

Article

Estimation in Partial Functional Linear Spatial Autoregressive Model

Yuping Hu ¹, Siyu Wu ¹, Sanying Feng ^{1,2,*} and Junliang Jin ¹

¹ School of Mathematics and Statistics, Zhengzhou University, Zhengzhou 450001, China; hyp@zzu.edu.cn (Y.H.); wwwrsy2020@163.com (S.W.); jinjunliang1319@163.com (J.J.)

² Henan Key Laboratory of Financial Engineering, Zhengzhou University, Zhengzhou 450001, China

* Correspondence: fsy5801@zzu.edu.cn

Received: 6 September 2020; Accepted: 24 September 2020; Published: 1 October 2020



Abstract: Functional regression allows for a scalar response to be dependent on a functional predictor; however, not much work has been done when response variables are dependence spatial variables. In this paper, we introduce a new partial functional linear spatial autoregressive model which explores the relationship between a scalar dependence spatial response variable and explanatory variables containing both multiple real-valued scalar variables and a function-valued random variable. By means of functional principal components analysis and the instrumental variable estimation method, we obtain the estimators of the parametric component and slope function of the model. Under some regularity conditions, we establish the asymptotic normality for the parametric component and the convergence rate for slope function. At last, we illustrate the finite sample performance of our proposed methods with some simulation studies.

Keywords: partial functional linear spatial autoregressive model; spatial autoregression; functional principal component analysis; instrument variable

1. Introduction

Over the last two decades, there has been an increasing interest in functional data analysis in econometrics, biometrics, chemometrics, and medical research, as well as other fields. Due to the infinite-dimensional nature of functional data, the classical methods for functional data are no longer applicable. There has been a large amount of work in function data analysis; see Ramsay and Silverman [1], Cardot et al. [2], Yao et al. [3], Lian and Li [4], Fan et al. [5], Feng and Xue [6], Kong et al. [7], and Yu et al. [8]. Some methods and theories on partial functional linear models have been proposed. For example, based on a two-stage nonparametric regression calibration method, Zhang et al. [9] discussed a partial functional linear model. Shin [10] proposed new estimators of the parameters and coefficient function of partial functional linear model. Lu et al. [11] considered quantile regression for the functional partially linear model. Yu et al. [12] proposed a prediction procedure for the partial functional linear quantile regression model. However, the aforementioned articles have a significant limitation. That is, they assumed that response variables are independence variables. However, in many fields, such as economics, finance, and environmental studies, sometimes response variables are dependence spatial variables. Therefore, it is of practical interest to develop more flexible approaches using a broader family of data.

There has been considerable work on dependence spatial variables. One useful approach in dealing with spatial dependence is the spatial autoregressive model, which adds a weighted average of nearby values of the dependent variable to the base set of explanatory variables. Theories and methods based on parametric spatial autoregressive models have been extensively studied in Cliff and Ord [13], Anselin [14], and Cressie [15]. Lee [16] proposed the quasi-maximum likelihood estimation.

Then, Su and Jin [17] extended the quasi-likelihood estimation method to partially linear spatial autoregressive models. Koch and Krisztin [18] developed the B-splines and genetic-algorithms method for partially linear spatial autoregressive models. Chen et al. [19] proposed a new estimation method based on the kernel estimation method. Du et al. [20] considered partially linear additive spatial autoregressive models, proposed the instrumental variable estimation method, and established the asymptotic normality for the parametric component.

It is a good idea to develop more flexible approaches using a broader family of data, where the limitation can, in principle, be easily solved by proposing a new model. Thus, in this paper, based on spatial variables and functional data, we combine the spatial autoregressive model and the partial functional linear model, and propose a partial functional linear spatial autoregressive model.

Let Y_i be a real-valued dependence spatial variable corresponding to the i th observation, Z_i be a p -dimensional vector of associated explanatory variables, for $i = 1, \dots, n$. $X_i(t)$ be zero mean random functions belonging to $L^2(\mathcal{T})$, and be independent and identically distributed, $i = 1, \dots, n$. For simplicity, we suppose throughout that $\mathcal{T} = [0, 1]$. The partial functional linear spatial autoregressive model is given by

$$Y_i = \rho \sum_{j=1}^n w_{ij} Y_j + Z_i^T \beta + \int_0^1 \gamma(t) X_i(t) dt + \varepsilon_i, \tag{1}$$

where w_{ij} is the (i, j) th element of a given $n \times n$ non-stochastic spatial weighting matrix W_n , such that $w_{ij} = 0$ for all $i = j$, W_n is a specified $n \times n$ spatial weight matrix. The definition of spatial weight matrix W_n is based on the geographic arrangement of the observations or contiguity. More generally, W_n matrices can be specified based on geographical distance decay, economic distance, and the structure of a social network. $\beta = (\beta_1, \dots, \beta_p)^T$ is a vector of p -dimensional unknown parameters, $\gamma(t)$ is a square integrable unknown slope function on $[0, 1]$, and ε_i are independent and identically distributed random errors with zero mean and finite variance σ^2 .

There are many methods which can be used to deal with functional data, such as the functional principal component method, spline methods, and the rough penalty method. Functional principal component analysis (FPCA) can analyse an infinite dimensional problem by a finite dimensional one—therefore, FPCA is popular and widely used by researchers. Dauxois et al. [21] investigated the asymptotic theory of FPCA. Cardot et al. [22] applied FPCA to estimate the slope function of the functional linear model. Hall and Horowitz [23] and Hall and Hosseini-Nasab [24] showed the optimal convergence rates of slope function based on the FPCA technique.

In this paper, we consider the estimating problem of the model (1). Based on FPCA and the instrumental variable estimation techniques, we obtain the estimators of the parameters and slope function of model (1) with the two-stage least squares method. Under some mild conditions, the rate of convergence and asymptotic normality of the resulting estimators are established. Finally, some simulation studies are carried out to assess the finite sample performance of the proposed method. The results are encouraging and show that all estimators perform well in finite samples. Overall, simulation experiments lend support to our asymptotic results.

