Statistica Sinica Preprint No: SS-2023-0415	
Title	Functional Linear Operator Quantile Regression for Sparse Longitudinal Data
Manuscript ID	SS-2023-0415
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202023.0415
Complete List of Authors	Xingcai Zhou, Tingyu Lai and Linglong Kong
Corresponding Authors	Linglong Kong
E-mails	lkong@ualberta.ca

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

Statistica Sinica

FUNCTIONAL LINEAR OPERATOR QUANTILE REGRESSION FOR SPARSE LONGITUDINAL DATA

Xingcai Zhou[1], Tingyu Lai[2] and Linglong Kong[3]

[1] *School of Statistics and Data Science, Nanjing Audit University, Nanjing 211815, China*

[2] *School of Mathematics and Statistics, Guangxi Normal University, Guilin 541006, China*

[3] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada*

Abstract: We propose a functional linear operator quantile regression (FLOQR) framework, which includes many important and useful functional data models, and devote to the new framework model for longitudinal data with the typically sparse and irregular designs. The non-smooth quantile loss and functional linear operator pose new challenges to functional data analysis for longitudinal data in both computation and theoretical development. To address the challenge, we propose the iterative surrogate least squares estimation approach for the FLOQR model, which transforms the response trajectories and establishes a new connection between FLOQR and functional linear operator model. In addition, we use Karhunen-Loève expansion to alleviate the problem of the nonexistence of the inverse of the covariance in the infinite-dimensional Hilbert space. Then, the approach is used to classic functional varying coefficient QR, functional linear QR, and functional varying coefficient QR with history index function for sparse longitudinal data by using functional principal components analysis through conditional expectation. The resulting technique is flexible and allows the prediction of an unobserved quantile response trajectory from sparse measurements of a predictor trajectory. Theoretically, we show that, after a constant number of iterations, the proposed estimator is asymptotic consistent for sparse designs. Moreover, asymptotic pointwise confidence bands are obtained for predicted quantile individual trajectories based on their asymptotic distributions. The proposed algorithms perform well in simulations, and are illustrated with longitudinal primary biliary liver cirrhosis data.

Key words and phrases: Functional data analysis, Quantile regression, Linear operator, Repeated measurements, Confidence band, Prediction, Sparse design.

1. Introduction

Functional data analysis has become increasingly useful in various fields. One well-known application is growth curve analysis in biology, medicine, and chemistry, see, for instance, Müller (2009), Ramsay and Silverman (2005) and Ferraty and Vieu (2006), and references therein. In these applications, especially in longitudinal studies, often a few repeated measurements can be obtained for each subject or item, owing to cost or logistical constraints that limit the number of measurements (Ji and Müller, 2017). Functional data analysis methodology has been shown useful to infer the general trend and covariance structure of trajectories, and the prediction of an unobserved response trajectory from some sparse and irregular functional or longitudinal data (Yao et al., 2005a,b; Li and Hsing, 2010; Jiang and Wang, 2011; Ji and Müller, 2017). On the other hand, there are some classic estimation methods and theoretical studies on functional data in reproducing kernel Hilbert space, for example, Sun et al. (2018); Cui et al. (2020); Lv et al. (2020); Zhang and Lian (2021); Li et al. (2021); Yang et al. (2021); Liu et al. (2024), but these results are generally not used for irregular longitudinal data analysis.

Recently, there has been increased interest in extending regression models to functional

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

data, where both the predictor and the response are random functions. We can abstract most functional regression models as functional linear operator model, which has form

$$\mathrm{E}[Y(t)|X(t)] = (\mathcal{L}_X\beta)(t), \tag{1.1}$$

where $Y(\cdot)$ and $X(\cdot)$ respectively are the square-integrable predictor and response trajectories on compact interval domain $\mathcal{T}$, and $(\mathcal{L}_X\beta)$ is a functional linear operator. The functional linear operator model (1.1) has been well studied under various regression models. For instance, $(\mathcal{L}_X\beta)(t) = \beta(t)X(t)$, a standard functional varying coefficient model (FVC) (Şentürk and Müller, 2010; Wu et al., 2010); $(\mathcal{L}_X\beta)(t) = \int_{\mathcal{S}} \beta_\tau(s,t)X(s)ds$, functional linear regression model (FLQ) (He et al., 2000; Yao et al., 2005b); $(\mathcal{L}_X\beta)(t) = \beta_\tau(t)\int_0^\Delta \gamma_\tau(s)X(t-s)ds$, functional varying coefficient models with history index (FVC-HI) (Şentürk and Müller, 2010), and so on. The operators $(\mathcal{L}_X\beta)$ in FVC, FLQ, and FVC-HI will be defined in detail in the next section. The model (1.1) can be regarded as an extension of multivariate data analysis where observations consist of vectors of finite dimension. That is, it is an extension from finite dimensions to the infinite-dimensional case. The extension to functional data is not obvious and requires tools from functional analysis. A basic problem for this extension is the inversion of linear operators. He et al. (2000) solved the problem well, who derived properties of functional linear regression modeling for random processes $\{(X(\cdot), Y(\cdot))\}$ under the appropriate conditions. But, it is changeling to our proposed functional linear operator quantile regression model (FLOQR), because these tools and analysis from functional analysis can not directly applied to our model.

In the paper, we proposed a FLOQR, that is, given a prespecified quantile $\tau \in (0,1)$, the conditional τth quantile of $Y(\cdot)$ given $X(\cdot)$,

$$Q_{Y|X}(t;\tau) = \alpha_\tau(t) + (\mathcal{L}_X\beta_\tau)(t), \tag{1.2}$$

where α_τ is a varying coefficient intercept function, and $\mathcal{L}_X\beta_\tau$ is a functional linear operator. A comprehensive survey of the theory of quantile regression and its applications can be found in Koenker (2005). For the functional linear quantile regression model, Kato (2012) considered an estimator for the slope function based on the principal component basis. Recently, Zhang et al. (2022) developed a novel spatial quantile function-on-scalar regression model, which studies the conditional spatial distribution of a high-dimensional functional response given scalar predictors. All investigations on functional quantile regression to date are for the case of completely observed trajectories. Most of these methods face severe challenges for our FLOQR with sparse and irregular designs.

We know that our FLOQR model falls within the framework of function-on-function vary-coefficient linear/quantile regression. For function-on-function vary-coefficient linear regression, Zhang and Chen (2007) constructed local polynomial kernel reconstruction-based estimators for the covariate effects and the covariance function, and derived their asymptotics, and then Zhang (2011) further investigated the F-type test for the general linear hypothesis and derived its asymptotic power. They are both for dense functional data; From functional principal component analysis, Yao et al. (2005b), Wu et al. (2010) and Şentürk and Müller (2010) provided a representation of varying-coefficient functions through suitable auto and cross-covariance of the underlying stochastic processes, which is particularly advantageous for sparse and irregular longitudinal data. **In our FLOQR, we adopt the estimation strategy**. In addition, Zhou and Wu (2010) and Gu et al. (2014) considered the construction of simultaneous confidence corridor for varying coefficient; Luo et al. (2016)

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

and Li et al. (2017) explored functional varying coefficient single index model; Recently, Chen et al. (2019) studied functional regression with random response curve and vector covariates by a supervised least squares estimation procedure after utilizing B-spline function to approximate the unknown functions; Boumahdi et al. (2023) constructed an estimator for the nonparametric function-on-function models with surrogate responses by kernel; Luo and Qi (2024) extended linear function-on-function regression to general nonlinear function-on-function regression, and developed a novel method to fit the model via the functional universal approximation theorem. But the above works mainly focus on mean regression. It is nonrobust for the typically skewed distribution of functional response curves or data containing some outliers. For this scenario, Liu et al. (2020a) performed a function-on-scalar quantile regression in a Bayesian framework; Beyaztas et al. (2025) introduced a function-on-function linear quantile regression model to characterize the entire conditional distribution of a functional response for a given function predictor by tensor cubic B-splines; Battagliola et al. (2025) considered estimation methodology on quantile regression for longitudinal data via spline representations, and only applied to feed intake of lactating sows without theoretical results; Zhou et al. (2023) proposed a functional response quantile regression model and developed a data-driven estimation procedure based on a local linear approximation; and further its empirical likelihood inference is given by Zhou et al. (2024). These function-on-function quantile regression models are special cases of our FLOQR. They did not study the function-to-function quantile model from the viewpoint of the intrinsic characteristics of functions but rather inferred unknown functions by simple kernel or B-spline approximations. They are generally not suitable for sparse and irregular longitudinal data.

Our proposed FLOQR model is a new framework of function-to-function quantile regres-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

sion models. It includes many important and useful models, for example, functional varying coefficient QR model (1.2) with (2.3), functional linear QR model (1.2) with (2.4), functional varying coefficient QR model with history index function (1.2) with (2.5), and so on. In addition, we also develop a unified estimation strategy motivated by the functional linear operator model (He et al., 2000). However, We devote to the new framework model for longitudinal data with typically sparse and irregular designs. It brings us several challenges in statistical estimation: (a) For estimating the FLOQR model, we need to introduce the non-smooth quantile loss function. Unlike the squared loss, the non-smooth quantile loss results in no closed-form solution for the estimate, which also causes some difficulties in both computation and theoretical development. (b) Under sparse data situations, the key for the functional approach is to target the covariance structure of X, the cross-covariance structure of X, and the τth quantile of Y for our FLOQR model. But the cross-covariance structure of the latter cannot be obtained because there is no data representation of α_τ and β_τ. (c) One main obstacle for the operator QR model involves the inversion of covariance operators. The inversion is not feasible in infinite dimensional Hilbert space. We need to adopt the representation of the regression coefficient functions in the functional linear operator model (1.1), which is outlined in He et al. (2000). However, such a straightforward representation from the functional linear operator model to the FLOQR model does not exist. To address these challenges, we propose the iterative surrogate least squares estimation (SLES) approach for the FLOQR model, which transforms the response trajectories and establishes a new connection between FLOQR (1.2) and functional linear operator model (1.1). However, the transformation cannot directly deal with heterogeneous error. In FLOQR model, we assume that the random error of the model is independent of covariates. The technique has also been

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

applied to distributed learning under quantile loss function (Chen et al., 2020).

The contributions of the article are as follows. First, we propose a new model framework for functional data: the FLOQR model, which includes many important functional quantile regression models. They are versatile and flexible analysis tools for relating longitudinal responses to longitudinal predictors. Second, we introduce a new technique, SLES with an iterative algorithm, for the FLOQR model, which transforms the FLOQR model into functional linear operator model; and then apply principal components analysis through conditional expectation to handle sparse and irregular longitudinal data for which the pooled time points are sufficiently dense. Third, theoretically, we show that, after a constant number of iterations, the proposed estimator is asymptotic consistency. Fourth, we respectively construct asymptotic pointwise confidence bands for predicted quantile response trajectories of the three functional QR models based on their asymptotic distribution.

The remainder of the article is organized as follows. In Section 2, we propose FLOQR model, give its estimation strategy and SLSE algorithm, and establish theoretical guarantees. In Section 3, we discuss the three specific functional quantile regression models under the framework of the FLOQR model, including the prediction of quantile response trajectories and the construction of pointwise confidence bands for individual quantile trajectories. Simulation studies that illustrate the usefulness of the proposed functional QR models and estimation strategies can be found in Section 4. Applications of the proposed functional QR models to two longitudinal data are presented in Section 5. Proofs and auxiliary results are compiled in the Supplementary material (SM).

2. Methodology

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

2.1 Functional linear operator quantile regression model

First, we propose our FLOQR model. We know that any random variable Y can be characterized by it cumulative distribution function $F_Y(y) = P(Y \le y)$, or equivalently, by its quantile function $Q_Y(\tau) = F_Y^{-1}(\tau) = \inf\{y : F(y) \ge \tau\}$, where τ is quantile level with $0 < \tau < 1$. The τ-th quantile $Q_Y(\tau)$ minimizes the expected loss

$$Q_Y(\tau) = \mathrm{argmin}_y \mathrm{E}\{\rho_\tau(Y - y)\}, \tag{2.1}$$

for the asymmetric loss $\rho_\tau(u) = u\{\tau - I(u < 0)\}$. For quantile regression, besides robust aspects, it may also help to derive some kind of confidence prediction regions based on quantiles. When the response trajectory Y is coupled with the predictor trajectory X. One is interested in studying the conditional quantile $Q_{Y|X}(\tau) = F_{Y|X}^{-1}(\tau)$. Since Y and X are observed sparsely and irregularly in our setting, we will take full advantage of the functional nature of the underlying data to model sparse longitudinal data via quantile regression (2.1). We thus propose the FLOQR model as

$$Q_{Y|X}(t;\tau) = \alpha_\tau(t) + (\mathcal{L}_X\beta_\tau)(t), \tag{2.2}$$

where α_τ is a varying coefficient intercept function, and $\mathcal{L}_X\beta_\tau$ is a functional linear operator. In this article, we consider the following three definitions of the operator $\mathcal{L}_X\beta_\tau$, but not limited to them, which are very rich in forms. Our definitions of the operator $\mathcal{L}_X\beta_\tau$ are

(i) functional varying coefficient QR model, $\mathcal{L}_X : L_2(\mathcal{T}) \to L_2(\mathcal{T})$ by

$$(\mathcal{L}_X\beta)(t) = \beta_\tau(t)X(t); \tag{2.3}$$

(ii) functional linear QR model, $\mathcal{L}_X : L_2(\mathcal{S} \times \mathcal{T}) \to L_2(\mathcal{T})$ by

$$(\mathcal{L}_X\beta)(t) = \int_{\mathcal{S}} \beta_\tau(s,t)X(s)ds; \tag{2.4}$$

(iii) functional varying coefficient QR with history index, $\mathcal{L}_X: L_2(\mathcal{T}) \to L_2(\mathcal{T})$ by

$$(\mathcal{L}_X\beta)(t) = \beta_\tau(t)\int_0^\Delta \gamma_\tau(s)X(t-s)ds, \tag{2.5}$$

where Δ is a length of sliding window, and $t \in \mathcal{T} = [\Delta, T]$ with a suitable $T > 0$. In section 3, we will focus on considering the three functional quantile regression models.

