Preprints.org

Statistical Inference for Interval-Valued Spatial Error Models

Li Guan , Hai Zhou , Wei Zhang *

Article

Statistical Inference for Interval-Valued Spatial Error Models

Li Guan, Haixin Zhou and Wei Zhang *

Faculty of Science, Beijing University of Technology, 100 Ping Le Yuan, Chao Yang District, Beijing 100124, China; guanli@bjut.edu.cn, 1564335zhx@sina.com

* Correspondence: weizhang0313@bjut.edu.cn

Abstract: In this paper, we shall introduce the interval-valued spatial error model. Based on the idea of least square method of single-valued case, we give the parameter estimator for interval-valued spatial error model. The theoretical properties of the proposed estimator are proved. Finally, we give the numerical analysis and a real example.

Keywords: interval-valued random variable; spatial error model; parameter estimator

1. Introduction

It is well known that classical linear regression model and time series models are most widely used in statistical inference, including medical treatment, education, finance, science, technology and many other fields. Most of the cases, these models are used for single-valued random variables. In the real world, there are plenty of random phenomenas cannot be characterized by single-valued random variables. Taking the price of a stock on a given day for example, it is clearly unreasonable to use a single-valued data to decribe the stock price. If only single-valued data such as stock closing price or opening price are used, the fluctuation information in the process of stock trading is ignored and the resulting analysis results provided to decision-makers are also one-sided. Moreover, people will pay more attention to the data in a certain range, such as the temperature for a given day, instead of knowing the temperature at a certain time of one day, people care more about the maximum and minimum temperature of one day. In economic forecasting, economists mostly give a prediction range of economic growth rate. In the process of medical impact diagnosis, the impact result is usually a two-dimensional plan, and it is not a single value. Therefor, the interval-valued data are more appropriate and valuable in these cases because they provide more information. Thus it is necessary to consider the interval-valued statistical models and statictical inference problems.

Interval-valued random variables are special set-valued random variables. In mid twentith century, Aumann and Debreu first used set-valued mapping when studying economic phenomena. Aumann [1] gave the integral of set-valued random variables in 1965. Hiai and Umegaki [6] gave the concept of conditional expectation of set-valued random variables in 1978. Lyashenko [11,12] discussed the properties of set-valued random variables in Euclidean space, introduced the definition of set-valued Gaussian random variables, and gave the definition of variance for set-valued random variables. Vitale [16] studied the properties of D_p distance. In 2005, Xuhua Yang and Shoumei Li [17] gave the definitions of variance and covariance for set-valued random variables under the D_p distance, and obtained excellent properties. In 2008, Blanco et. al. defined the variance and studied the properties of interval valued random variables under a new distance in [2]. Hess [5], Papageorgiou [14,15], Shoumei Li et. al. [7–9] explored the convergence theory of set-valued random variables under different conditions. Molchanov in [13], Shoumei Li et. al. in [10] systematically summarized the theory of set-valued random variables. The above research promoted the development of set-valued random variable theory.

For interval-valued statistical models, Billard and Diday [3] established a linear regression model by using the midpoint of interval-valued random variables in 2000. In 2002, Billard and Diday [4]

established linear regression models by using the two endpoints of interval-valued random variables respectively. In 2008, Lima Neto and de Carvalho [19] established linear regression models by using the center and radius of interval-valued random variables. In 2010, Lima Neto and de Carvalho [20] imposed non-negative constraints on the regression coefficients of radius on the basis of [19]. Wang [22] in 2012 proposed the complete information method to deal with the interval-valued linear regression model. Souza [23] introduced the parametrization method to linear regression model in 2017. In 2015, Wang Xun et. al. in [24] used set-valued theory to study linear regression problems, and gave the least square estimaor and the related properties. All the above research works are about the linear regression models of interval-valued random variables. The research on interval-valued spatial regression models and spatial error models are still blank.

As for the single-valued spatial error model, Anselin gave the maximum likelihood estimation method in [26] in 1988. Prucha proposed the generalized moment estimation method in [27] in 1999. In 2020, Yildirim [25] systematically summarized the methods of parameter estimation of spatial error model and proposed a new parameter estimation method based on likelihood equation. Many scholars have studied the classical linear regression and time series models of interval-valued random variables and achieved wonderful research results. We are considering the interval-valued spatial error models.

This paper attempts to extend the classical spatial error model to interval-valued case. The orginazation of this paper is as follows: in Section 2, we mainly introduce the notations and basic concepts of interval-valued random theory. In Section 3, we mainly discuss the interval-valued spatial error model, give the least square estimator of parameter obtain a series of digital characteristics and consistency of parameter estimation; In Section 4, the effectiveness of the method is illustrated by numerical simulation; In Section 5, gives an application of the model by studying the relationship between temperature and latitude in major cities of China.

2. Preliminaries on interval-valued random variables

2.1 d_p distance and D_p distance

Throughout this paper, we assume that $(\Omega, \mathcal{A}, \mu)$ is a complete probability space. $\mathbb{R}^d$ is a d-dimensional Euclidean space, $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ are the norm and inner product in $\mathbb{R}^d$ respectively, and the family of compact convex subsets in $\mathbb{R}^d$ is $\mathbf{K}_{kc}(\mathbb{R}^d)$. When $d = 1$, $\mathbb{R}^1$ is abbreviated as $\mathbb{R}$, then $\mathbf{K}_{kc}(\mathbb{R})$ is a family of nonempty bounded closed intervals in $\mathbb{R}$, that is

$$\mathbf{K}_{kc}(\mathbb{R}) = \{ A = [\underline{a}, \overline{a}] : -\infty < \underline{a} \leq \overline{a} < \infty, \underline{a}, \overline{a} \in \mathbb{R} \},$$

Where, $\underline{a}$ and $\overline{a}$ are the left and right endpoints of interval A respectively. In addition, interval A is also denoted as center radius form $A = (c_A; r_A)$, where $c_A = (\overline{a} + \underline{a})/2$ and $r_A = (\overline{a} - \underline{a})/2$ are the center and radius of interval A respectively. For any set A, B, the addition and multiplication operations are defined as:

$$A + B = \{a + b : a \in A, b \in B\},$$

$$kA = \{ka : a \in A\}, \forall k \in \mathbb{R}.$$

Interval is a special case of set, for $A = [\underline{a}, \overline{a}] = (c_1; r_1)$, $B = [\underline{b}, \overline{b}] = (c_2; r_2)$, the addition and multiplication operations are defined as:

$$A + B = [\underline{a} + \underline{b}, \overline{a} + \overline{b}] = (c_1 + c_2; r_1 + r_2),$$

$$kA = \begin{cases} [k\underline{a}, k\overline{a}], & k \geq 0 \\ [k\overline{a}, k\underline{a}], & k < 0 \end{cases} = (kc_1; |k|r_1).$$

Note that if set A does not degenerate to a point, $A - A = A + (-A) \neq \{0\}$. Then $\mathbf{K}_{kc}(\mathbb{R}^d)$ is not a linear space with respect to addition and multiplication.

For any set A, B in $\mathbf{K}_{kc}(\mathbb{R}^d)$, the subtraction operation is defined as: $A - B = \{a - b : a \in A, b \in B\}$. As a special case of set value, for interval $A = [\underline{a}, \overline{a}] = (c_1; r_1), B = [\underline{b}, \overline{b}] = (c_2; r_2)$, the definition of subtraction operation is derived as follows:

$$A - B = [\underline{a} - \underline{b}, \overline{a} - \overline{b}] = (c_1 - c_2; r_1 + r_2).$$

The support function of set $A \in \mathbf{K}_{kc}(\mathbb{R}^d)$ is defined as

$$s(x, A) = \sup_{a \in A} \langle x, a \rangle, \ x \in \mathbb{R}^d.$$

The d_p distance is defined as follows: for any $1 \leq p < \infty$, the d_p distance between set A and B is

$$d_p(A, B) = \Big[\int_{S^{d-1}} |s(x, A) - s(x, B)|^p \mathrm{d}\mu(x) \Big]^{\frac{1}{p}}, \ 1 \leq p < \infty,$$

where, S^{d-1} is the unit sphere of $\mathbb{R}^d$, μ is a measure on S^{d-1}, in particular, take $\mu(1) = \mu(-1) = 1$ on S^0. Further, from Yang and Li [17], $(\mathbf{K}_{kc}(\mathbb{R}^d), d_p)$ is a complete separable space. Specially, for interval $A = [\underline{a}, \overline{a}] = (c_A; r_A)$ and $B = [\underline{b}, \overline{b}] = (c_B; r_B)$, the d_p distance is

$$d_p(A, B) = \left(|\underline{b} - \underline{a}|^p + |\overline{b} - \overline{a}|^p \right)^{\frac{1}{p}}$$
$$= [((c_B - c_A) - (r_B - r_A))^p + ((c_B - c_A) + (r_B - r_A))^p]^{\frac{1}{p}}.$$

