Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 6s (2025)

Spatial Moving Average Polynomial Model on Irregular Lattice and its Full-Likelihood Based Implementation

Lidia Khonglam[1] Sanjeeva Kumar Jha[2]

[1]Department of Statistics, NEHU, Shillong- 793022, India. Email: lidiaklam123@gmail.com

[2]Department of Statistics, NEHU, Shillong- 793022, India. Email: skjha@nehu.ac.in

Article History:

Received: *20-10-2024*

Revised: *05-12-2024*

Accepted: *12-12-2024*

Abstract:

The simultaneous incorporation of a contiguity-based neighbourhood matrix of first and second order neighbors in the spatial moving average model on an irregular lattice can be specified in two different ways- as a quadratic polynomial or as a product of two linear polynomials. The present article concerns with the second order spatial moving average polynomial model of the latter form. Expression for the joint probability distribution of the response is derived together with expressions for the first and second order moments of the model using characteristic function. Model parameters are estimated using method of maximum likelihood and expression for their standard errors are obtained. In the absence of an analytical solution of the full likelihood function, optimization of the full likelihood function using Differential Evolution technique has been performed. Confidence interval for drawing inference has been constructed based on the derived standard error expression and bootstrapping method. Using likelihood ratio test statistic and bootstrap results, simultaneous confidence regions for the two spatial dependence parameters are also constructed. Implementation of the model has been executed on simulated data, and relevant comparisons made with linear models and first order spatial moving average model. The advantage of using the proposed model lies in its ability to provide statistically valid inference, as it affects producing spatial dependence of the first and the second lag orders in the residuals, insignificant.

Keywords: spatial moving average; spatial lattice; contiguity neigbourhood; maximum likelihood; Differential Evolution; simultaneous confidence region; bootstrap

1. Introduction

Economic, biological, health and ecological data on space commonly occur as aggregated data over geographically or administratively defined regions, forming regular or irregular spatial lattices. Spatial regression models, model for regression analysis of such spatially dependent data, incorporate spatial dependence in the concept of a contiguity neighborhood matrix constructed based on the shape of a spatial lattice. The prominent spatial regression models for aggregated data are the Conditional Autoregressive (CAR), Simultaneous Autoregressive (SAR) and the Spatial Moving Average (SMA) model. CAR was developed by Besag (1974) under the specification of a conditional probability on the response variable. The SAR and SMA model are specified by the distributional assumption on the error; hence provide joint probability of the response variable, Cressie (1993). SAR was developed by Whittle (1954) and SMA by Hainning (1978). SAR and CAR consider the observation of the response at the neighboring units of a particular region as an additional covariate and associated with it a spatial dependence parameter. The defined dependence is then incorporated into the mean structure of the linear models, similar to the time series autoregressive model. SAR functional form and CAR model's property, respectively analogue the autoregressive and Markov property of time series models (Wall, 2004; Cressie, 1993). SMA model on the other hand, resembled the time series moving average model. In this model, the disturbances at the neighboring units of a particular region are considered as an additional covariate with spatial dependence parameter associated with it. Implementation of CAR is more suitable with the Bayesian approach whereas SAR and SMA with likelihood based approaches. All three models have established its significance in enhancing the precision of the parameter estimates and capturing spatial dependence.

Aggregated data on spatial lattices exhibit the existence of various kind of spatial interaction which causes spatial dependence. The geographical location at which an event occurs is the main factor of data on space; structuring dependence based on it forms the basis of spatial regression models. It assumed that an observation at one location shows an effects level that are similar to those of its neighbouring units, Anselin et al., 1998. The mentioned spatial regression models usually, established effects level based on the first order neighbours of the observation at each region. In essence, spatial dependence is not confined to the first order neighbours only. Somewise the higher order neighbours may or may not have an effect on the observation in a region. Analysis using the usual spatial regression models may have a substantial impact on the accuracy of the estimates, if higher order neighbours have a significant effect on the data at a region. Therefore, it is important to take into account the effects level based on higher order neighbourhood structure. Moreover, the incorporation of a higher order neighbourhood structure in the model allows the examination of complicated interactions that may exist and, to account for its presence, which may not be possible with just the first order neighbourhood framework. This manuscript study the specification of spatial dependence based on the first and second order neighbours in the SMA model and provides its implementation.

The incorporation of a contiguity matrix of second order neighbors (W_2) into the SMA model with first order contiguity neighbors (W_1) can be specified analogous to the moving average time series models. It follows extending the linear polynomial $(I + \theta_1 W_1)$ associated with the error structure of SMA to the form $(I + \theta_1 W_1 + \theta_2 W_2)$, a quadratic polynomial as named by Lesage & Pace, 2009. Likelihood based implementation of the second order SMA specified hereby involves the evaluation of $|I + \theta_1 W_1 + \theta_2 W_2|$ which consist of both the first and second order spatial dependence parameters. The computation of $|I + \theta_1 W_1 + \theta_2 W_2|$ becomes complicated with higher order neighbourhood structure, as the contiguity matrix becomes less sparse with increasing order of the neigbourhood structure, Lesage and Pace, 2009. The author suggested factorizing $(I + \theta_1 W_1 + \theta_2 W_2)$ into a product of two linear polynomials i.e., $(I + \theta_1 W_1)(I + \theta_2 W_2)$ so as to ease the computational complexity. Thus, the second order SMA polynomial model studied in this article is of the form where $(I + \theta_1 W_1)(I + \theta_2 W_2)$ is associated with the error structure, as given in section 2. Theoretically, the specification provides plausible ranges of the two spatial dependence parameters.

The second order SMA polynomial model proposed is implemented under the likelihood paradigm. The maximum likelihood (ML) estimate of the parameters is derived as shown in Section 3. It may be noticed that the analytical expression of the parameters ML estimates are functions of the other estimates. In the case of σ^2, it is a function of $\hat{\theta}_1, \hat{\theta}_2, \text{ and } \hat{\beta}$, and for β, it is a function of $\hat{\theta}_1 \text{ and } \hat{\theta}_2$. Expressing the estimates of σ^2 and β accordingly as a function of $\theta_1 \text{ and } \theta_2$ in the likelihood function, ML estimates of $\theta_1 \text{ and } \theta_2$ may be initially obtained. ML estimates of β and σ^2 may be correspondingly attained. Thus, estimation is based on modified likelihoods. In this paper estimation is not based on modified likelihoods. Differential evolution (DE) algorithm, proposed by Price & Storn (1997) available in R package; an optimization method for any non-differentiable or non-linear function is used as the method of optimization of the full likelihood function of the study model. The DE optimization method consists of the initialization, mutation, recombination and selection steps. The optimization of the full likelihood function provides desirable properties of its estimates relative to consistency, asymptotic efficiency and asymptotic normality, as compared to the other likelihood based estimation. It also made the construction of the simultaneous confidence regions for the two spatial dependence parameters possible. The optimization method has been successfully used for the implementation of skew Normal SAR model, Jha S.K. et al. (2021) and SMA model of first order, Jha S.K. et al. (2023).

