Annals of the Institute of Statistical Mathematics (2021) 73:441–455
https://doi.org/10.1007/s10463-020-00756-1

Copula and composite quantile regression-based estimating equations for longitudinal data

Kangning Wang[1] · **Wen Shan**[1]

Received: 10 September 2019 / Revised: 8 February 2020 / Published online: 25 May 2020

Abstract
Composite quantile regression (CQR) is a powerful complement to the usual mean regression and becomes increasingly popular due to its robustness and efficiency. In longitudinal studies, it is necessary to consider the intra-subject correlation among repeated measures to improve the estimation efficiency. This paper proposes a new approach that uses copula to account for intra-subject dependence in CQR. By using the copula-based covariance matrix, efficient CQR estimating equations are constructed for the longitudinal data partial linear varying coefficient models. Our proposed new methods are flexible, and can provide efficient estimation. The properties of the proposed methods are established theoretically, and assessed numerically through simulation studies and real data analysis.

Keywords Composite quantile regression · Longitudinal data · Copula · Efficiency · Robustness

1 Introduction

Longitudinal data arises frequently from many subject-matter studies, such as medical and public health studies. For longitudinal data, subjects are often measured repeatedly over a given time period. Thus, observations from the same subject are correlated and those from different subjects are often independent. As we all known, the within-subject correlation plays an important role in improving estimation

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10463-020-00756-1) contains supplementary material, which is available to authorized users.

✉ Kangning Wang
wkn1986@126.com

Wen Shan
shanwen2019@126.com

[1] School of Statistics, Shandong Technology and Business University, No. 191, Binhai Middle Road, Laishan District, Yantai 264005, China

efficiency, it is an important problem to study how to use this within-subject correlation. Liang and Zeger (1986) proposed generalized estimating equations (GEE), which can incorporate the correlation by using a working correlation matrix. Liang and Zeger (1986) showed that the GEE estimators are still consistent even if the correlation matrix is misspecified. Recently, there is a huge literature devoted to studying the GEE method, e.g., Tian et al. (2015), Lai et al. (2012), Lv et al. (2015), Wang et al. (2012), Li et al. (2013), Lian et al. (2014), Zhao and Li (2013) and so on.

However, the GEE method is in principle very similar to the weighted least squares method, which does not possess robustness. Furthermore, in longitudinal data, one outlier in the subject level may generate a set of outliers due to repeated measurements. Hence, robustness is very important in longitudinal studies. To make GEE more robust, Fan et al. (2012), He et al. (2005), Qin and Zhu (2007) and Qin et al. (2009, 2012) all used the traditional robust M-estimations (e.g., Huber's estimation) on the Pearson residuals to dampen the effect of outliers in the response to obtain a robust GEE.

Although the Huber's score function is a robust modeling tool, it has limitation in estimation efficiency. To obtain a highly efficient and robust estimator, Zou and Yuan (2008) proposed composite quantile regression (CQR). It becomes a popular approach and has been extended to many fields. Kai et al. (2010, 2011) extended it to the nonparametric and semiparametric models, respectively. To further improve the estimation efficiency, Jiang et al. (2012) and Sun et al. (2013) proposed weighted CQR method for the independent data. Zhao et al. (2017) investigates CQR estimation on the basis of quadratic inference functions. Fan et al. (2018) extended it to the single index models with high covariates.

However, in CQR setting with longitudinal data, modeling the correlation structure to improve estimation efficiency is challenging. The main difficulty is how to construct the correlation structure of the score of the CQR loss function, which is may be different with the correlation structure of the within-subject random errors. To solve this issue, we propose a copula and CQR-based method, where copula functions are employed to accommodate the correlation structure of longitudinal data. Specifically, unbiased CQR estimating function are proposed, which can incorporate the correlation of CQR with longitudinal data by using the copula-based covariance matrix. Because the estimating functions are discontinuous, we further smooth them by using the induced smoothing method (Brown and Wang 2005). The asymptotic properties of the proposed new methods and resulting estimators are derived under some regularity conditions.