Second, we give the representations of predict and response functions in our model through functional principal components. The observed data consist of square integrable random predictor and response trajectories (X_i, Y_i), $i = 1, \cdots, n$, which are the realization of the underlying smooth random trajectory processes (X, Y). We usually refer to the arguments of $X(\cdot)$ and $Y(\cdot)$ as time or location, with compact interval domains $\mathcal{S}$ and $\mathcal{T}$ respectively. So $X \in L_2(\mathcal{S})$ and $Y \in L_2(\mathcal{T})$.The two smooth random trajectories have unknown smooth mean functions $\mu_X = EX(s)$, $\mu_Y = EY(t)$, and covariance functions $r_{XX}(s,t) = \mathrm{Cov}(X(s), X(t))$, $s,\ t \in \mathcal{S}$ and $r_{YY}(s,t) = \mathrm{Cov}(Y(s), Y(t))$, $s,\ t \in \mathcal{T}$, respectively. Similarly we can define covariance function $r_{XY}(s,t) = \mathrm{Cov}(X(s), Y(t))$, $s \in \mathcal{S}, t \in \mathcal{T}$. We have orthogonal expansions of r_{XX} and r_{YY} in terms of eigenfunctions ϕ_m and ψ_k with nonincreasing eigenvalues ρ_m and λ_k, that is,

$$r_{XX}(s_1, s_2) = \sum_{m=1}^{\infty} \rho_m \phi_m(s_1)\phi_m(s_2), \qquad r_{YY}(t_1, t_2) = \sum_{k=1}^{\infty} \lambda_k \psi_k(t_1)\psi_k(t_2),$$

for $s_1, s_2 \in \mathcal{S}$, $t_1, t_2 \in \mathcal{T}$. The Karhunen-Loève L_2 representations have

$$X(s) = \mu_X(s) + \sum_{m=1}^{\infty} \zeta_m \phi_m(s),\ s \in \mathcal{S};\ Y(t) = \mu_Y(t) + \sum_{k=1}^{\infty} \xi_k \psi_k(t),\ t \in \mathcal{T} \tag{2.6}$$

with a sequence of uncorrelated (independent in the Gaussian case) functional principal components ζ_m, (resp. ξ_k) with $\mathrm{E}[\zeta_m] = 0$ and $\mathrm{Var}[\zeta_m] = \rho_m$ for all m (resp. $E[\xi_k] = 0$ and $\mathrm{Var}[\xi_k] = \lambda_k$ for all k). The $\{(\rho_m, \phi_m)\}$ and $\{(\lambda_k, \psi_k)\}$ respectively are the pairs of eigenvalues and eigenfunctions of the covariance operators of X and Y.

2.2 Estimation strategy

Without loss of generality, assuming that $\alpha_\tau(t) = 0$ for all t, one may simplify the proposed FLOQR (1.2) to

$$Y(t) = (\mathcal{L}_X \beta_\tau)(t) + \varepsilon(t), \tag{2.7}$$

where we assume that $P(\varepsilon(t) < 0) = \tau$ for each $t \in \mathcal{T}$, and ε is independent of the X. Otherwise, we can shift the $\varepsilon(t)$ to be $\varepsilon(t) - F_t^{-1}(\tau)$, so that this assumption holds, where F_t is the cumulative distribution function of $\varepsilon(t)$. By the way, denote f_t to be the density of the noise $\varepsilon(t)$. In addition, if $\alpha_\tau(t) \neq 0$, we have a new functional operator $(\mathcal{L}_{\tilde{X}} \tilde{\beta}_\tau)(t)$ such that $Y(t) = \alpha_\tau(t) + (\mathcal{L}_X \beta_\tau)(t) + \varepsilon(t) = (\mathcal{L}_{\tilde{X}} \tilde{\beta}_\tau)(t) + \varepsilon(t)$, where $\tilde{X} = (1, X)$ and $\tilde{\beta}_\tau = (\alpha_\tau, \beta_\tau)^T$. Notice that $\mathcal{L}_{\tilde{X}} = (\mathcal{L}_I, \mathcal{L}_X)$, where $\mathcal{L}_I$ is identity operator. Obviously, $\mathcal{L}_{\tilde{X}}$ is a linear operator. For FLOQR model (2.7), the estimation of the common slope function β_τ can be transformed into the following stochastic optimization problem:

$$\beta_\tau^* = \mathrm{argmim}_{\beta_\tau} \mathrm{E} \int_{\mathcal{T}} \rho_\tau \left\{ Y(t) - (\mathcal{L}_X \beta_\tau)(t) \right\} dt. \tag{2.8}$$

Our motivation is to translate the quantile regression optimization to least squares estimation, and then take full advantage of least squares estimation. We know that there is a classic prediction method in a functional linear regression by estimating $E(Y|X)$ (Draper and Smith, 1998; Yao et al., 2005b). An important step is to estimate the regression coefficient function β. In the section, we explore the estimation strategy of FLOQR (1.2) by investigating the population least squares for functional linear operator mean regression model $\mathrm{E}[Y(t)|X(t)] = (\mathcal{L}_X \beta)(t)$. We seek the solution of the functional linear operator model by finding the function β^* which minimizes the squared distance, that is, $\beta^* = \mathrm{argmin}_\beta \mathrm{E}\|Y - \mathcal{L}_X \beta\|^2$. Let $\mathcal{L}_X^*$ is the adjoint operator of $\mathcal{L}_X$, $\Gamma_{XX} = \mathrm{E}[\mathcal{L}_X^* \mathcal{L}_X]$, and denote the range of the Γ_{XX}

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

as $\mathbb{R}(\Gamma_{XX})$. Then, $\beta = \Gamma_{XX}^{-1}\mathrm{E}[\mathcal{L}_X^* Y]$ exists and is the unique solution of the score equation $\Gamma_{XX}\beta = \mathrm{E}[\mathcal{L}_X^* Y]$ in $\mathbb{R}(\Gamma_{XX})$, and β^* can be expanded in the eigenbasis representation (He et al., 2000). For example, if $\mathcal{L}_X\beta$ is (2.4), β has the representation (He et al., 2000; Yao et al., 2005b), with convergence under mild regularity conditions:

$$\beta(s,t) = \sum_{m=1}^{\infty}\sum_{k=1}^{\infty}\frac{\mathrm{E}[\zeta_m\xi_k]}{\mathrm{E}[\zeta_m]}\phi_m(s)\psi_k(t). \tag{2.9}$$

This novel result, the eigenbasis representation of β^*, makes it possible to handle the sparsity and irregularity of the longitudinal data, and also to incorporate additional information that is inherent in the underlying covariance structure in the pivotal estimation step.

But the non-smooth quantile loss ρ_τ poses new challenges to the estimation of the quantile regression model because of its diamond-shaped polyhedral contours (Koenker, 2005). We do not obtain an elegant "closed-form" solution such as $\beta = \Gamma_{XX}^{-1}\mathrm{E}[\mathcal{L}_X^* Y]$ for mean regression $\mathrm{E}[Y(t)|X(t)] = (\mathcal{L}_X\beta)(t)$. Thus, we cannot obtain the eigenbasis representation of β^*. Moreover, some optimization algorithms are proposed to solve the estimation of quantile regression. As discussed in (Chen and Wei, 2005; Koenker, 2005), the optimization problem (2.1) can be written as a linear programming problem and solved by the simplex method or the interior point method. It is well-known that the simplex method is computationally demanding where the worst-case complexity increases exponentially with the data size. And while the interior point method also has shortcomings in dealing with modern scale data (Yang et al., 2014). These methods are mainly difficult to obtain the eigenbasis representation of β, so we cannot use the technical advantages of functional nature to model sparse longitudinal data.

Motivated by the connection between quantile regression and ordinary linear regression (Chen et al., 2020), we develop a unified estimation strategy for FLOQR based on the mean regression. Let $G(\beta_\tau; X, Y) = \int_\mathcal{T} \rho_\tau \{Y(t) - (\mathcal{L}_X\beta_\tau)(t)\}\, dt$. The subgradient and Hessian-type functions take the form of

$$g(\beta_\tau; X, Y) = \frac{\partial G(\beta_\tau; X, Y)}{\partial \beta_\tau} = \mathcal{L}_X^* \int_\mathcal{T} \{\mathbb{I}[Y(t) - (\mathcal{L}_X\beta_\tau)(t) \le 0] - \tau\} dt,$$

$$H(\beta_\tau) = \frac{\partial \mathrm{E} g(\beta_\tau; X, Y)}{\partial \beta_\tau} = \mathrm{E}\left[\int_\mathcal{T} \mathcal{L}_X^* \mathcal{L}_X f_t(\mathcal{L}_X(\beta_\tau - \beta_\tau^*))dt\right].$$

To solve the stochastic optimization (2.8), we begin with the Newton-Raphson iteration algorithm. Its population version is

$$\tilde{\beta}_\tau^{(1)} = \beta_\tau^{(0)} - H^{-1}(\beta_\tau^{(0)})\mathrm{E}[g(\beta_\tau^{(0)}; X, Y)], \tag{2.10}$$

where $\beta_\tau^{(0)}$ is an initial estimation. The algorithm (2.10) is not feasible because it contains β_τ in H that needs to be estimated. When $\beta_\tau^{(0)}$ is close to the true slope β_τ, we have

$$H(\beta_\tau^{(0)}) \approx H(\beta_\tau) \;\; = \;\; \mathrm{E}\left[\int_\mathcal{T} \mathcal{L}_X^* \mathcal{L}_X f_t(0)dt\right] = \int_\mathcal{T} \Gamma_{XX} f_t(0)dt,$$

which is estimable. It impels the following iteration

$$\beta_\tau^{(1)} = \beta_\tau^{(0)} - H^{-1}(\beta_\tau)\mathrm{E}[g(\beta_\tau^{(0)}; X, Y)]. \tag{2.11}$$

Under some regularity conditions, the Taylor expansion of $\mathrm{E}[g(\beta_\tau^{(0)}; X, Y)]$ at β_τ, $\mathrm{E}[g(\beta_\tau^{(0)}; X, Y)] = H(\beta_\tau)(\beta_\tau^{(0)} - \beta_\tau) + O(\|\beta_\tau^{(0)} - \beta_\tau\|_2^2)$. Thus, it is easy to see that

$$\begin{aligned} \|\beta_\tau^{(1)} - \beta_\tau\|_2 &= \left\|\beta_\tau^{(0)} - H^{-1}(\beta_\tau)\left[H(\beta_\tau)(\beta_\tau^{(0)} - \beta_\tau) + O(\|\beta_\tau^{(0)} - \beta_\tau\|_2^2)\right] - \beta_\tau\right\| \\ &= O(\|\beta_\tau^{(0)} - \beta_\tau\|_2^2), \end{aligned}$$

which implies that we can refine $\beta_\tau^{(1)}$ closer to the true β_τ by one step Newton-Raphson iteration (2.11) if $\beta_\tau^{(0)}$ is a consistent estimator of β_τ. But we find it is complex to compute $H^{-1}(\beta_\tau^{(0)})$ because of inverse operation $\left(\int_\mathcal{T} \Gamma_{XX} f_t^{(0)}(0)dt\right)^{-1}$. From (2.11), we find that the algorithm is similar to gradient descent. The $H^{-1}(\beta_\tau^{(0)})$ can be regarded as the learning rate of the gradient descent iterative algorithm. Therefore, for constructing our SLSE algorithm of FLOQR and computational simplicity, we replace $\left(\int_{\mathcal{T}} \Gamma_{XX} f_t^{(0)}(0)dt\right)^{-1}$ with $\Gamma_{XX}^{-1}\left(\int_{\mathcal{T}} f_t^{(0)}(0)dt\right)^{-1}$. Thus, it facilitates our algorithm. Therefore, the iteration (2.11) can be written as

$$\begin{aligned}
\beta_\tau^{(1)} &= \Gamma_{XX}^{-1}\left[\Gamma_{XX}\beta_\tau^{(0)} - \left(\int_{\mathcal{T}} f_t^{(0)}(0)dt\right)^{-1} \mathrm{E}[g(\beta_\tau^{(0)}; X, Y)]\right] \\
&= \Gamma_{XX}^{-1}\mathrm{E}\left\{\mathcal{L}_X^*\left[\mathcal{L}_X\beta_\tau^{(0)} - \left(\int_{\mathcal{T}} f_t^{(0)}(0)dt\right)^{-1}\int_{\mathcal{T}}\{\mathbb{I}[Y(t) - (\mathcal{L}_X\beta_\tau^{(0)})(t) \leq 0] - \tau\}dt\right]\right\}.
\end{aligned}$$

Define a new response trajectory

$$\widetilde{Y} = \mathcal{L}_X\beta_\tau^{(0)} - \left(\int_{\mathcal{T}} f_t^{(0)}(0)dt\right)^{-1}\int_{\mathcal{T}}\{\mathbb{I}[Y(t) - (\mathcal{L}_X\beta_\tau^{(0)})(t) \leq 0] - \tau\}dt,$$

then get

$$\beta_\tau^{(1)} = \Gamma_{XX}^{-1}\mathrm{E}[\mathcal{L}_X^*\widetilde{Y}], \tag{2.12}$$

which is the linear operator regression coefficient (He et al., 2000). It can be regarded as a surrogate least squares estimation (SLSE)

$$\beta_\tau^{(1)} = \mathrm{argmin}_{\beta_\tau}\mathrm{E}\|\widetilde{Y} - \mathcal{L}_X\beta_\tau\|_2^2. \tag{2.13}$$

Further, we have the eigenbasis representation of $\beta_\tau^{(1)}$ such as (2.9) based on functional principal components analysis. In the estimation strategy, there are still two issues: the initial estimation $\beta_\tau^{(0)}$ and the density $f_t(0)$. From a population perspective, we address the two issues. For $\beta_\tau^{(0)}$, we can take $\beta_\tau^{(0)} = \beta^* = \Gamma_{XX}^{-1}\mathrm{E}[\mathcal{L}_X^*Y]$, a function linear operator regression estimator. In practice, it works well. Kernel density estimator can be used to get $f_t(0)$, since $E[K_h\{Y(t) - \mathcal{L}_X\beta_\tau^*(t)\}] - f_t(0) = \frac{1}{2}h^2\mu_2(K)f_t''(0) + o(h^2)$ under mild regularity conditions, where $K(\cdot)$ is a kernel function, $K_h(\cdot) = h^{-1}K(\cdot/h)$ with bandwidth h, and $\mu_2(K) = \int z^2 K(z)dz$. See Fan and Gijbels (1996).

