Research Article

Integral Least-Squares Inferences for Semiparametric Models with Functional Data

Limian Zhao and Peixin Zhao

College of Mathematics and Statistics, Hechi University, Yizhou, Guangxi 546300, China

Correspondence should be addressed to Limian Zhao; limianzhao@126.com

The inferences for semiparametric models with functional data are investigated. We propose an integral least-squares technique for estimating the parametric components, and the asymptotic normality of the resulting integral least-squares estimator is studied. For the nonparametric components, a local integral least-squares estimation method is proposed, and the asymptotic normality of the resulting estimator is also established. Based on these results, the confidence intervals for the parametric component and the nonparametric component are constructed. At last, some simulation studies and a real data analysis are undertaken to assess the finite sample performance of the proposed estimation method.

1. Introduction

In the recent literature, there has been increased interest in regression modeling for functional data, where both the predictor and response are random functions. Compared with the discrete multivariate analysis, functional data analysis can take into account the smoothness of the high dimensional covariates and can suggest some new approaches to the problems that have not been discovered before. Examples of functional data can be found in different application fields such as biomedicine, economics, and archaeology (see Ramsay and Silverman [1]). Furthermore, the statistical analysis for the regression model with functional data also has been considered by many authors. For example, Ramsay and Silverman [2] studied the linear regression model with functional data. Ait-Saïdi et al. [3] proposed a cross-validated estimation procedure for the single-functional index model. Ferraty et al. [4] and Chen et al. [5] considered the inferences for single and multiple index functional regression models by using the functional projection pursuit regression technology. In addition, Ferraty and Vieu [6] and Rachdi and Vieu [7] considered the nonparametric regression modeling for functional data. More works for the functional data analysis can be found in [8–10] and among others.

However, the linear functional model, which assumes that the model satisfies the linear relationship between the response and the covariates, may be too restrictive. Then, the semiparametric model with functional data is a useful extension of functional linear regression models and functional nonparametric regression models. More specifically, let $X(t)$, $Z(t)$, and $Y(t)$ be continuous random functions of index t; then, the semiparametric regression model with functional data has the following structure:

$$Y(t) = X(t)^T \theta(t) + Z(t)^T \beta + \epsilon(t), \tag{1}$$

where $Y(t)$ is the response variable, $X(t)$ is the $p \times 1$ covariate vector, $Z(t)$ is the $q \times 1$ covariate vector, $\beta = (\beta_1, \ldots, \beta_q)^T$ is a vector of unknown parameters, $\theta(t) = (\theta_1(t), \ldots, \theta_p(t))^T$ is a vector of unknown function of t, and $\epsilon(t)$ is a zero-mean stochastic process. Here, without loss of generality, we assume that index t ranges over a nondegenerate compact interval such as $[0, 1]$.

Because the samples of response and covariate are functions of index t, the Euclidean distance cannot measure the distance between $Y(t)$ and $X(t)$. Hence, the ordinary least-squares method cannot be implemented directly. Recently, the studies for such semiparametric regression model with functional data have been considered by many papers (see Aneiros-Pérez and Vieu [11], Shin [12], and Lian [13]). In this paper, we provide additional positive results of the inferences for semiparametric models with functional data

and extend the application literature of the classical least-squares technology. More specifically, we propose an integral least-squares method for estimating the parametric components and the nonparametric component. Furthermore, the asymptotic normalities of the integral profile least-squares estimators are studied. Some simulation studies and a real data application imply that the proposed method is workable.

The rest of this paper is organized as follows. In Section 2, we introduce the integral least-squares based estimation procedure for the parametric components and the nonparametric components. The asymptotic distributions of these estimators are also derived under some regularity conditions. In Section 3, some simulations and a real data analysis are carried out to assess the performance of the proposed estimation method. The technical proofs of all asymptotic results are provided in the Appendix.

2. Estimation and Asymptotic Distributions

Suppose $\{X_i(t), Z_i(t), Y_i(t)\}_{i=1}^n$ is a random sample of size n. From (1), we have that

$$Y_i(t) = X_i^T(t)\theta(t) + Z_i(t)^T\beta + \epsilon_i(t), \quad i = 1, \ldots, n. \quad (2)$$

In this paper, we assume that, for given t, $(Y_i(t), X_i(t), Z_i(t))$ and $\epsilon_i(t)$ are i.i.d. for different i and $E\{\epsilon_i(t) \mid X_i(t), Z_i(t)\} = 0$.

Now, we in turn model (2); for t in a small neighborhood of t_0, $\theta_k(t)$ can be locally approximated by a linear function

$$\theta_k(t) = \theta_k(t_0) + \theta_k'(t_0)(t - t_0) \equiv a_k + b_k(t - t_0), \quad (3)$$
$$k = 1, \ldots, p.$$