The rest of the paper proceeds as follows. In Section 2, functional principal component analysis and the instrumental variable estimation method is proposed to estimate the partial functional linear spatial autoregressive regression model. In Section 3, the asymptotic properties are given. Some simulation studies are described in Section 4. Lastly, we conclude the paper in Section 5 with some future work.

2. Estimation Procedures

First, we introduce FPCA. Denote the covariance function of $X(t)$ by K_X . Then, by Mercer’s Theorem, we can obtain the spectral decomposition as $K_X(s, t) = \sum_{k=1}^{\infty} \lambda_k \phi_k(s) \phi_k(t)$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ are the eigenvalues of the linear operator associated with $K_X(s, t)$, and $\phi_k(t)$

are the corresponding eigenfunctions. By the Karhunen-Loève expansion, $X_i(t)$ can be represented as $X_i(t) = \sum_{j=1}^{\infty} \zeta_{ij} \phi_j(t)$, where the coefficients $\zeta_{ik} = \int X_i(t) \phi_k(t) dt$ are uncorrelated random variables with mean zero and variances $E(\zeta_{ik}^2) = \lambda_k$, also called the functional principal component scores. Expanded on the orthonormal eigenbasis $\{\phi_k(t)\}$, the slope function can be written as $\gamma(t) = \sum_{k=1}^{\infty} \gamma_k \phi_k(t)$. Based on the above FPCA, model (1) can be well-approximated by

$$Y_i \doteq \rho \sum_{j=1}^n w_{ij} Y_j + Z_i^T \beta + \sum_{j=1}^m \gamma_j \langle X_i, \phi_j \rangle + \varepsilon_i, \quad i = 1, \dots, n, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ represents the $L^2(\mathcal{T})$ inner product, $\gamma_j = \langle \gamma, \phi_j \rangle$, and m is sufficiently large.

The approximate model (2) naturally suggests the idea of principal components regression. However, in practice, ϕ_j are unknown and must be replaced by estimates in order to estimate β and γ_j ($j = 1, \dots, m$). For this purpose, we consider the empirical version of $K_X(s, t)$, which is given by

$$\widehat{K}_X(s, t) = \frac{1}{n} \sum_{i=1}^n X_i(s) X_i(t) = \sum_{j=1}^{\infty} \widehat{\lambda}_j \widehat{\phi}_j(s) \widehat{\phi}_j(t),$$

where $(\widehat{\lambda}_j, \widehat{\phi}_j)$ are pairs of eigenvalues and eigenfunctions for the covariance operator associated with \widehat{K}_X and $\widehat{\lambda}_1 \geq \widehat{\lambda}_2 \geq \dots \geq 0$. We take $(\widehat{\lambda}_j, \widehat{\phi}_j)$ as the estimator of (λ_j, ϕ_j) .

Replacing $\phi_j(t)$ by $\widehat{\phi}_j(t)$, model (2) can be written as

$$Y_i \doteq \rho \sum_{j=1}^n w_{ij} Y_j + Z_i^T \beta + \sum_{j=1}^m \gamma_j \langle X_i, \widehat{\phi}_j \rangle + \varepsilon_i, \quad i = 1, \dots, n. \tag{3}$$

Let $\mathbf{Y}_n = (Y_1, \dots, Y_n)^T$, $\mathbf{Z}_n = (Z_1, \dots, Z_n)^T$, $\mathbf{X}_n = (X_1(t), \dots, X_n(t))^T$, $\langle \mathbf{X}_n, \widehat{\phi}_j \rangle = \int_0^1 \widehat{\phi}_j(t) \mathbf{X}_n(t) dt = (\int_0^1 \widehat{\phi}_j(t) X_1(t) dt, \dots, \int_0^1 \widehat{\phi}_j(t) X_n(t) dt)^T$, and $\mathbf{\Pi} = (\langle \mathbf{X}_n, \widehat{\phi}_1 \rangle, \dots, \langle \mathbf{X}_n, \widehat{\phi}_m \rangle)$, $\boldsymbol{\alpha} = (\gamma_1, \dots, \gamma_m)^T$, $\boldsymbol{\varepsilon}_n = (\varepsilon_1, \dots, \varepsilon_n)^T$, model (3) can be written as matrix notation

$$\mathbf{Y}_n \doteq \rho \mathbf{W}_n \mathbf{Y}_n + \mathbf{Z}_n \beta + \mathbf{\Pi} \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_n. \tag{4}$$

Let $\mathbf{P} = \mathbf{\Pi}(\mathbf{\Pi}^T \mathbf{\Pi})^{-1} \mathbf{\Pi}^T$ denote the projection matrix onto the space spanned by $\mathbf{\Pi}$, and we obtain

$$(\mathbf{I} - \mathbf{P}) \mathbf{Y}_n \doteq \rho (\mathbf{I} - \mathbf{P}) \mathbf{W}_n \mathbf{Y}_n + (\mathbf{I} - \mathbf{P}) \mathbf{Z}_n \beta + (\mathbf{I} - \mathbf{P}) \boldsymbol{\varepsilon}_n. \tag{5}$$

Let $\mathbf{Q} = (\mathbf{W}_n \mathbf{Y}_n, \mathbf{Z}_n)$, $\boldsymbol{\theta} = (\rho, \beta^T)^T$, applying the two-stage least squares procedure proposed by Kelejian and Prucha [25], we propose the following estimator

$$\widehat{\boldsymbol{\theta}} = \left(\mathbf{Q}^T (\mathbf{I} - \mathbf{P}) \mathbf{M} (\mathbf{I} - \mathbf{P}) \mathbf{Q} \right)^{-1} \mathbf{Q}^T (\mathbf{I} - \mathbf{P}) \mathbf{M} (\mathbf{I} - \mathbf{P}) \mathbf{Y}_n, \tag{6}$$

where $\mathbf{M} = \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ and \mathbf{H} is matrix of instrumental variables. Moreover,

$$\widehat{\boldsymbol{\alpha}} = (\widehat{\gamma}_1, \dots, \widehat{\gamma}_m)^T = (\mathbf{\Pi}^T \mathbf{\Pi})^{-1} \mathbf{\Pi}^T (\mathbf{Y}_n - \mathbf{Q} \widehat{\boldsymbol{\theta}}). \tag{7}$$

Consequently, we use $\widehat{\gamma}(t) = \sum_{k=1}^m \widehat{\gamma}_k \widehat{\phi}_k(t)$ as the estimator of $\gamma(t)$.