We present the entire estimation procedure in Algorithm 1 from a population viewpoint.

Algorithm 1: SLSE algorithm for functional linear operator QR model (2.7).

Input: Initialize estimator $\beta_\tau^{(0)} = \Gamma_{XX}^{-1}\mathrm{E}[\mathcal{L}_X^* Y]$, kernel function $K(\cdot)$, bandwidth h, quantile level τ and the number of iterations K.

for $k = 1, 2, \cdots, \mathrm{K}$ **do**

Estimate $f_t^{(k)}(0) = \mathrm{E}[K_h\{Y(t) - \mathcal{L}_X\beta_\tau^{(k-1)}(t)\}]$ by kernel density estimation.

Compute

$$\widetilde{Y}^{(k)}(t) = \mathcal{L}_X\beta_\tau^{(k-1)} - \left(\int_\mathcal{T} f_t^{(k)}(0)dt\right)^{-1}\int_\mathcal{T}\{\mathbb{I}[Y(t) - (\mathcal{L}_X\beta_\tau^{(0)})(t) \le 0] - \tau\}dt.$$

Obtain $\beta_\tau^{(k)} = \Gamma_{XX}^{-1}\mathrm{E}[\mathcal{L}_X^*\widetilde{Y}^{(k)}]$.

end

Output: The final estimator $\beta_\tau^{(\mathrm{K})}$.

2.3 Theoretical guarantees

In the paper, the random processes we consider are square integrable, i.e., are in the L_2 space of square-integrable functions. For a compact interval $\mathcal{T}$, $L_2(\mathcal{T})$ is a Hilbert space when equipped with the inner product $\langle f, g\rangle = \int_\mathcal{T} f(t)g(t)dt$ for $f, g \in L_2(\mathcal{T})$ where dt is the Lebesgue measure, which generates the norm $\|\cdot\|$. If $\beta \in L_2(\mathcal{S}\times\mathcal{T})$, we define $\|\beta\|^2 = \int_\mathcal{T}\int_\mathcal{S}\beta^2(s,t)dsdt$. Note that our results can be easily extended to cover more general measures ν and scalar products in spaces $L_2(\mathcal{T};\nu)$. In our FLOQR model (2.7), assume that X and ε are uncorrelated.

In multivariate analysis, the unique minimizer of a linear regression model can be found when the covariance matrix of covariate is invertible, by classical least squares theory. How-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

ever, the theory cannot be simply extended to the functional setting. For (2.13), the solution is not unique even if $\widetilde{Y}$ is Y because there is a problem of the nonexistence of the inverse of covariance (operator) of X in the infinite-dimensional setting. The problem can be alleviated by a suited generalized inverse under the following conditions, which refers to the Karhunen-Loève decomposition (2.6).

Condition 1. The eigenvalues ρ_m of the covariance operator of X are positive, and X is square integrable (i.e. $E\|X\|^2 = E\int X^2(t)dt < \infty$).

Condition 2. The linear operator regression coefficient β_τ satisfies $\|\beta_\tau\| < \infty$.

Condition 3. The density function of the noise $f_t(\cdot)$ is bounded and Lipschitz continuous (that is, $|f_t(x) - f_t(y) \leq C_L|x-y|$ for any $x, y \in \mathbb{R}$ and some constant $C_L > 0$) for any $t \in \mathcal{T}$. Moreover, assume $f_t(0) > c > 0$ for some constant c and for any $t \in \mathcal{T}$.

Remark 2.1. *For Condition 1, $\rho_m > 0$ for $m \geq 1$ implies the inverse of the linear operator Γ_{XX} exists. A square integrable X guarantees $\sum_{m=1}^{\infty} \rho_m < \infty$. Let the covariance operator of X to be G, and $EX = 0$. Indeed, by the relation $G\phi_m = \rho_m \phi_m$, we have $\rho_m = \langle G\phi_m, \phi_m \rangle = \langle E[\langle X, \phi_m \rangle], \phi_m \rangle = E\left[\langle X, \phi_m \rangle^2\right]$. The eigenfunctions ϕ_m are orthogonal.By Parseval's equality,*

$$\sum_{m=1}^{\infty} \rho_m = \sum_{m=1}^{\infty} E\left[\langle X, \phi_m \rangle^2\right] = E\|X\|^2 < \infty.$$

Also see the definition of X on page 4 of He et al. (2000), or the random elements in $\mathcal{L}_2$ and the covariance operator in Subsection 2.3 in Horváth and Kokoszka (2012). In Condition 2, $\|\beta_\tau\| < \infty$ means β_τ is integrable in L_2 space. The condition ensures that β_τ in (2.13) has a representation similar to (2.9), and the representation converges uniformly. The condition

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

has been used in Şentürk and Müller (2010); Yao et al. (2005b); He et al. (2000). Condition 3 is a regular condition on the smoothness of the density function f_t. The condition about density function of the noise is required in quantile regression (Koenker, 2005; Kato, 2012; Liu et al., 2020b). For example, Kato (2012) and Liu et al. (2020b) assumed that density function is continuously differentiable and bounded; Zhou et al. (2023) required that density function is uniformly bounded away from zero and ∞, and has a continuous second-order derivative at t. These assumptions are stronger than Condition 3.

Theorem 2.1. *Assume that Conditions 1-3 hold. Then*

(a) $\beta_\tau^{(1)} = \Gamma_{XX}^{-1}\mathcal{L}_X^*\widetilde{Y}$ *exists and is the unique solution of (2.13) in* $\mathbb{R}(\Gamma_{XX})$;

(b) $\left\|\beta_\tau^{(1)} - \beta_\tau\right\| = O\left(\left\|\left[E^{1/2}(\mathcal{L}_X\mathcal{L}_X^*\mathcal{L}_X)\right](\beta_\tau^{(0)} - \beta_\tau^*)\right\|^2\right)$;

(c) $\left\|\beta_\tau^{(k)} - \beta_\tau\right\| = O\left(\|E(\mathcal{L}_X\mathcal{L}_X^*\mathcal{L}_X)\|^{2k-1}\left\|\beta_\tau^{(0)} - \beta_\tau^*\right\|^{2k}\right)$.

Remark 2.2. *From the perspective of population, we give the closed-form solution of quantile regression as ordinary linear regression by transforming the response trajectory Y to $\widetilde{Y}$. If we have a consistent estimator $\beta_\tau^{(0)}$, we can refine it by multiple rounds of iteration. In fact, we do not require good initialization since quantile loss is convex and Algorithm 1 is based on Newton-Raphson optimization, which enjoys quadratic converge guarantees under general conditions.*

3. Specific functional quantile regression for sparse longitudinal data

In this section, we will propose three specific functional quantile regression models for analyzing sparse and irregular data. The $\mathcal{L}_X\beta$ in FLOQR model (1.2) corresponds to (2.3)-(2.5) respectively. We call them functional varying coefficient QR model, functional linear QR

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

model, and functional varying coefficient QR model with history index. The following subsections will give their estimates, predictions, inferences, and asymptotic properties.

Although the population methodology was derived in the previous section, in practice they must be estimated from available observed data. Our observed predictor and response trajectories X_i and Y_i are measured at sparse and irregularly spaced locations or time points, and contaminated with additional measurement errors (Staniswalis and Lee, 1998; Rice and Wu, 2001; Yao et al., 2005b; Wu et al., 2010). To adequately reflect the sparse and irregular measurements, we assume that the random measurements of X_i (resp. Y_i) for the ith subject are at locations $S_{i1}, \cdots, S_{iL_i}$ (resp. $T_{i1}, \cdots, T_{iN_i}$), where L_i and N_i are assumed to be i.i.d. as L and N (which may be correlated, but are independent of all other random variables), respectively. Let U_{il} (resp. V_{ij}) be the observation of the random trajectory X_i (resp. Y_i) made at a random location S_{il} (resp. T_{ij}), contaminated with measurement errors $\epsilon_{X,il}$ (resp. $\epsilon_{Y,ij}$). Based on the Karhunen-Loève L_2 representations (2.6) of the trajectories X and Y, we may represent predictor and response observations for the subjects $i = 1, \cdots, n$ as follows:

$$U_{il} = X_i(S_{il}) + \epsilon_{X,il} = \mu_X(S_{il}) + \sum_{m=1}^{\infty} \zeta_{im}\phi_m(S_{il}) + \epsilon_{X,il},\ 1 \le l \le L_i, \tag{3.1}$$

$$V_{ij} = Y_i(T_{ij}) + \epsilon_{Y,ij} = \mu_Y(T_{ij}) + \sum_{k=1}^{\infty} \xi_{ik}\psi_k(T_{ij}) + \epsilon_{Y,ij},\ 1 \le j \le N_i, \tag{3.2}$$

where the errors $\epsilon_{X,il}$ and $\epsilon_{Y,ij}$ are assumed to be i.i.d. with mean zero and variance σ_X^2 and σ_Y^2, and independent of component scores ζ_{im} (resp ξ_{ik}) which have $\mathrm{E}\zeta_{im} = 0$, $\mathrm{E}\zeta_{im}^2 = \rho_m\rho$ and $\mathrm{E}[\zeta_{im}\zeta_{im'}] = 0$ for $m \neq m'$ (resp. $\mathrm{E}\xi_{ik} = 0$, $\mathrm{E}\xi_{ik}^2 = \lambda_k\rho$ and $\mathrm{E}[\xi_{ik}\xi_{ik'}] = 0$ for $k \neq k'$).

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

3.1 Functional varying coefficient QR model for sparse longitudinal data

In the subsection, we consider the scenario $(\mathcal{L}_X\beta_\tau)(t) = \beta_\tau(t)X(t)$ for $t \in \mathcal{T}$. The functional/longitudinal varying coefficient QR model is

$$Q_{Y|X}(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)X(t), \tag{3.3}$$

where it is assumed that the varying coefficient functions α_τ and β_τ are smooth for a given quantile level τ. We should optimize a QR problem via check function $\rho_\tau(\cdot)$. But it may be biased or inefficient in case of sparse, irregular, noise-corrupted measurements U_{ij} and V_{ij}.

3.1.1 Functional approach

Based on the estimation strategy in Section 2.2, given initial estimators $\alpha_\tau^{(0)}$ and $\beta_\tau^{(0)}$, we can translate the model (3.3) into the following functional linear operator regression

$$\mathrm{E}[\widetilde{Y}(t)|X(t)] = \alpha_\tau(t) + (\mathcal{L}_X\beta_\tau)(t), \tag{3.4}$$

where $(\mathcal{L}_X\beta_\tau)(t) = \beta_\tau(t)X(t)$ and

$$\widetilde{Y}(t) = Q_{Y|X}^{(0)}(t;\tau) - \left(\int_{\mathcal{T}} f_t^{(k)}(0)dt\right)^{-1} \int_{\mathcal{T}} \{\mathbb{I}[Y(t) - Q_{Y|X}^{(0)}(\tau;t) \le 0] - \tau\}dt$$

with $Q_{Y|X}^{(0)}(t;\tau) = \alpha_\tau^{(0)}(t) + \beta_\tau^{(0)}(t)X(t)$. Similarly, we have the covariance expansion of $\widetilde{Y}$

$$r_{\widetilde{Y}\widetilde{Y}}(t_1,t_2) = \mathrm{Cov}\left(\widetilde{Y}(t_1), \widetilde{Y}(t_2)\right) = \sum_{l=1}^{\infty} \kappa_l \varphi_l(t_1)\varphi_l(t_2),\ t_1,t_2 \in \mathcal{T}, \tag{3.5}$$

where the eigenvalues κ_l of the covariance operator are positive and ordered, $\kappa_1 > \kappa_2 > \cdots$; φ_l is the corresponding eigenfunction of κ_l; and the function principal components ς_l satisfy $\mathrm{E}[\varsigma_l] = 0$ and $\mathrm{Var}[\varsigma_l] = \kappa_l$ for all l. Let $\widetilde{Y}^c(t) = \widetilde{Y}(t) - \mathrm{E}[\widetilde{Y}(t)]$ and $X^c(t) = X(t) - \mu_X(t)$. The model (3.4) can be written as

$$\mathrm{E}[\widetilde{Y}^c(t)|X(t)] = \beta_\tau(t)X^c(t) \tag{3.6}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

with $\alpha(t) = \mathrm{E}[\widetilde{Y}(t)] - \beta_\tau(t)\mu_X(t)$. By our result (2.12), one gets the following population least squares representation

$$\begin{aligned}\beta_\tau^{(1)}(t) &= \operatorname{argmin}_{\beta_\tau}\mathrm{E}\|\widetilde{Y}^c(t) - \beta_\tau(t)X^c(t)\|_2^2 = \Gamma_{X^cX^c}^{-1}E\left(\mathcal{L}_{X^c}^*\widetilde{Y}^c\right)\\ &= \frac{\mathrm{Cov}(X^c(t), Y^c(t))}{\mathrm{Var}(X^c(t))} = \frac{\mathrm{Cov}(X(t), \widetilde{Y}(t))}{\mathrm{Var}(X(t))} = \frac{r_{X\widetilde{Y}}(t,t)}{r_{XX}(t)},\end{aligned} \quad (3.7)$$

with the cross-covariance function between X and $\widetilde{Y}$,

$$r_{X\widetilde{Y}}(s,t) = \sum_{m=1}^{\infty}\sum_{l=1}^{\infty}\mathrm{E}[\zeta_m\varsigma_l]\phi_m(s)\varphi_l(t). \quad (3.8)$$

Some details about the estimate for $\alpha_\tau^{(1)}(t)$ and $\beta_\tau^{(1)}(t)$ are given in Appendix S1.1 of Supplementary Material. We sketch the functional estimation approach for functional varying coefficient QR model (3.3) in Algorithm 2 by combining the above steps. Thus, the final estimators $\hat{\alpha}_\tau^{(\mathrm{K})}$ and $\hat{\beta}_\tau^{(\mathrm{K})}$ are obtained via Algorithm 2.