In particular, if $p = 2$,

$$d_2(A, B) = [(\underline{b} - \underline{a})^2 + (\overline{b} - \overline{a})^2]^{\frac{1}{2}}$$
$$= \left[((c_B - c_A) - (r_B - r_A))^2 + ((c_B - c_A) + (r_B - r_A))^2 \right]^{\frac{1}{2}}$$
$$= [2(c_B - c_A)^2 + (r_B - r_A)^2]^{\frac{1}{2}}.$$

Call set-valued mapping $F : \Omega \to \mathbf{K}_{kc}(\mathbb{R}^d)$ be a set-valued random variable, if for any closed sets $C \in \mathbf{K}_{kc}(\mathbb{R}^d)$,

$$F^{-1}(C) = \{\omega \in \Omega : F(\omega) \cap C \neq \varnothing\} \in \mathcal{A}.$$

Let $\mathcal{U}[\Omega, \mathbf{K}_{kc}(\mathbb{R}^d)]$ denote the family of set-valued random variables in $\mathbf{K}_{kc}(\mathbb{R}^d)$. The expression of D_p distance between set-valued random variables F_1 and F_2 is

$$D_p(F_1, F_2) = [\mathbb{E}d_p^p(F_1, F_2)]^{\frac{1}{p}}.$$

Similarly, for interval-valued random variables, the D_p distance between interval value random variables $F_1 = [\underline{f}_1, \overline{f}_1] = (c_{F_1}; r_{F_1})$ and $F_2 = [\underline{f}_2, \overline{f}_2] = (c_{F_2}; r_{F_2})$ is

$$D_p(F_1, F_2) = \left[\mathbb{E}(\underline{f}_2 - \underline{f}_1)^p + \mathbb{E}(\overline{f}_2 - \overline{f}_1)^p \right]^{\frac{1}{p}}$$
$$= \left[\mathbb{E}((c_{F_2} - c_{F_1}) - (r_{F_2} - r_{F_1}))^p + \mathbb{E}((c_{F_2} - c_{F_1}) + (r_{F_2} - r_{F_1}))^p \right]^{\frac{1}{p}}.$$

Further, from Yang and Li [17], $(\mathbf{K}_{kc}(\mathbb{R}^d), D_p)$ is a complete separable distance space. In particular, if $p = 2$,

$$
\begin{aligned}
D_2(F_1, F_2) &= \left[\mathrm{E}(\underline{f}_2 - \underline{f}_1)^2 + \mathrm{E}(\overline{f}_2 - \overline{f}_1)^2 \right]^{\frac{1}{2}} \\
&= \left[\mathrm{E}((c_{F_2} - c_{F_1}) - (r_{F_2} - r_{F_1}))^2 + \mathrm{E}((c_{F_2} - c_{F1}) + (r_{F_2} - r_{F_1}))^2 \right]^{\frac{1}{2}}.
\end{aligned}
$$

2.2 Moment of set-valued random variables

The expectation of set-valued random variable $F \in \mathcal{U}[\Omega, \mathbf{K}_{kc}(\mathbb{R}^d)]$ is given by Aumann in [1],

$$
\mathrm{E}[F] = \int_\Omega F \mathrm{d}^- = \left\{ \int_\Omega f \mathrm{d}^- : f \in S_F \right\},
$$

where S_F is the integrable selection set of F, that is,

$$
S_F = \{ f \in L^p[\Omega, \mathbb{R}^d] : f(\omega) \in F(\omega) \ a.e.(\mu) \}.
$$

Yang and Li in [17] introduced the variance and covariance of set-valued random variables based on D_p distance. For set-valued random variable $F \in \mathcal{U}[\Omega, \mathbf{K}_{kc}(\mathbb{R}^d)]$, the variance is defined as follows:

$$
\begin{aligned}
\mathrm{Var}(F) &= D_2^2(F, \mathrm{E}[F]) \\
&= \mathrm{E}\left[d_2^2(F, \mathrm{E}[F]) \right] \\
&= \mathrm{E}\left[\int_{S^{d-1}} (s(x, F) - s(x, \mathrm{E}[F]))^2 \mathrm{d}\mu(x) \right].
\end{aligned}
$$

For two set-valued random variables $F_1, F_2 \in \mathcal{U}[\Omega, \mathbf{K}_{kc}(\mathbb{R}^d)]$, the covariance is defined as follows

$$
\mathrm{Cov}(F_1, F_2) = \mathrm{E}\left[\int_{S^{d-1}} (s(x, F_1) - s(x, \mathrm{E}[F_1]))(s(x, F_2) - s(x, \mathrm{E}[F_2])) \mathrm{d}\mu(x) \right].
$$

If $F = (c_F; r_F)$ is an interval-valued random variable, then

$$
\begin{aligned}
\mathrm{Var}(F) =& \mathrm{E}[\underline{f} - \mathrm{E}[\underline{f}]]^2 + \mathrm{E}[\overline{f} - \mathrm{E}[\overline{f}]]^2 \\
=& \mathrm{E}\left[(c_F - \mathrm{E}[c_F]) - (r_F - \mathrm{E}[r_F]) \right]^2 + \mathrm{E}\left[(c_F - \mathrm{E}[c_F]) + (r_F - \mathrm{E}[r_F]) \right]^2.
\end{aligned}
$$

The covariance of interval-valued random variables $F_1, F_2 \in \mathcal{U}[\Omega, \mathbf{K}_{kc}(\mathbb{R})]$ is

$$
\begin{aligned}
\mathrm{Cov}(F_1, F_2) =& \mathrm{E}\left[(\underline{f}_1 - \mathrm{E}[\underline{f}_1])(\underline{f}_2 - \mathrm{E}[\underline{f}_2]) \right] + \mathrm{E}\left[(\overline{f}_1 - \mathrm{E}[\overline{f}_1])(\overline{f}_2 - \mathrm{E}[\overline{f}_2]) \right] \\
=& \mathrm{E}\left[(c_{F_1} - \mathrm{E}[c_{F_1}] - (r_{F_1} - \mathrm{E}[r_{F_1}]))(c_{F_2} - \mathrm{E}[c_{F_2}] - (r_{F_2} - \mathrm{E}[r_{F_2}])) \right] \\
&+ \mathrm{E}\left[(c_{F_1} - \mathrm{E}[c_{F_1}] + (r_{F1} - \mathrm{E}[r_{F_1}]))(c_{F_2} - \mathrm{E}[c_{F_2}] + (r_{F_2} - \mathrm{E}[r_{F_2}])) \right].
\end{aligned}
$$

Through calculation, we can easily have

$$
\begin{aligned}
\mathrm{Var}(F) &= 2\mathrm{E}[c_F - \mathrm{E}[c_F]]^2 + 2\mathrm{E}[r_F - \mathrm{E}[r_F]]^2 \\
&= 2\mathrm{Var}(c_F) + 2\mathrm{Var}(r_F),
\end{aligned}
$$

$$\mathrm{Cov}(F_1, F_2) = 2\mathrm{E}\left[(c_{F_1} - \mathrm{E}[c_{F_1}])(c_{F_2} - \mathrm{E}[c_{F_2}])\right] + 2\mathrm{E}\left[(r_{F_1} - \mathrm{E}[r_{F_1}])(r_{F_2} - \mathrm{E}[r_{F_2}])\right]$$
$$= 2\mathrm{Cov}(c_{F_1}, c_{F_2}) + 2\mathrm{Cov}(r_{F_1}, r_{F_2}).$$

The variance and covariance of interval-valued random variables will be used in Section 3. For more information about the variance and covariance of set-valued random variables, readers can refer to [17].