For given β, applying the integral least-squares method, we can get the weighted local integral least-squares estimator of $\{\theta_1(t), \ldots, \theta_p(t)\}$ by minimizing

$$\sum_{i=1}^n \int_0^1 \left\{ Y_i(s) - Z_i^T(s)\beta - \sum_{k=1}^p [a_k + b_k(s-t)] X_{ik}(s) \right\}^2 \times K_h(t-s)\, ds, \quad (4)$$

where $K_h(\cdot) = h^{-1}K(\cdot/h)$, $K(\cdot)$ is a kernel function, h is a bandwidth, and $X_{ik}(s)$ denote the kth component of $X_i(s)$. Let

$$D_{xy}(t) = \begin{pmatrix} \sum_{i=1}^n \int_0^1 X_i(s) Y_i(s) K_h(t-s)\, ds \\ \sum_{i=1}^n \int_0^1 X_i(s) Y_i(s) \dfrac{t-s}{h} K_h(t-s)\, ds \end{pmatrix},$$

$$D_{xz}(t) = \begin{pmatrix} \sum_{i=1}^n \int_0^1 X_i(s) Z_i^T(s) K_h(t-s)\, ds \\ \sum_{i=1}^n \int_0^1 X_i(s) Z_i^T(s) \dfrac{t-s}{h} K_h(t-s)\, ds \end{pmatrix}, \quad (5)$$

$$D_{xx}(t) = \begin{pmatrix} \sum_{i=1}^n \int_0^1 X_i(s) X_i^T(s) K_h(t-s)\, ds & \sum_{i=1}^n \int_0^1 X_i(s) X_i^T(s) \dfrac{t-s}{h} K_h(t-s)\, ds \\ \sum_{i=1}^n \int_0^1 X_i(s) X_i^T(s) \dfrac{t-s}{h} K_h(t-s)\, ds & \sum_{i=1}^n \int_0^1 X_i(s) X_i^T(s) \left(\dfrac{t-s}{h}\right)^2 K_h(t-s)\, ds \end{pmatrix}.$$

Then, the solution to (4) is given by

$$\widetilde{\theta}(t) = \left(I_p, 0_p\right) D_{xx}(t)^{-1} \left[D_{xy}(t) - D_{xz}(t)\beta\right], \quad (6)$$

where I_p is $p \times p$ identity matrix and 0_p is $p \times p$ zero matrix. Substituting (6) into (2), and by a simple calculation, we have

$$\check{Y}_i(t) = \check{Z}_i^T(t)\beta + \epsilon_i(t), \quad (7)$$

where

$$\check{Z}_i(t) = Z_i(t) - \widehat{\mu}(t)^T X_i(t), \quad (8)$$
$$\check{Y}_i(t) = Y_i(t) - X_i^T(t)\widehat{g}(t),$$

$$\widehat{\mu}(t) = \left(I_p, 0_p\right) D_{xx}(t)^{-1} D_{xz}(t), \quad (9)$$
$$\widehat{g}(t) = \left(I_p, 0_p\right) D_{xx}(t)^{-1} D_{xy}(t).$$

Applying the integral least-squares technology to linear model (7), we can get the integral least-squares estimator of β, say $\widehat{\beta}$, by minimizing

$$\sum_{i=1}^n \int_0^1 \left\{\check{Y}_i(t) - \check{Z}_i^T(t)\beta\right\}^2 dt. \quad (10)$$

Let

$$\widehat{\Gamma} = \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \check{Z}_i(t)\,\check{Z}_i(t)^T dt. \tag{11}$$

If the matrix $\widehat{\Gamma}$ is invertible, $\widehat{\beta}$ can be given by

$$\widehat{\beta} = \widehat{\Gamma}^{-1}\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \check{Z}_i(t)\,\check{Y}_i(t)\,dt. \tag{12}$$

Let

$$\Gamma = E\left\{\int_0^1 \left[Z(t) - \mu^T(t)X(t)\right]^{\otimes 2} dt\right\}, \tag{13}$$

$$B = E\left(\int_0^1 \left\{\left[Z(t) - \mu(t)^T X(t)\right]\epsilon(t)\right\} dN(t)\right)^{\otimes 2}, \tag{14}$$

where $A^{\otimes 2} = AA^T$, $\mu(t) = \Psi^{-1}(t)\Phi(t)$, $\Phi(t) = E\{X(t)Z^T(t) \mid t\}$, and $\Psi(t) = E\{X(t)X^T(t) \mid t\}$. The following result states the asymptotic normality of $\widehat{\beta}$.

Theorem 1. *Suppose that conditions (C1)–(C5) in the Appendix hold; then, one has*

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) \xrightarrow{\mathscr{L}} N(0, \Sigma), \tag{15}$$

where $\Sigma = \Gamma^{-1}B\Gamma^{-1}$.

In order to construct the confidence interval of β by Theorem 1, we give the estimator of Σ, say $\widehat{\Sigma} = \widehat{\Gamma}^{-1}\widehat{B}(\widehat{\beta})\widehat{\Gamma}^{-1}$, where $\widehat{\Gamma}$ is defined in (11) and

$$\widehat{B}(\widehat{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_i(\widehat{\beta})\,\widehat{\eta}_i^T(\widehat{\beta}). \tag{16}$$

Invoking $\|\widehat{\beta} - \beta\| = O_p(n^{-1/2})$, with the similar argument to Lemma A.6, we can prove that $\widehat{\Sigma}$ is a consistent estimator of Σ. Thus, by Theorem 1, we have

$$\widehat{\Sigma}^{-1/2}\sqrt{n}\left(\widehat{\beta} - \beta\right) \xrightarrow{\mathscr{L}} N(0, I_q), \tag{17}$$

where I_q is an identity matrix of order q. Therefore, the confidence region of β can be constructed by using (17).

Furthermore, substituting $\widehat{\beta}$ into (6), we can get the integral least-squares estimator of $\theta(t)$ as

$$\widehat{\theta}(t) = \left(I_p, 0_p\right) D_{xx}(t)^{-1}\left[D_{xy}(t) - D_{xz}(t)\widehat{\beta}\right]. \tag{18}$$

We state the asymptotic normality of $\widehat{\theta}(t)$ in the following theorem.