The main purpose of the paper is to study the implementation of the second order SMA polynomial model and to acknowledge its importance in modeling relationships in spatially dependent data. The remaining of the paper is organized as follows: Section 2 define the form of the second order SMA polynomial, Section 3 gives the ML derivation of the model, Section 4 provides the implementation of the model on simulated data and Section 5 presented the conclusion.

2. Models

Let $s_1, s_2, \ldots\ldots, s_n$ be the n-sites in the study region S which forms a regular or irregular spatial lattice. Let z_i and $x_i = (x_{i0}, x_{i1}, \ldots\ldots, x_{ik})^T$ be respectively the response and explanatory variables observed at each of the site $s_i \in S$; $i = 1,2,3,\ldots\ldots n$. Haining (1978) defined Spatial Moving Average (SMA) model as:

$$\mathbf{Z} = X\boldsymbol{\beta} + (I + \theta_1 W_1)\boldsymbol{\varepsilon} \tag{2.1}$$

Where $\varepsilon \sim N(0, \sigma^2 I)$; $\beta = (\beta_0, \beta_1, \ldots., \beta_k)$ is a vector of regression parameters, θ_1 is the unknown scalar spatial dependence parameter associated with the first order neighbors and W_1 is a symmetric contiguity matrix of first order neighbors with elements w_{ij}^1 defined as:

$$w_{ij}^1 = \begin{cases} 1 \text{ if } s_i \text{ and } s_j \text{ are first order neighbors} \\ 0, \text{ otherwise and if } s_i = s_j \end{cases} \tag{2.2}$$

2.1.Second order SMA Polynomial

Under the same notations as mentioned above, the second order SMA polynomial may be defined as:

$$\mathbf{Z} = X\boldsymbol{\beta} + (I + \theta_1 W_1)(I + \theta_2 W_2)\boldsymbol{\varepsilon} \tag{2.3}$$

Where $\varepsilon \sim N(0, \sigma^2 I)$; θ_2 is the unknown scalar spatial dependence parameter associated with the second order neighbors and W_2 is a symmetric contiguity neighborhood matrix of second order neighbors with elements w_{ij}^2 defined as:

$$w_{ij}^2 = \begin{cases} 1 \text{ if } s_i \text{ and } s_j \text{ are second order neighbors} \\ 0, \text{ otherwise and if } s_i = s_j \end{cases} \tag{2.4}$$

The probability density function of $\boldsymbol{\varepsilon}$, takes the form,

$$f(\varepsilon/0, \sigma^2 I) = \frac{e^{-\frac{\varepsilon^T \varepsilon}{2\sigma^2}}}{\sigma(2\pi)^{n/2}}$$

$$= \frac{e^{-\frac{\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]^T\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]}{2\sigma^2}}}{\sigma(2\pi)^{n/2}}$$

The corresponding kernel is,

$$K(Z) = e^{-\frac{\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]^T\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]}{2\sigma^2}}$$

Normalizing constant can be obtained as;

$$\int K(Z)dZ = \int e^{-\frac{\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]^T\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]}{2\sigma^2}}$$

Let, $(I + \theta_2 W_2)^{-1}(I + \theta_1 W_1)^{-1}(Z - X\beta)\sigma^{-1} = U$

$$\therefore Z = (I + \theta_1 W_1)(I + \theta_2 W_2)\sigma U + X\beta$$

The differential of Z with respect to U is $\sigma|(I + \theta_1 W_1)(I + \theta_2 W_2)|$

$$\therefore \int K(Z)dZ = \int \frac{e^{-\frac{U^T U}{2}}}{(2\pi)^{n/2}} \times \sigma|(I + \theta_1 W_1)(I + \theta_2 W_2)|(2\pi)^{n/2} dU$$

$$= \sigma(2\pi)^{n/2}|(I + \theta_1 W_1)(I + \theta_2 W_2)| \tag{2.5}$$

Equation (2.5) gives the normalizing constant for the joint distribution of Z as, $\sigma(2\pi)^{n/2}|(I + \theta_1 W_1)(I + \theta_2 W_2)|$. If $(I + \theta_1 W_1)$ and $(I + \theta_2 W_2)$ is a full rank matrix then the induced joint density function of Z takes the form;

$$f(Z) = \frac{1}{\sigma(2\pi)^{n/2}|(I + \theta_1 W_1)(I + \theta_2 W_2)|} \times \exp\left\{-\frac{\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]^T\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]}{2\sigma^2}\right\} \tag{2.6}$$

Equation (2.6) gives the required joint probability density function under the proposed second order SMA polynomial. Let the distribution be denoted henceforth as;

$$Z \sim 2^{nd} - order - SMA - polynomial\ (\beta, \sigma^2, \theta_1, \theta_2)$$

2.2.Characteristic function and moments:

The derivation of the characteristic function and the corresponding moments of the response vector 'Z' is given in **Appendix A**. The obtained characteristic function is:

$$\Psi_Z(t) = e^{it^T X\beta - \frac{t^T \omega t}{2}} \tag{2.7}$$

Where $\omega = \sigma^2 (I + \theta_1 W_1)(I + \theta_2 W_2)(I + \theta_2 W_2)^T (I + \theta_1 W_1)^T$

And the derived mean vector and variance-covariance matrix of 'Z' is:

$$E(Z) = X\beta \text{ and, } Cov\ (Z) = \sigma^2 (I + \theta_1 W_1)(I + \theta_2 W_2)(I + \theta_2 W_2)^T (I + \theta_1 W_1)^T$$

3. Likelihood function and Maximum Likelihood (ML) estimate

The log- likelihood function of second order SMA polynomial is:

$$\log\left(L\left(\beta, \sigma^2, \theta_1, \theta_2 / z\right)\right) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\sigma^2\right) - \log\ \left|(I + \theta_1 W_1)(I + \theta_2 W_2)\right|$$
$$- \frac{1}{2\sigma^2}\left\{(I + \theta_2 W_2)^{-1}(I + \theta_1 W_1)^{-1}(Z - X\beta)\right\}^T \left\{(I + \theta_2 W_2)^{-1}(I + \theta_1 W_1)^{-1}(Z - X\beta)\right\} \tag{3.1}$$