Algorithm 2: SLSE algorithm for functional linear operator QR model (3.3).

Input: Kernel function $K(\cdot)$, bandwidth h, quantile level τ and the number of iterations K.

Calculate mean function $\hat{\mu}_X$ and $\hat{\mu}_Y$, covariance surface $\hat{r}_{XX}$, cross-covariance surface $\hat{r}_{XY}$, eigenfunctions $\hat{\phi}_k$ and$\hat{\psi}_k$, and eigenvalues $\hat{\rho}_k$ and $\hat{\lambda}_k$ by Steps 1-2.

Initialize estimators $\hat{\alpha}_\tau^{(0)}$ and $\hat{\beta}_\tau^{(0)}$ by Step 3.

for $k = 1, 2, \cdots, \mathrm{K}$ **do**

Estimate $\hat{f}_t^{(k)}(0)$, $\hat{\omega}_i$, $\hat{U}_i$ and $\hat{V}_i$ for obtaining $\widetilde{V}_{ij}$ by Step 4.

Compute mean function $\hat{\mu}_{\widetilde{Y}}$ via Step 5, cross-covariance surface $\hat{r}_{X\widetilde{Y}}$ by Step 6.

Obtain $\hat{\alpha}_\tau^{(k)}$ and $\hat{\beta}_\tau^{(k)}$ by Step 7.

end

Output: The final estimators $\hat{\alpha}_\tau^{(\mathrm{K})}$ and $\hat{\beta}_\tau^{(\mathrm{K})}$.

3.1.2 Uniform consistency and prediction

Some assumptions and proofs are provided in Appendix S5 of SM. We first present the convergence rate for $\hat{\alpha}_\tau^{(1)}$ and $\hat{\beta}_\tau^{(1)}$ after one iteration.

Theorem 3.1. *Let* $\sup_{t\in\mathcal{T}}|\hat{\alpha}_\tau^{(0)}(t)-\alpha_\tau(t)| = O_p(a_n)$ *and* $\sup_{t\in\mathcal{T}}|\hat{\beta}_\tau^{(0)}(t)-\beta_\tau(t)| = O_p(a_n)$. *Under Conditions 1-3, and Assumptions (A1)-(A6) in Appendix S5 of SM, we have*

$$\sup_{t\in\mathcal{T}}|\hat{\alpha}_\tau^{(1)}(t)-\alpha_\tau(t)| = O_p\left\{\frac{1}{\sqrt{n}}\left[\frac{1}{b_X}+\frac{1}{b_Y}+\frac{1}{h_X^2}+\frac{1}{h_Y^2}+\frac{1}{h_1h_2}\right]+a_n^2\right\},$$
$$\sup_{t\in\mathcal{T}}\|\hat{\beta}_\tau^{(1)}(t)-\beta_\tau(t)| = O_p\left\{\frac{1}{\sqrt{n}}\left[\frac{1}{b_X}+\frac{1}{b_Y}+\frac{1}{h_X^2}+\frac{1}{h_Y^2}+\frac{1}{h_1h_2}\right]+a_n^2\right\}.$$

Theorem 3.2. *Let* $\sup_{t\in\mathcal{T}}|\hat{\alpha}_\tau^{(0)}(t)-\alpha_\tau(t)| = O_p(a_n)$ *and* $\sup_{t\in\mathcal{T}}|\hat{\beta}_\tau^{(0)}(t)-\beta_\tau(t)| = O_p(a_n)$. *Under Conditions 1-3, and Assumptions (A1)-(A6) in Appendix S5 in SM, we have for* k *iterations of Algorithm 2,*

$$\sup_{t\in\mathcal{T}}|\hat{\alpha}_\tau^{(k)}(t)-\alpha_\tau(t)| = O_p\left\{\frac{1}{\sqrt{n}}\left[\frac{1}{b_X}+\frac{1}{b_Y}+\frac{1}{h_X^2}+\frac{1}{h_Y^2}+\frac{1}{h_1h_2}\right]+a_n^{k+1}\right\},$$
$$\sup_{t\in\mathcal{T}}\|\hat{\beta}_\tau^{(k)}(t)-\beta_\tau(t)| = O_p\left\{\frac{1}{\sqrt{n}}\left[\frac{1}{b_X}+\frac{1}{b_Y}+\frac{1}{h_X^2}+\frac{1}{h_Y^2}+\frac{1}{h_1h_2}\right]+a_n^{k+1}\right\}. \tag{3.9}$$

Remark 3.1. *The bandwidths* h_1, h_2, b_X, *and* h_X *are used in estimation for mean and (cross)-covariance functions as described explicitly in Appendixes S1 and S5. It can be shown that when the iteration number* k *is sufficiently large, that is,* $k \geq \dfrac{\log(c_0 b_n)}{\log a_n} - 1, for\ some\ c_0 > 0,$ *where* $b_n = \dfrac{1}{\sqrt{n}}\left[\dfrac{1}{b_X}+\dfrac{1}{b_Y}+\dfrac{1}{h_X^2}+\dfrac{1}{h_Y^2}+\dfrac{1}{h_1h_2}\right]$, *which match the convergence rate of least square estimator (see Şentürk and Müller (2010)).*

Remark 3.2. *In fact, the estimation of quantile regression is a convex optimization problem because the check function* $\rho_\tau(\cdot)$ *is a convex loss. Therefore, its optimization solution is essentially independent of the initial value. Therefore, we adopt the typical initial estimator* $\hat{\beta}_\tau^{(0)} = \Gamma_{XX}^{-1}E[\mathcal{L}_X^* Y]$ *in practice. Due to technical limitations, we obtain the results such as Theorem*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

3.2, which depends on the accuracy of the initial estimate $O_p(a_n)$. For functional varying coefficient QR model (3.3), we observe the data $\{(T_{ij}, U_{ij}, V_{ij}) : i = 1, \cdots, n, j = 1, \cdots, N_i\}$. We can obtain $\hat{\beta}_\tau^{(0)}$ by minimizing the following locally weighted quantile regression loss function

$$\sum_{i=1}^{n}\sum_{j=1}^{N_i} \rho_\tau \left(V_{ij} - U_{ij}\left(a + b\left(\frac{T_{ij}-t}{h}\right)\right)\right) K\left(\frac{T_{ij}-t}{h_k}\right),$$

where K is a kernel function, and h is a bandwidth. Thus, we have $\hat{\beta}_\tau^{(0)}(t) = \hat{a}$. By Theorem 1 Yao (2007), we have $|\hat{\beta}_\tau^{(0)}(t) - \beta_\tau(t)| = O_p(a_n) \to 0$ with $a_n = (h\sum_{i=1}^n N_i)^{-1/2}$ under a suitable conditions. So, we can also take the above locally weighted quantile regression estimation $\hat{\beta}_\tau^{(0)}(t)$ as an alternative initial value, which meets the theoretical consistency condition.

One of our central tasks is to predict the tth quantile trajectory $Q_{Y^*|X^*}(t;\tau)$ of the response Y^* for a new subject from a sparse predictor trajectory X^*, we get from (3.3),

$$Q_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)\left(\mu_X(t) + \sum_{m=1}^{\infty}\zeta_m^*\phi_m(t)\right), \tag{3.10}$$

where $\zeta_m^* = \int_{\mathcal{T}}(X^*(t) - \mu_X(t))\,\phi_m(t)$ is the mth functional principal component of X^*. These quantities μ_X, ϕ_m, α_τ and β_τ in (3.10) can be estimated based on the data. Details can be found in Algorithm 2 and Appendix S1.1. The FPC scores ζ_m^* can be estimated by the traditional numerical integration, but the estimate can't provide reasonable approximations to ζ_m^* because of our sparse longitudinal data. For example, when one has only two observations per subject. Following (Yao et al., 2005b), we invoke Gaussian assumptions for the estimation of ζ_m^* to handle sparsity of data.

Let U_j^* be the jth measurement at location T_j^* made for the predictor function X_j^*, $j = 1, \cdots, N^*$, with a random number N^*. The observed data $U_j^* = X_j^* + \epsilon_{X,j}^*$. Assume that the FPC scores ζ_m^* and the measurement errors $\epsilon_{X,j}^*$ are jointly Gaussian. From Yao et al. (2005b,a), the best predictor of the scores ζ_m^*, given $\boldsymbol{U}^* = (U_1^*, \cdots, U_{N^*}^*)^T$, $\boldsymbol{T}^* = (T_1^*, \cdots, T_{N^*}^*)^T$ and their number N^*, is

$$\tilde{\zeta}_m^* = \rho_m \boldsymbol{\phi}_m^{*T} \boldsymbol{\Sigma}_{U^*}^{-1} (\boldsymbol{U}^* - \boldsymbol{\mu}_X^*), \tag{3.11}$$

where $\boldsymbol{\mu}_X^* = (\mu_X(T_1^*), \cdots, \mu_X(T_{N^*}^*))^T$, $\boldsymbol{\phi}_m^* = (\phi_m(T_1^*), \cdots, \phi_m(T_{N^*}^*))^T$, $\boldsymbol{\Sigma}_{U^*} = \mathrm{Cov}(\boldsymbol{U}^*|\boldsymbol{T}^*, N^*) = \mathrm{Cov}(\boldsymbol{X}^*|\boldsymbol{T}^*, N^*) + \sigma_X^2 I_{N^*}$, I_{N^*} being the $N^* \times N^*$ identity matrix. Further, the (j,l)th entry $(\boldsymbol{\Sigma}_{U^*})_{j,l} = r_{XX}(T_j, T_l) + \sigma_X^2 \delta_{jl}$ with $\delta_{jl} = 1$ if $j = l$ and 0 if $j \neq l$. By (3.11), we have FPC scores

$$\hat{\zeta}_m^* = \hat{\rho}_m \hat{\boldsymbol{\phi}}_m^{*T} \hat{\boldsymbol{\Sigma}}_{U^*}^{-1} (\boldsymbol{U}^* - \hat{\mu}_X^*), \tag{3.12}$$

where $\hat{\Sigma}_{U^*} = \hat{r}_{XX}(T_j, T_l) + \hat{\sigma}_X^2 \delta_{jl}$. The details on obtaining $\hat{\sigma}_X^2$, see Appendix S1.1. The predicted tth quantile trajectories are

$$\widehat{Q}_{Y^*|X^*}^{\mathrm{K},M}(t;\tau) = \hat{\alpha}_\tau^{\mathrm{K}}(t) + \hat{\beta}_\tau^{\mathrm{K}}(t) \left(\hat{\mu}_X(t) + \sum_{m=1}^{M} \hat{\zeta}_m^* \hat{\phi}_m(t) \right). \tag{3.13}$$

The Gaussian assumption, which is crucial in the sparse situation, allows us to obtain the best linear predictors (3.13) via the conditional expectation. The functional approach borrows strength from the entire observations and thus makes up the sparseness of individual trajectories.