Copulas have been widely used in longitudinal data analysis, e.g., Sun et al. (2008), Song (2000), Bai et al. (2014), Wang and Sun (2017) and Wang et al. (2018) and Noh et al. (2015) proposed a method for semiparametric quantile regression by modeling the joint distribution of the response and covariates through copulas. Shi and Frees (2010) considered a copula method for quantile regression by modeling the conditional marginals of the responses with an asymmetric Laplace (AL) distribution. Also, Fu and Wang (2016) utilized the Gaussian copula to explore the correlations in longitudinal data linear quantile regression. However, the main differences between our method and Fu and Wang (2016) are as follows. First, Fu and

Wang (2016) focused on simple quantile regression and Gaussian copula. While our method is built upon composite quantile regression, which not only possesses all merits of quantile regression but also has estimation efficiency gain over a single quantile regression. What is more, the correlation structure in longitudinal data composite quantile regression is different from that of simple quantile regression, how to incorporate the correlations in composite quantile regression with longitudinal data is an unknown issue. We propose constructing the correlation matrix via copula functions, and find that our method is quite insensitive to the choice of copula. Second, the method in Fu and Wang (2016) is for the simple parametric linear regression model, while we consider a more general semiparametric partial linear varying coefficient models. Obviously, the estimation method and asymptotic properties for the semiparametric models have essentially differences with that for the linear regression model.

The rest of this paper is organized as follows. In Sect. 2, we introduce the new copula and CQR-based estimating equations. Numerical studies and real data analysis are reported in Sect. 3. All the technical proofs and regularity conditions are provided in the supplementary file.

2 Copula and composite quantile regression-based estimating equations

Let $(Y_{ij}, \boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}), j = 1, \ldots, m_i$ be the jth observation for the ith subject, where Y_{ij} is response variable, $\boldsymbol{X}_{ij} = (X_{ij,1}, \ldots, X_{ij,p})^{\mathrm{T}} \in R^p$ and $\boldsymbol{Z}_{ij} = (Z_{ij,1}, \ldots, Z_{ij,q})^{\mathrm{T}} \in R^q$ are covariates. Without loss of generality, we assume $T_{ij} \in [0, 1]$ and consider a balanced design with $m_i = m$ being finite. Let $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{im})^{\mathrm{T}}$, $\boldsymbol{X}_i = \left(\boldsymbol{X}_{i1}, \ldots, \boldsymbol{X}_{im}\right)^{\mathrm{T}}$, $\boldsymbol{Z}_i = \left(\boldsymbol{Z}_{i1}, \ldots, \boldsymbol{Z}_{im}\right)^{\mathrm{T}}$ and $\boldsymbol{T}_i = (T_{i1}, \ldots, T_{im})^{\mathrm{T}}$, suppose that $\{\boldsymbol{Y}_i, \boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{T}_i, i = 1, \ldots, n\}$ is a random sample of $\{\boldsymbol{Y} = (Y_1, \ldots, Y_m)^{\mathrm{T}}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{T}\}$. The partial linear varying coefficient models for longitudinal data is given by

$$Y_{ij} = \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}(T_{ij}) + \epsilon_{ij}, \tag{1}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^{\mathrm{T}}$ is regression parameter vector with true value $\boldsymbol{\beta}_0$, and $\boldsymbol{\alpha}(\cdot) = (\alpha_1(\cdot), \ldots, \alpha_q(\cdot))^{\mathrm{T}}$ is function coefficient vector with true value $\boldsymbol{\alpha}_0(\cdot)$.

2.1 Estimating equations and main algorithm

Let $\boldsymbol{B}(t) = (B_{1,D}(t), \ldots, B_{d_n,D}(t))^{\mathrm{T}}$ be a set of B-spline basis of order $D + 1$ with K_n quasi-uniform internal knots, where $d_n = K_n + D + 1$. Then, $\alpha_k(T_{ij})$ can be approximated as $\alpha_k(T_{ij}) \approx \boldsymbol{B}^{\mathrm{T}}(T_{ij})\boldsymbol{\gamma}_k$, where $\boldsymbol{\gamma}_k = (\gamma_{1k}, \ldots, \gamma_{d_n k})^{\mathrm{T}}$. For more details about the construction of B-spline basis functions, one can refer to Schumaker (1981). Thus, model (1) can be approximated as

$$Y_{ij} \approx \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{B}_{ij}^{\mathrm{T}}\boldsymbol{\gamma} + \epsilon_{ij} = \boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\zeta} + \epsilon_{ij}, \tag{2}$$

where $\boldsymbol{B}_{ij} = (Z_{ij,1}\boldsymbol{B}^{\mathrm{T}}(T_{ij}), \ldots, Z_{ij,q}\boldsymbol{B}^{\mathrm{T}}(T_{ij}))^{\mathrm{T}}$, $\boldsymbol{D}_{ij} = (\boldsymbol{X}_{ij}^{\mathrm{T}}, \boldsymbol{B}_{ij}^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^{\mathrm{T}}, \ldots, \boldsymbol{\gamma}_q^{\mathrm{T}})^{\mathrm{T}}$ and $\boldsymbol{\zeta} = (\boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\gamma}^{\mathrm{T}})^{\mathrm{T}}$.

