Research Article

Estimation of Error Variance-Covariance Parameters Using Multivariate Geographically Weighted Regression Model

Sri Harini

Mathematics Department, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang, East Java, Indonesia
Correspondence should be addressed to Sri Harini; sriharini@mat.uin-malang.ac.id
Received 18 October 2019; Accepted 29 November 2019; Published 1 February 2020
Academic Editor: Abdel-Maksoud A. Soliman

Copyright © 2020 Sri Harini. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Multivariate Geographically Weighted Regression (MGWR) model is a development of the Geographically Weighted Regression (GWR) model that takes into account spatial heterogeneity and autocorrelation error factors that are localized at each observation location. The MGWR model is assumed to be an error vector \( \mathbf{\varepsilon} \) that distributed as a multivariate normally with zero vector mean and variance-covariance matrix \( \mathbf{\Sigma} \) at each location \( (u_i, v_i) \), which \( \mathbf{\Sigma} \) is sized \( q \times q \) for samples at the \( i \)-location. In this study, the estimated error variance-covariance parameters is obtained from the MGWR model using Maximum Likelihood Estimation (MLE) and Weighted Least Square (WLS) methods. The selection of the WLS method is based on the weighting function measured from the standard deviation of the distance vector between one observation location and another observation location. This test uses a statistical inference procedure by reducing the MGWR model equation so that the estimated error variance-covariance parameters meet the characteristics of unbiased. This study also provides researchers with an understanding of statistical inference procedures.

1. Introduction

In statistical inference, estimation of spatial data parameters using the GWR approach has been carried out by many researchers. According to [1], the GWR method is selected due to the weaknesses of the ordinary least square (OLS) parameter estimation results, where the variance error in the OLS model is still assumed to be fixed (homoscedasticity) and there is no dependency between errors (spatial effects) at each observation location. Spatial problems, specifically in parameter estimation has been studied by Cressie [2]. The author discussed spatial analysis in detail by using OLS and estimator of spatial regression models with the maximum likelihood estimation (MLE) methods. Yasin [3] proposed the GWR stepwise method in order to choose a significant variable. The selection of the stepwise GWR method reduces several predictor variables that are not significant to the response variable. A Mixed Geographically Weighted Regression model (MGWR) is a combination of linear regression and the GWR. A statistical test of MGWR models with the maximum likelihood ratio test (MLRT) method have been carried out by [1], Cressie [2] and Harini et al. [4]. By inference, the MLRT method can maximize the probability value of the resulting parameters. Furthermore, to complete the MGWR model, which in inference analysis, the first derivative analytical solution of the log-likelihood function is unavailable in closed form.

Harini and Purhadi [5] used the Matrix Laboratory algorithm approach, a high-level programming language based on numerical computational techniques to solve problems involving mathematical operations with database arrays and vector formulations. The advantage of this approach is the absence of variable dimension constraints. Referring to [4], Triyanto et al. [6] discussed the parameter estimation of the Geographically Weighted Multivariate Poisson Regression (GWMPR) model using the Maximum Likelihood Estimation (MLE) methods. The GWMPR is used to model the spatial data with response variables that are distributed Poisson.

Another problem that often arises in the GWR model is to validate hypothesis testing using statistical inference analysis because invalidating hypothesis test requires several stages of parameter estimation that cannot be done globally [7]. Therefore, the R and GWR4 programs can be used to check the validity level of hypothesis testing. The advantages of the
program parameter estimation results are both global and local and can be done together. Soemartojo et al. [8] analyzed the spatial heterogeneity problem of the GWR model using Weighted Least Squares (WLS) method with Gaussian kernel weight function. Spatial heterogeneity occurs because there is a strong dependence between one observation with other observations that are nearby (nearest neighboring) to cause spatial effects. The process of non-stationarity by applying an extended hyper-local GWR is examined by Comber et al. [9]. This model optimizes the covariates of each local regression simultaneously, to determine the local bandwidth specifications based on lots of data at each location and evaluates different bandwidths in each location to choose the right local regression model.