Let $G = (I + \theta_1 W_1)(I + \theta_2 W_2) = G_1 G_2$ where $G_1 = (I + \theta_1 W_1)\ and\ G_2 = I + \theta_2 W_2$

Then, the first partial derivative of (3.1) w.r.t the unknown parameters $\beta, \sigma^2, \theta_1, \theta_2$ respectively is,

$$\frac{\partial \log(L)}{\partial \beta} = \frac{1}{\sigma^2} X^T \left(G^{-1}\right)^T G^{-1}(Z - X\beta) \tag{3.2}$$

$$\frac{\partial \log(L)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\left[G^{-1}(Z - X\beta)\right]^T \left[G^{-1}(Z - X\beta)\right]$$
(3.3)

$$\frac{\partial \log(L)}{\partial \theta_1} = -\ tr\left(G^{-1} W_1 G_2\right) + \frac{1}{2\sigma^2}(Z - X\beta)^T \left\{\left(G^{-1}\right)^T G^{-1} W_1 G_2 G^{-1} + \left(G^{-1}\right)^T \left(G^{-1}\right)^T W_1 G_2 G^{-1}\right\}(Z - X\beta) \tag{3.4}$$

And,

$$\frac{\partial \log(L)}{\partial \theta_2} = -\ tr\left(G^{-1} W_2 G_1\right) + \frac{1}{2\sigma^2}(Z - X\beta)^T \left\{\left(G^{-1}\right)^T G^{-1} W_2 G_1 G^{-1} + \left(G^{-1}\right)^T \left(G^{-1}\right)^T W_2 G_1 G^{-1}\right\}(Z - X\beta) \tag{3.5}$$

At the maximum of $\beta, \sigma^2, \theta_1, \theta_2$ respectively, the expected values of partial derivatives (3.2) to (3.4) must be zero. The proofs for the first two are familiar; therefore we will show here only for the two spatial dependence parameters.

$$E\left[\frac{\partial \log(L)}{\partial \theta_1}\right] = -tr\left(G^{-1}W_1G_2\right) + \frac{1}{2\sigma^2}E\left[\varepsilon^T G^{-1}W_1G_2\varepsilon + \varepsilon^T\left(G^{-1}\right)^T W_1G_2\varepsilon\right]$$

$$= -tr\left(G^{-1}W_1G_2\right) + \frac{\left(\sigma^2 tr\left(G^{-1}W_1G_2\right) + \sigma^2 tr\left(\left(G^{-1}\right)^T W_1G_2\right)\right)}{2\sigma^2}$$

$$= -tr\left(G^{-1}W_1G_2\right) + tr\left(G^{-1}W_1G_2\right)$$

$$= 0$$

Similarly,

$$E\left(\frac{\partial \log(L)}{\partial \theta_2}\right) = -tr\left(G^{-1}W_2G_1\right) + \frac{1}{2\sigma^2}E\left[\varepsilon^T G^{-1}W_2G_1\varepsilon + \varepsilon^T\left(G^{-1}\right)^T W_2G_1\varepsilon\right]$$

$$= -tr\left(G^{-1}W_2G_1\right) + tr\left(G^{-1}W_2G_1\right)$$

$$= 0$$

The elements of the information matrix are:

$$-E\left[\frac{\partial^2 \log L}{\partial \beta\, \partial \beta^T}\right] = \frac{1}{\sigma^2}\left[G^{-1}X\right]^T\left[G^{-1}X\right] \tag{3.6}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \beta\, \partial \sigma^2}\right] = 0 \tag{3.7}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \beta\, \partial \theta_1}\right] = 0 \tag{3.8}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \beta\, \partial \theta_2}\right] = 0 \tag{3.9}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \sigma^2\, \partial \sigma^2}\right] = \frac{n}{2\sigma^4} \tag{3.10}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \sigma^2\, \partial \theta_1}\right] = \frac{tr\left(W_1G_2G^{-1}\right)}{\sigma^2} \tag{3.11}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \sigma^2\, \partial \theta_2}\right] = \frac{tr\left(W_2G_2G^{-1}\right)}{\sigma^2} \tag{3.12}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \theta_1^2}\right] = tr\left(W_1G_2G^{-1}\right)^2 + tr\left[\left(\left(W_1G_2\right)G^{-1}\right)^T\left(\left(W_1G_2\right)G^{-1}\right)\right] \tag{3.13}$$

$$-E\left[\frac{\partial^2 \log L}{\partial \theta_2^2}\right] = tr\left(W_2 G_1 G^{-1}\right)^2 + tr\left[\left(\left(W_2 G_1\right) G^{-1}\right)^T \left(\left(W_2 G_1\right) G^{-1}\right)\right] \quad (3.14)$$

$$-E\left[\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_2}\right] = 2tr\ \left(W_1 W_2\right) \quad (3.15)$$

The information matrix $I\left(\beta, \sigma^2, \theta_1, \theta_2\right)$ may then be expressed as:

$$\begin{bmatrix} \frac{1}{\sigma^2}\left[G^{-1}X\right]^T\left[G^{-1}X\right] & 0 & 0 & 0 \\ 0 & \frac{n}{2\sigma^4} & \frac{tr\left(W_1 G_2 G^{-1}\right)}{\sigma^2} & \frac{tr\left(W_2 G_1 G^{-1}\right)}{\sigma^2} \\ 0 & \frac{tr\left(W_1 G_2 G^{-1}\right)}{\sigma^2} & tr\left(W_1 G_2 G^{-1}\right)^2 + tr\left[\left(\left(W_1 G_2\right) G^{-1}\right)^T \left(\left(W_1 G_2\right) G^{-1}\right)\right] & 2tr\ \left(W_1 W_2\right) \\ 0 & \frac{tr\left(W_2 G_1 S^{-1}\right)}{\sigma^2} & 2tr\ \left(W_1 W_2\right) & tr\left(W_2 G_1 G^{-1}\right)^2 + tr\left[\left(\left(W_2 G_1\right) G^{-1}\right)^T \left(\left(W_2 G_1\right) G^{-1}\right)\right] \end{bmatrix} \quad (3.16)$$

The information matrix (3.16) has block-diagonality between group of parameters, similar to the first order SMA model as shown in Hepple (2003); therefore, the variance for $\hat{\beta}$ may be calculated using,

$$Var\ (\hat{\beta}) = \sigma^2[((I + \theta_2 W_2)^{-1}(I + \theta_1 W_1)^{-1}X)^T((I + \theta_2 W_2)^{-1}(I + \theta_1 W_1)^{-1}X)]^{-1} \quad (3.17)$$

With the ML estimate of θ_1 and θ_2. Further, the sub-matrix involving only σ^2, θ_1, θ_2 can be used to obtain the variance of θ_1 *and* θ_2 .