Assume that $0 < \tau_1 < \cdots < \tau_K < 1$ and $b(\tau_k)$ is the $100\tau_k\%$ conditional quantile of the random error ϵ_{ij}. For brevity, we use the equally spaced quantiles, i.e., $\tau_k = \frac{k}{1+K}$ for $k = 1, \ldots, K$, and assume the density function of ϵ_{ij} is nonvanishing everywhere, which implies that $b(\tau_k)$ is unique. It is easy to show that

$$E\Big\{\tau_k - I(Y_{ij} - \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0 - \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{ij}) \le b(\tau_k))|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}\Big\} = 0, \tag{3}$$

where $I(\cdot)$ is the indicator function. Then, motivated by (3) and by using the spline approximation (2), consistent estimators can be obtained by solving

$$\sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{D}_i^{\mathrm{T}}\boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}) = \boldsymbol{0}, \tag{4}$$

where $\boldsymbol{D}_i = (\boldsymbol{X}_i, \boldsymbol{B}_i)$, $\boldsymbol{B}_i = \big(\boldsymbol{B}_{i1}, \ldots, \boldsymbol{B}_{im}\big)^{\mathrm{T}}$ and $\boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}) = (S_{i1(k)}(\boldsymbol{\zeta}), \ldots, S_{im(k)}(\boldsymbol{\zeta}))^{\mathrm{T}}$ with $S_{ij(k)}(\boldsymbol{\zeta}) = \tau_k - I(Y_{ij} - \boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\zeta} \le b(\tau_k))$. However, Eq. (4) can not be directly used, because $b(\tau_k), k = 1, \ldots, K$ are unknown. Therefore, we find a proxy for it by replacing $b(\tau_k), k = 1, \ldots, K$ with their estimators. A simple way is ignoring the possible correlations by minimizing the following CQR objective function

$$(\hat{b}_1, \ldots, \hat{b}_K, \hat{\boldsymbol{\zeta}}) = \underset{b_1,\ldots,b_K,\boldsymbol{\zeta}}{\arg\min}\left\{\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m}\rho_{\tau_k}(Y_{ij} - b_k - \boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\zeta})\right\}, \tag{5}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$. Similar with the proof of Theorem 1 in Zou and Yuan (2008), we can prove that $\hat{b}_k - b(\tau_k) = O_p(\frac{1}{\sqrt{n}})$. Hence, it is easy to show that

$$\begin{aligned}
&E\Big\{I(Y_{ij} - \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0 - \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{ij})\\
&\quad\le b(\tau_k)) - I(Y_{ij} - \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0 - \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{ij}) \le \hat{b}_k)|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}\Big\}\\
&\quad= E\big\{I(\epsilon_{ij} \le b(\tau_k)) - I(\epsilon_{ij} \le \hat{b}_k)|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}\big\}\\
&\quad= P\big(\epsilon_{ij} \le b(\tau_k)|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}\big) - P\big(\epsilon_{ij} \le \hat{b}_k|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}\big) = o_p(1).
\end{aligned}$$

So solving Eq. (4) is asymptotically equivalent to solving

$$\sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{D}_i^{\mathrm{T}}\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}) = \boldsymbol{0}, \tag{6}$$

where $\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}) = (\hat{S}_{i1(k)}(\boldsymbol{\zeta}), \ldots, \hat{S}_{im(k)}(\boldsymbol{\zeta}))^{\mathrm{T}}$ with $\hat{S}_{ij(k)}(\boldsymbol{\zeta}) = \tau_k - I(Y_{ij} - \boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\zeta} \le \hat{b}_k)$.