Theorem 2. *Suppose that conditions (C1)–(C5) in the Appendix hold. For given t_0, then one has*

$$\sqrt{nh}\left\{\widehat{\theta}(t_0) - \theta(t_0) - b(t_0)\right\} \xrightarrow{\mathscr{L}} N(0, V(t_0)), \tag{19}$$

where $b(t_0) = (1/2)\theta''(t_0)h^2\int_0^1 t^2K(t)dt$, $V(t_0) = \Psi^{-1}(t_0)\sigma^2(t_0)\int_0^1 K^2(t)dt$.

Let

$$\widehat{\Psi}(t_0) = \frac{1}{nh}\sum_{i=1}^{n}\int_0^1 X_i(t)\,X_i(t)^T K_h(t - t_0)\,dt,$$

$$\widehat{b}(t_0) = \frac{1}{nh}\sum_{i=1}^{n}\int_0^1 \left[\widehat{\theta}(t) - \widehat{\theta}(t_0)\right]K_h(t - t_0)\,dt,$$

$$\widehat{\sigma}^2(t_0) = \frac{1}{nh}\sum_{i=1}^{n}\int_0^1 \left[Y_i(t) - Z_i(t)^T\widehat{\beta} - X_i(t)^T\widehat{\theta}(t)\right]^2$$

$$\times K_h(t - t_0)\,dt. \tag{20}$$

Using the law of large numbers, with the similar argument to Lemma A.2 and Proposition 4.1 of Xue and Zhu [14] and Lemma 1 of Wu et al. [15], it can be shown that $\widehat{\sigma}^2(t_0)$, $\widehat{\Psi}(t_0)$, and $\widehat{b}(t_0)$ are consistent estimators of $\sigma^2(t_0)$, $\Psi(t_0)$, and $b(t_0)$, respectively. Finally, we can obtain the estimator $\widehat{V}(t_0)$ by substituting $\Psi^{-1}(t_0)$ and $\sigma^2(t_0)$ in $V(t_0)$ with $\widehat{\Psi}^{-1}(t_0)$ and $\widehat{\sigma}^2(t_0)$, respectively. It can be shown that $\widehat{b}(t_0)$ and $\widehat{V}(t_0)$ are the consistent estimators of $b(t_0)$ and $V(t_0)$, respectively. By Theorem 2, we have that

$$\widehat{V}(t_0)^{-1/2}\sqrt{nh}\left\{\widehat{\theta}(t_0) - \theta(t_0) - \widehat{b}(t_0)\right\} \xrightarrow{\mathscr{L}} N(0, I_p), \tag{21}$$

where I_p is the unit matrix of order p.

Using (21), a pointwise confidence interval for $\theta_r(t_0)$ can be given by

$$\widehat{\theta}_r(t_0) - (nh)^{-1/2}\widehat{b}_r(t_0) \pm Z_{\alpha/2}(nh)^{-1/2}\widehat{v}_r(t_0),$$

$$r = 1, \ldots, p, \tag{22}$$

where $\widehat{b}_r(t_0)$ is the rth component of $\widehat{b}(t_0)$, $\widehat{v}_r(t_0)$ is the (r, r)th element of $\widehat{V}(t_0)$, and $Z_{\alpha/2}$ is the $1 - \alpha/2$ quantile value of the standard normal distribution.

3. Numerical Results

In this section, we conduct several simulation experiments to illustrate the finite sample performances of the proposed method and consider a real data set analysis for further illustration.

3.1. Simulation Studies. To evaluate the performance of the proposed method, we consider the following model:

$$Y(t) = \sin(2\pi t)X(t) + 2.5Z(t) + \epsilon(t), \tag{23}$$

where $\beta = 2.5$ and $\theta(t) = \sin(2\pi t)$. To perform the simulation, we generated $n = 30, 50, 100$ samples, respectively. The covariates $X(t)$ and $Z(t)$ are generated according to the model

$$X(t) = X_{01}(0.5 - t)^2 + X_{02}, \qquad Z(t) = Z_{01}(0.5 - t)^3 + Z_{02}, \tag{24}$$

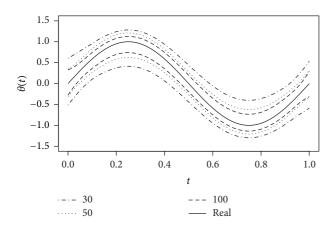where $X_{01} \sim U(5, 7)$, $X_{02} \sim N(0, 1.5)$, $Z_{01} \sim N(0, 5)$, and $Z_{02} \sim N(0, 1)$, respectively.

Figure 1: The pointwise confidence intervals for $\theta(t)$.

Figure 2: The corresponding coverage probabilities.

Table 1: The estimators and the 95% confidence intervals for β.

n	$\widehat{\beta}$	SD	(β_L, β_U)	Length	Coverage
30	2.472	0.166	(2.326, 2.618)	0.292	0.937
50	2.491	0.136	(2.389, 2.594)	0.205	0.945
100	2.502	0.096	(2.441, 2.562)	0.121	0.951

Table 2: The estimators and the 95% confidence intervals for β.