3. Interval-valued spatial error model

$$Y = X\beta + u, \quad u = \lambda w u + \varepsilon, \qquad |\lambda| < 1. \tag{3.1}$$

Consider the classical spatial error model with the following form, where X is the explanatory variable, Y is the explained variable, β is the unknown parameter, error term u, and ε are single point values, and W is a known $n \times n$ space weight matrix, λ is a spatial autoregressive coefficient parameter,
the error item $\varepsilon \sim N(0, \sigma^2 I_n)$, I_n is an identity matrix. By transforming, model (3.1) becomes,

$$(I_n - \lambda W)Y = (I_n - \lambda W)X\beta + \varepsilon, \tag{3.2}$$

denoted by

$$Y_\lambda = (I_n - \lambda W)Y,$$
$$X_\lambda = (I_n - \lambda W)X,$$

Model (3.2) can be expressed as

$$Y_\lambda = X_\lambda \beta + \varepsilon,$$
$$E(Y_\lambda) = X_\lambda \beta. \tag{3.3}$$

Now we extend the above classical single-valued model to interval-valued case.

Definition 3.1 If $Y_\lambda = (Y_{\lambda 1}, Y_{\lambda 2}, \cdots, Y_{\lambda n})^T$ is the n-dimensional vector of interval-valued observations, $X_\lambda = (x_{\lambda ij})_{n \times p}$ is the $n \times p$ single point valued design matrix, $\beta = (\beta_1, \beta_2, \cdots, \beta_p)$ is a p-dimensional interval-valued parameter vector, then model (3.3), is called interval-valued space error model.

Next, we give the algorithm for multiplication of the matrix and interval values.

Definition 3.2 Let $A_i = [\underline{a_i}, \overline{a_i}] = (c_i; r_i), i = 1, \cdots, p$ be the interval in $\mathbf{K}_{kc}(\mathbb{R})$, the interval value vector $A = (A_1, A_2, \cdots, A_p)^T$ is multiplied by any $n \times p$ dimensional matrix $(m_{ij})_{n \times p}, i = 1, 2, \cdots, n; j = 1, 2, \cdots, p$, the algorithm is defined as follows:

$$(m_{ij})_{n \times p} A = \begin{pmatrix} m_{11} A_1 + \cdots + m_{1p} A_p \\ \vdots \\ m_{n1} A_1 + \cdots + m_{np} A_p \end{pmatrix}$$
$$= \begin{pmatrix} m_{11}(c_1; r_1) + \cdots + m_{1p}(c_p; r_p) \\ \vdots \\ m_{n1}(c_1; r_1) + \cdots + m_{np}(c_p; r_p) \end{pmatrix}$$
$$= \begin{pmatrix} m_{11}[\underline{a_1}, \overline{a_1}] + \cdots + m_{1p}[\underline{a_p}, \overline{a_p}] \\ \vdots \\ m_{n1}[\underline{a_1}, \overline{a_1}] + \cdots + m_{np}[\underline{a_p}, \overline{a_p}] \end{pmatrix}.$$

For the general single-valued linear model, the idea of the least squares estimation method is to minimize the sum of the squares of the residuals. We shall use the same mathematical idea here. For

interval-valued spatial error model, the least square estimation of interval-valued unknown parameter β is to minimize $d_2^2(Y_\lambda, X_\lambda \beta)$ under the definition of d_2 distance

$$
d_2^2(Y_\lambda, x_\lambda \beta) = \sum_{i=1}^n d_2^2 \left(Y_{\lambda i}, x_{\lambda i1}\beta_1 + x_{\lambda i2}\beta_2 + \cdots + x_{\lambda ip}\beta_p\right)
$$

$$
= \sum_{i=1}^n \left[\left(c_{Y_{\lambda i}} - x_{\lambda i1}c_{\beta_1} - \cdots - x_{\lambda ip}c_{\beta_p}\right) \right.
$$

$$
\left. - \left(r_{Y_{\lambda i}} - |x_{\lambda i1}|r_{\beta_1} - \cdots - |x_{\lambda ip}|r_{\beta_p}\right) \right]^2
$$

$$
+ \sum_{i=1}^n \left[\left(c_{Y_{\lambda i}} - x_{\lambda i1}c_{\beta_1} - \cdots - x_{\lambda ip}c_{\beta_p}\right) \right.
$$

$$
\left. + \left(r_{Y_{\lambda i}} - |x_{\lambda i1}|r_{\beta_1} - \cdots - |x_{\lambda ip}|r_{\beta_p}\right) \right]^2
$$

$$
= 2 \sum_{i=1}^n \left[\left(c_{Y_{\lambda i}} - x_{\lambda i1}c_{\beta_1} - \cdots - x_{\lambda ip}c_{\beta_p}\right)^2 \right.
$$

$$
\left. + \left(r_{Y_{\lambda i}} - |x_{\lambda i1}|r_{\beta_1} - \cdots - |x_{\lambda ip}|r_{\beta_p}\right)^2 \right],
$$

where c_m and r_m represent the center and radius of interval value m respectively. The above formula is the quadratic function of c_{β_j} and r_{β_j}, and $d_2^2(Y_\lambda, X_\lambda \beta) \geq 0$, so there is a minimum value.

Next, calculate the partial derivatives of c_{β_j} and r_{β_j} respectively

$$
\begin{cases}
\frac{\partial d_2^2(Y_\lambda, X_\lambda \beta)}{\partial c_{\beta_j}} = 0 \\
\frac{\partial d_2^2(Y_\lambda, X_\lambda \beta)}{\partial r_{\beta_j}} = 0
\end{cases}
, j = 1, 2, \cdots, p.
$$

that is,

$$
\begin{cases}
\sum_{i=1}^n \left(c_{Y_{\lambda i}} - x_{\lambda i1}c_{\beta_1} - \cdots - x_{\lambda ip}c_{\beta_p}\right)(-x_{\lambda ij}) = 0 \\
\sum_{i=1}^n \left(r_{Y_{\lambda i}} - |x_{\lambda i1}|r_{\beta_1} - \cdots - |x_{\lambda ip}|r_{\beta_p}\right)(-|x_{\lambda ij}|) = 0.
\end{cases}
$$

The regular equation is:

$$
\begin{cases}
X_\lambda^T c_{Y_\lambda} = X_\lambda^T X_\lambda c_\beta \\
|X_\lambda|^T r_{Y_\lambda} = |X_\lambda|^T |X_\lambda| r_\beta,
\end{cases}
$$

where $|X_\lambda| = \left(|x_{ij}|\right)_{n \times p}$. The parameter estimation of the interval-valued spatial error model can be obtained by solving the regular equation. The following is the result about the rank of X_λ.

Lemma 3.3 If $rk(X) = p$, then $rk(X_\lambda) = p$.

Proof Since

$$
rk(X_\lambda) = rk((I_n - \lambda w)X)
$$
$$
\leq \min(rk(I_n - \lambda w), rk(X))
$$
$$
\leq rk(X)
$$
$$
= p
$$

and

$$
rk((I_n - \lambda w)X) \geq rk(I_n - \lambda w) + rk(X) - n = p,
$$

it has

$$rk((I_n - \lambda w)X) = rk(X_\lambda) = rk(X) = p.$$

The result is proved. □

Based on Lemma 3.3, suppose $rk(|X_\lambda|) = p$, then the estimator of interval-valued spatial error model can be obtained by solving the regular equation, which is shown in the following theorem.

Theorem 3.3 Under the condition of Lemma 3.3, the least squares estimation of interval-valued spatial error model is unique, which is denoted as

$$\widehat{\beta}_{LS}(\lambda) = ((X_\lambda^T X_\lambda)^{-1} X_\lambda^T c_{Y_\lambda}; (|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T r_{Y_\lambda})$$
$$= \Big((X^T (I_n - \lambda W)^T (I_n - \lambda W) X)^{-1} X^T (I_n - \lambda W)^T (I_n - \lambda W) c_Y;$$
$$(|X|^T |I_n - \lambda W|^T |I_n - \lambda W||X|)^{-1} |X|^T |I_n - \lambda W|^T |I_n - \lambda W| r_Y \Big).$$

After obtaining the estimation form of unknown parameter β, we then discuss the properties. First, consider the unbiasedness of $\widehat{\beta}_{LS}(\lambda)$.

Theorem 3.4 The least squares estimate $\widehat{\beta}_{LS}(\lambda)$ is an unbiased estimate of β.

