

Function-on-Partially Linear Functional Additive Models

Jinyou Huang, Shuang Chen*

School of Sciences, Hebei University of Technology, Tianjin, China

Email: 1324253788@qq.com, *shchen@bjpt.edu.cn

How to cite this paper: Huang, J.Y. and Chen, S. (2020) Function-on-Partially Linear Functional Additive Models. *Journal of Applied Mathematics and Physics*, 8, 1-9.

<https://doi.org/10.4236/jamp.2020.81001>

Received: December 1, 2019

Accepted: December 23, 2019

Published: December 26, 2019

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We consider a functional partially linear additive model that predicts a functional response by a scalar predictor and functional predictors. The B-spline and eigenbasis least squares estimator for both the parametric and the nonparametric components proposed. In the final of this paper, as a result, we got the variance decomposition of the model and establish the asymptotic convergence rate for estimator.

Keywords

Functional Data Analysis, Functional Principal Component Analysis, Partial Linear Regression Models, Penalized B-Splines, Variance Model

1. Introduction

Function data are infinite dimensional vectors in functional space, the study of functional data helps people to further understand data changes in finance, medicine, etc. The study of functional response can date back to the work Ramsay and Silverman (2005) [1] where the model the functional concurrent model considers current response relate to the current values of the covariates. Wong *et al.* (2018) [2] investigated a class of partially linear functional additive models (PLFAM) that predicts a scalar response by both parametric effects of a multivariate predictor and nonparametric effects of a multivariate functional predictor. Further an additive function-on-function regression is established using principal component basis functions and b-spline basis functions by Janet S. Kim *et al.* (2016) [3]. Luo, R., Qi, X. (2017) [4] consider functional linear regression models with a functional response and multiple functional predictors, with the goal of finding the best finite-dimensional approximation to the signal part of the response function. This paper extends function on function regression, con-

sidering functional partially linear additive models that predict a functional response by a scalar predictor and functional predictors. Specifically we consider the model

$$Y(t) = \mu(t)\gamma(t) + \sum_{j=1}^p \int_{\tau_x} \beta_j(s,t)x_j(s)ds + \varepsilon(t) \tag{1}$$

where $Y(t)$ is the response function defined in an interval τ_y , For convenience, we assume that the response has zero-mean, the predictive curves.

$x_1(s), x_2(s), \dots, x_p(s)$ are defined in τ_x . where τ_y and τ_x are compact intervals. $\mu(t)$ is scalar predictor defined in an interval τ_y . $\gamma(t)$ is scalar coefficient, $\varepsilon(t)$ is the noise function with mean zero and unknown autocovariance function $A(t, t')$ and is independent of the predictor and the response. In the previous papers, these focused on function-on-function, scalar-on-function, and scalar-mixed data. There is no paper devoted to function-on-mixed data. That is to say, the functional dependent variable analyzes the regression model of the mixed data. This paper considers the actual situation is closer to the actual demand. Considering that the real world data cannot be all the functional data, there will definitely be a part of the scalar in the model. From this perspective, expand the new model. This paper uses a different estimation method from the previous. The general method adopted by these articles is to use non-parametric methods such as spline function to approximate the model generation estimation parameters, and at the same time minimize the objective function. In the past, the model was generally complicated and inconvenient to calculate. They consider fitting with B-splines in three directions at the same time, and there will be over-fitting when doing so. In this paper, the problem is optimized by using the eigenbasis and orthogonal B-spline, that is, using the special feature base of $Y(t)$ and an orthogonal B-spline to fit the model. While ensuring the fitting effect, the complexity of the model is reduced, and the model is not over-fitting. We use the classical formula of probability to express the error term of the model from the perspective of variance, and the method of resampling is used to approximate the error of the model. At the same time, assume from the article hypothesis, we establish the asymptotic convergence rate for estimator.