Similar to Zhang and Shen [26], we next construct the instrument variables \mathbf{H} . In the first step, the following instrumental variables are obtained

$$\widetilde{\mathbf{H}} = (\mathbf{W}_n (\mathbf{I} - \rho \mathbf{W}_n)^{-1} (\mathbf{\Pi} \widehat{\boldsymbol{\alpha}}, \mathbf{Z}_n), \mathbf{Z}_n),$$

where $\bar{\rho}$ and $\bar{\alpha}$ are obtained by simply regressing Y_n on pseudo regressor variables $W_n Y_n, Z_n, \Pi$. In the second step, we use \tilde{H} to obtain the estimators $\bar{\alpha}$ and $\bar{\theta}$, and then we can construct the instrumental variables

$$H = (W_n(I - \bar{\rho}W_n)^{-1}(\Pi\bar{\alpha} + Z_n\bar{\beta}), Z_n).$$

To implement our estimation method, we need to choose m . Here, truncation parameter m is selected by AIC criterion. Specifically, we minimize

$$AIC(m) = \log \text{RSS}(m) + 2n^{-1}m, \tag{8}$$

where

$$\text{RSS}(m) = \sum_{i=1}^n \left\{ Y_i - \left(\hat{\rho} \sum_{j=1}^n w_{ij} Y_j + Z_i^T \hat{\beta} + \sum_{j=1}^m \hat{\gamma}_j(X_i, \hat{\phi}_j) \right) \right\}^2,$$

with $\hat{\rho}, \hat{\beta}$ and $\hat{\gamma}_j$ being the estimated value.

3. Asymptotic Properties

In this section, we discuss the asymptotic normality of $\hat{\theta}$ and the rate of convergence of $\hat{\gamma}(t)$. For convenience and simplicity, we let c denote a positive constant that may be different at each appearance. The following assumptions will be maintained throughout the paper.

Assumption 1. The matrix $I - \rho W_n$ is nonsingular with $|\rho| < 1$.

Assumption 2. The row and column sums of the matrices W_n and $(I - \rho W_n)^{-1}$ are bounded uniformly in absolute value for any $|\rho| < 1$.

Assumption 3. For matrix $S = W_n(I - \rho W_n)^{-1}$, there exists a constant ρ_c such that $\rho_c I - SS^T$ is a positive, semidefinite matrix.

Assumption 4. $\frac{1}{n} \tilde{Q}^T(I - P)M(I - P)\tilde{Q} \xrightarrow{P} \Sigma$ in probability for some positive definite matrix, where $\tilde{Q} = (S(Z_n\beta + \eta), Z_n)$, $\eta = (\int_0^1 \gamma(t)X_1(t)dt, \dots, \int_0^1 \gamma(t)X_n(t)dt)^T$.

Assumption 5. For matrix $\tilde{Q} = (S(Z_n\beta + \eta), Z_n)$, there exists a constant ρ_{c^*} such that $\rho_{c^*}I - \tilde{Q}\tilde{Q}^T$ is a positive semidefinite matrix.

Assumption 6. The random vector Z has bounded fourth moments.

Assumption 7. For any $c > 0$, there exists an $\epsilon > 0$, such that

$$\sup_{t \in [0,1]} [E\{|X(t)|^c\}] < \infty, \quad \sup_{s,t \in [0,1]} (E\{|s-t|^{-\epsilon}|X(s) - X(t)|^c\}) < \infty.$$

For each integer $r \geq 1$, $\lambda_k^{-r} E(\xi_k^{2r})$ is bounded uniformly in k .

Assumption 8. $X(t)$ is twice continuously differentiable on $[0, 1]$ with probability 1 and $\int E[X^{(2)}(t)]^4 dt < \infty$, $X^{(2)}(t)$ denotes the second derivative of $X(t)$.

Assumption 9. There exists some constants $a > 1$ and $b > a/2 + 1$, such that $\lambda_j - \lambda_{j+1} \geq Cj^{-a-1}$ and $|\gamma_j| \leq Cj^{-b}$ for $j \geq 1$.

Assumption 10. For truncation parameter m , we assume that $m = O(n^{1/(a+2b)})$.

Assumptions 1–3 impose restrictions on the spatial weighting matrix, and these restrictions are imposed for the spatial regression models (see Lee [16]; Zhang and Shen [26]; Du et al. [20]). Let the weighting matrix $W_n = I_D \otimes B_F$, where I_D is a D -dimensional unit matrix, $B_F = (I_F I_F^T - I_F)/(F - 1)$,

$\mathbf{1}_F$ is the F -dimensional unit vector, and \otimes is a Kronecker product, then weighting matrix W_n can satisfy Assumptions 1–3. Assumption 4 is used to represent the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$. Assumption 5 is required to ensure the identifiability of parameter $\boldsymbol{\theta}$. Assumption 6 is the usual condition for the proofs of asymptotic properties of the estimators. Assumptions 7–9 are regularity assumptions for functional linear models (see Hall and Hosseini-Nasab [24]), where a Gaussian process with Hölder continuous sample paths satisfies Assumption 7. Assumption 10 usually appears in functional linear regression (see Feng and Xue [6]; Shin [10]; Hall and Horowitz [23]).

The following Theorem 1 shows the asymptotic property of the estimator of the parameter vector $\boldsymbol{\theta} = (\rho, \boldsymbol{\beta}^T)^T$.

Theorem 1. *Under the Assumptions 1–10, then*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}^{-1}),$$

where $\hat{\boldsymbol{\theta}} = (\hat{\rho}, \hat{\boldsymbol{\beta}}^T)^T$ and “ \xrightarrow{D} ” denotes convergence in distribution.

