right]^t \mathbf{Z}' + \mathbf{\Sigma} \right\} \\ = \mathbf{Z}(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \mathbf{H} \right\} \\ \times \left[(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \right]^t \mathbf{Z}' + \lim_{n \rightarrow \infty} \frac{1}{n} \mathbf{\Sigma} \\ = \mathbf{Z}(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \times \mathbf{0} \\ \times \left[(\mathbf{I}_q - \lambda\mathbf{W})^{-1} \right]^t \mathbf{Z}' + \mathbf{0} \\ = \mathbf{0},$$

consequently,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{asy. var } \{\mathbf{g}\} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \\ = \mathbf{0},$$

and thus, $\lim_{n \rightarrow \infty} \text{asy. var } \{\hat{\boldsymbol{\delta}}\} = \mathbf{0}$.

This shows that $\hat{\boldsymbol{\delta}}$ is consistent estimator, and thus $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are also consistent estimators.

7 Illustration

Suppose that there are one dependent variable, say y , and three independent variables, say x_1, x_2, x_3 . It is observed in two time periods and in each time period consists of 10 locations. We can write an equation models as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (20)$$

$$\boldsymbol{\beta} = \lambda\mathbf{W}\boldsymbol{\beta} + \mathbf{u}, \quad (21)$$

where

$$\mathbf{y} = [\mathbf{y}_1 \ \mathbf{y}_2]', \text{ and } \mathbf{y}_i = [y_{i1} \ y_{i2} \ \dots \ y_{i10}]'$$

$$\text{for } i=1,2, \ \boldsymbol{\alpha} = [\mu \ \alpha_1 \ \alpha_2 \ \alpha_3]'$$

$$\mathbf{u} = [u_1 \ u_2 \ \dots \ u_{10}]', \ \boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \dots \ \beta_{10}]'$$

$$\mathbf{X}_{20 \times 10} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}, \ \mathbf{X}_i = \begin{bmatrix} 1 & x_{i11} & x_{i21} & x_{i31} \\ 1 & x_{i12} & x_{i22} & x_{i32} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{i10} & x_{i210} & x_{i310} \end{bmatrix},$$

$$\mathbf{Z}_{20 \times 10} = \begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{bmatrix}, \ \mathbf{Z}_i = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

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

for $j = 1, 2, \dots, 10$, $\boldsymbol{\varepsilon} = [\boldsymbol{\varepsilon}_1 \ \boldsymbol{\varepsilon}_2]^t$, and

$\boldsymbol{\varepsilon}_i = [\varepsilon_{i1} \ \varepsilon_{i2} \ \cdots \ \varepsilon_{i10}]^t$ for $i = 1, 2$, and

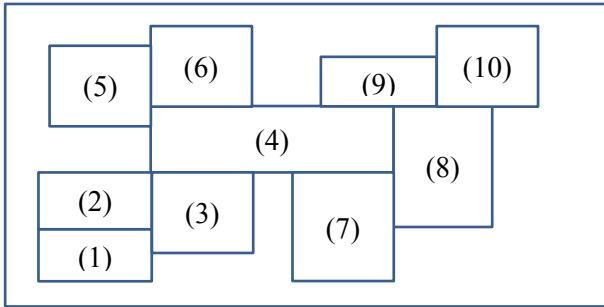


Fig. 1: Illustration of the 10 neighboring locations

Table 1 Data for dependent and independent variables

Time	Location	Variables			
		Dependent	Independent		
		Y	x_1	x_2	x_3
1	1	25	46	50	45
	2	28	41	54	43
	3	27	40	55	41
	4	26	42	57	42
	5	29	48	56	44
	6	28	47	55	45
	7	31	44	57	46
	8	33	45	54	47
	9	29	48	58	43
	10	27	45	51	44
2	1	20	51	64	50
	2	23	50	65	51
	3	21	58	67	57
	4	24	57	63	58
	5	22	56	64	61
	6	26	60	66	60
	7	29	61	67	59
	8	31	59	69	60
	9	28	60	70	62
	10	25	59	68	58

Note: data illustration

We first check spatial influence in the equation models. Based on Fig. 1, we have row-standardized spatial weight matrix as follows

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

The formulation of Moran Index is as follows

$$I = \frac{\sum_{j=1}^{10} \sum_{j^*=1}^{10} w_{jj^*} (y_j - \bar{y})(y_{j^*} - \bar{y})}{\sum_{j=1}^{10} (y_j - \bar{y})^2} = \frac{\mathbf{y}^* \mathbf{W} \mathbf{y}^*}{\mathbf{y}^* \mathbf{y}^*}$$

where $\bar{y} = \frac{1}{10} \sum_{i=1}^{10} y_i$, $\mathbf{y}^* = \mathbf{y} - \bar{y} \mathbf{1}$, and

$\mathbf{1}_{10 \times 1} = [1 \ 1 \ \cdots \ 1]^t$. As we know that the expected

value of I , $E(I)$, is $\frac{-1}{n-1}$.