2. Model

In this section, we consider the estimation of coefficients in a partial linear regression model of a function. In this paper, the model is estimated by using the eigenbasis spline and the B-spline function. Note ϕ_k is the corresponding eigenfunctions of $y(t)$, and

$$\int_0^1 \phi_n(t)\phi_m(t)dt = \begin{cases} 1, n = m \\ 0, n \neq m \end{cases} \tag{2}$$

Let

$$h(t) = \mu(t)\gamma(t) \tag{3}$$

we project $h(t)$ onto the eigenbasis of $y(t)$,

$$H_k = \int_{\tau_y} h(t) \phi_k(t) dt \tag{4}$$

then we can express $h(t)$ as

$$h(t) = \sum_{k=1}^{\infty} H_k \phi_k(t) \tag{5}$$

For the function coefficient term part of (1), we estimate it by means of B-spline and feature base. For convenience we write model (1) as

$$y(t) = \mu(t)\gamma(t) + \int_{\tau_x} X(s)^T \boldsymbol{\beta}(s,t) ds + \varepsilon(t) \tag{6}$$

where $X(s) = (x_1(s), \dots, x_2(s))^T$ is the p -dimensional vector of predictive functions, $\boldsymbol{\beta}(s,t) = (\beta_1(s,t), \dots, \beta_p(s,t))$ is the p -dimensional vector of coefficient functions. Let

$$F(s,t) = X(s)^T \boldsymbol{\beta}(s,t) \tag{7}$$

Transform $F(s,t)$ in the same way as $h(t)$

$$G_k(s) = \int_{\tau_y} F(s,t) \phi_k(t) dt \tag{8}$$

So we can put $F(s,t)$ similar expression of Karhunen-Loeve expansion

$$F(s,t) = \sum_{k=0}^{\infty} G_k(s) \phi_k(t) dt \tag{9}$$

In this paper, on spectral decomposition method projecting the corresponding function to the orthogonal eigenbasis of y , we estimate the parameter of the model can be performed without punished the complexity of the t direction. Generally, we preset the maximum variance number similar to the method in the functional principal component analysis, usually 85%. Next we will use the Orthogonal b-spline method (2010) [5] [6] to represent the parameter items in the s direction.

$$G_k(s) = \sum_{l=1}^M B_l(s) \theta_{l,k} \tag{10}$$

$$F(s,t) = \sum_{k=0}^{\infty} \sum_{l=1}^M B_l(s) \theta_{l,k} \phi_k(t) \tag{11}$$

where $B_l(s)$ are or-thogonalized B-spline bases of dimensions M , $\theta_{l,k}$ is unknown parameters, we get

$$y(t) = \sum_{k=1}^{\infty} H_k \phi_k(t) + \int_{\tau_x} \sum_{k=0}^{\infty} \sum_{l=1}^M B_l(s) \phi_k(t) ds + \varepsilon(t) \tag{12}$$

In the t direction perform truncation K for the preset maximum variance, we get

$$y(t) \approx \sum_{k=1}^K H_k \phi_k(t) + \int_{\tau_x} \sum_{k=0}^K \sum_{l=1}^M B_l(s) \theta_{l,k} \phi_k(t) ds \tag{13}$$

We denote $Z = (Z_1, Z_2, \dots, Z_p)$, Z_l as $\int_{\tau_x} B_l(s) ds$, let $\Phi_k = (\theta_{1k}, \theta_{2k}, \dots, \theta_{pk})$.

Then we can get the simplified expression of $y(t)$

$$y(t) \approx \sum_{k=1}^K H_k \phi_k(t) + \sum_{k=1}^k \mathbf{Z} \Phi_k \phi_k(t) \tag{14}$$

3. Parameter Estimation

We assume that with probability 1, the trajectory of x_n is contained in a Hilbert space χ_n , with inner product. We will focus on the case that χ_j 's are L^2 functional spaces and the inner products are $\langle f, g \rangle = \int f(t)g(t)dt$ for any $f, g \in \chi_n$. We use quadratic penalties for the direction s and control the roughness in the direction t by the preset number of orthogonal basis functions. so the loss function is as follow

$$L(H_k, \Phi_k) = \left(y(t) - \sum_{k=1}^K h_k \phi_k(t) - \sum_{k=1}^K \mathbf{Z}^T \Phi_k \phi_k(t) \right) \tag{15}$$