Proof of Theorem 1. Let $e_n = \boldsymbol{\eta} - \boldsymbol{\Pi}\boldsymbol{\alpha}$, then

$$Y_n = \mathbf{Q}\boldsymbol{\theta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_n = \mathbf{Q}\boldsymbol{\theta} + e_n + \boldsymbol{\Pi}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}_n.$$

By the definition of $\hat{\boldsymbol{\theta}}$, we have

$$\begin{aligned} & \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \\ &= \left(\mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\mathbf{Q} \right)^{-1} \mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})Y_n - \boldsymbol{\theta} \\ &= \left(\mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\mathbf{Q} \right)^{-1} \mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})[(\mathbf{I} - \mathbf{P})Y_n] - \boldsymbol{\theta} \\ &= \left(\mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\mathbf{Q} \right)^{-1} \mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})(\mathbf{Q}\boldsymbol{\theta} + e_n + \boldsymbol{\varepsilon}_n) - \boldsymbol{\theta} \\ &= \left(\mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\mathbf{Q} \right)^{-1} \mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})(e_n + \boldsymbol{\varepsilon}_n). \end{aligned}$$

First, consider $\mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\mathbf{Q}$. Recall that when $Y_n = (\mathbf{I} - \rho\mathbf{W}_n)^{-1}(\mathbf{Z}_n\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_n)$, it has

$$\begin{aligned} \mathbf{Q} &= (\mathbf{W}_n Y_n, \mathbf{Z}_n) \\ &= \left(\mathbf{W}_n(\mathbf{I} - \rho\mathbf{W}_n)^{-1}(\mathbf{Z}_n\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_n), \mathbf{Z}_n \right) \\ &= \left(\mathbf{W}_n(\mathbf{I} - \rho\mathbf{W}_n)^{-1}(\mathbf{Z}_n\boldsymbol{\beta} + \boldsymbol{\eta}), \mathbf{Z}_n \right) + \left(\mathbf{W}_n(\mathbf{I} - \rho\mathbf{W}_n)^{-1}\boldsymbol{\varepsilon}_n, \mathbf{0} \right) \\ &\triangleq \tilde{\mathbf{Q}} + \tilde{\boldsymbol{\varepsilon}}, \end{aligned}$$

where $\tilde{\mathbf{Q}} = (\mathbf{S}(\mathbf{Z}_n\boldsymbol{\beta} + \boldsymbol{\eta}), \mathbf{Z}_n)$, $\tilde{\boldsymbol{\varepsilon}} = (\mathbf{S}\boldsymbol{\varepsilon}_n, \mathbf{0})$, $\mathbf{S} = \mathbf{W}_n(\mathbf{I} - \rho\mathbf{W}_n)^{-1}$.

Hence, one has

$$\begin{aligned} & \mathbf{Q}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\mathbf{Q} \\ &= (\tilde{\mathbf{Q}} + \tilde{\boldsymbol{\varepsilon}})^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})(\tilde{\mathbf{Q}} + \tilde{\boldsymbol{\varepsilon}}) \\ &= \tilde{\mathbf{Q}}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\tilde{\mathbf{Q}} + \tilde{\boldsymbol{\varepsilon}}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\tilde{\boldsymbol{\varepsilon}} \\ &\quad + \tilde{\mathbf{Q}}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\tilde{\boldsymbol{\varepsilon}} + \tilde{\boldsymbol{\varepsilon}}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\tilde{\mathbf{Q}} \\ &\triangleq R_{11} + R_{12} + R_{13} + R_{14}, \end{aligned}$$

where

$$\begin{aligned} R_{11} &= \tilde{\mathbf{Q}}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\tilde{\mathbf{Q}}, \\ R_{12} &= \tilde{\boldsymbol{\varepsilon}}^T(\mathbf{I} - \mathbf{P})\mathbf{M}(\mathbf{I} - \mathbf{P})\tilde{\boldsymbol{\varepsilon}}, \end{aligned}$$

$$R_{13} = \tilde{Q}^T(I - P)M(I - P)\tilde{e},$$

and

$$R_{14} = \tilde{e}^T(I - P)M(I - P)\tilde{Q}.$$

By the properties of projection matrix and Assumption 3, we have

$$\begin{aligned} & E \left[\varepsilon_n^T S^T (I - P)M(I - P)S\varepsilon_n \right] \\ &= E \left(\text{trace}[\varepsilon_n^T S^T (I - P)M(I - P)S\varepsilon_n] \right) \\ &= E \left(\text{trace}[\varepsilon_n^T S^T (I - P)H(H^T H)^{-1}H^T (I - P)S\varepsilon_n] \right) \\ &= E \left(\text{trace}[(H^T H)^{-\frac{1}{2}}H^T (I - P)S\varepsilon_n \varepsilon_n^T S^T (I - P)H(H^T H)^{-\frac{1}{2}}] \right) \\ &\leq \sigma^2 \rho_c E \left(\text{trace}[(H^T H)^{-\frac{1}{2}}H^T (I - P)H(H^T H)^{-\frac{1}{2}}] \right) \\ &\leq \sigma^2 \rho_c E \left(\text{trace}[(H^T H)^{-\frac{1}{2}}H^T H(H^T H)^{-\frac{1}{2}}] \right) \\ &\stackrel{\Delta}{=} O(1). \end{aligned}$$

Hence, we have

$$\varepsilon_n^T S^T (I - P)M(I - P)S\varepsilon_n = O_p(1).$$

Then, we get that

$$R_{12} = (S\varepsilon_n, \mathbf{0})^T (I - P)M(I - P)(S\varepsilon_n, \mathbf{0}) = O_p(1).$$

By straightforward algebra, one has $E(R_{13}) = 0$. In addition, based on Assumption 3,

$$\begin{aligned} & E \left(\|\tilde{Q}^T (I - P)M(I - P)S\varepsilon_n\|^2 \right) \\ &= E \left(\text{trace}[\tilde{Q}^T (I - P)M(I - P)S\varepsilon_n \varepsilon_n^T S^T (I - P)M(I - P)\tilde{Q}] \right) \\ &\leq \sigma^2 \rho_c E \left(\text{trace}[\tilde{Q}^T (I - P)M(I - P)M(I - P)\tilde{Q}] \right) \\ &\leq \sigma^2 \rho_c E \left(\text{trace}[\tilde{Q}^T M\tilde{Q}] \right) \\ &= \sigma^2 \rho_c E \left(\text{trace}[\tilde{Q}^T H(H^T H)^{-1}H^T \tilde{Q}] \right) \\ &= \sigma^2 \rho_c E \left(\text{trace}[(H^T H)^{-\frac{1}{2}}H^T \tilde{Q}\tilde{Q}^T H(H^T H)^{-\frac{1}{2}}] \right) \\ &\leq \sigma^2 \rho_c \rho_c^* E \left(\text{trace}[(H^T H)^{-\frac{1}{2}}H^T H(H^T H)^{-\frac{1}{2}}] \right) \\ &\stackrel{\Delta}{=} O(1), \end{aligned}$$

Therefore, we have $R_{13} = O_p(1)$. Similarly, we have $R_{14} = O_p(1)$.