Theorem 3.3. *Let* $\sup_{t\in\mathcal{T}} |\hat{\alpha}_\tau^{(0)}(t) - \alpha_\tau(t)| = O_p(a_n)$ *and* $\sup_{t\in\mathcal{T}} |\hat{\beta}_\tau^{(0)}(t) - \beta_\tau(t)| = O_p(a_n)$, *with* $0 \le a_n < 1$. *Under Conditions 1-3, and Assumptions (A1)-(A7) in the Appendix S5, given* N^* *and* $\boldsymbol{T}^*$, *for all* $t \in \mathcal{T}$, *the predicted* τ*th quantile response trajectories in the FVCQR model (3.3) satisfy*

$$\lim_{n\to\infty} \widehat{Q}_{Y^*|X^*}^{\mathrm{K},M}(t;\tau) = \widetilde{Q}_{Y^*|X^*}(t;\tau), \quad \text{in proabaility,}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

for the target trajectory $\widetilde{Q}_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)\left(\mu_X(t) + \sum_{m=1}^{\infty}\tilde{\zeta}_m^*\phi_m(t)\right)$*, with* $M = M(n) \to \infty$ *as* $n \to \infty$*, and the iteration number* K *enough large.*

3.1.3 Asymptotic pointwise confidence bands for quantile response trajectories

We next construct asymptotic confidence bands for the τ-th quantile response trajectories $Q_{Y^*|X^*}$ of a new subject, conditional on the sparse, irregular noisy measurements of X^*. With $\boldsymbol{\zeta}_M^* = (\zeta_1^*, \cdots, \zeta_M^*)$, $\tilde{\boldsymbol{\zeta}}_M^* = \left(\tilde{\zeta}_1^*, \cdots, \tilde{\zeta}_M^*\right)$, where $\tilde{\zeta}_m^*$ is as in (3.11), the $M \times N^*$ matrix $\boldsymbol{H} = \mathrm{Cov}(\boldsymbol{\zeta}_M^*, \boldsymbol{U}^*|N^*, \boldsymbol{T}^*) = (\rho_1\phi_1^*, \cdots, \rho_M\phi_M^*)^T$, and thus covariance matrix of $\tilde{\boldsymbol{\zeta}}_M^*$ is $\mathrm{Cov}(\tilde{\boldsymbol{\zeta}}_M^*|N^*, \boldsymbol{T}^*) = \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{U}^*}^{-1}\boldsymbol{H}^T$. We observe that $\tilde{\boldsymbol{\zeta}}_m^* = \mathrm{E}\left[\boldsymbol{\zeta}_M^*|\boldsymbol{U}^*, N^*, \boldsymbol{T}^*\right]$, which implies that $\tilde{\boldsymbol{\zeta}}_M^*$ is the project of $\boldsymbol{\zeta}_M^*$ on the space spanned by the linear functions of $\boldsymbol{U}^*$ given N^* and $\boldsymbol{T}^*$. Hence,

$$\begin{aligned}\mathrm{Cov}\left(\tilde{\boldsymbol{\zeta}}_M^* - \boldsymbol{\zeta}_M^*|N^*, \boldsymbol{T}^*\right) &= \mathrm{Cov}\left(\boldsymbol{\zeta}_M^*|N^*, \boldsymbol{T}^*\right) - \mathrm{Cov}\left(\tilde{\boldsymbol{\zeta}}_M^*|N^*, \boldsymbol{T}^*\right)\\ &= \boldsymbol{D} - \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{U}^*}^{-1}\boldsymbol{H}^T \equiv \boldsymbol{\Omega}_M,\end{aligned}$$

where $\boldsymbol{D} = \mathrm{diag}(\rho_1, \cdots, \rho_M)$. Under Gaussian assumptions and conditional on N^* and $\boldsymbol{T}^*$,

$$\tilde{\boldsymbol{\zeta}}_M^* - \boldsymbol{\zeta}_M^* \sim N(0, \boldsymbol{\Omega}_M).$$

Denote $\hat{\boldsymbol{\Omega}}_M = \hat{\boldsymbol{D}} - \hat{\boldsymbol{H}}\hat{\Sigma}_{\boldsymbol{U}^*}^{-1}\hat{\boldsymbol{H}}^T$, where $\hat{\boldsymbol{D}} = \mathrm{diag}(\hat{\rho}_1, \cdots, \hat{\rho}_M)$ and $\hat{\boldsymbol{H}} = (\hat{\rho}_1\hat{\phi}_1^*, \cdots, \hat{\rho}_M\hat{\phi}_M^*)^T$. Define $\boldsymbol{\phi}_\tau(t) = (\beta_\tau(t)\phi_1(t), \cdots, \beta_\tau(t)\phi_M(t))^T$, and $\hat{\boldsymbol{\phi}}_\tau^{\mathrm{K},M}(t) = \left(\hat{\beta}_\tau^{\mathrm{K}}(t)\hat{\phi}_1(t), \cdots, \hat{\beta}_\tau^{\mathrm{K}}(t)\hat{\phi}_M(t)\right)^T$ as the estimate of $\boldsymbol{\phi}_\tau(t)$, for $t \in \mathcal{T}$. Write the prediction (3.13) in the vector form $\widehat{Q}_{Y^*|X^*}^{\mathrm{K},M}(t;\tau) = \hat{\alpha}_\tau^{\mathrm{K}}(t) + \hat{\tilde{\boldsymbol{\zeta}}}_M^{*T}\hat{\boldsymbol{\phi}}_t^{\mathrm{K},M}$. The following result facilitates the construction of pointwise confidence intervals for the τth quantile response $Q_{Y^*|X^*}(t;\tau)$ at predictor level X^*.

Theorem 3.4. *Let* $\sup_{t\in\mathcal{T}}|\hat{\alpha}_\tau^{(0)}(t) - \alpha_\tau(t)| = O_p(a_n)$ *and* $\sup_{t\in\mathcal{T}}|\hat{\beta}_\tau^{(0)}(t) - \beta_\tau(t)| = O_p(a_n)$*, with* $0 \le a_n < 1$*. Under Conditions 1-3, and Assumptions (A1)-(A8)(i) in the Appendix S5,*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

given N^* and $\boldsymbol{T}^*$, for a given $\tau \in (0,1)$, all $t \in \mathcal{T}$, $x \in \mathbb{R}$, predicted response trajectories

$$\lim_{n\to\infty} P\left(\frac{\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau)}{\sqrt{\hat{\omega}^{\mathrm{K},M}_{\tau}(t)}} \le x\right) = \Phi(x),$$

where $\hat{\omega}^{\mathrm{K},M}_{\tau}(t) = \left(\hat{\boldsymbol{\phi}}^{\mathrm{K},M}_{\tau}(t)\right)^T \hat{\boldsymbol{\Omega}}_M \hat{\boldsymbol{\phi}}^{\mathrm{K},M}_{\tau}(t)$, which is a estimate of $\omega^M_\tau(t) = \left(\boldsymbol{\phi}^M_\tau(t)\right)^T \boldsymbol{\Omega}_M \boldsymbol{\phi}^M_\tau(t)$, and $\Phi(\cdot)$ denotes Gaussian cumulative distribution function and $M(n) \to \infty$ as $n \to \infty$.

It follows from Theorem 3.4 that, ignoring truncation bias resulting from truncation at M and negligible optimization error by K iterations of Algorithm 2, the $(1-\alpha)100\%$ asymptotic pointwise confidence interval for the τth quantile response at predictor level X^* is given by

$$\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) \pm \Phi(1-\alpha/2)\sqrt{\hat{\omega}^{\mathrm{K},M}_{\tau}(t)}.$$

3.2 Functional linear QR model

In the subsection, we consider the scenario $(\mathcal{L}_X\beta)(t) = \int_{\mathcal{S}} \beta_\tau(s,t)X(s)ds$ for $s \in \mathcal{S}$ and $t \in \mathcal{T}$. Thus, the functional linear QR model is

$$Q_{Y|X}(t;\tau) = \alpha_\tau(t) + \int_{\mathcal{S}} \beta_\tau(s,t)X(s)ds, \tag{3.14}$$

where the bivariate regression function $\beta_\tau(s,t)$ is smooth and square integrable, that is, $\int_{\mathcal{T}}\int_{\mathcal{S}} \beta_\tau(s,t)dsdt < \infty$. From the definition of the operator $\mathcal{L}_X$, it is easy to see that the adjoint of $L^*_X : L_2(\mathcal{S}) \to L_2(\mathcal{T},\mathcal{S})$, given by

$$(\mathcal{L}^*_X z)(s,t) = X(s)z(t), \text{ for all } z \in L_2(\mathcal{S}).$$

For this model (3.14), we have a linear integral operator $\Gamma_{XX} : L_2(\mathcal{S},\mathcal{T}) \to L_2(\mathcal{S}\times\mathcal{T})$ as $(\Gamma_{XX}\beta_\tau)(s,t) = \int_{\mathcal{S}} r_{XX}(s,w)\beta_\tau(w,t)dw$. It is easy to see that $\Gamma_{XX} = \mathrm{E}[\mathcal{L}^*_X\mathcal{L}_X]$. Notice that the observations of the model (3.14) are (S_{il}, U_{il}) and (T_{ij}, V_{ij}), $l = 1, \cdot, L_i$, $j = 1, \cdots, N_i$, $i = 1, \cdots, n$. Herein after, we still use symbols of Subsection 3.1.

3.2.1 Functional Approach

From the estimation strategy in Section 2.2, the functional linear QR model (3.14) can be translated into the following functional linear operator regression model

$$\mathrm{E}[\widetilde{Y}(t)|X(t)] = \alpha_\tau(t) + (\mathcal{L}_X\beta)(t), \tag{3.15}$$

where $(\mathcal{L}_X\beta)(t) = \int_{\mathcal{S}} \beta_\tau(s,t)X(s)ds$

$$\widetilde{Y}(t) = Q^{(0)}_{Y|X}(t;\tau) - \left(\int_{\mathcal{T}} f_t^{(k)}(0)dt\right)^{-1} \int_{\mathcal{T}} \{\mathbb{I}[Y(t) - Q^{(0)}_{Y|X}(t;\tau) \le 0] - \tau\}dt,$$

$$Q^{(0)}_{Y|X}(t;\tau) = \alpha^{(0)}_\tau(t) + \int_{\mathcal{S}} \beta^{(0)}_\tau(s,t)X(s)ds.$$

The model (3.15) can be rewritten as

$$\mathrm{E}[\widetilde{Y}(t)|X(t)] = \mu_{\widetilde{Y}}(t) + \int_{\mathcal{S}} \beta_\tau(s,t)X^c(s)ds,$$

where $\mathrm{E}[\widetilde{Y}(t)] = \mu_{\widetilde{Y}}(t) = \alpha_\tau(t) + \int_{\mathcal{S}} \beta_\tau(s,t)\mu_X(s)ds$. By our result (2.12), one gets

$$\begin{aligned} \beta^{(1)}_\tau(s,t) &= \Gamma^{-1}_{X^cX^c} E\left(\mathcal{L}^*_{X^c}\widetilde{Y}^c\right) \\ &= \sum_{m=1}^\infty \sum_{l=1}^\infty \frac{\mathrm{E}[\zeta_m \varsigma_l]}{\mathrm{E}[\zeta_m^2]}\phi_m(s)\varphi_l(t), \end{aligned} \tag{3.16}$$

where ς_l and φ_l are defined in (3.5), and ζ_m and ϕ_m are given in Section 2.1. $r_{\tilde{Y}\tilde{Y}}$ and $r_{X\tilde{Y}}$ are defined in (3.5) and (3.8), respectively.

Here, we present some main results of the model. Its function approach similar to subsection 3.1.1, and some detains including subsections 3.2.2 and 3.2.3 are given in Appendix S1.2 of SM.

3.2.2 Uniform consistency and prediction

We first present the consistency of the regression function estimates $\hat{\alpha}^{(\mathrm{K})}_\tau$ and $\hat{\beta}^{(\mathrm{K})}_\tau$, and then the prediction of the τth quantile response function $Q_{Y|X}(t;\tau)$.

Theorem 3.5. *Let* $\sup_{t\in\mathcal{T}}|\hat{\alpha}_\tau^{(0)}(t)-\alpha_\tau(t)|=O_p(a_n)$ *and* $\sup_{(s,t)\in\mathcal{S}\times\mathcal{T}}|\hat{\beta}_\tau^{(0)}(s,t)-\beta_\tau(s,t)|=O_p(a_n)$. *Under Conditions 1-3, and Assumptions (A1)-(A6) and (B1)-(B2) in Appendix S5 of SM, we have for* k *iterations of Algorithm 2 in SM,*

$$\begin{aligned}\sup_{(s,t)\in\mathcal{S}\times\mathcal{T}}\left|\hat{\alpha}_\tau^{(k)}(s,t)-\hat{\alpha}_\tau(s,t)\right|&=O_p\left(\delta_{1n}+\delta_{2n}+a_n^{k+1}\right),\\ \sup_{(s,t)\in\mathcal{S}\times\mathcal{T}}\left|\hat{\beta}_\tau^{(k)}(s,t)-\hat{\beta}_\tau(s,t)\right|&=O_p\left(\delta_{1n}+\delta_{2n}+a_n^{k+1}\right),\end{aligned}\tag{3.17}$$

where δ_{1n} *and* δ_{2n} *are defined in Assumptions (B1)-(B2) in Appendix S5, respectively.*

Remark 3.3. *The rate of convergence in 3.5 depends on special properties of processes* X *and* $\tilde{Y}$. *The rates* δ_{1n} *and* δ_{2n} *are similar to the* ς_n *in (B) and* ϑ_n *in (41) of Yao et al. (2005b), respectively. As Yao et al. (2005b) said, due to the sparsity of the data, fast rates of convergence cannot be expected in this ambient, in contrast to the case where entire trajectories are measured or are densely sampled.*

We establish the consistency of the τth quantile prediction $\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau)$, which is defined in S1.2.3 of SM.

Theorem 3.6. *Let* $\sup_{t\in\mathcal{T}}|\hat{\alpha}_\tau^{(0)}(t)-\alpha_\tau(t)|=O_p(a_n)$ *and* $\sup_{(s,t)\in\mathcal{S}\times\mathcal{T}}|\hat{\beta}_\tau^{(0)}(s,t)-\beta_\tau(s,t)|=O_p(a_n)$ *with* $0\le a_n<1$. *Under Conditions 1-3, and Assumptions (A1)-(A7) and (B1)-(B2) in the Appendix S5, given* L^* *and* $\boldsymbol{S}^*$, *for all* $t\in\mathcal{T}$, *the predicted* τ*th quantile response trajectories in the functional linear QR model (3.14) satisfy*

$$\lim_{n\to\infty}\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau)=\widetilde{Q}_{Y^*|X^*}(t;\tau),\quad \text{in proabaility},$$

for the target trajectory $\widetilde{Q}_{Y^*|X^*}(t;\tau)=\alpha_\tau(t)+\sum_{m=1}^\infty\sum_{k=1}^\infty\frac{\tilde{\sigma}_{mk}}{\rho_m}\tilde{\zeta}_m^*\varphi_k(t)$, *with* $M=M(n)\to\infty$ *and* $K=K(n)$ *as* $n\to\infty$, *and the iteration number* K *enough large .*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

3.2.3 Asymptotic pointwise confidence bands for quantile response trajectories

We here construct asymptotic confidence bands for the τth quantile response trajectory $Q_{Y^*|X^*}(t;\tau)$ of a new subject. We can construct the pointwise confidence bands for the τth quantile response $Q_{Y^*|X^*}(t;\tau)$ at predictor level X^* by the following theorem.