β	Estimator	Confidence interval	Interval length
β_1	−1.758	(−1.907, −1.609)	0.298
β_2	−0.661	(−0.743, −0.579)	0.164

We use the Epanechnikov kernel function $K(t) = 0.75(1 - t^2)_+$ and use the cross-validation method to determine bandwidth h. Let $\widehat{\theta}_{-i}(\cdot)$ and $\widehat{\beta}_{-i}$ be the integral least-squares estimators of $\theta(\cdot)$ and β, respectively, which are computed with all of the measurements but not the ith observation. Define the integral least-squares cross-validation function

$$\mathrm{CV}(h) = \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \left\{ Y_i(t) - X_i(t)^T \widehat{\theta}_{-i}(t) - Z_i(t)^T \widehat{\beta}_{-i} \right\}^2. \tag{25}$$

The cross-validation bandwidth is the one that minimizes (25); that is,

$$\widehat{h}_{\mathrm{CV}} = \arg \min_h \mathrm{CV}(h). \tag{26}$$

For the parametric component β, the average and standard deviation of the estimator $\widehat{\beta}$, based on 1000 simulations, are reported in Table 1. In addition, the average length and coverage probability of the confidence interval $(\widehat{\beta}_L, \widehat{\beta}_U)$, with a nominal level $1 - \alpha = 95\%$, are computed with 1000 simulation runs. The results are also summarized in Table 1.

For the nonparametric component $\theta(t)$, the average pointwise confidence intervals, based on 1000 simulations, with a nominal level $1 - \alpha = 95\%$ are presented in Figure 1, and the corresponding coverage probabilities are presented in Figure 2.

Table 1 shows that, for the parametric component, our method can give a shorter confidence interval and the corresponding coverage probability is close to real nominal level. Figures 1 and 2 show that the average interval length decreases as the sample size increases, while the corresponding coverage probability increases. In addition, we can see that, for the nonparametric component, the proposed estimation method works well except for boundary points.
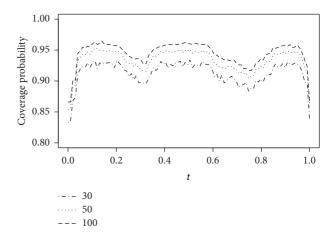
3.2. Application to Spectrometric Curves Data.

In this section, we present an application of the proposed estimation method to spectrometric curves data. This original data comes from a quality control problem in the food industry. This data set concerns a sample of finely chopped meat, and each food sample contains finely chopped pure meat with different fat, protein, and moisture (water) contents. The sample size of this data set is $n = 240$, and, for each food sample, the functional data consist of 100 channel spectrum of absorbances, which were recorded on the Tecator Infratec Food and Feed Analyzer working in the wavelength range 850–1050 nm by the near infrared transmission (NIT) principle. Because of the fineness of the grid, we can consider each subject as a continuous curve. Thus, each spectrometric analysis can be summarized by some continuous curves giving the observed absorbance as function of the wavelength. More details of the data can be found in Ferraty and Vieu [16].

The aim is to find the relationship between the percentage of fat content $Y(t)$ and the corresponding percentages of protein content $X_1(t)$, the moisture content $X_2(t)$, and the spectrometric curve t. The results, obtained by Aneiros-Pérez and Vieu [11], indicate that there is a strong linear relationship between the fat content and the protein and moisture contents, but the spectrometric curve t has a functional effect on the fat content. Hence, we consider the following semiparametric model:

$$Y(t) = \theta(t) + \beta_1 X_1(t) + \beta_2 X_2(t) + \varepsilon. \tag{27}$$

We computed the estimators of the parametric components β_1 and β_2 and the nonparametric component $\theta(\cdot)$ by using the proposed integral least-squares method. The results for the parametric components are reported in Table 2, and the results for the nonparametric components are reported in Figure 3, where the solid curve is the estimator of $\theta(\cdot)$
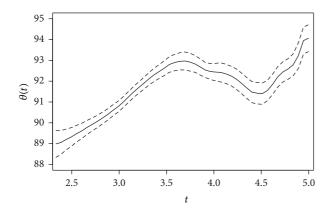
FIGURE 3: The estimator and 95% pointwise confidence interval for $\theta(t)$.

and the dashed curve is the pointwise confidence interval of $\theta(\cdot)$. From Table 2, we can see that there is a significant negative correlation relationship between the fat content and the protein and moisture contents. In addition, Figure 3 indicates that the baseline function $\theta(\cdot)$ really varies over the spectrometric curve t. This finding basically agrees with what was discovered by Aneiros-Pérez and Vieu [11].

Appendix

Proof of Theorems

For convenience and simplicity, let c denote a positive constant which may be different value at each appearance throughout this paper. Before we prove our main theorems, we list some regularity conditions which are used in this paper.

- (C1) The bandwidth satisfies $h = cn^{-1/5}$, for some constant $c > 0$.

- (C2) The kernel $K(\cdot)$ is a symmetric probability density function, and $\int t^4 K(t)dt < \infty$.

- (C3) $\theta(t)$, $\Phi(t)$, and $\Psi(t)$ are twice continuously differentiable on $(0, 1)$.

- (C4) $\sup_{0 \leq t \leq 1} E(\epsilon(t)^4 \mid t) < \infty$, $\sup_{0 \leq t \leq 1} E(X_r(t)^4 \mid t) < \infty$, and $E(\epsilon(t)^4 \mid t)$ and $E(X_r(t)^4 \mid t)$ are continuous at t, $r = 1, \ldots, p$, where $X_r(t)$ is the rth component of $X(t)$.