At the maximum, the partial derivatives of equation (3.2) and (3.3) equals zero. Therefore,

$$\hat{\beta} = \left[X^T G^{-1^T} G^{-1} X\right]^{-1} X^T G^{-1^T} G^{-1} Z \quad (3.18)$$

And, $\hat{\sigma}^2 = n^{-1} D^{*^T} G^{-1^T} G^{-1} D^*$ where, $D^* = Z - X\hat{\beta}$ (3.19)

Under the specified values of θ_1 *and* θ_2, the estimates of β can be obtained as a linear model, using the Generalised Least Squares (GLS) equation. The log likelihood function in equation (3.1) after replacing β by $\hat{\beta}$ of (3.18) and σ^2 by $\hat{\sigma}^2$ of (3.19) becomes,

$$\log\left(L(\beta, \theta_1, \theta_2/Z)\right) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(n^{-1} D^{*^T} G^{-1^T} G^{-1} D^*\right) - \log|G| - \frac{1}{2}$$
$$= -\frac{n}{2}\left[\log\left((2\pi)n^{-1}\right) + n^{-1}\right] - \frac{n}{2}\log\left[|G|^{2/n} D^{*^T} G^{-1^T} G^{-1} D^*\right] \quad (3.20)$$

Thus, numerical optimization of (3.20) gives estimates of θ_1 *and* θ_2, after which $\hat{\beta}$ *and* $\hat{\sigma}^2$ could be obtained using (3.18) and (3.19) respectively. The present paper however, uses the technique of DE algorithm for optimization of the full likelihood in (3.1). The method does not necessarily require the step wise derivation shown for obtaining the ML estimates of the parameters. It simultaneously estimates all the parameters; hence allows one to construct confidence intervals based on bootstrap method and for construction of simultaneous confidence region for the two or more spatial dependence parameters.

3.1. Parameter spaces for $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$: Hepple, 2003 mentioned that the structure of W_1 and its eigen values determine the range of θ_1 in the first order SMA. The condition arises from $|I+\theta_1 W_1|$ that appears in the likelihood function of the model, as it should be positive (Anselin, 1988; Mur et al, 2006). Its range should be between the negative of the reciprocal of the largest and the smallest eigen values of W_1. Proceeded in a similar manner, the feasible parameter spaces for the two spatial dependence parameters of the second order SMA polynomial may be obtained theoretically. The jacobian of the model under study is:

$$\begin{aligned}|G| &= |(I+\theta_1 W_1)(I+\theta_2 W_2)| \\ &= |I+\theta_1 W_1||I+\theta_2 W_2| \\ &= \prod_{i=1}^{R}(1+\theta_1\lambda_i)\prod_{i=1}^{R}(1+\theta_2\delta_i) > 0\end{aligned}$$

Thus, $\prod_{i=1}^{R}(1+\theta_1\lambda_i) > 0$ and, $\prod_{j=1}^{R}(1+\theta_2\delta_j) > 0$

$\Rightarrow \frac{1}{\lambda_{max}} < \theta_1 < \frac{1}{\lambda_{min}}$ and, $\frac{1}{\delta_{max}} < \theta_2 < \frac{1}{\delta_{min}}$

Where $\lambda_{\max}$ and $\lambda_{\min}$ correspondingly represents the maximum and minimum eigen values associated with the weighted contiguity neighbors of first orders; $\delta_{\max}$ and $\delta_{\min}$ respectively are the maximum and minimum eigen values of the contiguity neighbors of second orders. In the present analysis, the range of θ_1 is (-1, 1.6108) and for θ_2 is (-1, 1.9195).

4. Implementation

Simulated data is used for implementation of the study second order SMA polynomial. The New York shape file, which has 281 sub-regions available in 'spData' R-package, is used for constructing the spatial contiguity matrix of first and second order neighborhoods required for implementation. A vector of random error $\varepsilon \sim N(0,25)$ and a vector of explanatory variable, $X_2 \sim uniform((20,30))$ each having dimensions 281 are generated. X_1 is a unitary vector with 281 dimension, such that $X=[X_1 \quad X_2]$. W_1 *and* W_2 are 281×281 spatial contiguity matrix of first and second order neighborhoods respectively. Then, with $\beta_1 = 20$, $\beta_2 = 10$, $\theta_1 = 0.3$, $\theta_2 = 0.1$ a response variable Z with 281 dimensions is obtained using (2.3).

With the response variable, Z and the explanatory variable X_2 simulated as above, analysis under the linear model assumption signifies the presence of a significant spatial dependence of the first two lag orders in the residuals. The Moran's I value for the first lag order is 0.55045 and that of second lag order is 0.17189 with both p-value less than 0.05. Therefore, this simulated data is considered fit for implementation of second order SMA polynomial.

Optimization of the full-likelihood function is carried out using the 'DEoptim' technique for obtaining the required parameter ML estimates. In running the program, convergence of the optimized log likelihood function and its corresponding parameters with respect to the number of iterations is firstly assessed, and the results are presented in Fig. 1 and Fig. 2 accordingly. Both figures show that the overall convergence is at about 70 iterations. The current analysis reports result of the parameter estimated in running the program for 200 iterations.

Results from second order SMA polynomial: The obtained ML estimate of the parameters, its standard error and 95 per cent exact confidence interval is presented in Table 1. The variance for constructing the 95 per cent exact confidence interval for the covariates is obtained using (3.16) and, for the two spatial dependence parameters using the inverse of the information sub-matrix involving σ^2, θ_1 *and* θ_2. The table also, presents bootstrap results based on 1000 resampled values, and its respective density plots are presented in Fig 3.

From table 1, it is observed that both the regression parameters β_0 *and* β_1 are significant at 0.05 levels of significance. The estimated spatial dependence is 1.07464 and 0.76755 for the first and second lag orders respectively; both being significant at 5 per cent significance level. Bootstrap confidence interval also gives similar results for the regression parameters and the two spatial dependences. The estimated global variance under bootstrapping is 21.88125 and under the ML estimate is 25.05253.