Proof By Theorem 3.3,

$$E(\widehat{\beta}_{LS}(\lambda) = ((X_\lambda^T X_\lambda)^{-1} X_\lambda^T E[c_{Y_\lambda}]; (|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T E[r_{Y_\lambda}])$$
$$= E((X^T (I_n - \lambda W)^T (I_n - \lambda W) X)^{-1} X^T (I_n - \lambda W)^T (I_n - \lambda W) c_Y;$$
$$(|X|^T |I_n - \lambda W|^T |I_n - \lambda W||X|)^{-1} |X|^T |I_n - \lambda W|^T |I_n - \lambda W| r_Y)$$
$$= ((X^T (I_n - \lambda W)^T (I_n - \lambda W) X)^{-1} X^T (I_n - \lambda W)^T (I_n - \lambda W) E(c_Y);$$
$$(|X|^T |I_n - \lambda W|^T |I_n - \lambda W||X|)^{-1} |X|^T |I_n - \lambda W|^T |I_n - \lambda W| E(r_Y))$$
$$= ((X^T (I_n - \lambda W)^T (I_n - \lambda W) X)^{-1} X^T (I_n - \lambda W)^T (I_n - \lambda W) X c_\beta;$$
$$(|X|^T |I_n - \lambda W|^T |I_n - \lambda W||X|)^{-1} |X|^T |I_n - \lambda W|^T |I_n - \lambda W||X| r_\beta)$$
$$= (c_\beta; r_\beta) = \beta.$$

The result is proved. □

For the interval-valued spatial error model, when $rk(X_\lambda) = rk(|X_\lambda|) = p$, the covariance of $\widehat{\beta}_{LS}(\lambda)$ can be obtained, as shown in the following result.

Theorem 3.5 If $rk(X_\lambda) = rk(|X_\lambda|) = p, E(Y_\lambda) = X_\lambda \beta, Cov(c_{Y_\lambda}) = c_{\sigma^2} I_n, Cov(r_{Y_\lambda}) = r_{\sigma^2} I_n$, then the covariance matrix of $\widehat{\beta}_{LS}(\lambda)$ is

(1) $i \neq j$,

$$Cov\left(\widehat{\beta}_{LS}^{(i)}(\lambda), \widehat{\beta}_{LS}^{(j)}(\lambda) \right) = 2c_{\sigma^2} \left((X_\lambda^T X_\lambda)^{-1} X_\lambda^T \right)_{(i)} \left(X_\lambda (X_\lambda^T X_\lambda)^{-1} \right)_{(j)}$$
$$+ 2r_{\sigma^2} \left((|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T \right)_{(i)} \left(|X_\lambda|(|X_\lambda|^T |X_\lambda|)^{-1} \right)_{(j)},$$

(2) $i = j$,

$$Cov\left(\widehat{\beta}_{LS}^{(i)}(\lambda), \widehat{\beta}_{LS}^{(j)}(\lambda) \right) = 2c_{\sigma^2} \left(X^T (I_n - \lambda W)^T (I_n - \lambda W) X \right)^{-1}$$
$$+ 2r_{\sigma^2} \left(|X|^T |I_n - \lambda W|^T |I_n - \lambda W||X| \right)^{-1},$$

where $\widehat{\beta}_{LS}^{(i)}(\lambda), \widehat{\beta}_{LS}^{(j)}(\lambda)$ represent the ith and jth element of $\widehat{\beta}_{LS}(\lambda)$ respectively, and $A_{(i)}, A_{(j)}$ represent the ith, jth rows of matrix A respectively.

Proof For the ith and jth element of $\widehat{\beta}_{LS}(\lambda)$, if $i \neq j$, it has

$$
\begin{aligned}
Cov\left(\widehat{\beta}_{LS}^{(i)}(\lambda), \widehat{\beta}_{LS}^{(j)}(\lambda)\right) &= Cov\Bigg\{ \left(((X_\lambda^T X_\lambda)^{-1} X_\lambda^T)_{(i)} c_{Y_\lambda}; ((|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T)_{(i)} r_{Y_\lambda}\right), \\
&\qquad\qquad \left(((X_\lambda^T X_\lambda)^{-1} X_\lambda^T)_{(j)} c_{Y_\lambda}; ((|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T)_{(j)} r_{Y_\lambda}\right) \Bigg\} \\
&= 2Cov\left(((X_\lambda^T X_\lambda)^{-1} X_\lambda^T)_{(i)} c_{Y_\lambda}, ((X_\lambda^T X_\lambda)^{-1} X_\lambda^T)_{(j)} c_{Y_\lambda}\right) \\
&\quad + 2Cov\left(((X_\lambda^T X_\lambda)^{-1} X_\lambda^T)_{(i)} r_{Y_\lambda}, ((X_\lambda^T X_\lambda)^{-1} X_\lambda^T)_{(j)} r_{Y_\lambda}\right) \\
&= 2\left((X_\lambda^T X_\lambda)^{-1} X_\lambda^T\right)_{(i)} Cov(c_{Y_\lambda}) \left((X_\lambda^T X_\lambda)^{-1} X_\lambda^T\right)_{(j)}^T \\
&\quad + 2\left((|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T\right)_{(i)} Cov(c_{Y_\lambda}) \left((|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T\right)_{(j)}^T \\
&= 2c_{\sigma^2}\left((X_\lambda^T X_\lambda)^{-1} X_\lambda^T\right)_{(i)} \left(X_\lambda (X_\lambda^T X_\lambda)^{-1}\right)_{(j)} \\
&\quad + 2r_{\sigma^2}\left((|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T\right)_{(i)} \left(|X_\lambda| (|X_\lambda|^T |X_\lambda|)^{-1}\right)_{(j)}.
\end{aligned}
$$

When $i = j$, it has

$$
\begin{aligned}
Cov\left(\widehat{\beta}_{LS}^{(i)}(\lambda), \widehat{\beta}_{LS}^{(i)}(\lambda)\right) &= 2c_{\sigma^2}\left((X_\lambda^T X_\lambda)^{-1} + 2r_{\sigma^2}(|X_\lambda|^T |X_\lambda|)^{-1}\right) \\
&= 2c_{\sigma^2}\left(X^T (I_n - \lambda W)^T (I_n - \lambda W) X\right)^{-1} \\
&\quad + 2r_{\sigma^2}\left(|X|^T |I_n - \lambda W|^T |I_n - \lambda W| |X|\right)^{-1}.
\end{aligned}
$$

The result is proved. $\square$

Next we discuss the estimation of error ε and error variance. We mainly consider the expectation and covariance of interval-valued error estimation.

Theorem 3.6 The error estimator $\widehat{\varepsilon}$ can be obtained from $Y_\lambda - X_\lambda \widehat{\beta}_{LS}$, and its expectation and variance are as follows:

(1) $E(\widehat{\varepsilon}) = 0$,

(2) $Cov(\widehat{\varepsilon}) = 2c_{\sigma^2}(I_n - P_{X_\lambda}) + 2r_{\sigma^2}(I_n + P_{|x_\lambda|})$,

where $P_{X_\lambda} = X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T, P_{|X_\lambda|} = |X_\lambda|(|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T$.