If $I > E(I)$ then there is a spatial influence in the equation models. We calculate from data in Table 1 and obtain the result, that is $I = 0.4861$ and $E(I) = -0.0526$. Therefore, there is spatial influence in the equation models.

We then estimate spatial autocorrelation by means of (18). By using \mathbf{W} matrix, we have the acceptable spatial autocorrelation parameter, that is $-1.6242 < \lambda < 1$. By means of the procedures, we have:

- $\hat{\boldsymbol{\alpha}} = [26.6000 \ 0.1212 \ -0.2098 \ -0.1663]^t$ and $\hat{\boldsymbol{\beta}} = [9.8840 \ 13.6888 \ \cdots \ 14.0666]^t$
- $\lambda = \text{seq}(-1.6142, 0.99, 0.01)$.
- Based on $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$ and λ obtained in steps 1 and 2, we set all the values in (18).
- We find $\hat{\lambda} = 0.9258$ that gives a largest value $\log L^{con}$ and it is to be an optimum $\hat{\lambda}$. The graph of function of lambda is given in Fig. 2.

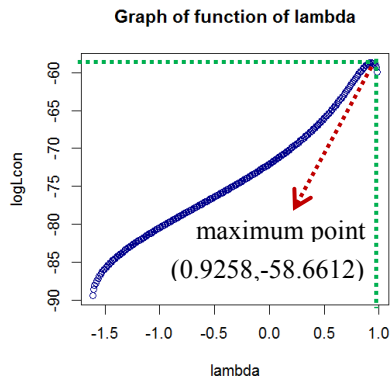


Fig. 2: Graph of function of lambda

Based on (15) and (16), we obtain $\hat{\sigma}_\varepsilon^2 = 1.2446$ and $\hat{\sigma}_u^2 = 9.0974$, respectively. We then continue to (19), and obtain $[\hat{\alpha}^* \ \hat{\beta}^*]^t$ as follows

$$\hat{\alpha}^* = \begin{bmatrix} 41.1846 \\ 0.0840 \\ -0.1833 \\ -0.1505 \end{bmatrix}; \hat{\beta}^* = \begin{bmatrix} -4.9496 \\ -1.7210 \\ -2.8114 \\ -1.4684 \\ -1.0963 \\ 0.1822 \\ 3.6002 \\ 5.3940 \\ 2.2486 \\ -0.5914 \end{bmatrix}$$

Estimation of the model is as follows

$$\hat{y}_j = 41.1846 + 0.0840x_1 - 0.1833x_2 - 0.1505x_3 + 0.9258\mathbf{w}_j^t \hat{\beta}^*$$

for $j = 1, 2, \dots, 10$.

8 Conclusion

This research emphasizes an assumption that a random error variance of sampling (sampling variance) and a random error variance between regions (component variance) are unknown.

In this research, we restrict to find an optimum spatial autocorrelation coefficient estimate by means of a sequence of the acceptable spatial autocorrelation coefficients that gives a largest value of concentrated log-likelihood function.

The parameters estimates, $\hat{\alpha}^*$ and $\hat{\beta}^*$, obtained in this research are consistent estimators and called

linear mixed model parameters estimators with spatial influence among regions based on an optimum estimate of autocorrelation coefficient.

In this paper, we use the numerical procedures to find the spatial autocorrelation estimator. In the next research, we motivate you to find the closed-form of its estimator.

References:

- [1] Fischer, F., Miller, G. J. and Sidney, M. S., *Handbook of Public Policy Analysis: Theory, Politics and Methods*, CRC Press, London, 2007.
- [2] Memon, M. A., Ting, H., Cheah, J. H., Ramayah, T., Chuah, F. and Cham, T. H., Sample size for survey research: review and recommendations, *Journal of Applied Structural Equation Modeling* 4(2), 2020, pp. i-xx.
- [3] Kothari, C. R., *Research Methodology: Methods and Techniques, 2nd revised ed.*, New Age International Publishers, New Delhi, 2004.
- [4] Ali, M., Sholihah, P. I., Ahmed, K. and Prabandari, S. P., Small data and big data: combination make better decision, *International Journal of Research in Management, Economics and Commerce* 06(10), 2016, pp. 1-6.
- [5] Fabrizi, E., Ferrante, M. R. and Pacei, S., Small area estimation of average household income based on unit level models for panel data, *Survey Methodology* 33(2), 2007, pp. 187-198.
- [6] Talakua, M. W. and Patty, G. G., Small area estimation untuk pendugaan tingkat partisipasi angkatan kerja di Provinsi Maluku dengan pendekatan kernel-bootstrap, *Jurnal Ilmu Matematika dan Terapan* 11(1), 2016, pp. 17-23.
- [7] Kurniawan, A., Elmira, E., Arfyanto, H., Anbarani, M. D., Rizky, M., Saputri, N. S., Izzati, R. A. and Ruhmaniyati, *Testing Small Area Estimation (SAE) Method for Generating Nutrition Maps in Indonesia: Rokan Hulu District*, Edited by Dhanita Putri Sarahtika, The SMERU Research Institute, Jakarta, 2019.
- [8] Elbers, C., Lanjouw, J. O. and Lanjouw, P., Micro-level estimation of poverty and inequality, *Econometrica* 71(1), 2003, pp. 355-364.
- [9] Rao, J. N. K., *Small Area Estimation*, John Wiley & Sons, Inc., New Jersey, 2003.
- [10] Fay, R. E. and Herriot, R. A., Estimates of income for small places to census data: an application of James-Stein procedures, *Journal*

of the American Statistical Association 74(366), 1979, pp. 269-277.

- [11] Anselin, L., *Spatial Econometrics: Methods and Models*, Springer, USA, 1988.
- [12] Fischer, M. M. and Wang, J., *Spatial Data Analysis: Models, Methods and Techniques*, Springer, New York, 2011.
- [13] Salvati, N., *Small area estimation by spatial models: the spatial empirical best linear unbiased prediction (spatial EBLUP)*, Working Paper 2004/03, "G. Parenti" Department of Statistics, University of Florence, Italy, 2004.
- [14] Pratesi, M. and Salvati, N., Small area estimation in the presence of correlated random area effects, *Journal of Official Statistics* 25(1), 2009, pp. 37-53.
- [15] Singh, B. B., Shukla, G. K. and Kundu, D., Spatio-temporal models in small area estimation, *Survey Methodology* 31(2), 2005, pp. 183-195.
- [16] McCulloch, C.E. and Searle, S.R., *Generalized, Linear, and Mixed Models*, John Wiley and Sons, Inc., Canada, 2001.
- [17] Ward, M.D. and Gleditsch K.S., *Spatial Regression Models*. Sage Publications, Inc., Los Angeles, 2008.
- [18] Keisuke, K., *Introduction to Spatial Econometric Analysis: Creating spatially lagged variables in Stata*, RIETI Technical Paper Series 16-T-001, Japan, 2016.
- [19] Anselin, L. and Hudak, A. K., Spatial econometrics in practice: a review of software options, *Regional Science and Urban Economics* 22, 1992, pp. 509-536.
- [20] LeSage, J. and Pace, R. K., *Introduction to Spatial Econometrics*, Taylor and Francis Group, LLC., New York, 2009.
- [21] Myers, R. H. and Milton, J. S., *A First Course in the Theory of Linear Statistical Models*, PWS-KENT Publishing Company, Boston, 1991.
- [22] Sirait, T., Sumertajaya, I. M., Mangku, I. W., Asra, A. and Siregar, H., Multivariate three-stage least squares fixed effect panel simultaneous models and estimation of their parameters, *Far East Journal of Mathematical Sciences (FJMS)* 102(7), 2017a, pp. 1503-1521.
- [23] Anselin, L., Bera, A. K., Florax, R. and Yoon, M. J., Simple diagnostic test for spatial dependence, *Regional Science and Urban Economics* 26, 1996, pp. 77-104.
- [24] Sirait, T., Sumertajaya, I. M., Mangku, I. W., Asra, A. and Siregar, H., 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), 2017b, pp. 2941-2970.
- [25] Sirait, T., Sumertajaya, I. M., Mangku, I. W., Asra, A. and Siregar, H., *Simultaneous Equation Models for Spatial Panel Data with Application to Klein's Model*, Dissertation, Bogor Agricultural University, 2018.
- [26] Sirait, T., Multivariate spatial autoregressive three-stage least squares fixed effect panel simultaneous models and estimation of their parameters, *WSEAS Transactions on Mathematics* 18, 2019, pp. 307-318.
- [27] Sirait, T., Multivariate general spatial three-stage least squares fixed effect panel simultaneous models and estimation of their parameters, *WSEAS Transactions on Mathematics* 19, 2020, pp. 373-383.
- [28] Mood, A. M., Graybill, F. A. and Boes, D.C., *Introduction to the Theory of Statistics, 3rd ed.*, McGraw-Hill, Inc., Auckland, 1974.