Obviously, Eq. (6) do not incorporate the correlation structure, the estimation efficiency may not be satisfactory. Then, in order to take account of the correlation within subject, motivated by the idea of GEE (Liang and Zeger 1986) and Jung (1996), we propose the following CQR estimating functions as

$$\boldsymbol{U}(\boldsymbol{\zeta}) = \sum_{k=1}^{K}\sum_{i=1}^{n} \boldsymbol{D}_i^{\mathrm{T}} \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1} \hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}), \tag{7}$$

where $\boldsymbol{V}_{i(k)} = \mathrm{Cov}\{\boldsymbol{S}^0_{i(k)}|\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{T}_i\}$ with $\boldsymbol{S}^0_{i(k)} = (S^0_{i1(k)}, \ldots, S^0_{im(k)})^{\mathrm{T}} = (\tau_k - I(Y_{i1} - \boldsymbol{X}_{i1}^{\mathrm{T}}\boldsymbol{\beta}_0 - \boldsymbol{Z}_{i1}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{i1}) - b(\tau_k) \le 0), \ldots, \tau_k - I(Y_{im} - \boldsymbol{X}_{im}^{\mathrm{T}}\boldsymbol{\beta}_0 - \boldsymbol{Z}_{im}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{im}) - b(\tau_k) \le 0))^{\mathrm{T}}$, $\boldsymbol{\Lambda}_{i(k)} = \mathrm{diag}(s_{i1(k)}, \ldots, s_{im(k)})$ with $s_{ij(k)} = f_j(\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0 + \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{ij}) + b(\tau_k)|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij})$, which measures the dispersion of $Y_{ij} - \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0 - \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{ij}) - b(\tau_k)$, and $f_j(\cdot|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij})$ is the conditional density function of Y_{ij}.

Next, we will model the conditional covariance matrix $\boldsymbol{V}_{i(k)}$ via copulas. Suppose that $F(y_1, \ldots, y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t})$ is the joint distribution function of $(Y_1, \ldots, Y_m)^{\mathrm{T}}$ given $\boldsymbol{X} = \boldsymbol{x}$, $\boldsymbol{Z} = \boldsymbol{z}$ and $\boldsymbol{T} = \boldsymbol{t}$, with continuous conditional marginal distributions $F_1(\cdot|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}), \ldots, F_m(\cdot|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t})$. Note that $F_i(Y_i|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t})$ is uniform distribution on [0, 1], then by Sklar's theorem, there exists an unique copula C on $[0, 1]^m$ such that,

$$\begin{aligned} &F(y_1, \ldots, y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}) \\ &\quad = P(Y_1 \le y_1, \ldots, Y_m \le y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}) \\ &\quad = P\big\{F_1(Y_1|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}) \le F_1(y_1|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}), \ldots, F_m(Y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}) \\ &\quad \le F_m(y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t})|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}\big\} \\ &\quad = C(F_1(y_1|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}), \ldots, F_m(y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t})|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}). \end{aligned} \tag{8}$$

For model parsimony, we consider a parametric copula function C, i.e., we consider the following simplified copula model

$$F(y_1, \ldots, y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}) = C(F_1(y_1|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}), \ldots, F_m(y_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}); \boldsymbol{\theta}_0). \tag{9}$$

Obviously, for $(u_1, \ldots, u_m) \in [0, 1]^m$, the above copula model can be rewritten as $C(u_1, \ldots, u_m; \boldsymbol{\theta}_0) = F(F_1^{-1}(u_1|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}), \ldots, F_m^{-1}(u_m|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t})|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t})$, e.g., if $F(\cdot)$ is $N(\boldsymbol{0}, \boldsymbol{\Sigma})$, $C(u_1, \ldots, u_m; \boldsymbol{\Sigma}) = F(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_m); \boldsymbol{\Sigma})$ is Gaussian copula. The simplified copula assumption in model (9) has been commonly used in the copula literature for modeling multivariate distributions, e.g., Haff et al. (2010), Smith et al. (2010) and Wang et al. (2018). Haff et al. (2010) and Wang et al. (2018) also showed that the simplified copula serves as a good approximation even when the simplifying assumption is far from being satisfied.

Note that, for model (1) and any $\tau \in (0, 1)$,

$$F_j^{-1}(\tau|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}) = \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0 + \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{ij}) + F_{\epsilon_j}^{-1}(\tau|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}),$$

where $F_{\epsilon_j}(\cdot|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij})$ is conditional distribution of ϵ_{ij}. Let $u_{ij} = F_j(Y_{ij}|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij})$, which follows $U(0,1)$, then in the above formula, replacing τ with u_{ij} can deduce

$$Y_{ij} = \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0 + \boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{\alpha}_0(T_{ij}) + F_{\epsilon_j}^{-1}(u_{ij}|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}).$$