We use the least squares penalty to punish the curvature of the functional parameter, considering the penalty term

$$\iint \left(\frac{\partial^2 \mathbf{F}(s,t)}{\partial s^2} \right)^2 dsdt = \iint \left(\frac{\partial^2 \sum_{k=1}^K \mathbf{G}_k(s) \phi_k(t)}{\partial s^2} \right)^2 dsdt = \sum_{k=1}^K \int \left(\frac{\partial^2 \mathbf{G}_k(s)}{\partial s^2} \right)^2 ds \tag{16}$$

$$\sum_{k=1}^K \int \left(\frac{\partial^2 \mathbf{G}_k(s)}{\partial s^2} \right)^2 ds = \lambda \sum_{k=1}^K \int \left(\frac{\partial^2 \mathbf{B}_l(s)}{\partial s^2} \right)^2 \theta_{l,k}^2 ds = \lambda \Phi_k^T \mathbf{P} \Phi_k \tag{17}$$

where $\mathbf{P} = \sum_{l=1}^M \int \frac{\partial^2 \mathbf{B}_l(s)}{\partial s^2} ds$. Then we can get the penalized criterion

$$\left\| y(t) - \sum_{k=1}^K H_k \phi_k(t) + \sum_{k=1}^k \mathbf{Z}^T \Phi_k \phi_k(t) \right\|^2 + \lambda \Phi_k \mathbf{P} \Phi_k \tag{18}$$

Similarly $\mathbf{F}(s,t)$ and $h(t)$ to process $y(t)$. We have Karhunen-Loeve expansion on $y(t)$

$$y(t) = \sum_{k=1}^K \xi_k \phi_k(t) + e_t \tag{19}$$

where e_t is zero-mean error, and the variables ξ_k are the FPCA (Functional principal component) scores of Y . And note that $\int \phi_k(t)^2 dt = 1$, so the criterion can be simplify written as

$$\sum_{k=1}^K \left\| \xi_k - H_k + \mathbf{Z}^T \Phi_k \right\|^2 + \lambda \Phi_k \mathbf{P} \Phi_k \tag{20}$$

So the estimation process for the parameters H_k, Φ_k is as follows.

1) We get the function curve $\hat{y}(t)$ for smooth sample points, and we de-mean $\hat{Y}(t)$ for get a curve $\hat{y}^c(t)$ with a mean of 0. The point K is determined by the pre-specified percentage of variance.

2) We use FPCA(Functional Principal Component) to estimate eigenbasis $\hat{\phi}_k$, and we project $\hat{y}(t)$ onto the corresponding eigenbasis $\hat{\phi}_k$ to obtain an estimation of the principal component score $\hat{\xi}_k = \int_{\tau,y} \hat{y}^c(t) \hat{\phi}_k(t) dt$.

3) We get a two-step estimation for H_k, Φ_k

a) Obtain estimation of $\Phi_k, k = 1, 2, 3, \dots, K$, of the coefficients by minimizing the penalized criterion(19) with respect to Φ_k and H_k , The estimated value we get for Φ_k is

$$\hat{\Phi}_k = \frac{\mathbf{Z}(\hat{\xi}_k - H_k)}{\mathbf{Z}\mathbf{Z}^T + \lambda\mathbf{P}} \tag{21}$$

b) Obtain estimation of H_k by minimizing

$$\arg \min \sum_{k=1}^K \|\hat{\xi}_k - H_k - \mathbf{Z}^T \hat{\Phi}_k\| \tag{22}$$

So the estimation of Φ_k is $\hat{\Phi}_k$ and H_k is \hat{H}_k . We can get

$$\hat{y}(t) = \sum_{k=1}^K \hat{\phi}_k(t) (\hat{H}_k + \mathbf{Z}^T \hat{\Phi}_k) \tag{23}$$

4. Error Variance Decomposition

In this section, we get error variance decomposition. Let $A(t, t')$ be the variance function of $\varepsilon(t)$.we estimate $A(t, t')$.