Combining the convergence rates of R_{12} , R_{13} and R_{14} , we have

$$Q^T(I - P)M(I - P)Q = R_{11} + O_p(1).$$

Now, we consider $Q^T(I - P)M(I - P)e_n$. Obviously,

$$\begin{aligned} & Q^T(I - P)M(I - P)e_n \\ &= \tilde{Q}^T(I - P)M(I - P)e_n + \tilde{e}^T(I - P)M(I - P)e_n \\ &= \tilde{Q}^T(I - P)M(I - P)e_n + \varepsilon_n^T S^T (I - P)M(I - P)e_n \\ &\stackrel{\Delta}{=} R_{21} + R_{22}, \end{aligned}$$

where

$$R_{21} = \tilde{Q}^T(I - P)M(I - P)e_n,$$

$$R_{22} = \varepsilon_n^T S^T(I - P)M(I - P)e_n.$$

By Lemma 1 of Hu et al. [27], we have

$$\begin{aligned} \|e_n\| &= \left\| \sum_{j=1}^{\infty} \gamma_j \langle X_N, \phi_j \rangle - \sum_{j=1}^m \gamma_j \langle X_N, \hat{\phi}_j \rangle \right\| \\ &= \left\| \sum_{j=1}^m \gamma_j \langle X_N, \phi_j - \hat{\phi}_j \rangle + \sum_{j=m+1}^{\infty} \gamma_j \langle X_N, \phi_j \rangle \right\| \\ &\leq \left\| \sum_{j=1}^m \gamma_j \langle X_N, \phi_j - \hat{\phi}_j \rangle \right\| + \left\| \sum_{j=m+1}^{\infty} \gamma_j \langle X_N, \phi_j \rangle \right\|. \end{aligned}$$

By Lemma 1(b) of Kong et al. [7] with the help of Assumptions 7–9, we have

$$\|\hat{\phi}_j - \phi_j\| = O_p(n^{-\frac{1}{2}j}).$$

By Assumptions 7 and 9, one has

$$\begin{aligned} \left\| \sum_{j=1}^m \gamma_j \langle X_i, \phi_j - \hat{\phi}_j \rangle \right\|^2 &\leq \|X_i(t)\|^2 \left\| \sum_{j=1}^m (\phi_j - \hat{\phi}_j) \gamma_j \right\|^2 \\ &= O_p\left(\sum_{j=1}^m j^{-b} j n^{-\frac{1}{2}}\right)^2 \\ &= O_p(n^{-1} m^{4-2b}). \end{aligned}$$

By Assumptions 9–10, one has

$$E \left| \sum_{j=m+1}^{\infty} \gamma_j \langle X_i, \phi_j \rangle \right| \leq \sum_{j=m+1}^{\infty} |\gamma_j| E |\langle X_i, \phi_j \rangle| = O(m^{-b+\frac{1}{2}}).$$

Thus, we have

$$\begin{aligned} \left\| \sum_{j=m+1}^{\infty} \gamma_j \langle X_i, \phi_j \rangle \right\|^2 &= O_p(m^{1-2b}), \\ \|e_n\|^2 &\leq n \cdot \left\| \sum_{j=1}^m \gamma_j \langle X_i, \phi_j - \hat{\phi}_j \rangle \right\|^2 + n \cdot \left\| \sum_{j=m+1}^{\infty} \gamma_j \langle X_i, \phi_j \rangle \right\|^2 \\ &= O_p(m^{4-2b}) + O(nm^{1-2b}) \\ &= o_p(n). \end{aligned}$$

Combining this with Assumption 4, we have

$$\begin{aligned} E(\|R_{21}\|^2) &= E \left(\text{trace}[e_n^T(I - P)M(I - P)\tilde{Q}\tilde{Q}^T(I - P)M(I - P)e_n] \right) \\ &\leq \rho_c^* E \left(\text{trace}[e_n^T(I - P)M(I - P)e_n] \right) \\ &\leq \rho_c^* E \left(\text{trace}[e_n^T e_n] \right) \\ &= o_p(n). \end{aligned}$$

Thus, we can get $R_{21} = o_p(\sqrt{n})$. Similarly, we have $R_{22} = o_p(\sqrt{n})$.

Then, we can find

$$\begin{aligned} & \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \\ &= \sqrt{n} (R_{11} + O_p(1))^{-1} \mathbf{Q}^T (\mathbf{I} - \mathbf{P}) \mathbf{M} (\mathbf{I} - \mathbf{P}) (\mathbf{e}_n + \boldsymbol{\varepsilon}_n) \\ &= \left(\frac{R_{11}}{n} + O_p(n^{-1}) \right)^{-1} \left(\frac{\tilde{\mathbf{Q}}^T (\mathbf{I} - \mathbf{P}) \mathbf{M} (\mathbf{I} - \mathbf{P}) \boldsymbol{\varepsilon}_n}{\sqrt{n}} + o_p(1) \right). \end{aligned}$$

Invoking the central limit theorem and Slutsky’s theorem, we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}^{-1}).$$

□

Rate of convergence of the slope function $\hat{\gamma}(t) = \sum_{k=1}^m \hat{\gamma}_k \hat{\phi}_k(t)$ is given in the following theorem.

Theorem 2. Under the Assumptions 1–10, then

$$\|\hat{\gamma}(t) - \gamma(t)\|^2 = O_p(n^{-\frac{2b-1}{a+2b}}).$$

The proof of Theorem 2 follows the proof of Theorem 2 of Shin [10], so we omitted it here.