For a given multivariate copula $C(\cdot; \boldsymbol{\theta}_0)$, denote $\gamma_{ijl(k)}$ as the conditional covariance of $S^0_{ij(k)}$ and $S^0_{il(k)}$, by some calculation, we have that

$$
\begin{aligned}
&\gamma_{ijl(k)} \\
&= P(\epsilon_{ij} - b(\tau_k) \le 0, \epsilon_{il} - b(\tau_k) \le 0 | \boldsymbol{X}_{ij}, \boldsymbol{X}_{il}, \boldsymbol{Z}_{ij}, \boldsymbol{Z}_{il}, T_{ij}, T_{il}) - \tau_k^2 \\
&= P(F_{\epsilon_j}^{-1}(u_{ij} | \boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}) \le b(\tau_k), F_{\epsilon_l}^{-1}(u_{il} | \boldsymbol{X}_{il}, \boldsymbol{Z}_{il}, T_{il}) \\
&\le b(\tau_k) | \boldsymbol{X}_{ij}, \boldsymbol{X}_{il}, \boldsymbol{Z}_{ij}, \boldsymbol{Z}_{il}, T_{ij}, T_{il}) - \tau_k^2 \\
&= P(u_{ij} \le \tau_k, u_{il} \le \tau_k | \boldsymbol{X}_{ij}, \boldsymbol{X}_{il}, \boldsymbol{Z}_{ij}, \boldsymbol{Z}_{il}, T_{ij}, T_{il}) - \tau_k^2 \\
&= P(F_j(Y_{ij} | \boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}) \le \tau_k, F_l(Y_{il} | \boldsymbol{X}_{il}, \boldsymbol{Z}_{il}, T_{il}) \\
&\le \tau_k | \boldsymbol{X}_{ij}, \boldsymbol{X}_{il}, \boldsymbol{Z}_{ij}, \boldsymbol{Z}_{il}, T_{ij}, T_{il}) - \tau_k^2 \\
&= C_{jl}(\tau_k, \tau_k; \boldsymbol{\theta}_0) - \tau_k^2,
\end{aligned}
\tag{10}
$$

where $C_{jl}(\tau_k, \tau_k; \boldsymbol{\theta}_0)$ is induced by setting the jth and lth elements of $C(\cdot; \boldsymbol{\theta}_0)$ to τ_k and the rest to 1. Then, the resulting copula-based covariance matrix of $\boldsymbol{S}_{i(k)}^0$ is

$$
\boldsymbol{V}_{i(k)}(\boldsymbol{\theta}_0) = \left(C_{jl}(\tau_k, \tau_k; \boldsymbol{\theta}_0) - \tau_k^2 \right)_{j,l=1}^m.
$$

For example, if we choose Gaussian copula, the resulting copula-based covariance matrix is $\boldsymbol{V}_{i(k)} = (\Phi_2(\Phi^{-1}(\tau_k), \Phi^{-1}(\tau_k); \rho_{jl}) - \tau_k^2)_{j,l=1}^m$, where $\Phi_2(\cdot, \cdot; \rho_{jl})$ is the standardized bivariate normal distribution with correlation coefficient ρ_{jl}, and $\Phi^{-1}(\cdot)$ is the inverse function of the univariate standardized normal distribution. The correlation coefficient ρ_{jl} depends on the working correlation structures.

Furthermore, we need to estimate $\boldsymbol{\theta}_0$, which can be obtained by maximizing the following copula log-likelihood

$$
l(\boldsymbol{\theta}) = \sum_{i=1}^n \log[c(u_{i1}, \ldots, u_{im}; \boldsymbol{\theta})], \tag{11}
$$

where $c(\cdot; \boldsymbol{\theta})$ is the density associated with the copula $C(\cdot; \boldsymbol{\theta})$ and $u_{ij} = \{v \in (0,1) : F_{\epsilon_j}^{-1}(v | \boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}) + \boldsymbol{X}_{ij}^{\mathrm{T}} \boldsymbol{\beta}_0 + \boldsymbol{Z}_{ij}^{\mathrm{T}} \boldsymbol{\alpha}_0(T_{ij}) = Y_{ij}\} = P(Y_j \le Y_{ij} | \boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij}) = F_j(Y_{ij} | \boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}, T_{ij})$. However, the $u_{ij}, i = 1, \ldots, n, j = 1, \ldots, m$ are unknown, we will first obtain their estimators $\tilde{u}_{ij}$, then by maximizing the copula likelihood, we can obtain the estimator $\hat{\boldsymbol{\theta}}$. Finally, we obtain an efficient estimators of $\boldsymbol{\beta}_0$ and $\boldsymbol{\alpha}_0(\cdot)$ by incorporating the copula-based correlation structure $\boldsymbol{V}_{i(k)}(\hat{\boldsymbol{\theta}}) = (C_{jl}(\tau_k, \tau_k; \hat{\boldsymbol{\theta}}) - \tau_k^2)_{j,l=1}^m$ in the estimating Eq. (7). The details are as follows.