1) obtain residuals by fit (14), $e_j = y(t_j) - \hat{y}(t_j)$.

2) Use FPCA for the residual function, to approximate the infinite covariance function using the estimated finite covariance function.

Let x_{new}, μ_{new} be new observation independent of X and $\varepsilon(t)$, and fit function-on-partially linear functional additive models,then we measure the uncertainty in the prediction by predicting error $\hat{y}_{new}(t) - y_{new}(t)$.Base on Ruppert *et al.* (2009) [7]

$$\text{var}\{\hat{y}_{new}(t) - y_{new}(t)\} = \text{var}\{\hat{y}_{new}(t)\} + \text{var}\{\varepsilon_{new}(t)\} \tag{24}$$

We estimate the expression of $\text{var}\{\varepsilon_{new}(t)\}$ by estimating $\text{var}\{\varepsilon(t)\}$ Then we estimate $\text{var}\{\hat{y}_{new}(t)\}$ using the classical variance formula

$$\text{var}\{\hat{\varepsilon}_0(t)\} = E_{\pi} [\text{var}\{\hat{\varepsilon}_0(t) | \pi\}] + \text{var}_{\pi} [E\{\hat{\varepsilon}_0(t) | \pi\}] \tag{25}$$

In this paper, we assume that π is a known parameter set containing estimated arbitrary parameter values and corresponding parameter variance and covariance functions. So we will expand $\text{var}\{\hat{\varepsilon}_0(t) | \pi\}$ of the formula (24) as follows

$$\text{var}\{\hat{y}_{new}(t) | \pi\} = \text{var} \sum_{k=1}^K \hat{\phi}_k(t) (\hat{H}_{new,k} + \mathbf{Z}_{new}^T \hat{\Phi}_k) \tag{26}$$

If we write $\frac{\mathbf{Z}}{\mathbf{Z}^T \mathbf{Z} + \lambda \mathbf{P}}$ as m_{λ} , $\mathbf{Z}^T m_{\lambda}$ is recorded as J_{λ} , then

$$\hat{\Phi}_k = \frac{\mathbf{Z}(\hat{\xi}_k - H_k)}{\mathbf{Z}^T \mathbf{Z} + \lambda \mathbf{P}} = m_{\lambda} (\hat{\xi}_k - H_k) \text{ and}$$

$\hat{H}_k + \mathbf{Z}_{new}^T \hat{\Phi}_k = \hat{H}_k + \mathbf{Z}_{new}^T m_{\lambda} (\hat{\xi}_k - \hat{H}_k) = (1 - J_{\lambda}) \hat{H}_k + J_{\lambda} \hat{\xi}_k$. So we expand the expression (25) as follow

$$\begin{aligned} \text{var} \{ \hat{y}_{new}(t) | \pi \} &= \sum_{k=1}^K \hat{\phi}_k(t) \text{var} \left(\hat{H}_k + \mathbf{Z}_{new}^T \hat{\Phi}_k \right) \hat{\phi}_k(t) \\ &+ \sum_{k \neq k'} \hat{\phi}_k(t) \text{cov} \left(\hat{H}_k + \mathbf{Z}_{new}^T \hat{\Phi}_k, \hat{H}_{k'} + \mathbf{Z}_{new}^T \hat{\Phi}_{k'} \right) \hat{\phi}_{k'}(t) \end{aligned} \tag{27}$$

It can be seen that all the variances and covariance functions in the expression are represented by the variance and covariance functions of the existing estimates. Then we focus on the second item of expression (24), the second term is $\text{var}_{\pi} \left[E \{ \hat{\epsilon}_0(t) | \pi \} \right]$. This paper hopes to obtain an approximation form of the finite term of the variance function under finite samples by resampling the bootstrap method. We perform the bootstrap algorithm through the following steps;