Theorem 3.7. *Let* $\sup_{t\in\mathcal{T}} |\hat{\alpha}_\tau^{(0)}(t) - \alpha_\tau(t)| = O_p(a_n)$ *and* $\sup_{(s,t)\in\mathcal{S}\times\mathcal{T}} |\hat{\beta}_\tau^{(0)}(s,t) - \beta_\tau(s,t)| = O_p(a_n)$ *with* $0 \le a_n < 1$*. Under Conditions 1-3, and Assumptions (A1)-(A7), (A8)(ii) and (B1)-(B2) in Appendix S5 of SM, given* N^* *and* $\boldsymbol{T}^*$*, for a given* $\tau \in (0,1)$*, all* $t \in \mathcal{T}$*,* $x \in \mathbb{R}$*,*

$$\lim_{n\to\infty} P\left(\frac{\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau)}{\sqrt{\hat{\omega}_\tau^{\mathrm{K},M,K}(t)}} \le x\right) = \Phi(x),$$

where $\hat{\omega}_\tau^{\mathrm{K},M,K}(t)$ *is a estimate of* $\omega_\tau^{M,K}(t) = \left(\boldsymbol{\varphi}_t^K\right)^T \boldsymbol{P}_{M,K}^T \boldsymbol{\Omega}_M \boldsymbol{P}_{M,K} \boldsymbol{\varphi}_t^K$ *and* $M(n) \to \infty$ *as* $n \to \infty$*.*

Theorem 3.7 shows that asymptotic distribution of $\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau)$ conditional on L^* and $\boldsymbol{S}^*$ can be approximated by $N\left(0, (\hat{\boldsymbol{\varphi}}_t^K)^T \hat{\boldsymbol{P}}_{M,K}^T \hat{\boldsymbol{\Omega}}_M \hat{\boldsymbol{P}}_{M,K} \hat{\boldsymbol{\varphi}}_t^K\right)$. As a consequence, the $(1-\alpha)100\%$ asymptotic pointwise confidence interval for the τth quantile response at predictor level X^* is given by

$$\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau) \pm \Phi(1-\alpha/2)\sqrt{(\hat{\boldsymbol{\varphi}}_t^K)^T (\hat{\boldsymbol{P}}^{\mathrm{K}}_{M,K})^T \hat{\boldsymbol{\Omega}}_M \hat{\boldsymbol{P}}^{\mathrm{K}}_{M,K} \hat{\boldsymbol{\varphi}}_t^K}.$$

For the 3rd specific **functional varying coefficient QR model with history index**, its main contents, including models, estimation methods of functional approach and asymptotic properties, are presented in Appendix S1.3 because of page limitations.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

4. Numerical results

We study the performance of our algorithm in two functional linear operator quantile regression models: Functional varying coefficient QR model and functional linear QR model. The main purpose is to illustrate the robustness of the algorithm in dealing with sparse and irregular longitudinal data, overcoming non-smooth optimization, and in the face of heavy-tailed and outlier data. The following examples respectively come from Şentürk and Müller (2010); Yao et al. (2005b), and some modifications have been made to fit our model.

For each example, we employ 200 Monte Carlo runs. The average mean squared error (MSE) is used to demonstrate superior performance, which is defined as $\mathrm{MSE} = \int_{\mathcal{T}} \left(\hat{\beta}(t) - \beta(t)\right)^2 dt$ for the estimator of $\beta(t)$ on $\mathcal{T}$, and $\mathrm{MSE} = \int_{\mathcal{T}} \int_{\mathcal{S}} \left(\hat{\beta}(s,t) - \beta(s,t)\right)^2 dsdt$. for the estimator of bivariate function $\beta(s,t)$ on $\mathcal{S} \times \mathcal{T}$. In our simulation, take $\mathcal{T} = \mathcal{S}$.

Example 1. Functional varying coefficient QR model The response trajectories are generated from the model

$$Y_i(t) = \beta_0(t) + \beta_1(t) X_i(t) + W_{\tau,i}(t), t \in \mathcal{T}, i = 1, \cdots, n,$$

for a given quantile level τ; $\beta_0(t) = t$, $\beta_1(t) = \sin(\pi t)$; The predictor trajectories X_i are generated from mean function $\mu_X(t) = t + \sin(t)$, covariance function constructed from two eigenfunctions $\phi_1(t) = \frac{1}{\sqrt{2}} \cos(\pi t)$ and $\phi_2(t) = \frac{1}{\sqrt{2}} \sin(\pi t)$ with two corresponding eigenvalues $\rho_1 = 2$ and $\rho_2 = 1$, and its functional principal components $\zeta_{im} \sim N(0, \rho_m)$ for $m = 1, 2$; $W_{\tau,i}(t) = \phi_1(t)(W_i - q_\tau(W_i))$, where W_i with three different settings of random errors: $W_i \sim 0.5N(0,1)$ for normal data, $W_i \sim 0.01\mathrm{Cauchy}(0,1)$ for symmetric heavy-tailed data, and $W_i \sim 0.01\chi^2(2)$ (Chi-square distribution with 2 degrees of freedom) for skewed heavy-tailed data, and $q_\tau(W_i)$ is the τ-th quantile of the distribution of W_i; So that $P\left(W_{\tau,i}(t) \leq 0\right) = \tau$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

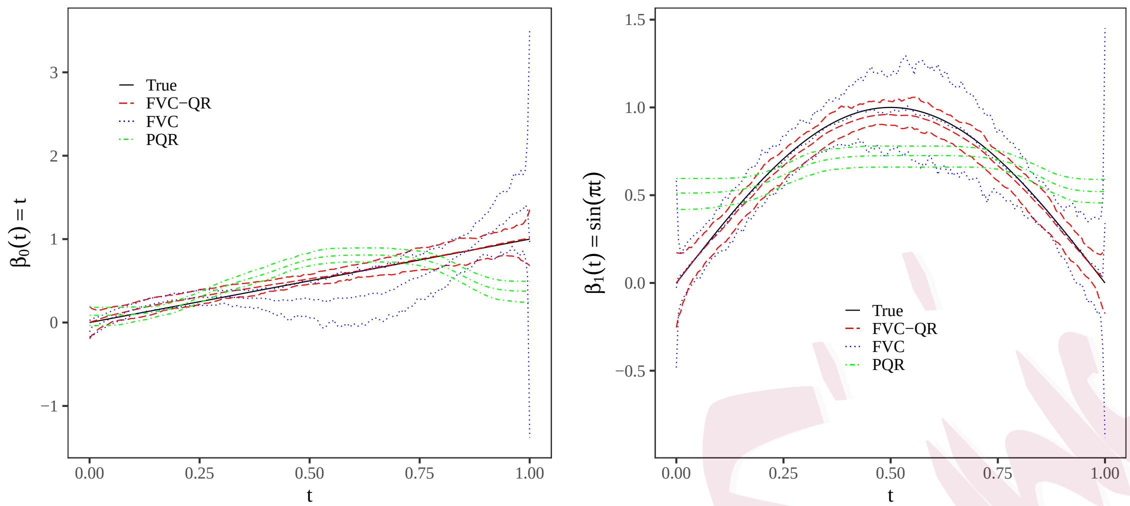


Figure 1: Simulation results of Example 1 for the estimators of cross-sectional mean varying coefficient functions and 95% pointwise bands when $\tau = 0.5$ and $n = 50$ with normal errors.

under the setting of $W_{\tau,i}(t)$. In addition, $W_{\tau,i}(t)$ is independent of $X_i(t)$. The predictor and response trajectories are observable at T_{ij} and are contaminated with measurement errors, that is, $U_{ij} = X_i(T_{ij}) + \varepsilon_{ij}, V_{ij} = Y_i(T_{ij}) + \epsilon_{ij}$, where ε_{ij}, ϵ_{ij} are i.i.d. errors with $\varepsilon_{ij} \sim N(0, 0.1^2)$, $\epsilon_{ij} \sim N(0, 0.1^2)$.

The sample sizes $n = 50$ and 100, and quantile levels $\tau = 0.10, 0.25, 0.5, 0.75$ and 0.90 are considered in the simulation. We randomly sample the number of measurements of each subject from $\{3, 4, 5\}$ with equal probability, and then the locations of the measurements $T_{ij} \sim U[0, 1]$. This is a very sparse design because there are at most five observations available for each subject. We choose some suitable bandwidths for smoothing mean and auto(cross)-covariance surfaces after a large-scale cross-validation experiment based on a generalized cross-validation procedure. We compare our method (FVC-QR) with the method proposed in Şentürk and Müller (2010), denoted as FVC, and the PQR proposed by Andriyana et al.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

Table 1: MSEs (SEs) of estimation of FVC, PQR and FVC-QR for Example 1.

Scenario	size	τ	FVC	PQR	FVC-QR	FVC	PQR	FVC-QR
				β_0			β_1	
Normal	50	0.10	1.1273 (3.96)	0.3180 (0.03)	0.3228 (0.13)	0.7190 (2.12)	0.2643 (0.02)	0.0933 (0.04)
		0.25	0.7067 (1.62)	0.2716 (0.03)	0.1695 (0.07)	0.6524 (2.01)	0.2687 (0.02)	0.0927 (0.04)
		0.50	0.8013 (3.21)	0.2209 (0.03)	0.1029 (0.03)	0.7184 (4.20)	0.2743 (0.02)	0.0879 (0.03)
		0.75	0.6736 (1.37)	0.1939 (0.03)	0.2280 (0.06)	0.4818 (0.80)	0.2815 (0.02)	0.0906 (0.03)
		0.90	0.6036 (0.89)	0.1980 (0.03)	0.4044 (0.14)	0.4319 (0.56)	0.2843 (0.02)	0.0911 (0.04)
	100	0.10	0.5518 (2.44)	0.3371 (0.02)	0.3287 (0.11)	0.2622 (1.32)	0.2342 (0.01)	0.0635 (0.01)
		0.25	0.3660 (0.98)	0.2898 (0.02)	0.1627 (0.05)	0.2788 (1.25)	0.2332 (0.01)	0.0628 (0.02)
		0.50	0.3648 (0.76)	0.2481 (0.02)	0.0770 (0.02)	0.2124 (0.48)	0.2352 (0.01)	0.0641 (0.01)
		0.75	0.3156 (0.37)	0.2247 (0.02)	0.2090 (0.05)	0.1779 (0.38)	0.2382 (0.01)	0.0637 (0.02)
		0.90	0.4097 (1.04)	0.2330 (0.02)	0.3843 (0.13)	0.1907 (0.54)	0.2393 (0.01)	0.0673 (0.02)
Cauchy	50	0.10	0.5279 (1.71)	0.2249 (0.03)	0.3425 (0.87)	0.5603 (2.65)	0.2696 (0.02)	0.0740 (0.09)
		0.25	0.3791 (0.99)	0.2209 (0.02)	0.1524 (0.14)	0.4649 (2.49)	0.2685 (0.02)	0.0738 (0.08)
		0.50	0.6629 (3.86)	0.2178 (0.02)	0.1435 (0.67)	0.5322 (2.92)	0.2691 (0.02)	0.0882 (0.37)
		0.75	0.3322 (0.62)	0.2165 (0.03)	0.1953 (0.17)	0.2977 (0.53)	0.2701 (0.02)	0.0786 (0.16)
		0.90	0.3041 (0.58)	0.2107 (0.02)	0.3094 (0.18)	0.2924 (0.64)	0.2706 (0.02)	0.0579 (0.06)
	100	0.10	0.3803 (0.27)	0.2500 (0.02)	0.2423 (0.12)	0.2050 (0.22)	0.2329 (0.01)	0.1912 (0.07)
		0.25	0.4486 (0.56)	0.2452 (0.02)	0.1478 (0.16)	0.2392 (0.38)	0.2353 (0.01)	0.0743 (0.18)
		0.50	6.4763 (85.45)	0.2418 (0.02)	1.6133 (21.59)	5.6255 (76.37)	0.2337 (0.01)	1.4449 (19.65)
		0.75	0.4871 (0.63)	0.2431 (0.02)	0.1975 (0.21)	0.2737 (0.44)	0.2347 (0.01)	0.0864 (0.21)
		0.90	0.3734 (0.17)	0.2400 (0.02)	0.3133 (0.18)	0.1921 (0.18)	0.2335 (0.01)	0.0505 (0.07)
$\chi^2(2)$	50	0.10	1.6821 (11.06)	0.2247 (0.02)	0.2172 (0.12)	1.1342 (7.00)	0.2688 (0.02)	0.0798 (0.03)
		0.25	0.5056 (0.65)	0.2220 (0.02)	0.1128 (0.04)	0.3612 (0.48)	0.2659 (0.02)	0.0783 (0.03)
		0.50	1.1587 (6.58)	0.2206 (0.02)	0.0994 (0.03)	0.8679 (5.46)	0.2679 (0.02)	0.0765 (0.03)
		0.75	0.5978 (0.94)	0.2201 (0.02)	0.1838 (0.05)	0.5430 (1.61)	0.2665 (0.02)	0.0772 (0.04)
		0.90	0.5653 (0.85)	0.2166 (0.02)	0.3028 (0.12)	0.4946 (1.57)	0.2666 (0.02)	0.0804 (0.04)
	100	0.10	0.3272 (0.74)	0.2477 (0.02)	0.2060 (0.12)	0.1980 (0.65)	0.2322 (0.01)	0.0514 (0.02)
		0.25	0.3455 (1.12)	0.2475 (0.02)	0.1019 (0.04)	0.1820 (0.64)	0.2333 (0.01)	0.0515 (0.02)
		0.50	0.2539 (0.10)	0.2452 (0.02)	0.0683 (0.02)	0.1291 (0.07)	0.2345 (0.01)	0.0498 (0.01)
		0.75	0.2444 (0.05)	0.2427 (0.02)	0.1569 (0.04)	0.1202 (0.04)	0.2342 (0.01)	0.0530 (0.02)
		0.90	0.4102 (1.22)	0.2389 (0.02)	0.2638 (0.13)	0.2140 (0.70)	0.2320 (0.01)	0.0518 (0.02)

(2014), which applied P-splines method to quantile regression in varying coefficient models for longitudinal data. The FVC and PQR are implemented by the *FCReg* function in R package fdapace and the *AHeVT* function in R package QRegVCM, respectively. MSEs and standard errors (SEs) of estimation are listed in Table 1. The estimators of cross-sectional

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

mean varying coefficient functions and 95% pointwise bands are presented in Figure 1 when $\tau = 0.5$ and $n = 100$ with normal errors. From Table 1 and Figure 1, we observe that our FVC-QR performs better in almost all the scenarios because of smaller MSEs, and FVC-QR and PQR are more robust because of smaller SEs than FVC; By constrast, PQR is more robust than our FVC-QR, but the MSEs of PQR is poor, especially for β_1.