4. Simulation Study

In this section, we conduct simulation studies to assess the finite sample performance of the proposed estimation method. The data $\{Y_i\}$ are generated from the following model

$$Y_n = \rho W_n Y_n + Z_{n1} \beta_1 + Z_{n2} \beta_2 + \int_0^1 \gamma(t) X_n(t) dt + \varepsilon_n,$$

where $Z_{n1} = (Z_{11}, Z_{21}, \dots, Z_{n1})^T$, $Z_{n2} = (Z_{12}, Z_{22}, \dots, Z_{n2})^T$, Z_{i1} and Z_{i2} are independent and following uniform distributions on $[-1, 1]$ and $[0, 1]$ respectively, for $i = 1, 2, \dots, n$, $\beta_1 = 1, \beta_2 = -1$, $\gamma(t) = \sqrt{2} \sin(\pi t/2) + 3\sqrt{2} \sin(3\pi t/2)$, $\varepsilon_n \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$.

We suppose the functional predictors can be expressed as $X_i(t) = \sum_{j=1}^{50} U_{ij} v_j(t)$, where U_{ij} are independently distributed as the normal with mean 0 and variance $\lambda_j = ((j - 0.5)\pi)^{-2}$, $v_j(t) = \sqrt{2} \sin((j - 0.5)\pi t)$. For the actual observations, we assume that they are realizations of $\{X_i(\cdot)\}$ at an equally spaced grid of 100 points in $[0, 1]$. As we have said in Section 2, the truncation parameters m are selected by AIC criterion in our simulation. Similar to Lee [16] and Case [28], we focus on the spatial scenario with R number of districts, q members in each district, and with each neighbor of a member in a district given equal weight, that is, $W_n = \mathbf{I}_R \otimes B_q$, where $B_q = (\mathbf{I}_q \mathbf{I}_q^T - \mathbf{I}_q) / (q - 1)$, \mathbf{I}_q is the q -dimensional unit vector, and \otimes is a Kronecker product. Some simulation studies are examined with different values of R for 50 and 70, q for 2, 5, and 8, and σ^2 for 0.25 and 1. For comparison, three different values $\rho = 0.2, 0.5, 0.7$ are considered, which represent spatial dependence of the responses from weak to strong. $\rho = 0.2$ represents weak spatial dependence, and $\rho = 0.5$ represents mild spatial dependence, whereas $\rho = 0.7$ represents relatively strong spatial dependence.

Throughout the simulations, for different scalar parameters ρ, β_1 and β_2 , we use the average bias, standard deviation (SD) as a measure of parametric estimation accuracy. The performance of the estimator of the slope function $\gamma(t)$ is assessed using the square root of average squared errors (RASE), defined as

$$\text{RASE} = \left\{ \frac{1}{N} \sum_{l=1}^N [\hat{\gamma}(t_l) - \gamma(t_l)]^2 \right\}^{1/2},$$

where $\{t_l, l = 1, \dots, N\}$ are the regular grid points at which the function $\hat{\gamma}(t)$ is evaluated. In our simulation, $N = 200$ is used.

The sample size is $n = Rq$. We use 1000 Monte Carlo runs for estimation assessment, and then summarize the results in Tables 1–3 and Figures 1 and 2. Tables 1–3 list average Bias and SD of the estimators of ρ , β_1 , and β_2 , and average RASE of the estimator of $\gamma(t)$ in the 1000 replications. Figures 1 and 2 present the average estimate curves of $\gamma(t)$.

From Tables 1–3 and Figures 1 and 2 we can see that: (1) The biases of $\hat{\rho}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ are fairly small for almost all cases. (2) The standard deviation of $\hat{\rho}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ decrease as either R or q increases. (3) The RASEs of $\gamma(t)$ are small for all cases and decrease as sample size n increases or σ^2 decreases, and it can be concluded that the estimate curves fit better to the corresponding true line, which coincides with what was discovered from Figures 1 and 2. Overall, the simulation results suggest that the proposed estimation procedure is effective for the partial functional linear spatial autoregressive model.

Table 1. Simulation results for $\rho = 0.2$.

σ^2	R	Est	$q = 2$			$q = 5$			$q = 8$		
			Bias	SD	RASE	Bias	SD	RASE	Bias	SD	RASE
0.25	50	$\hat{\rho}$	-1.0×10^{-4}	0.022		-1.4×10^{-4}	0.023		-0.001	0.021	
		$\hat{\beta}_1$	-0.002	0.045		-6.5×10^{-4}	0.028		4.0×10^{-4}	0.022	
		$\hat{\beta}_2$	0.003	0.048		-0.001	0.034		-0.002	0.030	
		$\hat{\gamma}(t)$			0.400			0.255			0.197
	70	$\hat{\rho}$	-1.4×10^{-4}	0.018		-0.001	0.019		-9.2×10^{-5}	0.019	
		$\hat{\beta}_1$	5.1×10^{-4}	0.038		-2.2×10^{-4}	0.023		1.5×10^{-4}	0.018	
$\hat{\beta}_2$		7.8×10^{-4}	0.042		4.5×10^{-4}	0.029		-1.8×10^{-4}	0.026		
	$\hat{\gamma}(t)$			0.345			0.211			0.168	
1	50	$\hat{\rho}$	-5.4×10^{-4}	0.084		-0.006	0.092		-0.009	0.086	
		$\hat{\beta}_1$	-0.012	0.179		-0.005	0.113		6.6×10^{-4}	0.087	
		$\hat{\beta}_2$	0.016	0.188		-0.006	0.136		-0.009	0.118	
		$\hat{\gamma}(t)$			0.601			0.386			0.296
	70	$\hat{\rho}$	-0.001	0.070		-0.009	0.075		-0.004	0.075	
		$\hat{\beta}_1$	-5.8×10^{-4}	0.151		-0.002	0.092		1.6×10^{-4}	0.072	
$\hat{\beta}_2$		0.003	0.166		-3.7×10^{-4}	0.117		-0.003	0.104		
	$\hat{\gamma}(t)$			0.517			0.317			0.251	

Table 2. Simulation results for $\rho = 0.5$.