The observed response and the estimated response values under the SMA polynomial of second order are further presented in a choropleth map as shown in Fig 4. It may be noticed from the figure that second order SMA polynomial properly capture the variability in the dataset as ranges of the response values is reduced by 15 per cent from the observed values. The observed response ranges from 212 to 329 whereas the estimated response ranges between 219 and 320.

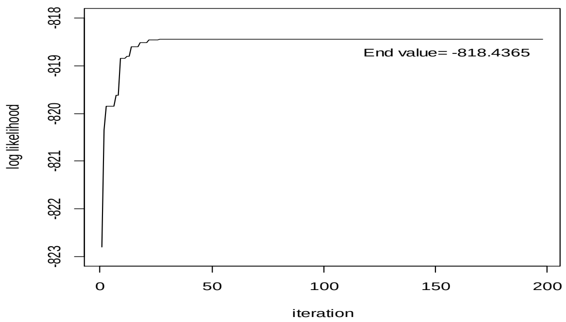


Fig 1: Convergence plot of the optimized log-likelihood function

Table 1: Parameter estimates, its standard error and 95 per cent confidence interval

	Estimates (se)	Exact 95 per cent CI	Bootstrap		
			Mean	se	95 per cent CI
β_0	18.708 (2.174)	(14.447, 22.969)	18.348	3.695	(11.203, 25.689)
β_1	10.022 (0.008)	(10.007, 10.037)	10.039	0.126	(9.795, 10.290)
θ_1	1.075 (0.009)	(1.057, 1.093)	1.143	0.111	(0.913, 1.352)
θ_2	0.768 (0.116)	(0.539, 0.996)	0.869	0.222	(0.401, 1.268)
σ	5.005		4.678		
AIC	1646.873				

**se* stands for standard error and *CI* for confidence interval*

Simultaneous Confidence region between θ_1 and θ_2: The simultaneous confidence region for the two spatial dependence parameters based on likelihood ratio (LR) test statistic and bootstrap results is constructed. The simultaneous confidence region based on LR test is obtained by firstly generating θ_1 *and* θ_2 within (-1, 1.611) and (-1, 1.919) respectively. The likelihood values with the different combinations of θ_1 *and* θ_2 is calculated; then only those combinations of θ_1 *and* θ_2 which gives the likelihood value greater than or equal to the maximized log likelihood values at 95 per cent confidence level is taken. Hence, the LR based simultaneous confidence region is as presented in Fig 5 (a). Bootstrap based confidence region is constructed by taking the lower 95 per cent values of the Mahalanobis distance between the bootstrap vector values of θ_1 *and* θ_2 and the region is as presented in Fig 5(b).

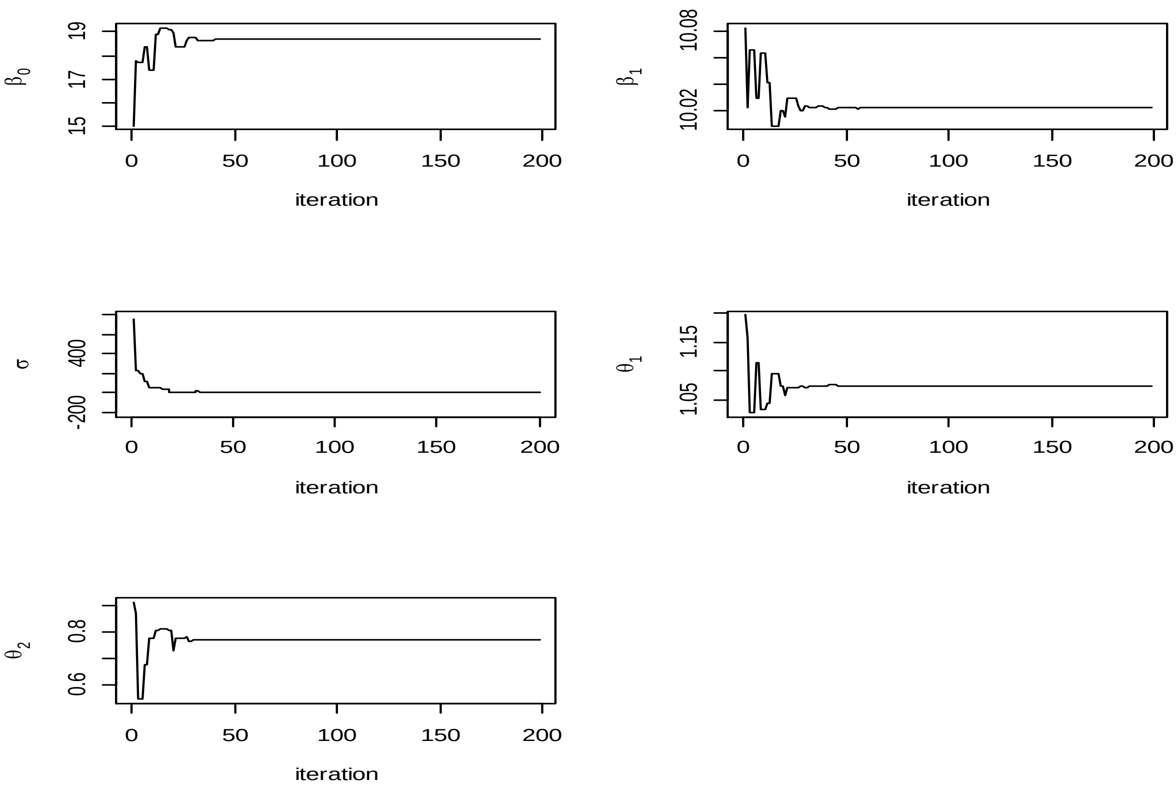


Fig 2: Convergence plot of the parameter estimate

From both figures in Fig 5, it may be noticed that there is a positive correlation between θ_1 *and* θ_2. The calculated correlation coefficient between them is 0.249 for LR test and 0.376 for bootstrap values with both p-values being less than the 0.05 significance level. The 95 per cent confidence interval based on LR statistic is (0.929, 1.184) for θ_1 and for θ_2 is (0.426, 1.049). Based on bootstrap values, the 95 per cent confidence interval is (0.913, 1.352) and (0.401, 1.268) for θ_1 *and* θ_2 respectively.