Example 2. Functional linear QR model The response trajectories are generated from the model

$$Y_i(t) = \int_0^{10} \beta(s,t) X_i(s) ds + W_{\tau,i}(t), \quad t \in [0, 10], i = 1, \cdots, n,$$

for a given quantile level τ, where the regression function is $\beta(s,t) = \sum_{k=1}^{2} \sum_{m=1}^{2} b_{km} \phi_m(s) \phi_k(t)$, $t, s \in [0, 10]$, with pre-given $b_{11} = 2, b_{12} = 2, b_{21} = 1$ and $b_{22} = 2$; The predictor trajectories X_i, are also generated from mean $\mu_X(t) = t + \sin(t)$, covariance function constructed from two eigenfunctions $\phi_1(t) = \frac{1}{\sqrt{5}} \cos(\pi t/10)$ and $\phi_2(t) = \frac{1}{\sqrt{5}} \sin(\pi t/10)$ with two corresponding functional principal components $\zeta_1 \sim 0.5N(1, 1) + 0.5N(-1, 1)$ and $\zeta_2 \sim 0.5N(1/\sqrt{2}, 1/2) + 0.5N(-1/\sqrt{2}, 1/2)$; It implies that the corresponding eigenvalues $\rho_1 = 0.5$ and $\rho_2 = 0.25$; $W_{\tau,i}(t) = \phi_1(t)(W_i - q_\tau(W_i))$, where W_i with three different settings of random errors: $W_i \sim N(0, 0.1)$ for normal data, $W_i \sim 0.5\text{Cauchy}(0, 1)$ for symmetric heavy-tailed data and $W_i \sim 0.5\chi^2(2)$ for skewed heavy-tailed data, so that $P\left(W_{\tau,i}(t) \le 0\right) = \tau$ under the setting of $W_{\tau,i}(t)$; $W_{\tau,i}(t)$ is independent of $X_i(t)$. The predictor and response trajectories are observable at T_{ij} and are contaminated with measurement errors, that is, $U_{ij} = X_i\left(T_{ij}\right) + \varepsilon_{ij}, V_{ij} = Y_i\left(T_{ij}\right) + \epsilon_{ij}$, where ε_{ij}, ϵ_{ij} are i.i.d. errors with $\varepsilon_{ij}, \epsilon_{ij} \sim N(0, 0.1^2)$.

The settings on sample size, quantile levels, and sparsity measurements are the same as the ones of Example 1, including the choice of bandwidths for smoothing mean and auto (cross)-covariance surfaces. We compare our method (FL-QR) with the approach proposed

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

Table 2: MSEs (SEs) of estimations based on FLR, RFL-QR and our FL-QR for Example 2.

Scenario	size	τ	FLR	RFL-QR	FL-QR (our)
Normal	50	0.10	0.3241 (0.19)	0.3628 (0.06)	0.3273 (0.08)
		0.25	0.3193 (0.08)	0.3679 (0.04)	0.3265 (0.04)
		0.50	0.3101 (0.07)	0.3622 (0.05)	0.3224 (0.04)
		0.75	0.3188 (0.10)	0.3621 (0.04)	0.3238 (0.06)
		0.90	0.3079 (0.06)	0.3614 (0.05)	0.3184 (0.04)
	100	0.10	0.3058 (0.07)	0.3539 (0.03)	0.3152 (0.05)
		0.25	0.2991 (0.04)	0.3543 (0.04)	0.3122 (0.03)
		0.50	0.3078 (0.05)	0.3610 (0.04)	0.3187 (0.04)
		0.75	0.3020 (0.05)	0.3537 (0.04)	0.3150 (0.05)
		0.90	0.3035 (0.08)	0.3552 (0.03)	0.3139 (0.06)
Cauchy	50	0.10	0.6548 (1.75)	0.4247 (0.07)	0.4141 (0.49)
		0.25	0.7077 (2.51)	0.4431 (0.20)	0.4406 (0.61)
		0.50	1.9190 (14.44)	0.6068 (1.95)	0.5879 (2.42)
		0.75	0.5016 (0.50)	0.4180 (0.03)	0.3928 (0.21)
		0.90	0.7006 (1.52)	0.4275 (0.08)	0.4269 (0.35)
	100	0.10	0.5666 (1.13)	0.4171 (0.02)	0.3993 (0.28)
		0.25	0.5516 (0.98)	0.4200 (0.06)	0.4375 (0.60)
		0.50	0.7418 (3.90)	0.4332 (0.17)	0.4405 (0.91)
		0.75	0.9880 (4.54)	0.4281 (0.14)	0.5322 (1.34)
		0.90	0.4696 (0.36)	0.4163 (0.02)	0.3780 (0.09)
$\chi^2(2)$	50	0.10	0.3311 (0.03)	0.3698 (0.03)	0.3267 (0.02)
		0.25	0.3363 (0.04)	0.3757 (0.04)	0.3323 (0.03)
		0.50	0.3339 (0.04)	0.3716 (0.03)	0.3284 (0.03)
		0.75	0.3347 (0.04)	0.3711 (0.03)	0.3294 (0.04)
		0.90	0.3353 (0.04)	0.3733 (0.04)	0.3304 (0.03)
	100	0.10	0.3302 (0.04)	0.3624 (0.03)	0.3246 (0.04)
		0.25	0.3320 (0.04)	0.3657 (0.03)	0.3241 (0.03)
		0.50	0.3258 (0.03)	0.3631 (0.03)	0.3195 (0.02)
		0.75	0.3330 (0.04)	0.3674 (0.04)	0.3231 (0.03)
		0.90	0.3291 (0.04)	0.3639 (0.03)	0.3222 (0.03)

in Yao et al. (2005b), called FLR, and RFL-QR in Beyaztas and Shang (2023), which used robust functional principal components (Bali et al. (2011)) for function-on-function linear quantile regression. The FLR and RFL-QR are implemented by the *FLM1* in R package

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

fdapace and *rob.ff.reg* function in R package robflreg, respectively. MSEs (SEs) of estimation based on FLR, RFL-QR and FL-QR are listed in Table 5, and boxplots of MSEs are presented in Figures S1 and S2 in S6 of SM. We see that our FL-QR is more accurate and robust for the errors of Cauchy distribution and $\chi^2(2)$ distribution, and it is not bad for the errors of normal distribution.

Based on Examples 1 and 2, our algorithms work well for typically sparse and irregular designs in longitudinal data, which provide more accurate and robust statistical results.

5. Applications to Longitudinal Primary biliary cirrhosis data

The Mayo Clinic established a database of primary biliary cirrhosis (PBC), which was collected between January 1974 and May 1984. A complete follow up to July 1986, was attempted on all patients. By this data, 125 of the 312 had died, only 11 deaths were not attributable to PBC; only 19 had undergone liver transplantation and 8 were lost to follow up. Appendix D of Fleming and Harrington (1991) contains this survival data. Fleming and Harrington (1991) puts down in writing that PBC is a rare but fatal chronic liver disease of unknown cause because the prevalence of the disease has been estimated to be 50 cases-per-million population. Physiological and demographic characteristics of the patients, such as platelet count, albumin, prothrombin time, age, and sex, were measured at six months, one year, and annually thereafter post-diagnosis. In addition, measurement times T_{ij} per individual are different, and measurement values are sparse and irregular with unequal numbers of measurements per patient because some patients missed some of their scheduled visits. This database is a valuable resource to liver specialists, which also has been studied by Yao et al. (2005b). We will compare our method with the one of Yao et al. (2005b) based on the

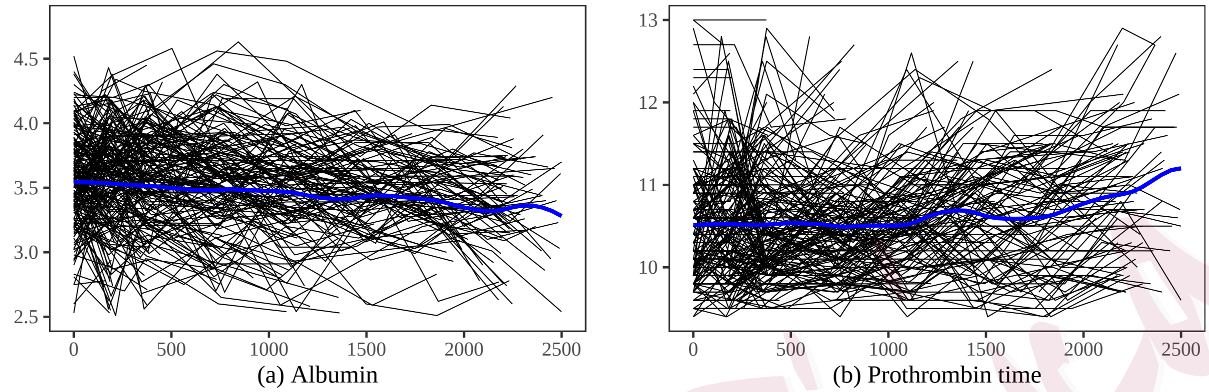
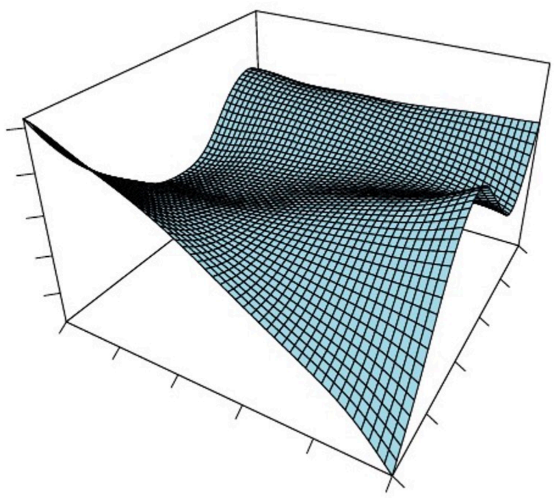


Figure 2: Patients' trajectories of albumin and prothrombin time. Thick solid line is the corresponding smooth estimate of mean function.

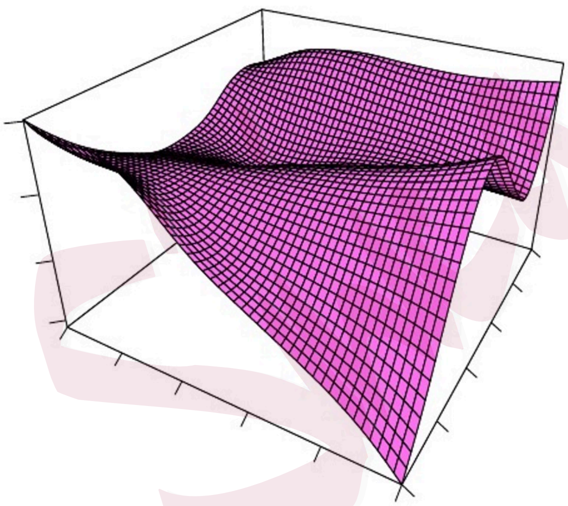
PBC dataset.