σ^2	R	Est	$q = 2$			$q = 5$			$q = 8$		
			Bias	SD	RASE	Bias	SD	RASE	Bias	SD	RASE
0.25	50	$\hat{\rho}$	-2.0×10^{-4}	0.017		-1.3×10^{-4}	0.015		-8.0×10^{-4}	0.014	
		$\hat{\beta}_1$	-0.002	0.046		-6.3×10^{-4}	0.029		4.6×10^{-4}	0.022	
		$\hat{\beta}_2$	0.003	0.051		-0.001	0.035		-0.002	0.030	
		$\hat{\gamma}(t)$			0.401			0.255			0.197
	70	$\hat{\rho}$	-1.9×10^{-4}	0.014		-8.9×10^{-4}	0.013		-7.2×10^{-5}	0.012	
		$\hat{\beta}_1$	6.0×10^{-4}	0.039		-1.0×10^{-4}	0.023		1.6×10^{-4}	0.018	
$\hat{\beta}_2$		6.0×10^{-4}	0.045		3.0×10^{-4}	0.030		-2.0×10^{-4}	0.027		
	$\hat{\gamma}(t)$			0.346			0.211			0.168	
1	50	$\hat{\rho}$	-0.002	0.066		-0.004	0.062		-0.006	0.056	
		$\hat{\beta}_1$	-0.010	0.182		-0.004	0.113		0.001	0.088	
		$\hat{\beta}_2$	0.013	0.200		-0.008	0.141		-0.010	0.121	
		$\hat{\gamma}(t)$			0.611			0.387			0.297
	70	$\hat{\rho}$	-0.002	0.055		-0.006	0.051		-0.003	0.049	
		$\hat{\beta}_1$	6.2×10^{-4}	0.153		-0.001	0.093		3.8×10^{-4}	0.073	
$\hat{\beta}_2$		6.7×10^{-4}	0.178		-0.002	0.121		-0.004	0.107		
	$\hat{\gamma}(t)$			0.525			0.317			0.251	

Table 3. Simulation results for $\rho = 0.7$.

σ^2	R	Est	$q = 2$		$q = 5$			$q = 8$			
			Bias	SD	RASE	Bias	SD	RASE	Bias	SD	RASE
0.25	R = 50	$\hat{\rho}$	-1.9×10^{-4}	0.012		-9.4×10^{-5}	0.010		-5.0×10^{-4}	0.009	
		$\hat{\beta}_1$	-0.001	0.047		-6.1×10^{-4}	0.029		5.1×10^{-4}	0.022	
		$\hat{\beta}_2$	0.003	0.053		-0.001	0.036		-0.002	0.031	
		$\hat{\gamma}(t)$			0.402			0.255			0.197
	R = 70	$\hat{\rho}$	-1.7×10^{-4}	0.010		-5.7×10^{-4}	0.008		-5.0×10^{-5}	0.007	
		$\hat{\beta}_1$	7.0×10^{-4}	0.040		-2.0×10^{-5}	0.023		1.6×10^{-4}	0.018	
$\hat{\beta}_2$		4.4×10^{-4}	0.047		1.9×10^{-4}	0.031		-2.2×10^{-4}	0.027		
	$\hat{\gamma}(t)$			0.348			0.211			0.168	
1	R = 50	$\hat{\rho}$	-0.003	0.046		-0.003	0.039		-0.004	0.035	
		$\hat{\beta}_1$	-0.009	0.187		-0.004	0.114		0.002	0.088	
		$\hat{\beta}_2$	0.010	0.210		-0.009	0.145		-0.011	0.123	
		$\hat{\gamma}(t)$			0.622			0.389			0.298
	R = 70	$\hat{\rho}$	-0.002	0.038		-0.004	0.032		-0.002	0.030	
		$\hat{\beta}_1$	0.002	0.157		-6.9×10^{-4}	0.093		5.5×10^{-4}	0.073	
$\hat{\beta}_2$		-0.002	0.187		-0.003	0.124		-0.004	0.109		
	$\hat{\gamma}(t)$			0.534			0.318			0.251	

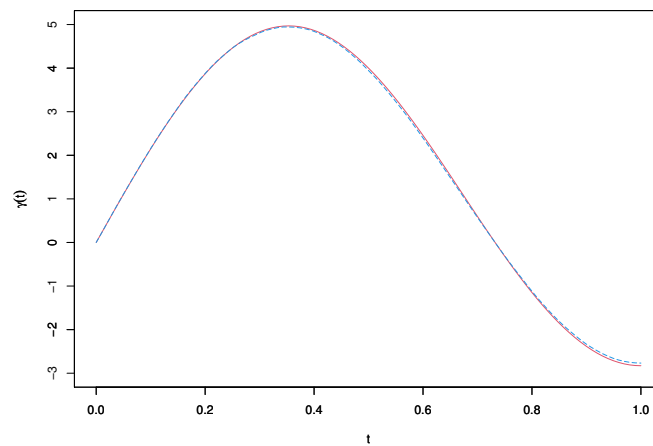


Figure 1. Simulation result of $\hat{\gamma}(t)$ when $\rho = 0.2$, $R = 50$, $q = 5$, $\sigma^2 = 0.25$. The solid curve denotes the true curve, the dash curve denotes its estimate.

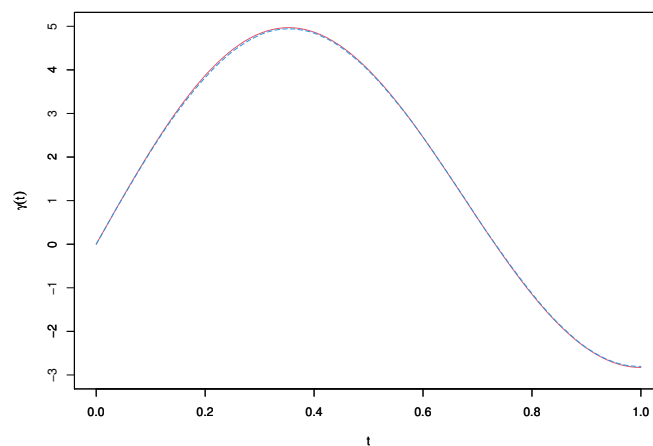


Figure 2. Simulation result of $\hat{\gamma}(t)$ when $\rho = 0.7$, $R = 70$, $q = 5$, $\sigma^2 = 1$. The solid curve denotes the true curve, the dash curve denotes its estimate.