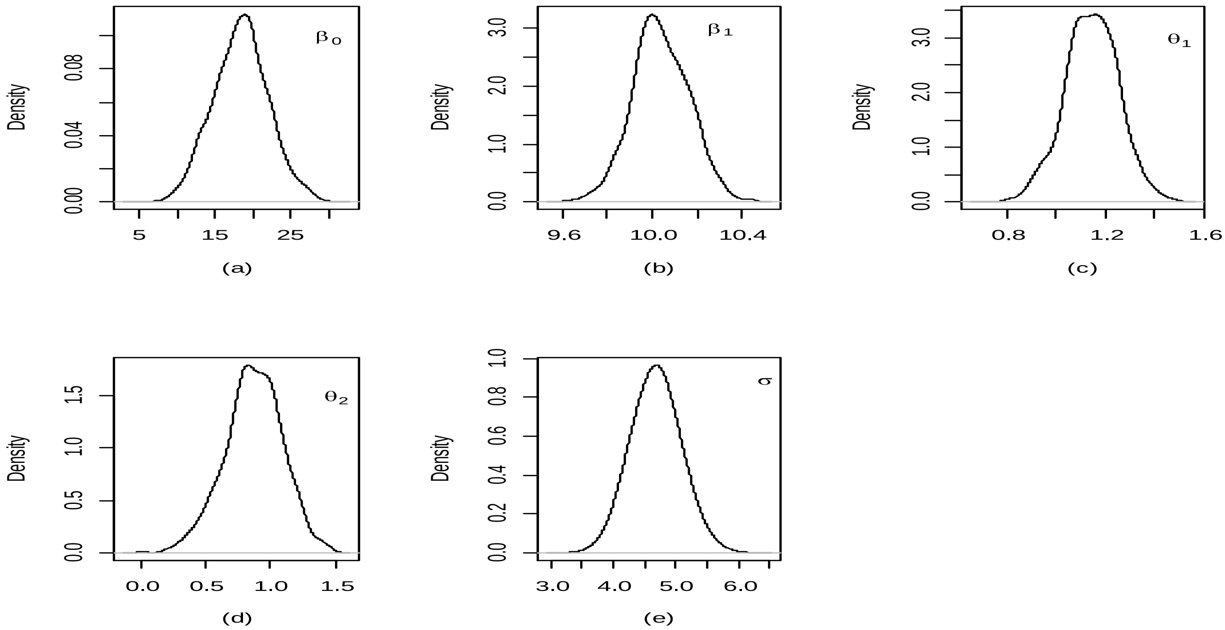
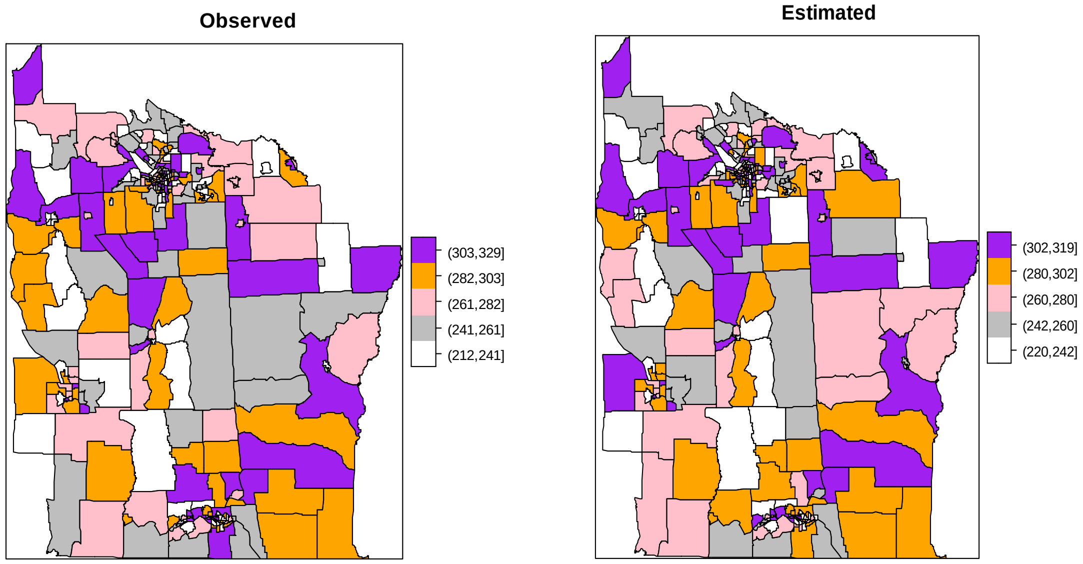
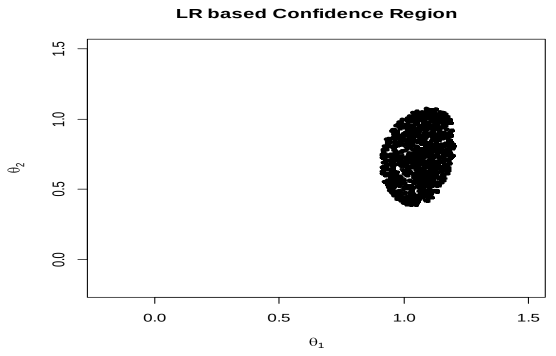


Fig 3: Bootstrap result density plots

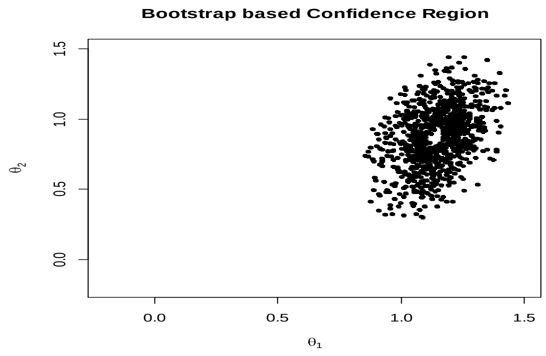


(a) (b)

Fig 4: (a) Observed response, and (b) estimated response under second order SMA polynomial



(a)



(b)

Fig 5: Simultaneous confidence region between θ_1 *and* θ_2 based on LR test and Bootstrap

Table 2: Parameter estimates with its 95 per cent confidence interval under the linear model, first order SMA and second order SMA polynomial

	Linear models	First order SMA	Second order SMA polynomial
β_0	17.579 (11.139, 24.019)	19.205 (14.891, 23.518)	18.7083 (14.447, 22.969)
β_1	10.074 (9.819, 10.328)	10.00709 (9.992, 10.023)	10.02187 (10.007, 10.037)
θ_1	——	1.007568 (0.959, 1.056)	1.075 (1.057, 1.092)
θ_2	——	——	0.76755 (0.539, 0.996)
σ^2	39.163	25.417	25.053
AIC	1832.099	1667.936	1646.873