In this research, we focus on the form and properties of the estimated error variance-covariance parameters of the MGWR model using the MLE and WLS methods. This test uses statistical inference procedures to obtain the estimated error variance-covariance parameters that meet the unbiased nature.

2. Theoretical GWR

Supporting theories for completing this research refer to the Geographically Weighted Regression (GWR) [1] and Statistics for Spatial Data [2].

3. Methods of MGWR

The MGWR method refers to [4] and [10].

4. Results

The MGWR is the development of a multivariate linear model with known location information. In the multivariate spatial linear model, the relationship between the response variable \( Y_1, Y_2, ..., Y_q \) and the predictor variable \( X_1, X_2, ..., X_p \) at the \( i \)th location is given by

\[
Y_i = \beta_{0i}(u_i, v_i) + \beta_{1i}(u_i, v_i)X_{i1} + \beta_{2i}(u_i, v_i)X_{i2} + ... + \beta_{pi}(u_i, v_i)X_{pi} + \epsilon_i, \quad h = 1, 2, ..., q \text{ and } i = 1, 2, ..., n. \tag{1}
\]

The assumptions used in the MGWR model are error vector (\( \epsilon \)) with multivariate normal distributions with zero vectors \( \mu \) and variance-covariance matrix (\( \Sigma \)) at each location \((u_i, v_i)\), which the size of \( \Sigma \) is \( q \times q \) for the samples at the \( i \)th location.

\[
\Sigma(u_i, v_i) = \begin{bmatrix}
\sigma_{11}(u_i, v_i) & \sigma_{12}(u_i, v_i) & \cdots & \sigma_{1q}(u_i, v_i) \\
\sigma_{21}(u_i, v_i) & \sigma_{22}(u_i, v_i) & \cdots & \sigma_{2q}(u_i, v_i) \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{q1}(u_i, v_i) & \sigma_{q2}(u_i, v_i) & \cdots & \sigma_{qq}(u_i, v_i)
\end{bmatrix}. \tag{2}
\]

From Equation (2), the estimation of the variance-covariance error matrix parameters \( \hat{\Sigma}(u_i, v_i) \) is observed at each study location using the MLE and WLS methods. To get the estimation of the variance-covariance matrix parameter \( \hat{\Sigma}(u_i, v_i) \), the parameter estimation is determined at one the \( j \)th location (\( \hat{\sigma}_{ij}(u_i, v_i) \)) as follows:

\[
\hat{\sigma}_{ij}(u_i, v_i) = \frac{\sum_{i=1}^{n} w_{ij}(u_j, v_j)(Y_{ih} - (\hat{\beta}_{0i}(u_j, v_j) + \sum_{k=1}^{p} \hat{\beta}_{ki}(u_j, v_j)X_{ij}))^2}{n}
\]

\[
= \frac{(Y_h - X\hat{\beta}_h(u_j, v_j))^T W(u_j, v_j)(Y_h - X\hat{\beta}_h(u_j, v_j))}{n}
\]

\[
= SSE(u_j, v_j),
\]

The vector error at the location \((u_i, v_i)\) can be stated as follows:

\[
\epsilon_h = Y_h - \tilde{Y}_h = (I - S)Y_h, \tag{4}
\]

where \( I \) is the matrix identity with order \( n \) and \( S \) is the symmetric matrix sized \( n \times n \),

\[
S_{(error)} = \begin{bmatrix}
X_i^T(X^TW(u_j, v_j)X)^{-1}X^TW(u_i, v_i) \\
X_i^T(X^TW(u_j, v_j)X)^{-1}X^TW(u_j, v_j) \\
\vdots \\
X_n^T(X^TW(u_j, v_j)X)^{-1}X^TW(u_n, v_n)
\end{bmatrix}. \tag{5}
\]

About the local character of the MGWR model (3), the sum of square error (SSE) and the estimated parameters of the error variance-covariance can be determined.

**Proposition 1.** If SSE the location of \((u_j, v_j)\) the MGWR model is \( \epsilon^T(u_j, v_j) \epsilon(u_j, v_j) \) \((u_j, v_j) \in (u_j, v_j)\), then it can be determined \( SSE_h \) and the expectation value \( SSE_e \).