5. Conclusions

In this paper, we proposed a partial functional linear spatial autoregressive model to study the link between a scalar dependence spatial response variable and explanatory variables containing both multiple real-valued scalar variables and a functional predictor. We then used functional principal component basis and an instrumental variable to estimate the parametric vector and slope function based on the two-stage least squares procedure. Under some mild conditions, we obtained the asymptotic normality of estimators of a parametric vector. Furthermore, the rate of convergence of the proposed estimator of slope function was also established. The simulation studies demonstrate that the proposed method performs satisfactorily and the theoretical results are valid.

There are some interesting future directions. In this paper, we only considered the estimation of the unknown parametric vector and slope function, which does not present a way to test for the effects of the covariates, an important aspect of any statistical analysis. In the future, we would like to be able to identify the model structure by testing for the main effects of the scalar predictors and the functional predictor. Another interesting direction can be to extend our new procedure to the generalized partial functional linear spatial autoregressive model.

Author Contributions: Conceptualization, Y.H.; Software, S.W.; methodology, Y.H. and S.F.; writing—original draft preparation, Y.H. and J.J.; writing—review and editing, S.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Social Science Foundation of China (18BTJ021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ramsay, J.O.; Silverman B.W. *Functional Data Analysis*; Springer: New York, NY, USA, 1997.
2. Cardot, H.; Ferraty, F.; Sarda, P. Spline estimators for the functional linear model. *Statist. Sin.* **2003**, *13*, 571–592.
3. Yao, F.; Müller, H.G.; Wang, J.L. Functional linear regression analysis for longitudinal data. *Ann. Statist.* **2005**, *33*, 2873–2903. [[CrossRef](#)]
4. Lian, H.; Li, G. Series expansion for functional sufficient dimension reduction. *J. Multivariate Anal.* **2014**, *124*, 150–165. [[CrossRef](#)]
5. Fan, Y.; James, G.M.; Radchenko, P. Functional additive regression. *Ann. Statist.* **2015**, *43*, 2296–2325. [[CrossRef](#)]
6. Feng, S.Y.; Xue, L.G. Partially functional linear varying coefficient model. *Statistics* **2016**, *50*, 717–732. [[CrossRef](#)]
7. Kong, D.; Xue, K.; Yao, F.; Zhang, H.H. Partially functional linear regression in high dimensions. *Biometrika* **2016**, *103*, 147–159. [[CrossRef](#)]
8. Yu, P.; Du, J.; Zhang, Z.Z. Varying-coefficient partially functional linear quantile regression models. *J. Korean Statist. Soc.* **2017**, *46*, 462–475. [[CrossRef](#)]
9. Zhang, D.; Lin, X.; Sowers, M.F. Assessing the effects of reproductive hormone profiles on bone mineral density using functional two-stage mixed models. *Biometrics* **2007**, *63*, 351–362. [[CrossRef](#)] [[PubMed](#)]
10. Shin, H. Partial functional linear regression. *J. Statist. Plann. Inference* **2009**, *139*, 3405–3418. [[CrossRef](#)]
11. Lu, Y.; Du, J.; Sun, Z. Functional partially linear quantile regression model. *Metrika* **2014**, *77*, 317–332. [[CrossRef](#)]
12. Yu, P.; Zhang, Z.Z.; Du, J. A test of linearity in partial functional linear regression. *Metrika* **2016**, *79*, 953–969. [[CrossRef](#)]
13. Cliff, A.; Ord, J.K. *Spatial Autocorrelation*; Pion: London, UK, 1973.
14. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer: Dordrecht, The Netherlands, 1988.
15. Cressie, N. *Statistics for Spatial Data*; John Wiley & Sons: New York, NY, USA, 1993.
16. Lee, L.F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial econometric models. *Econometrica* **2004**, *72*, 1899–1926. [[CrossRef](#)]
17. Su, L.; Jin, S. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *J. Econom.* **2010**, *157*, 18–33. [[CrossRef](#)]

18. Koch, M.; Krisztin, T. Applications for asynchronous multi-agent teams in nonlinear applied spatial econometrics. *J. Internet Technol.* **2014**, *12*, 1007–1014.
19. Chen, J.; Wang, R.; Huang, Y. Semiparametric spatial autoregressive model: A two-step bayesian approach. *Ann. Public Health Res.* **2015**, *2*, 1012–1024.
20. Du, J.; Sun, X.Q.; Cao, R.Y.; Zhang, Z.Z. Statistical inference for partially linear additive spatial autoregressive models. *Spat. Statist.* **2018**, *25*, 52–67. [[CrossRef](#)]
21. Dauxois, J.; Pousse, A.; Romain, Y. Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *J. Multivariate Anal.* **1982**, *12*, 136–154. [[CrossRef](#)]
22. Cardot, H.; Ferraty, F.; Sarda, P. Functional linear model. *Statist. Probab. Lett.* **1999**, *45*, 11–22. [[CrossRef](#)]
23. Hall, P.; Horowitz, J.L. Methodology and convergence rates for functional linear regression. *Ann. Statist.* **2007**, *35*, 70–91. [[CrossRef](#)]
24. Hall, P.; Hosseini-nasab, M. Theory for high-order bounds in functional principal components analysis. *Math. Proc. Cambridge Philos. Soc.* **2009**, *146*, 225–256. [[CrossRef](#)]
25. Kelejian, H.H.; Prucha, I.R. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *J. Real Estate Financ.* **1998**, *17*, 99–121. [[CrossRef](#)]
26. Zhang, Y.Q.; Shen, D.M. Estimation of semi-parametric varying-coefficient spatial panel data models with random-effects. *J. Statist. Plann. Inference* **2015**, *159*, 64–80. [[CrossRef](#)]
27. Hu, Y.P.; Xue, L.G.; Zhao, J.; Zhang, L.Y. Skew-normal partial functional linear model and homogeneity test. *J. Statist. Plann. Inference* **2020**, *204*, 116–127. [[CrossRef](#)]
28. Case, A.C. Spatial patterns in household demand. *Econometrica* **1991**, *59*, 953–965. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).