As Yao et al. (2005b), the longitudinal measurements albumin (mg/dl) and prothrombin times (seconds) are regarded as predictor and response trajectories, respectively. In the experiment, our data includes 222 female patients (while data in Yao et al. (2005b) includes 137 female patients), and the measurements of albumin level and prothrombin time before 2500 days. The number of measurements of the albumin and prothrombin time ranged from 1 to 10, and their medians are 5 for each patient. Patients' trajectories of albumin and prothrombin time are presented in Figure 2. From the smooth estimators of two mean functions, we observe that albumin level is generally relatively stable with a slight decrease over late measurement time; However, the prothrombin time is relatively stable in the early stage and monotonically increases in the later stage. It indicates that albumin level and prothrombin time have an opposite tendency, especially in the later measurement time. We also explore the dynamic relationship of the sparse longitudinal albumin and prothrombin

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415



(a) FLR



(b) FL-QR

Figure 3: Estimated regression function for PBC data where the predictor (albumin) time is s (in days), and the response (prothrombin) time is t (in days). (a) FLR method; (b) FL-QR method with $\tau = 0.5$.

times by functional linear QR model (FL-QR): $\mathcal{L}_X: \ L_2(\mathcal{S} \times \mathcal{T}) \rightarrow L_2(\mathcal{T})$ by

$$(\mathcal{L}_X \beta)(t) = \int_{\mathcal{S}} \beta_\tau(s,t) X(s) ds, \quad \mathcal{S}, \mathcal{T} = [0, 2500],$$

compared with function linear regression model (FLR) (Yao et al., 2005b). The estimates of the regression function β based FLR, and FL-QR with $\tau = 0.50$ (that is, Function Linear Median Regression, FL-MR) are displayed in Figure 3. The cases of $\tau = 0.1$ and $\tau = 0.9$ are similar to that of that FL-MR model. The shapes of these two figures are basically the same, but the interval changes are different, $-4.62e-3 \leq \hat{\beta} \leq 5.42e-3$ for FLR and $-4.24e-3 \leq \hat{\beta} \leq 2.03e-3$ for FL-MR. It shows the fluctuation of estimated regression

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

function $\hat{\beta}$ of our FL-MR is smaller than that of FLR, because of robustness of FL-MR. It sees that the information inherent in these two figures about $\hat{\beta}$ is similar, but the obtained information of Y based on estimators of FLR and FL-QR are completely different. Even for the FL-QR method, they also are different, such as the quantiles of Y of the high (e.g. $\tau = 0.9$) and/or low (e.g. $\tau = 0.1$) levels; See Figure S4. Their shapes of Figure 3 imply that, for the prediction of early prothrombin times, the negative effect of early albumin levels continues to weaken as time t increases in about $0 \leq t \leq 1000$; when $t > 1000$, it becomes a positive effect and continues to increase until about $t = 1500$; afterwards, its effect continues to decline until about $t = 2000$ and then rebounds, but late albumin levels in about $1500 \leq t \leq 2500$ substantially contribute positively; whereas the prediction of late prothrombin times generally contributes positively, but with highly positive weighting of early and later levels and negligible positive weighting of intermediate levels. These findings are different from the ones in Yao et al. (2005b), which only used 137 female patients.

For evaluating performance of our FL-QR and FLR, we also give the curve of estimated pointwise functional coefficients of determination $R_Q^2(t)$ based on FL-QR with the definition

$$R_Q^2(t) = \frac{\text{Var}[Q_{Y|X}(Y(t;\tau) \mid X)]}{\text{Var}[Y(t)]},$$

and compare with that of determination $R_M^2(t)$ based on FLR with the similar definition

$$R_M^2(t) = \frac{\text{Var}(E[Y(t) \mid X])}{\text{Var}(Y(t))}.$$

They are displayed in Figure S3 in S7 of SM, indicating that the dynamics of albumin in FL-QR are more capable of explaining the total variation of prothrombin time trajectories over a more time range (from 0 to 1975 days), than the one in FLQ. In addition, it indicates generally stronger linear association at intermediate days (1000 to 2000 days) compared to

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0415

the earlier days (0 to 500 days) and later days (2250 to 2500 days).

Last, we reconstruct mean trajectories of prothrombin times by using FLR and quantile trajectories of prothrombin times by applying FL-QR with the levels of quantile $\tau = 0.1, 0.5$ and 0.9, which is presented in Figure S4 in S7 of SM. We see that these trajectories have the same growth mode; mean and quantile with $\tau = 0.5$ trajectories of prothrombin times are almost identical, which implies that the conditional distributions of prothrombin time given albumin at each day don't skew; our FL-QR can capture lower (e.g. $\tau = 0.1$) and upper (e.g. $\tau = 0.9$) conditional quantiles of the trajectories of prothrombin time, which cannot be characterized by analyzing the conditional mean of FLR model alone.

6. Concluding remarks

The proposed method mainly faces two practical problems: one is that the longitudinal data is a sparse and irregular measurement, and the number of repeated measurements of each subject is relatively small or even two or three; the other is that many important features of the jointly distribution of the response and predictor trajectories cannot be captured with conditional mean models alone, especially mean models cannot depict lower and higher conditional quantile of the response trajectories, or the longitudinal data are typically skewed or data contain some outliers. This type of data is very common in practice, and we give a unified methodological framework: functional linear operator quantile regression for sparse longitudinal data.For functional linear operator quantile regression, we need to obtain the estimation of the model in an infinite-dimensional Hilbert space and deal with non-smooth quantile loss, which brings some new challenges in both computation and theoretical development. We develop an iterative surrogate least squares estimation via functional principal components analysis through conditional expectation. The technique is flexible and allows the prediction of an unobserved quantile response trajectory from sparse measurements of a predictor trajectory. However, our proposed iterative SLES method exists a major limitation for the FLOQR model. That is, our transformation from the FLOQR model into functional linear operator model cannot directly deal the FLOQR model with heterogeneous error. In addition, we also assume that the random error of the FLOQR model is independent of covariates. In future, we will further explore the FLOQR with heterogeneous errors depending on covariates of interest.

Supplementary Material

The online Supplementary Material includes all proofs, technical details and additional experimental results.

Acknowledgments

Dr. Zhou was supported by the National Natural Science Foundation of China (12171242, 12371267). Dr. Lai was supported by the National Natural Science Foundation of China under Grant 12271014. Dr. Kong was partially supported by grants from the Canada CIFAR AI Chairs program, the Alberta Machine Intelligence Institute (AMII), the Natural Sciences and Engineering Council of Canada (NSERC), and the Canada Research Chair program from NSERC.

References

Andriyana, Y., I. Gijbels, and A. Verhasselt (2014). P-splines quantile regression estimation in varying coefficient models. *Test 23*, 153–194.

Bali, J. L., G. Boente, D. E. Tyler, and J.-L. Wang (2011). Robust functional principal components: A projection-pursuit approach. *Ann. Statist. 39*(6), 2852–2882.

Battagliola, M., H. Srensen, A. Tolver, and A. Staicu (2025). Quantile regression for longitudinal functional data with application to feed intake of lactating sows. *Journal of Agricultural, Biological and Environmental Statistics 30*, 211–230.

Beyaztas, U. and H. L. Shang (2023). Robust functional linear regression models. *The R Journal 15*(1), 212–233.

Beyaztas, U., H. L. Shang, and S. Saricam (2025). Penalized function-on-function linear quantile regression. *Computational Statistics 40*, 301–329.

Boumahdi, M., I. Ouassou, and M. Rachdi (2023). Estimation in nonparametric functional-on-functional models with surrogate responses. *J. Mult. Anal. 198*(105231), 1–20.

Chen, C. and Y. Wei (2005). Computational issues for quantile regression. *Indian J. Stat. 67*, 399–417.

Chen, X., H. Li, H. Liang, and H. Lin (2019). Functional response regression analysis. *J. Mult. Anal. 169*, 218–233.

Chen, X., W. D. Liu, X. J. Mao, and Z. Y. Yang (2020). Distributed high-dimensional regression under a quantile loss function. *Journal of Machine learning Research 21*, 1–43.

Şentürk, D. and H.-G. Müller (2010). Functional varying coefficient models for longitudinal data. *J. Am. Statist. Assoc. 105*, 1256–1264.

Cui, X., H. Lin, and H. Lian (2020). Partially functional linear regression in reproducing kernel hilbert spaces. *Computational Statistics and Data Analysis 150*, 1–13.

Draper, N. R. and H. Smith (1998). *Applied Regression Analysis*. 3rd ed. Wiley, New York.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modelling and Its Applications*. London: Chapman & Hall.

Ferraty, F. and P. Vieu (2006). *Nonparametric Functional Data Analysis*. Springer: New York.

Fleming, T. and D. Harrington (1991). *Counting Processes and Survival Analysis*. Wiley, New York.

Gu, L., L. Wang, W. K. Hädle, and L. Yang (2014). A simultaneous confidence corridor for varying coefficient regression with sparse functional data. *Test 23*, 806–843.

He, G., H. G. Müller, and L. Wang (2000). Extending correlation and regression from multivariate to functional data. In M. L. Puri (Ed.), *Asymptotics in Statistics and Probability*, pp. pp. 301–315. VSP International Science Publishers.

Horváth, L. and P. Kokoszka (2012). *Inference for Functional Data with Applications*. New York: Springer.

Ji, H. and H. G. Müller (2017). Optimal designs for longitudinal and functional data. *J. R. Statist. Soc. B 79*, 859–876.

Jiang, C. R. and J. L. Wang (2011). Functional single index models for longitudinal data. *Ann. Statist. 39*, 362–388.

Kato, K. (2012). Estimation in functional linear quantile regression. *Ann. Statist. 40*, 3108–3136.

Koenker, R. (2005). *Quantile Regression*. Cambridge: Cambridge Univ. Press.

Li, J., C. Huang, and H. Zhu (2017). A functional varying-coefficient single-index model for functional response data. *J. Am. Statist. Assoc. 112*(519), 1169–1181.

Li, R., W. Lu, Z. Zhu, and H. Lian (2021). Optimal prediction of quantile functional linear regression in reproducing kernel hilbert spaces. *Journal of Statistical Planning and Inference 211*, 162–170.

Li, Y. and T. Hsing (2010). Uniform convergence rates for nonparametric regression and principal component analysis in function/longitudinal data. *Ann. Statist. 38*, 3321–3351.

Liu, J., R. Li, and H. Lian (2024). Distributed estimation of functional linear regression with functional responses. *Metrika 87*, 21–30.

Liu, Y., M. Li, and J. Morris (2020a). Function-on-scalar quantile regression with application to mass spectrometry proteomics data. *Annal of Applied Statistics 14*, 521–541.

Liu, Y., M. Li, and J. Morris (2020b). On function-on-scalar quantile regression. *arXiv 2002.03355v1*.

Luo, R. and X. Qi (2024). General nonlinear function-on-function regression via functional universal approximation. *Journal of Computaional and Graphical Statistics 33*, 578–587.

Luo, X., L. Zhu, and H. Zhu (2016). Single-index varying coefficient model for functional responses. *Biometrics 72*(4), 1275–1284.

Lv, S., Z. Fan, H. Lian, T. Suzuki, and K. Fukumizu (2020). A reproducing kernel hilbert space approach to high dimensional partially varying coefficient model. *Computational Statistics and Data Analysis 152*, 1–28.

Müller, H. G. (2009). Functional modeling of longitudinal data. In *Longitudinal Data Analysis, Handbooks of Modern Statistical Methods*, pp. 223–252. NewYork: Wiley.

Ramsay, J. O. and B. W. Silverman (2005). *Functional Data Analysis (2nd ed)*. New York: Springer.

Rice, J. and C. Wu (2001). Nonparametric mixed effects models for unequally sampled noisy curves. *Biometrics 57*, 253–259.

Staniswalis, J.-G. and J. J. Lee (1998). Nonparametric regression analysis of longitudinal data. *J. Amer. Statist. Assoc. 93*, 1403–1418.

Sun, X., P. Du, X. Wang, and P. Ma (2018). Optimal penalized function-on-function regression under a reproducing kernel hilbert space framework. *J. Am. Statist. Assoc. 113*, 1601–1611.

Wu, Y. C., J. Q. Fan, and H. G. Müller (2010). Varying-coefficient functional linear regression. *Bernoulli 16*, 730–758.

Yang, G., H. Lin, and H. Lian (2021). Minimax rate in prediction for functional principal component regression. *Communications in Statistics - Theory and Methods 50*, 1240–1249.

Yang, J., X. Meng, and M. W. Mahoney (2014). Quantile regression for large-scale applications. *SIAM J. Sci. Comput. 36*, S78–S110.

Yao, F. (2007). Asymptotic distributions of nonparametric regression estimators for longitudinal or functional data. *J. Mult. Anal. 98*, 40–56.

Yao, F., H. G. Müller, and J. L. Wang (2005a). Functional data analysis for sparse longitudinal data. *J. Am. Statist. Assoc. 100*, 577–590.

Yao, F., H. G. Müller, and J. L. Wang (2005b). Functional linear regression analysis for longitudinal data. *Ann. Statist. 33*, 2873–2903.

Zhang, J. T. (2011). Statistical inference for linear models with functional responses. *Statist. Sinica 21*, 1431–1451.

Zhang, J. T. and J. Chen (2007). Statistical inferences for functional data. *Ann. Statist. 35*, 1052–1079.

Zhang, Y. and H. Lian (2021). Sketched quantile additive functional regression. *Neurocomputing 461*, 17–26.

Zhang, Z., X. Wang, L. Kong, and H. Zhu (2022). High-dimensional spatial quantile function-on-scalar regression. *J. Am. Statist. Assoc. 117*, 1563–1578.

Zhou, X., D. Kong, A. B. Kashlak, L. Kong, R. Karunamuni, and H. Zhu (2023). Functional response quantile regression model. *Statistica Sinica 33*, 2643–2667.

Zhou, X., D. Kong, M. S. Pietrosanu, L. Kong, and R. J. Karunamuni (2024). Empirical likelihood m-estimation for the varying-coefficient model with functional response. *Scandinavian Journal of Statistics 51*, 1357–1387.

Zhou, Z. and W. B. Wu (2010). Simultaneous inference of linear models with time varying coefficients. *J. R. Statist. Soc. B 72*, 513–531.

Xingcai Zhou, School of Statistics and Data Science, Nanjing Audit University, Nanjing, China. xczhou@nuaa.edu.cn

Tingyu Lai, School of Mathematics and Statistics, Guangxi Normal University, Guilin, China. statisticslai@163.com

Linglong Kong, Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada. E-mail: lkong@ualberta.ca