Proof

(1) Since

$$
\begin{aligned}
\widehat{\varepsilon} &= Y_\lambda - X_\lambda \widehat{\beta}_{LS}(\lambda) \\
&= (c_{Y_\lambda}; r_{Y_\lambda}) - X_\lambda \left((X_\lambda^T X_\lambda)^{-1} X_\lambda^T c_{Y_\lambda}; (|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T r_{Y_\lambda}\right) \\
&= (c_{Y_\lambda}; r_{Y_\lambda}) - \left(X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T c_{Y_\lambda}; |X_\lambda|(|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T r_{Y_\lambda}\right) \\
&= (c_{Y_\lambda}; r_{Y_\lambda}) - (c_{Y_\lambda}; r_{Y_\lambda}) \\
&= (0; 2r_{Y_\lambda}),
\end{aligned}
$$

it has

$$E(\widehat{\varepsilon}) = E(Y_\lambda - X_\lambda \widehat{\beta}_{LS}(\lambda))$$

$$= \left((X_\lambda c_\beta; |X_\lambda| r_\beta) - (X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T X_\lambda c_\beta; |X_\lambda| (|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T |X_\lambda| r_\beta) \right)$$

$$= (X_\lambda c_\beta; |X_\lambda| r_\beta) - (X_\lambda c_\beta; |X_\lambda| r_\beta)$$

$$= (0; 2|X_\lambda| r_\beta).$$

(2) On the other hand,

$$\widehat{\varepsilon} = Y_\lambda - X_\lambda \widehat{\beta}_{LS}(\lambda)$$

$$= (c_{Y_\lambda}; r_{Y_\lambda}) - X_\lambda \left((X_\lambda^T X_\lambda)^{-1} X_\lambda^T c_{Y_\lambda}; (|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T r_{Y_\lambda} \right)$$

$$= \left((I_n - X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T) c_{Y_\lambda}; (I_n + |X_\lambda| (|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T) r_{Y_\lambda} \right)$$

$$= \left((I_n - P_{X_\lambda}) c_{Y_\lambda}; (I_n + P_{|X_\lambda|}) r_{Y_\lambda} \right).$$

Then the ith element of $\widehat{\varepsilon}$ is

$$\left((I_n - P_{X_\lambda})_{(i)} c_{Y_\lambda}; (I_n + P_{|X_\lambda|})_{(i)} r_{Y_\lambda} \right).$$

Thus when $i \neq j$,

$$Cov(\widehat{\varepsilon}_i, \widehat{\varepsilon}_j) = Cov\Bigg\{ \left((I_n - P_{X_\lambda})_{(i)} c_{Y_\lambda}; (I_n + P_{|X_\lambda|})_{(i)} r_{Y_\lambda} \right),$$

$$\left((I_n - P_{X_\lambda})_{(j)} c_{Y_\lambda}; (I_n + P_{|X_\lambda|})_{(j)} r_{Y_\lambda} \right) \Bigg\}$$

$$= 2Cov\left((I_n - P_{X_\lambda})_{(i)} c_{Y_\lambda}, (I_n - P_{X_\lambda})_{(j)} c_{Y_\lambda} \right) +$$

$$2Cov\left((I_n + P_{|X_\lambda|})_{(i)} r_{Y_\lambda}, (I_n + P_{|X_\lambda|})_{(j)} r_{Y_\lambda} \right)$$

$$= 2(I_n - P_{X_\lambda})_{(i)} Cov(c_{Y_\lambda})(I_n - P_{X_\lambda})_{(j)} +$$

$$2(I_n - P_{|X_\lambda|})_{(i)} Cov(r_{Y_\lambda})(I_n + P_{|X_\lambda|})_{(j)},$$

where $A_{(i)}, A_{(j)}$ respectively represent the ith and jth rows of matrix A. When $i = j$,

$$Cov(\widehat{\varepsilon}) = 2(I_n - P_{X_\lambda}) Cov(c_{Y_\lambda})(I_n - P_{X_\lambda}) +$$

$$2(I_n - P_{|X_\lambda|}) Cov(r_{Y_\lambda})(I_n + P_{|X_\lambda|})$$

$$= 2c_{\sigma^2}(I_n - P_{X_\lambda}) + 2r_{\sigma^2}(I_n + P_{|X_\lambda|}).$$

The result is proved. $\square$

Next, we consider the estimation of $c_{\sigma^2} = Cov(c_{Y_\lambda})$ and $r_{\sigma^2} = Cov(r_{Y_\lambda})$. Denote $\widehat{c}_\varepsilon = (I_n - P_{X_\lambda}) c_{Y_\lambda}, \widehat{r}_\varepsilon = (I_n + P_{X_\lambda}) r_{Y_\lambda}$.

Theorem 3.7 $\widehat{c}_{\sigma^2} = \frac{\widehat{c}_\varepsilon^T \widehat{c}_\varepsilon}{n-p}$ and $\widehat{r}_{\sigma^2} = \frac{\widehat{c}_\varepsilon^T \widehat{c}_\varepsilon}{n+p}$ are unbiased estimators of c_{σ^2} and r_{σ^2} respectively.

Proof Since $(I_n - P_{X_\lambda})$ is an idempotent matrix, it has

$$\widehat{c}_\varepsilon^T \widehat{c}_\varepsilon = \left((I_n - P_{X_\lambda}) c_{Y_\lambda} \right)^T \left((I_n - P_{X_\lambda}) c_{Y_\lambda} \right)$$

$$= c_{Y_\lambda}^T (I_n - P_{X_\lambda}) c_{Y_\lambda}.$$

So

$$E[\widehat{c}_\varepsilon^T \widehat{c}_\varepsilon] = E\left[c_{Y_\lambda}^T (I_n - P_{X_\lambda}) c_{Y_\lambda}\right]$$
$$= (X_\lambda c_\beta)^T (I_n - P_{X_\lambda})(X_\lambda c_\beta) + tr(I_n - P_{X_\lambda}) Cov(c_{Y_\lambda})$$
$$= c_{\sigma^2} tr(I_n - P_{X_\lambda})$$
$$= c_{\sigma^2}(n - p).$$

Then the estimator of c_{σ^2} is gived as

$$\widehat{c}_{\sigma^2} = \frac{\widehat{c}_\varepsilon^T \widehat{c}_\varepsilon}{n - p}.$$

So

$$E(\widehat{c}_{\sigma^2}) = E\left(\frac{\widehat{c}_\varepsilon^T \widehat{c}_\varepsilon}{n - p}\right) = \frac{1}{n - p} c_{\sigma^2}(n - p) = c_{\sigma^2}.$$

Since $(I_n + P_{|X_\lambda|})$ is an idempotent matrix, so

$$\widehat{r}_\varepsilon^T \widehat{r}_\varepsilon = \left((I_n + P_{|X_\lambda|}) r_{Y_\lambda}\right)^T \left((I_n + P_{|X_\lambda|}) r_{Y_\lambda}\right)$$
$$= r_{Y_\lambda}^T (I_n + P_{|X_\lambda|}) r_{Y_\lambda}.$$

Furthermore

$$E[\widehat{r}_\varepsilon^T \widehat{r}_\varepsilon] = E\left[r_{Y_\lambda}^T (I_n + P_{|X_\lambda|}) r_{Y_\lambda}\right]$$
$$= (|X_\lambda| r_\beta)^T (I_n + P_{|X_\lambda|})(|X_\lambda| r_\beta) + tr(I_n + P_{|X_\lambda|}) Cov(r_{Y_\lambda})$$
$$= r_{\sigma^2} tr(I_n + P_{|X_\lambda|})$$
$$= r_{\sigma^2}(n + p).$$

The estimator of r_{σ^2} is given as

$$\widehat{r}_{\sigma^2} = \frac{\widehat{r}_\varepsilon^T \widehat{r}_\varepsilon}{n + p}.$$

So

$$E(\widehat{r}_{\sigma^2}) = \frac{1}{n - p} r_{\sigma^2}(n - p) = r_{\sigma^2}.$$

The result is proved. □

In the following, we discuss the independence of $\widehat{\beta}_{LS} = (\widehat{c}_\beta; \widehat{r}_\beta)$ and $\widehat{\sigma^2} = (\widehat{c}_{\sigma^2}; \widehat{r}_{\sigma^2})$.

Theorem 3.8 $\widehat{c}_{\sigma^2}$ and $\widehat{c}_\beta$ are independent, $\widehat{r}_{\sigma^2}$ and $\widehat{r}_\beta$ are independent.

Proof Since

$$\widehat{\sigma^2} = (\widehat{c}_{\sigma^2}; \widehat{r}_{\sigma^2})$$
$$= \left(\frac{c_{Y_\lambda}^T (I_n - P_{X_\lambda}) c_{Y_\lambda}}{n - p}; \frac{r_{Y_\lambda}^T (I_n + P_{|X_\lambda|}) r_{Y_\lambda}}{n - p}\right),$$

$$\widehat{\beta}_{LS}(\lambda) = (\widehat{c}_\beta; \widehat{r}_\beta)$$
$$= \left((X_\lambda^T X_\lambda)^{-1} X_\lambda^T c_{Y_\lambda}; (|X_\lambda|^T |X_\lambda|)^{-1} |X_\lambda|^T r_{Y_\lambda}\right),$$

it can be seen that $\widehat{c}_{\sigma^2}$ is the quadratic form of c_{Y_λ}, $\widehat{c}_\beta$ is the linear form of c_{Y_λ}, and $c_{Y_\lambda} \sim N(0, c_{\sigma^2} I_n)$.