**Proof.** To get the SSE from the MGWR model using squaring (4) at the location to \((u_j, v_j)\) is:

\[
SSE = \epsilon_h^T(u_j, v_j) \epsilon_h(u_j, v_j)
\]

\[
= (I - S)Y_h^T (I - S)Y_h
\]

\[
= Y_h^T (I - S)^T (I - S)Y_h, \tag{6}
\]

where

\[
E(\epsilon_h(u_j, v_j)) = E(Y_h - \tilde{Y}_h)
\]

\[
= X^T \beta(u_j, v_j) - X^T \tilde{\beta}(u_j, v_j) = 0, \tag{7}
\]
and variance error is

\[
\text{Var}(\varepsilon_h(u_j, v_j)) = E \left[ \left( \varepsilon_h(u_j, v_j) - E(\varepsilon_h(u_j, v_j)) \right)^T \left( \varepsilon_h(u_j, v_j) - E(\varepsilon_h(u_j, v_j)) \right) \right]
\]

Based on (8), then (6) can be described as follows:

\[
SSE_h(u_j, v_j) = E_h(u_j, v_j) \varepsilon_h(u_j, v_j)
\]

From Equation (9), we can find the expected value \(E_h(u_j, v_j)\) as follows:

\[
E(SSE_h(u_j, v_j)) = E \left( E_h(u_j, v_j) (I - S)^T (I - S) \varepsilon_h(u_j, v_j) \right)
\]

Furthermore, \(SSE_h(E_h(u_j, v_j))\) is searched using (9), we obtain

\[
SSE_h(E_h(u_j, v_j)) = \left( E_h(u_j, v_j) (I - S)^T (I - S) \varepsilon_h(u_j, v_j) \right)^T
\]

**Theorem 1.** If \(SSE_h\) is given by Proposition 1 and the estimation of variance \(\hat{\sigma}_{h}^2(u_j, v_j)\) is given by Proposition 2, the estimated variance-covariance error of the MGWR model is given as follows:

\[
\hat{\sigma}_{h}^2(u_j, v_j) = \left( \frac{1}{n} \sum_{h=1}^{n} \frac{Y_{h}^T (Y_{h}^T - X_{h}^T \hat{\beta}_{h})}{n} \right)
\]

**Proof.** From Equation (1) of the MGWR model,

\[
Y_{hi} = \beta_{hi}(u_i, v_i) + \sum_{k=1}^{p} \beta_{hk}(u_i, v_i)X_{ki} + \varepsilon_{hi}.
\]

To determine \(SSE_h(E_h(u_j, v_j))\) at each location \((u_j, v_j)\), it can be approached using Equation (5),
\[ e_h W(u_j, v_j) e_h = \left( Y_h - X \hat{\beta}_h(u_j, v_j) \right)^T \cdot W(u_j, v_j) \left( Y_h - X \hat{\beta}_h(u_j, v_j) \right) \]

\[ E\left( e_h W(u_j, v_j) e_h \right) = E\left( Y_h - X \hat{\beta}_h(u_j, v_j) \right)^T \cdot W(u_j, v_j) \left( Y_h - X \hat{\beta}_h(u_j, v_j) \right) \]

\[ SSE_h E_h(u_j, v_j) = \left( Y_h - X \hat{\beta}_h(u_j, v_j) \right)^T \cdot W(u_j, v_j) \left( Y_h - X \hat{\beta}_h(u_j, v_j) \right), \]

and

\[ \sigma_{bh}^2(u_j, v_j) = \frac{SSE_h E_h(u_j, v_j)}{n}. \]

Based on Propositions 1 and 2, the theorems of estimation parameter variance-covariance error matrix for MGWR model are determined.

**Theorem 2.** If \( E(SSE_h(u_j, v_j)) \) satisfies Proposition 1 and \( E(SSE_h E_h(u_j, v_j)) \) satisfies Proposition 2, the estimated parameter variance-covariance errors matrix of the MGWR model are \( \hat{\sigma}_{bh}^2(u_j, v_j) = \frac{SSE_h E_h(u_j, v_j)}{n} \cdot \frac{1}{(I - S)^T (I - S)} \) and \( E(\hat{\sigma}_{bh}^2(u_j, v_j)) = \sigma_{bh}^2(u_j, v_j) \).