Step 1. Let $0 < \varsigma_1 < \cdots < \varsigma_{\kappa_n} < 1$ be a grid of quantile levels, where $\varsigma_k = \frac{k}{\kappa_n + 1}$. For $k = 1, \ldots, \kappa_n$, obtain $(\tilde{b}(\varsigma_k), \tilde{\boldsymbol{\zeta}}(\varsigma_k)) = \underset{b(\varsigma_k), \boldsymbol{\zeta}(\varsigma_k)}{\arg\min} \sum_{i=1}^n \sum_{j=1}^m \rho_{\varsigma_k}(Y_{ij} - b(\varsigma_k) - \boldsymbol{D}_{ij}^{\mathrm{T}} \boldsymbol{\zeta}(\varsigma_k))$. Note that $\tilde{b}(\varsigma_k) + \boldsymbol{D}_{ij}^{\mathrm{T}} \tilde{\boldsymbol{\zeta}}(\varsigma_k)$ and $\tilde{b}(\varsigma_{k+1}) + \boldsymbol{D}_{ij}^{\mathrm{T}} \tilde{\boldsymbol{\zeta}}(\varsigma_{k+1})$ are estimators of ς_k and ς_{k+1} conditional quantiles, respectively. Thus, if $Y_{ij} \in [\tilde{b}(\varsigma_k) + \boldsymbol{D}_{ij}^{\mathrm{T}} \tilde{\boldsymbol{\zeta}}(\varsigma_k), \tilde{b}(\varsigma_{k+1}) + \boldsymbol{D}_{ij}^{\mathrm{T}} \tilde{\boldsymbol{\zeta}}(\varsigma_{k+1}))$, we think that $u_{ij} \in (\varsigma_k, \varsigma_{k+1})$. Then, we estimate u_{ij} by weighted average of ς_k and ς_{k+1}, i.e.,

$$
\tilde{u}_{ij} = (1 - \tilde{\alpha}_{ij}) \varsigma_k + \tilde{\alpha}_{ij} \varsigma_{k+1},
$$

where

$$\tilde{\alpha}_{ij} = \frac{Y_{ij} - \tilde{b}(\varsigma_k) - \boldsymbol{D}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{\zeta}}(\varsigma_k)}{(\tilde{b}(\varsigma_{k+1}) - \tilde{b}(\varsigma_k)) + (\boldsymbol{D}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{\zeta}}(\varsigma_{k+1}) - \boldsymbol{D}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{\zeta}}(\varsigma_k))}$$

and κ_n is chosen to be $[4 + 3n^{0.4}]$ as suggested by Wang et al. (2018), where $[a]$ denotes the integer part of a.

Step 2. Obtain estimator of $\boldsymbol{\theta}_0$ by maximizing pseudo-copula log-likelihood, i.e.,

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^{n} \log[c(\tilde{u}_{i1}, \ldots, \tilde{u}_{im}; \boldsymbol{\theta})] \right\}.$$

Step 3. In (7), after replacing $\boldsymbol{V}_{i(k)}$ and $\boldsymbol{\Lambda}_{i(k)}$ by their estimators $\boldsymbol{V}_{i(k)}(\hat{\boldsymbol{\theta}})$ and $\hat{\boldsymbol{\Lambda}}_{i(k)}$ respectively, we can obtain the estimator $\hat{\boldsymbol{\zeta}}^{o} = (\hat{\boldsymbol{\beta}}^{oT}, \hat{\boldsymbol{\gamma}}^{oT})^{\mathrm{T}}$ by solving

$$\hat{\boldsymbol{U}}(\boldsymbol{\zeta}) = \sum_{k=1}^{K}\sum_{i=1}^{n} \boldsymbol{D}_i^{\mathrm{T}}\hat{\boldsymbol{\Lambda}}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\hat{\boldsymbol{\theta}})\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}) = \boldsymbol{0}.$$

In our implementation, $\hat{\boldsymbol{\Lambda}}_{i(k)} = \mathrm{diag}(\hat{s}_{i1(k)}, \ldots, \hat{s}_{im(k)})$ is obtained by the quotient estimation method of Hendricks and Koenker (1992), i.e.,

$$\hat{s}_{ij(k)} = \frac{2h}{(\tilde{b}(\tau_k + h) - \tilde{b}(\tau_k - h)) + (\boldsymbol{D}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{\zeta}}(\tau_k + h) - \boldsymbol{D}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{\zeta}}(\tau_k - h))},$$

where h is a positive bandwidth such that $h \to 0$ as $n \to \infty$, and we choose $h = 1.57n^{-1/3}\{1.5\phi^2[\Phi^{-1}(\tau_k)]/(2[\Phi^{-1}(\tau_k)]^2 + 1)\}^{1/3}$ as Hall and Sheather (1988).