According to the independence theorem of quadratic form and linear form of normal variables, it is necessary to prove that they are independent of each other, that is, the product of linear part, variance part and quadratic part of normal variables is 0. Then

$$(X_\lambda^T X_\lambda)^{-1} X_\lambda^T c_{\sigma^2} I_n (I_n - P_{X_\lambda}) = c_{\sigma^2} I_n ((X_\lambda^T X_\lambda)^{-1} X_\lambda^T - (X_\lambda^T X_\lambda)^{-1} X_\lambda^T P_{X_\lambda}) = 0.$$

Similarly, $\widehat{r}_{\sigma^2}$ is the quadratic form of r_{Y_λ}, $\widehat{r}_\beta$ is the linear form of r_{Y_λ}, and $r_{Y_\lambda} \sim N\left(0, r_{\sigma^2} I_n\right)$, then

$$
\begin{aligned}
(|X_\lambda|^T|X_\lambda|)^{-1}|X_\lambda|^T r_{\sigma^2} I_n (I_n - P_{|X_\lambda|}) = {} & r_{\sigma^2} I_n((|X_\lambda|^T|X_\lambda|)^{-1}|X_\lambda|^T - \\
& (|X_\lambda|^T|X_\lambda|)^{-1}|X_\lambda|^T P_{|X_\lambda|}) \\
= {} & 0.
\end{aligned}
$$

Thus $\widehat{c}_{\sigma^2}$ and $\widehat{c}_\beta$ are independent, $\widehat{r}_{\sigma^2}$ and $\widehat{r}_\beta$ are independent. $\square$

Theorem 3.9 In the sense of D_2 distance, the sufficient condition for the strong consistentancy of $\widehat{\beta}_{LS}(\lambda)$ for estimating β is:

$$
\lim_{n \to \infty} (S_n^{-1} + |S_n|^{-1}) = 0
$$

where, $S_n = X_\lambda^T X_\lambda$, $|S_n| = |X_\lambda|^T|X_\lambda|$.

Proof According to Theorem 3.4, $\widehat{\beta}_{LS}(\lambda)$ is an unbiased estimate of β, namely,

$$
E(\widehat{\beta}_{LS}(\lambda)) = \beta.
$$

Moreover,

$$
\begin{aligned}
Var(\widehat{\beta}_{LS}(\lambda)) = {} & 2c_{\sigma^2}(X^T(I_n - \lambda W)^T(I_n - \lambda W)X)^{-1} + \\
& 2r_{\sigma^2}(|X|^T|I_n - \lambda W|^T|I_n - \lambda W||X|)^{-1}),
\end{aligned}
$$

From $\lim_{n \to \infty}(S_n^{-1} + |S_n|^{-1}) = 0$, it has

$$
\begin{aligned}
\lim_{n \to \infty} Var(\widehat{\beta}_{LS}(\lambda)) = {} & \lim_{n \to \infty}(2c_{\sigma^2}(X^T(I_n - \lambda W)^T(I_n - \lambda W)X)^{-1} + \\
& 2r_{\sigma^2}(|X|^T|I_n - \lambda W|^T|I_n - \lambda W||X|)^{-1}) \\
= {} & \lim_{n \to \infty}(2c_{\sigma^2}S_n^{-1} + 2r_{\sigma^2}|S_n|^{-1}) \\
= {} & 0.
\end{aligned}
$$

Thus

$$
\lim_{n \to \infty} Var(\widehat{\beta}_{LS}(\lambda)) = \lim_{n \to \infty} D_2^2(\widehat{\beta}_{LS}(\lambda), E(\widehat{\beta}_{LS}(\lambda))) = 0.
$$

Therefore, in the sense of D_2 metric, $\widehat{\beta}_{LS}(\lambda)$ is a strong consistent estimate of β. $\square$

4. Numerical simulation

In this part, the parameter estimation process of interval-valued spatial error model is further explored by numerical simulation. Based on the d_p distance of the interval value, mean square error of the estimator is calculated to measure the goodness of the estimation.

Based on Equation 3.3, the interval-valued spatial error model in matrix form is expanded as follows:

$$
\begin{pmatrix} y_{1\lambda} \\ y_{2\lambda} \\ \vdots \\ y_{n\lambda} \end{pmatrix} = \begin{pmatrix} 1 & x_{1\lambda} \\ 1 & x_{2\lambda} \\ \vdots & \vdots \\ 1 & x_{n\lambda} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix},
$$

where

$$
\begin{aligned}
Y_\lambda &= (y_{1\lambda}, y_{2\lambda}, \cdots, y_{n\lambda}) = (I_n - \lambda W)\, Y \\
X_\lambda &= (x_{1\lambda}, x_{2\lambda}, \cdots, x_{n\lambda}) = (I_n - \lambda W)\, X.
\end{aligned}
$$

I_n is the unit matrix of dimension n, λ is the spatial autocorrelation coefficient. and W as the spatial weight matrix. Using the first-order adjacency method, assuming that n samples are arranged in one font, the spatial weight matrix can be written as follows:

$$W = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

The true values of the given interval value parameters are $\beta_1 = [1,2] = (1.5; 0.5), \beta_2 = [1.5, 2.5] = (2; 0.5)$ respectively. Then,

$$y_{i\lambda} = \beta_1 + x_{i\lambda}\beta_2 + \varepsilon_i$$
$$= (1.5 + 2x_{i\lambda} + c_{\varepsilon_i}; 0.5 + 0.5x_{i\lambda} + r_{\varepsilon_i}),$$

where, the error term follows the normal distribution, that is $c_{\varepsilon_i}, r_{\varepsilon_i} \sim N\left(0, 0.3^2\right)$.

First, take $n = 100$, the explanatory variable X is generated according to the rules of $x_i = 0.5 + 0.1i, i = 1, 2, \cdots, 100$, and a set of values will be obtained in each simulation experiment.

The scatter point in Figure 1 is the $y_{i\lambda}$ data generated by a simulation experiment, and the two fitting lines are the corresponding interval value spatial error model function:

$$Y_\lambda = [1.0653, 2.0171] + [1.4955, 2.5009]X_\lambda.$$

Repeat the above process for 500 times to obtain the average value of $\widehat{\beta} = (\widehat{\beta}_1, \widehat{\beta}_2)$ as follows:

$$\bar{\widehat{\beta}} = ([1.0004, 2.0019], [1.4999, 2.4999]).$$

Next, the mean square error MSE of the parameter estimation obtained by the model is calculated as one of the criteria to measure the goodness of the estimation. The calculation method is based on interval value d_p distance:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} d_2^2\left(\beta, \widehat{\beta}\right)$$
$$= \frac{1}{500}\sum_{i=1}^{500} d_2^2\left(\beta, \widehat{\beta}\right)$$

By calculation, the MSE of samples $\widehat{\beta}_1, \widehat{\beta}_2$ are 0.0165 and 0.0006 respectively.
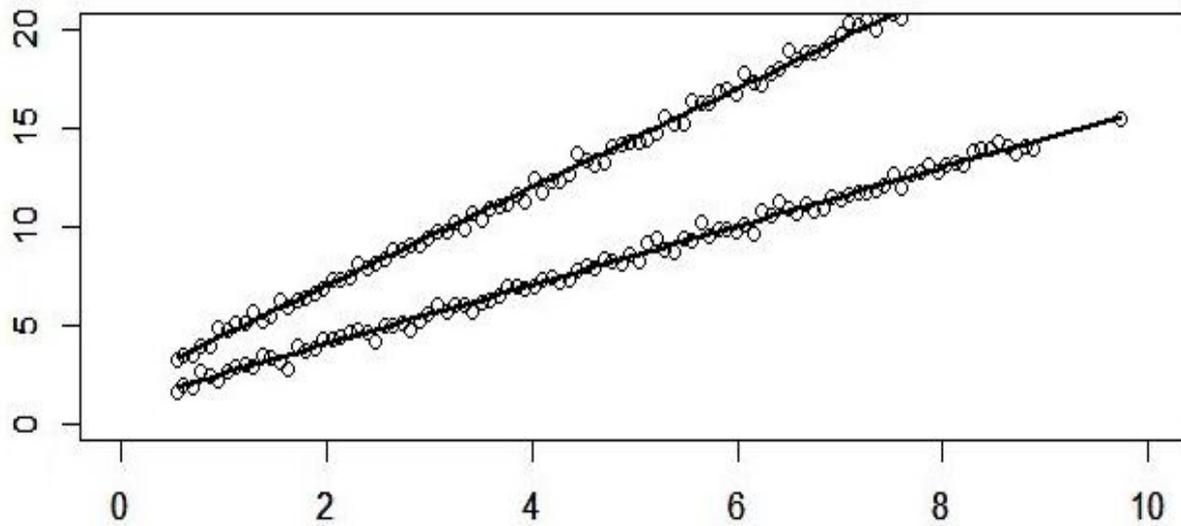
Figure 1. numerical simulation of interval valued spatial error model

Similarly, set $n = 200, n = 300$, repeat the above simulation process for 500 times respectively, and obtain the average value of $\widehat{\beta} = (\widehat{\beta}_1, \widehat{\beta}_2)$ and the sample mean square error MSE respectively. The specific simulation results are summarized in Table 1 and Table 2.