Table 3: Moran's I statistic of the residual with its p-values under the Linear model, first order SMA and second order SMA polynomial

	Moran's I statistic of the residuals		
	Linear model	First order SMA	Second order SMA polynomial
First lag order	0.550	0.103	0.067
(p-value)	(<0.05)	(0.004)	(0.102)
Second lag order	0.172	0.107	-0.002
(p-value)	(<0.05)	(<0.05)	(0.956)

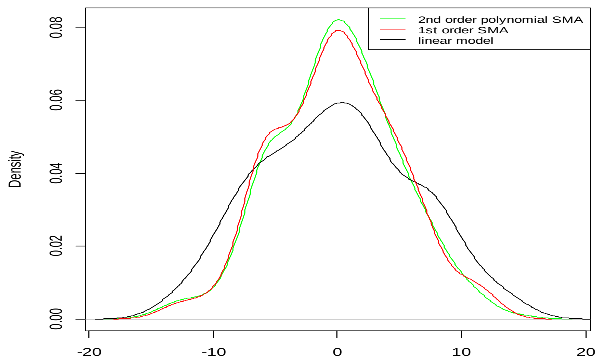
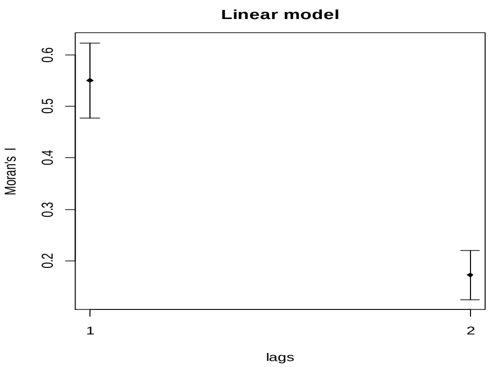
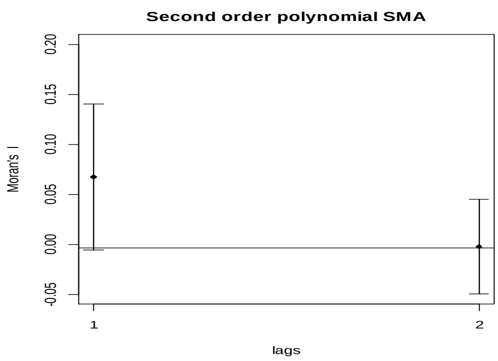


Fig 6: Residuals plot under linear models, first order SMA, and SMA polynomial of second order

Comparison of results: The simulated data generated is further analyzed under the linear models and first order SMA, and its results are presented in Table 2. It is observed that all three models give similar results in term of the regression parameters, both being significant at 95 per cent confidence interval. The overall gain in using the second order SMA polynomial rather than the first order SMA model and linear model is reflected in the AIC value. The AIC value for the proposed model is 1646.873 whereas for the linear model is 1832.099, and 1667.936 for SMA model of first order. The estimated variance is reduced by only 1 per cent when using the second order SMA polynomial.



(a) (b)



(c)

Fig 7: Moran’s I statistic with its 95 per cent Confidence interval of the residuals under the (a) linear model, (b) first order SMA, (c) Second order SMA polynomial

Residual density plot under the second order SMA polynomial overlaid with those of the linear model and first order SMA is given in Fig 6. The first order SMA and second order SMA polynomial shows nearly same pattern in achieving smoothness as compared to linear model. The Shapiro-wilks normality test statistic provides a p-value of 0.073, 0.489 and 0.805 for the residuals under the linear model, first order SMA and second order SMA polynomial model respectively. Table 3 provides results of the analytical test of the presence of a significant spatial dependence of the first and second lag orders in the residuals. Correspondingly, Fig 7 provides a diagrammatical representation of the result with its 95 per cent confidence interval. From both, the table and figure we may be clearly observed that the residuals from the linear model and first order SMA exhibit a significant spatial dependence of both lag orders, whereas those from the second order SMA polynomial do not.

5. Conclusion

The paper studies the aspect of modeling relationships between spatially dependent data through the simultaneous incorporation of the first and second order contiguity neighbourhood matrix in SMA, with the model named as second order SMA polynomial. The model provides a way of modeling, in cases where spatial dependence could not be properly explained in a single neighbourhood structure framework. The main achievement of the paper involves the successful implementation of the second order SMA polynomial through the full likelihood optimization using the DE method; construction of confidence intervals for statistical inferences based on the derived standard error expression and bootstrapping; and, the construction of the simultaneous confidence regions for the two spatial dependence parameters based on LR test and bootstrapping. The 95 percent bootstrap based confidence intervals are wider as compared to the exact 95 percent confidence intervals.

In the present analysis, the incorporation of the second order spatial dependence parameters reduces the error variance by a small amount only, as compared to the first order SMA model. Hence, not so much change could be observed in the parameters confidence intervals of both models. Nevertheless, the efficiency of the model is reflected in its AIC values. The potential of the second order SMA polynomial to apprehend spatial dependence of both first and second lag orders, validates the precision of the inferences drawn about the data using the second order SMA polynomial.

Acknowledgement:

We would like to thank the R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. The URL address is https://www.R-project.org. We extend our sincere gratitude to Bivand et al. (2018), Bivand et al. (2021), and Mullen et al. (2011) for developing the useful R packages 'spdep', 'spData', and 'DEoptim', respectively.

Conflict of interest: There is no conflict of interest among the authors.

Appendix A

If $Z \sim 2^{nd} - order - SMA - polynomial\ (\beta, \sigma^2, \theta_1, \theta_2)$, then its characteristic function is,

$$\psi_Z(t) = E\left(e^{it^T Z}\right)$$

$$= \frac{e^{it^T Z} e^{-\frac{\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]^T\left[(I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\right]}{2\sigma^2}}}{\sigma(2\pi)^{n/2}|(I+\theta_1 W_1)(I+\theta_2 W_2)|} dZ$$

Letting, $u = (I+\theta_2 W_2)^{-1}(I+\theta_1 W_1)^{-1}(Z-X\beta)\sigma^{-1}$

$\Rightarrow Z = \sigma(I+\theta_1 W_1)(I+\theta_1 W_1) + X\beta$

And, $J = \frac{dZ}{du} = \sigma|(I+\theta_1 W_1)(I+\theta_2 W_2)|$

$$\therefore \psi_Z(t) = \int \frac{e^{it^T X\beta}}{(2\pi)^{n/2}} \times e^{it^T \sigma(I+\theta_1 W_1)(I+\theta_1 W_1)u - \frac{u^T u}{2}} du$$

$$= \frac{e^{it^T X\beta}}{(2\pi)^{n/2}} \int e^{\frac{[u^T - 2it^T \sigma(I+\theta_1 W_1)(I+\theta_2 W_2)]u}{2}} du$$