**Proof.** Based on Proposition 1 and 2, the estimated error variance-covariance parameters from the MGWR model are:

\[ \text{Var}(e_h(u_j, v_j), e_h(u_j, v_j)) = E\left( e_h^T(u_j, v_j) e_h(u_j, v_j) \right) \]

\[ \hat{\sigma}_{bh}^2(u_j, v_j) = E\left( e_h^T(u_j, v_j) e_h(u_j, v_j) \right) \]

\[ \sigma_{bh}^2(u_j, v_j) = \frac{SSE_h(u_j, v_j)}{(n - 2tr(S) + tr(S^T S))}, \]

and \( \hat{\sigma}_{bh}^2(u_j, v_j) = \frac{SSE_h E_h(u_j, v_j)}{(n - 2tr(S) + tr(S^T S))} \).

By using the characteristics of the matrix \( (I - S)^T (I - S) \), \( E(\hat{\sigma}_{bh}^2(u_j, v_j)) \), and \( E(\hat{\sigma}_{bh}^2(u_j, v_j)) \) can be determined to satisfy the unbiased.

**Theorem 3.** If \( \hat{\sigma}_{bh}^2(u_j, v_j) = \frac{SSE_h E_h(u_j, v_j)}{tr(1 - S)^T (I - S)} \) is an unbiased estimator \( \sigma_{bh}^2(u_j, v_j) \), then \( E(\hat{\sigma}_{bh}^2(u_j, v_j)) \), and \( E(\hat{\sigma}_{bh}^2(u_j, v_j)) \) can be determined to satisfy the unbiased.

**Proof.**

\[ E(\hat{\sigma}_{bh}^2(u_j, v_j)) = E\left( \frac{SSE_h(u_j, v_j)}{tr(1 - S)^T (I - S)} \right) \]

\[ = \frac{1}{tr(1 - S)^T (I - S)} E(SSE_h(u_j, v_j)) \]

\[ = \frac{1}{tr(1 - S)^T (I - S)} tr(1 - S)^T (I - S) \sigma_{bh}^2(u_j, v_j) \]

\[ = \sigma_{bh}^2(u_j, v_j), \]

and in the same way, we obtain

\[ E(\hat{\sigma}_{bh}^2(u_j, v_j)) = E\left( \frac{SSE_h E_h(u_j, v_j)}{tr(1 - S)^T (I - S)} \right) = \sigma_{bh}^2(u_j, v_j), \]

where \( \hat{\sigma}_{bh}^2(u_j, v_j) \) and \( \hat{\sigma}_{bh}^2(u_j, v_j) \) are an estimate of the unbiased error variance-covariance matrix for \( \sigma_{bh}^2(u_j, v_j) \) and \( \sigma_{bh}^2(u_j, v_j) \).

By using Theorem 1, 3, an unbiased estimate is obtained from the variance-covariance error matrix \( \Sigma(u_j, v_j) \) at the \( f \)-th location as follows:

\[ \hat{\Sigma}(u_j, v_j) = \left[ \begin{array}{cccc} \hat{\sigma}_1^2(u_j, v_j) & \hat{\sigma}_{12}(u_j, v_j) & \cdots & \hat{\sigma}_{1q}(u_j, v_j) \\ \hat{\sigma}_{21}(u_j, v_j) & \hat{\sigma}_2^2(u_j, v_j) & \cdots & \hat{\sigma}_{2q}(u_j, v_j) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{q1}(u_j, v_j) & \hat{\sigma}_{q2}(u_j, v_j) & \cdots & \hat{\sigma}_q^2(u_j, v_j) \end{array} \right]. \]