- 1) For $q = 1, 2, 3, \dots, Q$.
- 2) Resample the samples to get Q groups different sample.
- 3) Smooth the resample points and get the corresponding curve $X_j^q(s), \mu^q(t), y^q(t)$.
- 4) Obtain the de-mean $(\hat{y}^c)^q(t)$ and the corresponding feature basis function $\hat{\phi}_k^q$, and obtain the corresponding principal component score $\hat{\xi}_k^q$.
- 5) Fit the function-on-partially linear functional additive models to obtain the corresponding parameter estimation.
- 6) Get prediction function $\hat{y}_{new}^q(t)$ to any new observation curve $x_{new}(t), y_{new}(t)$ and calculate $\hat{A}^q(t, t')$.
- 7) Get $\text{var} \{ \hat{y}_{new}^q(t) | \pi \}$ by expression (24).

Thus the approximate value of the finite term of $A(t, t')$ is found

$$\begin{aligned} \hat{A}(t, t') &= \frac{\sum_{q=1}^Q \hat{A}^q(t, t')}{Q} \text{, and the variance estimate of the error} \\ &\widehat{\text{var}} \{ \hat{y}_{new}(t) - y_{new}(t) \} \\ &= \frac{\sum_{q=1}^Q \widehat{\text{var}} \{ y_{new}^q(t) | \pi^q \}}{Q} + \frac{\sum_{q=1}^Q \{ y_{new}^q(t) - \bar{y}_{new}(t) \}^2}{Q} + \frac{\sum_{q=1}^Q \hat{A}^q(t, t')}{Q} \end{aligned} \tag{28}$$

5. Basic Assumptions and Convergence Properties

In this section, we study the asymptotic properties of $F(t), h(t)$, where

$$\hat{h}(t) = \sum_{k=1}^K \hat{H}_k \hat{\phi}_k(t) \tag{29}$$

$$\begin{aligned} F(t) &= \int_{\tau_x} F(s, t) ds = \int_{\tau_x} \sum_{k=1}^K G(s) \phi_k(t) ds \\ &= \int_{\tau_x} \sum_{k=1}^K \sum_{l=1}^M B_l(s) \theta_{l,k} \phi_k(t) ds \\ &= \sum_{k=1}^K \mathbf{Z} \Phi_k \phi_k(t) \end{aligned} \tag{30}$$

5.1. Basic Assumption

- 1) For the square integrable random function $\tilde{X}(s)$ satisfies

- $E \|\tilde{X}(s)\|^q < \infty, 2 < q < \infty$;
- 2) For each k , there is $E[\tilde{U}_{ik}] \leq c\lambda_k^2$, where c is a constant and c can take different values in different expressions;
- 3) For the eigenvalue λ_k , satisfied $c^{-1}k^{-s} \leq \lambda_k \leq ck^{-s}, k \geq 1, s > 1$;
- 4) For the coefficient H_k , assume there is a constant t , and $t > \frac{s}{3} + 1$, which make H_k satisfied $H_k < ck^{-s}$;
- 5) For the parameter K , assume $K = \frac{1}{n^{s+3t}}$
- 6) The number of knots $N_n = n^r$, with $\frac{1}{2q+1} \leq r \leq \frac{1}{2}$.
- 7) The distribution of $F(t)$ and $h(t)$ is absolutely continuous and its density is bounded on τy ;
- 8) The joint density function of $x(s)$ is bounded on τx ;
- 9) Covariance function $\text{cov}(y(t), y(t'))$ of $y(t)$ is bounded on interval τy

Conditions (1)-(5) are common in functional PCA literature and conditions (6)-(9) are common in nonparametric regression literature.