The difficulty of solving estimating equations $\hat{\boldsymbol{U}}(\boldsymbol{\zeta}) = \boldsymbol{0}$ lies in that $\hat{\boldsymbol{U}}(\boldsymbol{\zeta})$ is non-convex, noncontinuous and not differentiable. To overcome these difficulties, we apply the induced smoothing method introduced in Brown and Wang (2005) to solve it. Specifically, we assume $\boldsymbol{\Gamma} \sim N(\boldsymbol{0}, \boldsymbol{I}_{p+qd_n})$ and $\boldsymbol{\Omega}$ is a $(p + qd_n) \times (p + qd_n)$ positive definite matrix, the induced smoothing estimating functions of $\hat{\boldsymbol{U}}(\boldsymbol{\zeta})$ can be naturally defined as

$$\begin{aligned}\tilde{\boldsymbol{U}}(\boldsymbol{\zeta}) &= E_{\boldsymbol{\Gamma}}\left[\hat{\boldsymbol{U}}(\boldsymbol{\zeta} + \boldsymbol{\Omega}^{1/2}\boldsymbol{\Gamma})\right] \\ &= E_{\boldsymbol{\Gamma}}\left[\sum_{k=1}^{K}\sum_{i=1}^{n} \boldsymbol{D}_i^{\mathrm{T}}\hat{\boldsymbol{\Lambda}}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\hat{\boldsymbol{\theta}})\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta} + \boldsymbol{\Omega}^{1/2}\boldsymbol{\Gamma})\right] \\ &= \sum_{k=1}^{K}\sum_{i=1}^{n} \boldsymbol{D}_i^{\mathrm{T}}\hat{\boldsymbol{\Lambda}}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\hat{\boldsymbol{\theta}})\tilde{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}),\end{aligned} \tag{12}$$

$\tilde{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}) = (\tilde{S}_{i1(k)}(\boldsymbol{\zeta}), \ldots, \tilde{S}_{im(k)}(\boldsymbol{\zeta}))^{\mathrm{T}}$, $\tilde{S}_{ij(k)}(\boldsymbol{\zeta}) = \tau_k + \Phi(\frac{Y_{ij} - \boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\zeta} - \hat{b}_k}{r_{ij}}) - 1$, where $\Phi(\cdot)$ represents the standard normal cumulative distribution function, $r_{ij} = [\boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\Omega}\boldsymbol{D}_{ij}]^{1/2}$. Then, solve equation $\tilde{\boldsymbol{U}}(\boldsymbol{\zeta}) = \boldsymbol{0}$, we can obtain the estimator $\hat{\boldsymbol{\zeta}}^{s} = (\hat{\boldsymbol{\beta}}^{sT}, \hat{\boldsymbol{\gamma}}^{sT})^{\mathrm{T}}$.

Now we describe the algorithm of solving $\tilde{\boldsymbol{U}}(\boldsymbol{\zeta}) = \boldsymbol{0}$. Because $\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})$ are smoothing functions of $\boldsymbol{\zeta}$, the Newton-Raphson iteration algorithm can be written as follows.

Step 3.1: Given initial values $\boldsymbol{\zeta}^{(0)}$ obtained from (5), and $\boldsymbol{\Omega}^{(0)} = \frac{1}{n}\boldsymbol{I}_{p+qd_n}$.
Step 3.2: For a given $\boldsymbol{\zeta}^{(t)}$ and $\boldsymbol{\Omega}^{(t)}$, update $\boldsymbol{\zeta}^{(t+1)}$ and $\boldsymbol{\Omega}^{(t+1)}$ by

$$\boldsymbol{\zeta}^{(t+1)} = \boldsymbol{\zeta}^{(t)} - \tilde{\boldsymbol{D}}^{-1}\big(\boldsymbol{\zeta}^{(t)}, \boldsymbol{\Omega}^{(t)}\big)\tilde{\boldsymbol{U}}\big(\boldsymbol{\zeta}^{(t)}, \boldsymbol{\Omega}^{(t)}\big),$$