Since the variance-covariance error matrix \( \Sigma(u_j, v_j) \) satisfies the unbiased nature, then in the same way in other locations, it also meets the unbiased nature. Mathematically, the estimation of the variance-covariance matrix parameters \( \Sigma \) at the location to \( (u_i, v_i) \) can be stated as follows:

\[ \hat{\Sigma}(u_i, v_i) = \left[ \begin{array}{cccc} \hat{\sigma}_1^2(u_i, v_i) & \hat{\sigma}_{12}(u_i, v_i) & \cdots & \hat{\sigma}_{1q}(u_i, v_i) \\ \hat{\sigma}_{21}(u_i, v_i) & \hat{\sigma}_2^2(u_i, v_i) & \cdots & \hat{\sigma}_{2q}(u_i, v_i) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{q1}(u_i, v_i) & \hat{\sigma}_{q2}(u_i, v_i) & \cdots & \hat{\sigma}_q^2(u_i, v_i) \end{array} \right]. \]

Thus, it is proven that if \( \hat{\Sigma}(u_j, v_j) \) as an unbiased estimate of the variance-covariance error matrix \( \Sigma(u_j, v_j) \), then \( \hat{\Sigma}(u_i, v_i) \) is also an unbiased estimate of the variance-covariance error matrix \( \Sigma(u_i, v_i) \).

5. Conclusion

This research concludes that the MGWR model using MLE and WLS methods is suitable to obtain the estimated error variance-covariance parameters. The results prove that \( \hat{\Sigma}(u_j, v_j) \) is an unbiased estimate of the variance-covariance
error matrix $\Sigma(u_i, v_j)$. Since $\Sigma(u_i, v_j)$ is an unbiased estimate, then $\tilde{\Sigma}(u_i, v_j)$ is also an unbiased estimate of the variance-covariance error matrix $\Sigma(u_i, v_j)$ at all locations.

**Data Availability**

The authors declare that all of data is original and there is no data from others publication.

**Conflicts of Interest**

The authors declare that there is no conflict of interests regarding the publication of this article.

**Acknowledgments**

We would like to express our sincere gratitude to the Research Sub-Directorate, Community Development and Scientific Publications of the Directorate General of Islamic Higher Education (Dirjen DIKTIS) for providing funds for this research in 2018. Research and Community Service Institutions provide funding support with this publication.

**References**

[1] A. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression*, John Wiley and Sons, UK, 2002.

[2] N. Cressie, *Statistics for Spatial data*, John Wiley and Sons, New York, NY, USA, 2nd edition, 2015.

[3] H. Yasin, “Selection of variables in the geographically weighted regression model,” *Media Statistika*, vol. 4, no. 2, pp. 63–72, 2011.

[4] S. Harini, M. M. Purhadi, and S. Sunaryo, “Statistical test for multivariate geographically weighted regression model using the method of maximum likelihood ratio test,” *International Journal of Applied Mathematics & Statistics*, vol. 29, no. 5, pp. 110–115, 2012.

[5] S. Harini and Purhadi, "Parameter estimation of multivariate geographically weighted regression model using matrix laboratory," in *International Conference on Statistics in Science, Business and Engineering (ICSSBE)*, IEEE, Langkawi, Malaysia, 2012.

[6] P. Triyanto, W. O. Bambang, and W. P. Santi, "Parameter estimation of geographically weighted multivariate poisson regression," *Applied Mathematical Sciences*, vol. 9, no. 82, pp. 4081–4093, 2015.

[7] N. E. Syerrina, "A statistical analysis for geographical weighted regression," vol. 169, in *IOP Conference Series: Earth and Environmental Science*, p. 012105, IOP, Malaysia, 2018.

[8] S. Soemartojo, R. Ghaisani, T. Siswantining, and R. M. Shahab, "Parameter estimation of geographically weighted regression (GWR) model using weighted least square and its application," *AIP Conference Proceedings*, 2018.

[9] A. Comber, Y. Wang, and Y. Lü, Y. Xingchang, “Hyper-local geographically weighted regression: extending GWR through local model selection,” *Journal of Spatial Information Science*, vol. 17, pp. 63–84, 2018.

[10] S. Harini, M. M. Purhadi, and S. Sunaryo, “Linear model parameter estimator of spatial multivariate using restricted maximum likelihood,” *Journal of Mathematics and Technology*, pp. 56–61, 2010.