Table 1. Average of $\widehat{\beta} = (\widehat{\beta}_1, \widehat{\beta}_2)$

$\lambda = 0.0727$	n=100	n=200	n=300
$\widehat{\beta}_1$	[1.0004,2.0019]	[1.0023,1.9991]	[0.9988,2.0032]
$\widehat{\beta}_2$	[1.4999,2.4999]	[1.4998,2.5002]	[1.5001,2.4999]

Table 2. Sample mean square error MSE

$\lambda = 0.0727$	n=100	n=200	n=300
$MSE(\widehat{\beta}_1)$	0.01651	0.00723	0.00462
$MSE(\widehat{\beta}_2)$	0.00060	0.00007	0.00002

It can be seen from the simulation results that when $n = 100$, the obtained parameter estimation is very close to the real value. With the increase of sample size, the obtained estimation is closer to the real value, and the sample mean square error MSE is smaller and smaller.

5. Empirical analysis

5.1 Data preparation

In this paper, we select the index data of 31 cities from 31 provinces and autonomous regions in China (not considered temporarily due to geographical factors, Hong Kong, Macao and Taiwan) to illustrate the application of interval-valued spatial error model in practice. This section mainly studies factors such as latitude, maximum and minimum temperatures and temperature difference. Due to the vast territory of China, the temperature and latitude of different cities in a region are different. Therefore, the provincial capital cities in various regions are selected as the research object. The temperature data is the minimum and maximum temperature on July 8, 2021. The temperature data comes from Baidu weather forecast, and the latitude data comes from convenient query website(https://jingweidu.bmcx.com/), the data indicators are summarized in Table 3.

Table 3. Data and indicators

Region	Minimum temperature	Maximum temperature	Latitude
Hefei	24	29	31.79
Beijing	22	33	40.22
Chongqing	25	34	29.4
Fuzhou	27	38	26.05
Lanzhou	20	36	36.1
Guangzhou	27	34	23.16
Nanning	25	33	22.78
Guiyang	21	29	26.68
Haikou	26	33	20.02
Shijiazhuang	24	37	38.04
Haerbin	20	25	45.55
Zhengzhou	26	37	34.72
Wuhan	27	33	30.58
Changsha	25	33	28.26
Nanjing	26	29	31.33
Nanchang	28	35	28.55
Changchun	20	27	43.83
Shenyang	20	27	41.81
Huhehaote	19	31	40.81
Yinchuan	20	35	38.47
Xining	14	29	36.65
Xian	25	36	34.23
Jinan	25	33	36.55
Shanghai	26	32	31.41
Taiyuan	19	32	37.94
Chengdu	23	29	30.66
Tianjin	24	34	39.72
Wulumuqi	25	33	43.36
Lasa	12	23	29.65
Kunming	18	27	24.89
Hangzhou	27	35	30.21

Before modeling and analysis, spatial autocorrelation test is carried out for the data. First, the spatial autocorrelation test is based on the spatial weight matrix. In this paper, the distance based representation is selected for the selection of the spatial weight matrix. The spatial weight matrix is obtained by the open source software geoda. After obtaining the spatial weight matrix, the matrix is standardized. Of course, there are many representation methods of spatial weight matrix, and distance representation can also be selected, which will not be repeated here. Spatial autocorrelation is based on the explained variable. The explained variable in this paper consists of the interval value of the lowest temperature and the highest temperature. Therefore, in the spatial autocorrelation test, the general linear model can be used to model the upper and lower endpoints of interval values respectively, and the center values of the highest and lowest temperatures can be used for spatial autocorrelation test.

One of the main methods of spatial autocorrelation test is to conduct global or local Moran's I test. As can be seen from the table below, the global Moran's I indexes are 0.4193 and 0.3013 respectively, and the tested p values are less than the significant level of 0.05. Therefore, the original hypothesis is rejected and

it is considered that the maximum and minimum temperatures of 31 provinces, cities and autonomous regions in China have a certain spatial autocorrelation.

Table 4. Global Moran's I test results

Statistic	Minimum temperature	Maximum temperature
Moran's I	0.4193	0.3013
p-value	0.000014	0.000985

Figure 2 and Figure 3 are the scatter diagrams of global Moran's I coefficient. It can be seen that the highest and lowest temperatures of 31 regions in China have positive autocorrelation, that is, the trend of high high and low low.
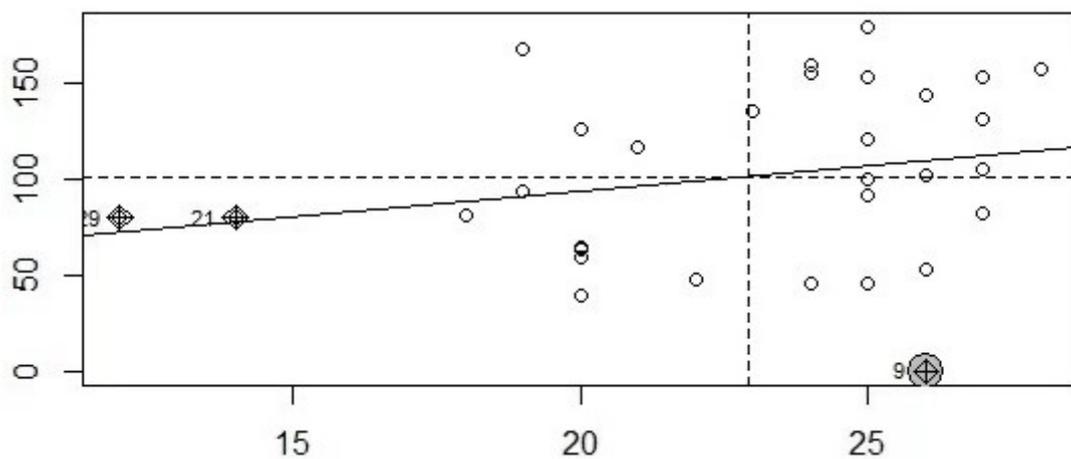
Figure 2. Scatter plot of global Moran's I coefficient of minimum temperature
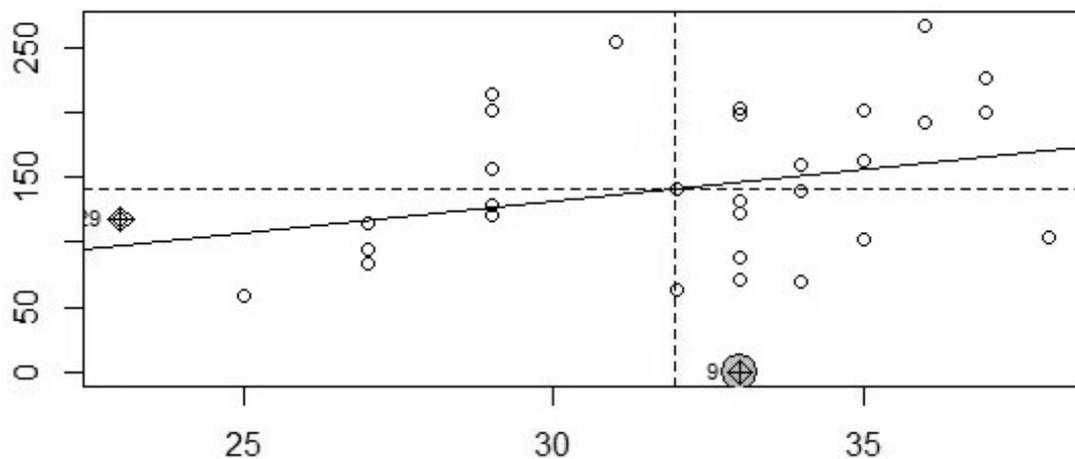