Again letting, $y^T = u^T - it^T\sigma(I+\theta_1 W_1)(I+\theta_2 W_2)$

$$\therefore u^T = y^T + it^T\sigma(I+\theta_1 W_1)(I+\theta_2 W_2)$$

And, $u = y + i\sigma(I+\theta_2 W_2)^T(I+\theta_1 W_1)^T t$

The differential of 'u' with respect to 'y' is 1.

$$\psi_Z(t) = \frac{e^{it^T X\beta}}{(2\pi)^{n/2}} \int e^{-\frac{1}{2}[y^T - it^T\sigma(I+\theta_1 W_1)(I+\theta_2 W_2)][y - i\sigma(I+\theta_1 W_1)(I+\theta_2 W_2)t]} dy$$

$$= \frac{e^{it^T X\beta}}{(2\pi)^{n/2}} \int e^{-\frac{y^T y}{2}} e^{\frac{i^2\sigma^2 t^T (I+\theta_1 W_1)(I+\theta_2 W_2)(I+\theta_2 W_2)^T (I+\theta_1 W_1)^T t}{2}} dy$$

$$= e^{it^T X\beta + \frac{i^2\sigma^2 t^T (I+\theta_1 W_1)(I+\theta_2 W_2)(I+\theta_2 W_2)^T (I+\theta_1 W_1)^T t}{2}} \int \frac{e^{-\frac{y^T y}{2}}}{(2\pi)^{n/2}} dy$$

$$= e^{it^T X\beta + \frac{i^2\sigma^2 t^T (I+\theta_1 W_1)(I+\theta_2 W_2)(I+\theta_2 W_2)^T (I+\theta_1 W_1)^T t}{2}} \times 1$$

$$= e^{it^T X\beta + \frac{i^2\sigma^2 t^T (I+\theta_1 W_1)(I+\theta_2 W_2)(I+\theta_2 W_2)^T (I+\theta_1 W_1)^T t}{2}}$$

Let, $\omega = t^T(I+\theta_1 W_1)(I+\theta_2 W_2)(I+\theta_2 W_2)^T(I+\theta_1 W_1)^T t$

$\therefore \psi_Z(t) = e^{it^T X\beta - \frac{t^T \omega t}{2}}$

Now,

$$\psi_Z'(t) = e^{it^T X\beta - \frac{t^T\omega t}{2}} iX\beta - e^{it^T X\beta - \frac{t^T\omega t}{2}} \omega t$$

$$\psi_Z''(t) = e^{it^T X\beta - \frac{t^T\omega t}{2}} iX\beta(iX\beta - \omega t) - \omega e^{it^T X\beta - \frac{t^T\omega t}{2}} - \omega t(iX\beta - \omega t) e^{it^T X\beta - \frac{t^T\omega t}{2}}$$

Therefore,

$$E(Z) = (-i)\psi_Z'(t)\big|_{t=0}$$

$$= (-i)iX\beta$$

$$= -(-1)X\beta$$

$$= X\beta$$

$$E(Z^2) = (-i)^2 \psi_Z''(t)\big|_{t=0}$$

$$= (-1)(i^2 X\beta\beta^T X^T - \omega)$$

$$= X\beta\beta^T X^T + \omega$$

Thus, $Cov\ (Z) = E(Z^2) - [E(Z)]^2$

$$= X\beta\beta^T X^T + \omega - X\beta\beta^T X^T$$

$$= \omega$$

$$= \sigma^2(I + \theta_1 W_1)(I + \theta_2 W_2)(I + \theta_2 W_2)^T(I + \theta_1 W_1)^T$$

Refrences

[1] Anselin, L. Bera, A.K. (1998).Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics: Textbooks and Monographs*. **155, 237-289.**

[2] Anselin, L. (1988). *Spatial econometrics: methods and models*. Springer Science & Business Media.

[3] Besag, J.(1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series (Methodological)*. **36, 192-236.**

[4] Bivand, R. S. Wong, D. W.S. (2018). Comparing implementations of global and local indicators of spatial association test. *TEST* . **27(3) , 716-748**. https://doi.org/10.1007/s11749-018-0599-x.

[5] Bivand R, Nowosad J, Lovelace R. (2021) spData:Datasets for Spatial Analysis. R package version 0.3.10. 2021. https://CRAN.R-project.org/package=spData.

[6] Cressie, N. (1993). *Statistics for spatial data*. New York: John Wiley and sons.

[7] Hainning, R.P. (1978). The moving average model for spatial interaction. *Transactions of the institute of British geographers.* **3(2), 202-225.**

[8] Hepple, L.W. (2003). Bayesian and maximum likelihood of the linear model with spatial moving average disturbances. http://www.ggy.bris.ac.uk/personal/LesHepple/bayesian.pdf.

[9] Jha, S.K. Begum, R. (2023). A proposed weighting scheme for spatial moving average model in irregular lattice. *Journal of the Indian Society for Probability and Statistics*. **24, 357-375**.

[10] Jha, S.K. Singh, N.V. (2021). A skew-normal spatial simultaneous autoregressive model and its implementation. *Sankhya A.* **85(1), 306-323.** https://doi.org/10.1007/s13171-021-00246-3.

[11] Lesage and Pace. (2009). *Introduction to spatial econometrics.* Statistics: Textbooks and Monographs. https://www.researchgate.net/publication/339938974.

[12] Mur, J. Angulo, A.M. (2006). Clues for discriminating between moving average and autoregressive process in spatial processes. *Verlag Springer*. **9, 273-298**.

[13] Mullen, K. Ardia, D. Gil, D. Windover, D and Cline, J. (2011). 'DEoptim': An R Package for Global Optimization by Differential Evolution. *Journal of Statistical Software.* **40(6), 1-26.** https://doi.org/10.18637/jss.v040.i06.

[14] Storn, R. Price, K. (1997). Differential Evolution- A simple and efficient Heuristic for global optimization over continuous spaces. *Journal of Global optimization*. **11, 341-359.**

[15] Whittle P. (1954). On stationary process in the plane. *Biometrika*. **41, 434-449**.

[16] Wall, M.M. (2004). A close look at the spatial structure implied by the CAR and SAR models. *Journal of Statistical Planning and Inference.* **121, 311-324**.