- (C5) For given t, $\Psi(t)$ is positive definite matrix.

The proofs of Theorems 1 and 2 rely on the following lemmas.

Lemma A.1. Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ $(b_1 \geq b_2 \geq , \ldots, \geq b_n)$ be two sequences of real numbers. Let $S_k = \sum_{i=1}^{k} a_i$; then, one has

$$\left| \sum_{i=1}^{n} a_i b_i \right| \leq c \max_{1 \leq i \leq n} |b_i| \max_{1 \leq i \leq n} |S_i|. \tag{A.1}$$

Lemma A.2. Let e_i, $i = 1, \ldots, n$, be a sequence of multi-independent random variate with $E(e_i) = 0$ and $E(e_i^2) < c < \infty$. Then, one has

$$\max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} e_i \right| = O_p \left(\sqrt{n} \log n \right). \tag{A.2}$$

Further, let $(j_1, j_2, \ldots, j_n)$ be a permutation of $(1, 2, \ldots, n)$. Then, we have

$$\max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} e_{j_i} \right| = O_p \left(\sqrt{n} \log n \right). \tag{A.3}$$

Lemma A.3. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be i.i.d. random vectors, where Y_i is scalar random variables. Further, assume that $E|Y_1|^s < \infty$, $\sup_x \int |y|^s f(x, y)dy < \infty$, where $f(\cdot, \cdot)$ denotes the joint density of (X, Y). Let $K(\cdot)$ be a bounded positive function with a bounded support, satisfying a Lipschitz condition. Then,

$$\sup_x \left| \frac{1}{n} \sum_{i=1}^{n} \left\{ K_h \left(X_i - x \right) Y_i - E \left[K_h \left(X_i - x \right) Y_i \right] \right\} \right|$$
$$= O_p \left(\left\{ \frac{\log(1/h)}{nh} \right\}^{1/2} \right) \tag{A.4}$$

provided that $n^{2\delta - 1} h \to \infty$ for some $\delta < 1 - s^{-1}$.

Lemma A.4. Suppose that conditions (C1)–(C5) hold. Then, one has

$$\sup_{0 < t < 1} \left\| \widehat{\mu}(t) - \Psi^{-1}(t) \Phi(t) \right\| = O_p(C_n),$$
$$\sup_{0 < t < 1} \left\| \widehat{g}(t) - \Psi^{-1}(t) \Phi(t) \beta - \theta(t) \right\| = O_p(C_n), \tag{A.5}$$

where $\widehat{\mu}(t)$, $\widehat{g}(t)$ are defined in (9) and $C_n = \{\log(1/h)/nh\}^{1/2} + h^2$.

Proof. Let

$$S_{nl}(t) = \sum_{i=1}^{n} \int_0^1 X_i(s) X_i(s)^T \left(\frac{s-t}{h} \right)^l K_h(s-t) \, ds, \tag{A.6}$$
$$l = 0, 1, 2;$$

then, a simple calculation yields

$$E\left\{ S_{nl}(t) \right\} = n\Psi(t) \int_0^1 s^l K(s) \, ds + o(1), \quad l = 0, 1, 2. \tag{A.7}$$

Note that

$$D_{xx}(t) = \begin{pmatrix} S_{n0}(t) & S_{n1}(t) \\ S_{nl}(t) & S_{n2}(t) \end{pmatrix}; \tag{A.8}$$

together with (A.6), (A.7), by using Lemma A.3, we obtain

$$
D_{xx}(t) = n\Psi(t) \otimes \begin{pmatrix} 1 & 0 \\ 0 & \int_0^1 s^2 K(s)\,ds \end{pmatrix} \{1 + O_p(C_n)\},
$$

(A.9)

uniformly for $t \in (0,1)$, where $\otimes$ is the Kronecker product. By using the same argument, we have

$$
D_{xz}(t) = n\Phi(t) \otimes (1,0)^T \{1 + O_p(C_n)\}
$$

(A.10)

uniformly for $t \in (0,1)$. Combining (A.9), (A.10), and (9) yields

$$
\hat{\mu}(t) = \Psi^{-1}(t)\Phi(t)\{1 + O_p(C_n)\},
$$

(A.11)

uniformly for $t \in (0,1)$. With the similar proof to (A.11), we have

$$
\sum_{i=1}^n \int_0^1 X_i(s) Y_i(s) \left(\frac{s-t}{h}\right)^l K_h(t-s)\,ds
$$

$$
= \sum_{i=1}^n \int_0^1 X_i(s)\left[X_i^T(s)\theta(s) + Z_i^T(s)\beta + \epsilon_i(s)\right]
$$

$$
\times \left(\frac{s-t}{h}\right)^l K_h(t-s)\,ds
$$

$$
= n\{\Psi(t)\theta(t) + \Phi(t)\beta\}\int_0^1 s^l K(s)\,ds \{1 + O_p(C_n)\},
$$

$$
l = 0, 1.
$$

(A.12)

Combining (A.12) with (5), we can prove that

$$
D_{xy}(t) = n\left[\Psi(t)\theta(t) + \Phi(t)\beta\right] \otimes (1,0)^T \{1 + O_p(C_n)\},
$$

(A.13)

uniformly for $t \in (0,1)$.