Figure 3. Scatter plot of global Moran's I coefficient of maximum temperature

5.2 Parameter estimation of interval-valued spatial error model

After data preparation and spatial autocorrelation test, the interval-value spatial error model is established. Among them, the dependent variable is air temperature, the interval value data is composed of the lowest and highest air temperatures, and the explanatory variable is latitude. The parameter estimation is carried out by the parameter estimation method of interval-value spatial error model in

Section 3. It is assumed that the air temperature and latitude conform to the interval-valued linear model, that is:

$$E(y_{i\lambda}) = \beta_1 + x_{i\lambda}\beta_2$$

As the method mentioned in Section 4, by the minimum temperature and the maximum temperature we can obtained $\lambda(\lambda_1 = 0.1187, \lambda_2 = 0.1429)$ respectively. Then take the average value as the $\lambda(\lambda = \frac{\lambda_1+\lambda_2}{2} = 0.1308)$ of the interval-valued spatial error model. GeoDa software is used to obtain the spatial weight matrix of 31 regions. Take the inverse square of the matrix elements, and then standardize to obtain the final spatial weight matrix..

The parameter estimation results are shown in Table 5, which shows that:

$$\hat{\beta} = (\hat{\beta_1}, \hat{\beta_2}) = ([24.7501, 29.5478], [-0.1681, -0.0619])$$

Table 5. Interval valued parameter estimation results

$\widehat{\beta_1}$	$\widehat{\beta_2}$
[24.7501,29.5478]	[-0.1681,-0.0619]

The final interval-valued spatial error model is:

$$Y_\lambda = [24.7501, 29.5478] + [-0.1681, -0.0619]X_\lambda.$$

In Figure 4, the abscissa and ordinate are latitude and temperature (maximum temperature and minimum temperature) respectively, it can be seen that the temperature and latitude of the 31 cities in China are negatively correlated. With the increasing of latitude, the temperature has a certain downward trend. At the same time, it can also be seen that the temperature difference (the difference between the maximum temperature and the minimum temperature) tends to expand with the increase of latitude, which is also consistent with the large diurnal temperature difference in northwest and northeast of China. And small temperature difference between day and night in central region of China, southeast and southwest of China.
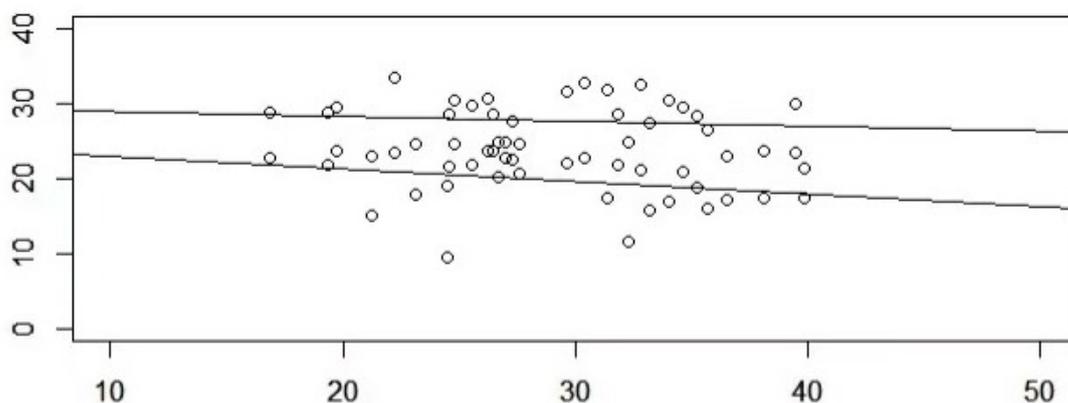

Figure 4. Interval value spatial error model of temperature and latitude

References

1. Aumann R J. Integrals of set-valued functions[J]. Journal of Mathematical Analysis and Applications, 1965, 12(1): 1-12.
2. Blanco Fernandez A, Corral N, Gonzalez-Redriguez G, Lubiano M A. Some properties of the dk-variance for interval-valued sets[J]. 2008,D.Dubois et al. (Eds.): Soft Methods for Hand. Var. And Imprecision,ASC48:331-337.
3. Billard L, Diday E. Regression analysis for interval-valued data. Conference of the International Federation of Classifification Societies[C]. Springer-Verlag, 2000: 369-374.
4. Billard L, Diday E. Symbolic regression analysis[J]. Studies in Classification Data Analysis and Knowledge Organization, 2002, 281- 288.
5. Hess C. On multivalued martingales whose values may be unbounded: martingale selectors and Mosco convergence[J]. Journal of Multivariate Analysis, 1991, 39: 175-201.
6. Hiai F, Umegaki H. Integrals, conditional expectations, and martingales of multivalued functions[J]. Journal of Multivariate Analysis, 1977, 7(1): 149-182.
7. Li S, Li J, Li X. Stochastic integral with respect to set-valued square integrable martingales [J]. Journal of Mathematical Analysis and Applications, 2010, 370: 659-671.
8. Li S, Ogura Y. Convergence of set valued sub- and supermartingales in the KuratowskiMosco sense[J]. The Annuals of Probability, 1998, 26: 1384-1402.
9. Li S, Ogura Y. Convergence of set-valued and fuzzy-valued martingales[J]. Fuzzy sets and systems, 1999, 101: 453-461.
10. Li S, Ogura Y, Kreinovich V. Limit Theorems and Applications of Set-Valued Random Variables[M]. Netherlands: Kluwer academic publishers(Springer), 2002.
11. Lyashenko N N. Limit theorems for sums of independent compact random subsets of a Euclidean space[J]. Journal of Mathematical Sciences, 1982, 20(3): 2187-2196.
12. Lyashenko N N. Statistics of random compacts in Euclidean space[J]. Journal of Mathematical Sciences, 1983, 21(1): 76-92.
13. Molchanov I S. Theory of Random Sets[M](Springer), 2005.
14. Papageorgiou N S. On the theory of Banach space valued multifunction. 2. set valued martingales and set valued measures[J]. Journal of Multivariate Analysis, 1985, 17: 207- 227.
15. Papageorgiou N S.On the conditional expectation and convergence properties of random sets[J]. Transactions of the American Mathematical Society, 1995, 347: 2495-2515.
16. Vitale R. Lp metrics for compact, convex sets[J]. Journal of Approximation Theory, 1985, 45(3): 280-287.
17. Yang X, Li S. The Dp-metric space of set-valued random variables and its application to covariances[J]. International Journal of Innovative Computing, Information and Control, 2005, 1: 73-82.
18. Zhang Wenxiu, Li Shoumei, Wang Zhenpeng, Gao Yong. Introduction to set-valued stochastic processes[M].Beijing: Science Press,2007.
19. Lima Neto E, de Carvalho F, Centre and range method for fitting a linear regression model to symbolic interval data[J] .Computational Statistics and Data Analysis,2008:52, 1500-1515.
20. Lima Neto E, de Carvalho F, Constrained linear regression models for symbolic interval-valued variables[J], Computational Statistics and Data Analysis, 2010:54,333-347.
21. Tang Nana. Linear regression and autoregressive time series models with constraints [D], Beijing University of technology,2017.
22. Wang H, Guan R, Wu J. Linear regression of interval-valued data based on complete information in hypercubes[J]. Journal of systems science and systems engineering (English Edition),2012, 21(4): 422-442.
23. Souza L, Souza R, Amaral G, Filho T. A parametrized approach for linear regression of interval data[J]. Knowledge-Based Systems, 2017, 131: 149-159.
24. Wang X, Li S, Denoeux T. Interval-valued linear model[J]. International journal of computational intelligence systems, 2015, 8(1): 114-127.
25. Yildirim V , Kantar Y M . Robust estimation approach for spatial error model[J]. Journal of Statistical Computation and Simulation, 2020, 90(3):1-21.

26.   Anselin L. Spatial Econometrics: Methods and Models [M],1988.

27.   Prucha K I R . A generalized moments estimator for the autoregressive parameter in a spatial model[J]. International Economic Review, 2010, 40(2):509-533.