5.2. Related Lemma

Lemma 1. *If $\tilde{F}(t)$ is spline expansion of $F(t)$, it is known by Chen [8] that there is $n > 2, h = \frac{1}{N}$, we can get*

$$\|\tilde{F}(t) - F(t)\|^2 = O(h^2) < O(h^n) \tag{31}$$

5.3. Convergence Properties

Consider the *model* (6)

$$y(t) = \mu(t)\gamma(t) + \int_{\tau x} \mathbf{X}(s)^\top \boldsymbol{\beta}(s, t) ds + \varepsilon(t) \tag{32}$$

Use the principal component of $y(t)$ for the model (6) to represent the following expression

$$\begin{aligned} \sum_{k=1}^{\infty} \tilde{\xi}_k \phi_k(t) &= \sum_{k=1}^{\infty} H_k \phi_k(t) + \sum_{k=1}^{\infty} \mathbf{Z}^\top \boldsymbol{\Phi}_k \phi_k(t) + \varepsilon(t) \\ &= \sum_{k=1}^{\infty} (H_k, \boldsymbol{\Phi}_k) (1, \mathbf{Z})^\top \phi_k(t) + \varepsilon(t) \\ &= \sum_{k=1}^{\infty} D_k \mathbf{Z}_0 \phi_k(t) + \varepsilon(t) \end{aligned} \tag{33}$$

where $D_k = (H_k, \boldsymbol{\Phi}_k)$ and $\mathbf{Z}_0 = (1, \mathbf{Z})$. We can get an estimation by least squares estimation of $\hat{D}_k = (\hat{H}_k, \hat{\boldsymbol{\Phi}}_k) = \frac{\mathbf{Z}_0 \hat{\xi}_k}{\mathbf{Z}_0 \mathbf{Z}_0^\top}$, let $\tilde{D}_k = (H_k, \tilde{\boldsymbol{\Phi}}_k)$, where

$$\tilde{F}(t) = \sum_{k=1}^K \tilde{\boldsymbol{\Phi}}_k \mathbf{Z}^\top \hat{\phi}_k(t) \text{ is spline expansion of } F(t).$$

Theorem 2. Under the conditions of(1)-(5), we can get

$$\|\hat{h}(t) - h(t)\|^2 = O\left(n^{\left(\frac{3t-1}{s+3t}\right)}\right), n \rightarrow \infty \tag{34}$$

Proof. Use Cauchy-Buniakowsky-Schwarz Inequality we can get

$$\begin{aligned} \|\hat{h}(t) - h(t)\|^2 &= \left\| \sum_{k=1}^K \hat{H}_k \hat{\phi}_k(t) - \sum_{k=1}^{\infty} H_k \phi_k(t) \right\|^2 \\ &\leq 2 \left\| \sum_{k=1}^K \hat{H}_k \hat{\phi}_k(t) - \sum_{k=1}^K H_k \phi_k(t) \right\|^2 + 2 \left\| \sum_{k=K+1}^{\infty} H_k \phi_k(t) \right\|^2 \\ &\leq 4 \left\| \sum_{k=1}^K \hat{H}_k (\hat{\phi}_k(t) - \phi_k(t)) \right\|^2 + 4 \left\| \sum_{k=1}^K (\hat{H}_k - H_k) \hat{\phi}_k(t) \right\|^2 + 2 \left\| \sum_{k=K+1}^{\infty} H_k \phi_k(t) \right\|^2 \\ &\leq 8K \sum_{k=1}^K (\hat{H}_k)^2 \|\hat{\phi}_k(t) - \phi_k(t)\|^2 + 4 \sum_{k=1}^K (\hat{H}_k - H_k)^2 + 2 \sum_{k=K+1}^{\infty} H_k^2 \end{aligned} \tag{35}$$

Assume from the article hypothesis, we have

$$\|\hat{h}(t) - h(t)\|^2 = O\left(n^{\left(\frac{3t-1}{s+3t}\right)}\right), n \rightarrow \infty \tag{36}$$

□

Theorem 3. Available under the conditions of(6)-(9), and let $\tilde{F}(t)$ is spline expansion of $F(t)$ we can get

$$\|\hat{F}(t) - F(t)\|_2^2 = O\left(\frac{MK}{n} + \rho^2\right) \tag{37}$$

Proof.

$$\|\hat{F}(t) - F(t)\|^2 \leq 2\|\hat{F}(t) - \tilde{F}(t)\| + 2\|\tilde{F}(t) - F(t)\| \tag{38}$$