Invoking (A.9) and (A.13), it is easy to show that

$$
\hat{g}(t) = (I_p, 0_p)[D_{xx}(t)]^{-1} D_{xy}(t)
$$

$$
= \left[\theta(t) + \Psi^{-1}(t)\Phi(t)\beta\right]\{1 + O_p(C_n)\}.
$$

(A.14)

This completes the proof of Lemma A.4. □

Lemma A.5. *Suppose that conditions* (C1)–(C5) *hold. Then, one has that*

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^n \hat{\eta}_i(\beta) \xrightarrow{\mathscr{L}} N(0, B),
$$

(A.15)

where B *is defined by* (14) *and* $\hat{\eta}_i(\beta) = \int_0^1 \check{Z}_i(t)[\check{Y}_i(t) - \check{Z}_i(t)^T\beta]dt$.

Proof. By (8), a simple calculation yields

$$
\check{Z}_i(t) = \left[Z_i(t) - \mu(t)^T X_i(t)\right] + \left[\mu(t) - \hat{\mu}(t)\right]^T X_i(t),
$$

$$
\check{Y}_i(t) - \check{Z}_i(t)^T\beta = X_i(t)^T\left[\theta(t) - \hat{g}(t) + \hat{\mu}(t)\beta\right] + \varepsilon_i(t),
$$

(A.16)

where $\mu(t) = \Psi(t)^{-1}\Phi(t)$. Then, we have

$$
\hat{\eta}_i(\beta)
$$

$$
= \int_0^1 \left[Z_i(t) - \mu(t)^T X_i(t)\right]\varepsilon_i(t)\,dt
$$

$$
+ \int_0^1 \left[\mu(t) - \hat{\mu}(t)\right]^T X_i(t)\varepsilon_i(t)\,dt
$$

$$
+ \int_0^1 \left[Z_i(t) - \mu(t)^T X_i(t)\right]X_i(t)^T\left[\theta(t) - \hat{g}(t) + \hat{\mu}(t)\beta\right]dt
$$

$$
+ \int_0^1 \left[\mu(t) - \hat{\mu}(t)\right]^T X_i(t) X_i(t)^T\left[\theta(t) - \hat{g}(t) + \hat{\mu}(t)\beta\right]dt
$$

$$
\equiv J_{i1} + J_{i2} + J_{i3} + J_{i4}.
$$

(A.17)

Hence, we get

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^n \hat{\eta}_i(\beta)
$$

$$
= \frac{1}{\sqrt{n}}\sum_{i=1}^n J_{i1} + \frac{1}{\sqrt{n}}\sum_{i=1}^n J_{i2} + \frac{1}{\sqrt{n}}\sum_{i=1}^n J_{i3} + \frac{1}{\sqrt{n}}\sum_{i=1}^n J_{i4}
$$

$$
\equiv J_1 + J_2 + J_3 + J_4.
$$

(A.18)

Note that J_{i1}, $i = 1,\ldots,n$, are i.i.d. It is easy to show that $E(J_{i1}) = 0$, $\mathrm{Var}(J_{i1}) = B + o(1)$. Using the central limit theorem, we have

$$
J_1 \xrightarrow{\mathscr{L}} N(0, B).
$$

(A.19)

Hence, to prove this lemma, we only need to prove $J_\nu \xrightarrow{P} 0$, $\nu = 2, 3, 4$. Now, we deal with J_2. Let $b_{rs}(t)$ be the (r,s) component of $[\mu(t) - \hat{\mu}(t)]^T$ and let $a_{is}(t)$ be the sth component of $X_i(t)\epsilon_i(t)$, $s = 1, \ldots, p$. Moreover, let

$$
\int_0^1 b_{rs_0}(t) a_{is_0}(t)\,dt = \max_{1 \le s \le p}\left\{\int_0^1 b_{rs}(t) a_{is}(t)\,dt\right\}.
$$

(A.20)

Let $J_{2,r}$ be the rth component of J_2. By Lemmas A.1–A.4, we obtain

$$
\begin{aligned}
|J_{2,r}| &= \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{n} \sum_{s=1}^{p} \int_0^1 b_{rs}(t) a_{is}(t)\, dt \right| \\
&\leq \frac{p}{\sqrt{n}} \left| \sum_{i=1}^{n} \int_0^1 b_{rs_0}(t) a_{is_0}(t)\, dt \right| \\
&\leq \frac{c}{\sqrt{n}} \sup_{0 \leq t \leq 1} \left| b_{rs_0}(t) \right| \max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} \int_0^1 a_{is_0}(t)\, dt \right| \\
&= \frac{c}{\sqrt{n}} O_p(C_n) O_p(\sqrt{n} \log n) = o_p(1).
\end{aligned}
\tag{A.21}
$$

That is, $J_2 \xrightarrow{P} 0$. By Lemma A.4, a simple calculation yields

$$
\sup_{0<t<1} \|\theta(t) - \widehat{g}(t) + \widehat{\mu}(t)\beta\| = O_p(C_n).
\tag{A.22}
$$

Together with $E\{[Z_i(t) - \mu(t)^T X_i(t)] X_i(t)^T\} = 0$, and similar to the proof of J_2, we have $J_3 \xrightarrow{P} 0$. In addition, by Lemma A.4, we have $\|J_4\| \leq O_p(\sqrt{n} C_n^2) = o_p(1)$. This completes the proof of Lemma A.5. □