Consider the latter item in the text, because $\tilde{F}(t)$ is a spline expansion of $F(t)$, by lemma we have

$$\|\tilde{F}(t) - F(t)\| = O(h^2) \tag{39}$$

Now consider the former item

$$\begin{aligned} \|\hat{F}(t) - \tilde{F}(t)\|^2 &= \left\| \sum_{i=1}^k Z^T \hat{\Phi}_k \hat{\phi}_k(t) - \sum_{i=1}^k Z^T \tilde{\Phi}_k \hat{\phi}_k(t) \right\|^2 \\ &\leq \sum_{i=1}^k \|Z^T \hat{\Phi}_k - Z^T \tilde{\Phi}_k\|^2 \leq \sum_{i=1}^k \|\hat{\Phi}_k - \tilde{\Phi}_k\|^2 \end{aligned} \tag{40}$$

Then consider

$$\sum_{i=1}^k \|\hat{\Phi}_k - \tilde{\Phi}_k\|^2 \leq \sum_{i=1}^k \left\| \frac{Z_0 \mathcal{E}}{Z_0 Z_0^T} \right\|^2 \leq \sum_{i=1}^k \left\| \frac{Z_0 \mathcal{E}}{n} \right\|^2 \tag{41}$$

Since $E[B_i(s_i)B_j(s_j)\varepsilon_i\varepsilon_j] = 0, i \neq j$,

$$E\left[\frac{Z_0 \mathcal{E}}{n}\right] = \frac{1}{n^2} E[\varepsilon Z_0^T Z_0 \varepsilon] = \frac{1}{n^2} \sum_{i=1}^M \sum_{i=1}^n E[B_i(s_i)\varepsilon_i]^2 \leq \frac{M}{n} \tag{42}$$

Get $\|\hat{F}(t) - F(t)\|^2 = O\left(\frac{MK}{n} + \rho^2\right)$ by the above formula. \square

6. Conclusion

This paper is a model extension of the function-to-function regression. The scalar variable is added to the dependent variable, which extends the application scope of the model. For real-life data, real data should include scalar data and function data, and the model used in this paper can be better explained. In this paper, the model is estimated by using the eigenbasis spline function and the orthogonal B-spline function. When the loss function is punished, the complexity of the t direction is controlled by controlling the preset principal component in the t direction, which reduces the complexity. It is more practical. At the same time, the variance decomposition of the error under the finite sample size is given. The approximation is performed by the resampling bootstrap method. Finally, the convergence properties of the estimated parameters are studied.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Billheimer, D. (2010) Functional Data Analysis, 2nd Edition Edited by J. O. Ramsay and B. W. Silverman. *Biometrics*, **63**, 300-301. https://doi.org/10.1111/j.1541-0420.2007.00743_1.x
- [2] Wong, R.K.W., Li, Y. and Zhu, Z. (2018) Partially Linear Functional Additive Models for Multivariate Functional Data. *Journal of the American Statistical Association*.
- [3] Kim, J.S., Staicu, A.M., Maity, A., *et al.* (2017) Additive Function-on-Function Regression. *Journal of Computational and Graphical Statistics*, **27**. <https://doi.org/10.1080/10618600.2017.1356730>
- [4] Luo, R. and Qi, X. (2016) Function-on-Function Linear Regression by Signal Compression. *Journal of the American Statistical Association*.
- [5] Chen, G.L. and Wang, Z.J. (2010) The Multivariate Partially Linear Model with B-Spline. *Chinese Journal of Applied Probability and Statistics*.
- [6] Benko, M., Hardle, W. and Kneip, A. (2009) Common Functional Principal Components. *Annals of Statistics*, **37**, 1-34. <https://doi.org/10.1214/07-AOS516>
- [7] Ruppert, D., Wand, M.P. and Carroll, R.J. (2009) Semiparametric Regression during 2003C2007. *Electronic Journal of Statistics*, **3**, 1193-1256. <https://doi.org/10.1214/09-EJS525>
- [8] Chen, T.P. (1985) Convergence and Asymptotic Expansion of Spline Functions. *Chinese Science Bulletin*, **30**, 1361-1364.