Lemma A.6. *Suppose that conditions* (C1)–(C5) *hold; then, one has*

$$
\widehat{\Gamma} \xrightarrow{P} \Gamma,
\tag{A.23}
$$

where $\widehat{\Gamma}$ and Γ are defined in (11) *and* (13)*, respectively.*

Proof. Combining (8) and (11), a simple calculation yields

$$
\begin{aligned}
\widehat{\Gamma} &= \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \left[Z_i(t) - \mu(t)^T X_i(t) \right] \left[Z_i(t) - \mu(t)^T X_i(t) \right]^T dt \\
&\quad \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \left[Z_i(t) - \mu(t)^T X_i(t) \right] X_i(t)^T \left[\mu(t) - \widehat{\mu}(t) \right] dt \\
&\quad \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \left[\mu(t) - \widehat{\mu}(t) \right]^T X_i(t) \left[Z_i(t)^T - X_i(t)^T \mu(t) \right] dt \\
&\quad \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \left[\mu(t) - \widehat{\mu}(t) \right]^T X_i(t) X_i(t)^T \left[\mu(t) - \widehat{\mu}(t) \right] dt \\
&\equiv A_1 + A_2 + A_3 + A_4.
\end{aligned}
\tag{A.24}
$$

By the law of large numbers, we can derive that $A_1 \xrightarrow{P} \Gamma$. We now show $A_2 \xrightarrow{P} 0$. From Lemma A.4, we have that

$$
\sup_{0<t<1} \|\widehat{\mu}(t) - \mu(t)\| = O_p(C_n).
\tag{A.25}
$$

Together with $E\{[Z_i(t) - \mu(t)^T X_i(t)] X_i(t)^T\} = 0$, using the similar argument to Lemma A.5, we can prove $A_2 \xrightarrow{P} 0$. Similarly, we can prove that $A_3 \xrightarrow{P} 0$. In addition, by (A.25), it is easy to show that $\|A_4\| = O_p(C_n^2)$. This completes the proof. □

Proof of Theorem 1. From (12), we can derive that

$$
\begin{aligned}
\widehat{\beta} &= \widehat{\Gamma}^{-1} \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \check{Z}_i(t) \check{Y}_i(t)\, dt \\
&= \widehat{\Gamma}^{-1} \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \check{Z}_i(t) \check{Z}_i(t)^T \beta\, dt \\
&\quad + \widehat{\Gamma}^{-1} \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \check{Z}_i(t) \left[\check{Y}_i(t) - \check{Z}_i(t)^T \beta \right] dt \\
&= \beta + \widehat{\Gamma}^{-1} \frac{1}{n} \sum_{i=1}^{n} \widehat{\eta}_i(\beta).
\end{aligned}
\tag{A.26}
$$

Together with Lemmas A.5 and A.6, it is easy to show that

$$
\sqrt{n}(\widehat{\beta} - \beta) = \widehat{\Gamma}^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \widehat{\eta}_i(\beta) \xrightarrow{\mathscr{L}} N(0, \Sigma).
\tag{A.27}
$$

This completes the proof of Theorem 1. □

Proof of Theorem 2. For t in a small neighborhood of t_0, such that $|t - t_0| \leq h$, by using Taylor expansion, we can get

$$
\theta(t) = \theta(t_0) + \theta'(t_0)(t - t_0) + \frac{1}{2}\theta''(t_0)(t - t_0)^2 + O_p(h^3).
\tag{A.28}
$$

Together with (5) and (1), we can derive that

$$
\begin{aligned}
D_{xy}(t_0) &= D_{xx}(t_0) \begin{pmatrix} \theta(t_0) \\ h\theta'(t_0) \end{pmatrix} + \frac{1}{2}h^2\theta''(t_0) T_{xx}(t_0) \\
&\quad + D_{xz}(t_0)\beta + D_{x\epsilon}(t_0) + O_p(nh^3),
\end{aligned}
\tag{A.29}
$$

where

$$
T_{xx}(t_0) = \begin{pmatrix} \sum_{i=1}^{n} \int_0^1 X_i(t) X_i^T(t) \left(\dfrac{t-t_0}{h}\right)^2 K_h(t-t_0)\, dt \\ \sum_{i=1}^{n} \int_0^1 X_i(t) X_i^T(t) \left(\dfrac{t-t_0}{h}\right)^3 K_h(t-t_0)\, dt \end{pmatrix},
$$

$$
D_{x\epsilon}(t_0) = \begin{pmatrix} \sum_{i=1}^{n} \int_0^1 X_i(t) \epsilon_i(t) K_h(t-t_0)\, dt \\ \sum_{i=1}^{n} \int_0^1 X_i(t) \epsilon_i(t) \left(\dfrac{s-t_0}{h}\right) K_h(t-t_0)\, dt \end{pmatrix}.
\tag{A.30}
$$

Invoking (18), (A.9), and (A.29), it is easy to show that

$$
\begin{aligned}
\widehat{\theta}(t_0) &= \theta(t_0) + \frac{1}{2}h^2\theta''(t_0)(I_p, 0_p) D_{xx}^{-1}(t_0) T_{xx}(t_0) \\
&\quad + (I_p, 0_p) D_{xx}^{-1}(t_0) D_{xz}(t_0)(\widehat{\beta} - \beta) \\
&\quad + (I_p, 0_p) D_{xx}^{-1}(t_0) D_{x\epsilon}(t_0) + O_p(h^3).
\end{aligned}
\tag{A.31}
$$

Let $\Delta_n(t_0) = n^{-1}D_{x\epsilon}(t_0)$; we can prove that $E\{\Delta_n(t_0)\} = 0$ and $\mathrm{Var}\{\Delta_n(t_0)\} = (nh)^{-1}\Delta(t_0)$, where

$$
\Delta(t_0)
$$

$$
= \sigma^2(t_0)\, f(t_0)\, \Psi(t_0) \otimes \begin{pmatrix} \int_0^1 K^2(t)\, dt & \int_0^1 tK^2(t)\, dt \\ \int_0^1 tK^2(t)\, dt & \int_0^1 t^2 K^2(t)\, dt \end{pmatrix}. \tag{A.32}
$$

By the central limit theorem, we have that

$$
\sqrt{nh}\Delta_n(t_0) \xrightarrow{\mathscr{L}} N(0, \Delta(t_0)). \tag{A.33}
$$

Combining this with (A.9), we can prove that

$$
\sqrt{nh}\,(I_p, 0_p)\, D_{xx}^{-1}(t_0)\, D_{x\epsilon}(t_0)
$$

$$
= (I_p, 0_p)\, nD_{xx}^{-1}(t_0)\, \sqrt{nh}\Delta_n(t_0) \xrightarrow{\mathscr{L}} N(0, V(t_0)). \tag{A.34}
$$

Using the similar argument to (A.9) in Lemma A.4, we can get

$$
T_{xx}(t) = n\Psi(t_0)\, h^2 \otimes \left(\int_0^1 t^2 K(t)\, dt, \int_0^1 t^3 K(t)\, dt \right)^T
$$

$$
\times \{1 + O_p(C_n)\}. \tag{A.35}
$$

Together with (A.9) and $\widehat{\beta} - \beta = O_p(n^{-1/2})$, it is easy to prove that

$$
\frac{1}{2}h^2\theta''(t_0)\,(I_0, 0_p)\, D_{xx}^{-1}(t_0)\, T_{xx}(t_0) = b(t_0) + O_p\left(h^2 C_n\right),
$$

$$
(I_0, 0_p)\, D_{xx}^{-1}(t_0)\, D_{xz}(t_0)\,(\widehat{\beta} - \beta) = O_p\left(n^{-1/2}\right). \tag{A.36}
$$

Invoking (A.31), (A.34), and (A.36), a simple calculation yields

$$
\sqrt{nh}\left(\widehat{\theta}(t_0) - \theta(t_0) - b(t_0)\right) \xrightarrow{\mathscr{L}} N(0, V(t_0)). \tag{A.37}
$$

This completes the proof of Theorem 2. □

Conflict of Interests

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

Acknowledgments

References

[1] J. O. Ramsay and B. W. Silverman, *Applied Functional Data Analysis: Methods and Case Studies*, Springer Series in Statistics, Springer, Berlin, Germany, 2002.

[2] J. O. Ramsay and B. W. Silveman, *Functional Data Analysis*, Springer, New York, NY, USA, 1997.

[3] A. Ait-Saïdi, F. Ferraty, and P. Vieu, "Cross-validated estimations in the single-functional index model," *Statistics*, vol. 42, no. 6, pp. 475–494, 2008.

[4] F. Ferraty, A. Goia, E. Salinelli, and P. Vieu, "Functional projection pursuit regression," *TEST*, vol. 22, no. 2, pp. 293–320, 2013.

[5] D. Chen, P. Hall, and H. Müller, "Single and multiple index functional regression models with nonparametric link," *The Annals of Statistics*, vol. 39, no. 3, pp. 1720–1747, 2011.

[6] F. Ferraty and P. Vieu, "Nonparametric models for functional data, with application in regression, time-series prediction and curve discrimination," *Journal of Nonparametric Statistics*, vol. 16, no. 1-2, pp. 111–125, 2004.

[7] M. Rachdi and P. Vieu, "Nonparametric regression for functional data: automatic smoothing parameter selection," *Journal of Statistical Planning and Inference*, vol. 137, no. 9, pp. 2784–2801, 2007.

[8] M. Escabias, A. Aguilera, and M. Valderrama, "Functional PLS logit regression model," *Computational Statistics & Data Analysis*, vol. 51, no. 10, pp. 4891–4902, 2007.

[9] F. Yao, H. Müller, and J. Wang, "Functional linear regression analysis for longitudinal data," *The Annals of Statistics*, vol. 33, no. 6, pp. 2873–2903, 2005.

[10] J. T. Zhang and J. Chen, "Statistical inferences for functional data," *The Annals of Statistics*, vol. 35, no. 3, pp. 1052–1079, 2007.

[11] G. Aneiros-Pérez and P. Vieu, "Semi-functional partial linear regression," *Statistics and Probability Letters*, vol. 76, no. 11, pp. 1102–1110, 2006.

[12] H. Shin, "Partial functional linear regression," *Journal of Statistical Planning and Inference*, vol. 139, no. 10, pp. 3405–3418, 2009.

[13] H. Lian, "Functional partial linear model," *Journal of Nonparametric Statistics*, vol. 23, no. 1, pp. 115–128, 2011.

[14] L. Xue and L. Zhu, "Empirical likelihood for a varying coefficient model with longitudinal data," *Journal of the American Statistical Association*, vol. 102, no. 478, pp. 642–654, 2007.

[15] C. O. Wu, C. T. Chiang, and D. R. Hoover, "Asymptotic confidence regions for kernel smoothing of a varying-coefficient model with longitudinal data," *Journal of the American Statistical Association*, vol. 93, no. 444, pp. 1388–1402, 1998.

[16] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis*, Springer, New York, NY, USA, 2006.