$$\boldsymbol{\Omega}^{(t+1)} = \Big[\tilde{\boldsymbol{D}}^{-1}\big(\boldsymbol{\zeta}^{(t)}, \boldsymbol{\Omega}^{(t)}\big)\Big]\mathrm{cov}\big[\tilde{\boldsymbol{U}}\big(\boldsymbol{\zeta}^{(t)}, \boldsymbol{\Omega}^{(t)}\big)\big]\Big[\tilde{\boldsymbol{D}}^{-1}\big(\boldsymbol{\zeta}^{(t)}, \boldsymbol{\Omega}^{(t)}\big)\Big]^{\mathrm{T}},$$

where $\tilde{\boldsymbol{D}}\big(\boldsymbol{\zeta}^{(t)}, \boldsymbol{\Omega}^{(t)}\big) = -\sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{D}_i^{\mathrm{T}}\hat{\boldsymbol{\Lambda}}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\hat{\boldsymbol{\theta}})\tilde{\boldsymbol{\Lambda}}_{i(k)}\boldsymbol{D}_i$,

$$\tilde{\boldsymbol{\Lambda}}_{i(k)} = \mathrm{diag}\left(\phi\left(\frac{Y_{i1} - \boldsymbol{D}_{i1}^{\mathrm{T}}\boldsymbol{\zeta}^{(t)} - \hat{b}_k}{r_{i1}^{(t)}}\right)/r_{i1}^{(t)}, \ldots, \phi\left(\frac{Y_{im} - \boldsymbol{D}_{im}^{\mathrm{T}}\boldsymbol{\zeta}^{(t)} - \hat{b}_k}{r_{im}^{(t)}}\right)/r_{im}^{(t)}\right),$$

$r_{ij}^{(t)} = [\boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\Omega}^{(t)}\boldsymbol{D}_{ij}]^{1/2}$, $\phi(\cdot)$ is the standard normal density function and

$$\mathrm{cov}\big[\tilde{\boldsymbol{U}}\big(\boldsymbol{\zeta}^{(t)}, \boldsymbol{\Omega}^{(t)}\big)\big] = \sum_{k,k'=1}^{K}\sum_{i=1}^{n}\boldsymbol{D}_i^{\mathrm{T}}\hat{\boldsymbol{\Lambda}}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\hat{\boldsymbol{\theta}})\tilde{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}^{(t)})\tilde{\boldsymbol{S}}_{i(k')}^{\mathrm{T}}(\boldsymbol{\zeta}^{(t)})\boldsymbol{V}_{i(k')}^{-1}(\hat{\boldsymbol{\theta}})\hat{\boldsymbol{\Lambda}}_{i(k')}\boldsymbol{D}_i.$$

Step 3.3: Iterate *Step 3.2* until convergence.

In the above procedure, if $\|\boldsymbol{\zeta}^{(t+1)} - \boldsymbol{\zeta}^{(t)}\| < 10^{-6}$, we stop the iteration. For computation convenience, we use cubic splines (i.e., $D = 3$) and choose the optimal K_n as the minimizer to the following Schwarz-type information criterion

$$\mathrm{SIC}(K_n) = \log\left\{\frac{1}{n}\sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{S}_{i(k)}^{\mathrm{T}}(\hat{\boldsymbol{\zeta}}^s)\boldsymbol{V}_{i(k)}^{-1}(\hat{\boldsymbol{\theta}})\boldsymbol{S}_{i(k)}(\hat{\boldsymbol{\zeta}}^s)\right\} + \frac{\log n}{2n}\{p + qd_n\}, \quad (13)$$

where $\hat{\boldsymbol{\zeta}}^s$ is the resulting estimator with K_n.

2.2 Asymptotic properties

Then, we will investigate the asymptotic properties of $\hat{\boldsymbol{\beta}}^o$, $\hat{\boldsymbol{\beta}}^s$, $\hat{\alpha}_k^o(t) = \boldsymbol{B}(t)^{\mathrm{T}}\hat{\boldsymbol{\gamma}}_k^o$ and $\hat{\alpha}_k^s(t) = \boldsymbol{B}(t)^{\mathrm{T}}\hat{\boldsymbol{\gamma}}_k^s$, $k = 1, \ldots, q$. For convenience, write $\boldsymbol{B} = \big(\boldsymbol{B}_1^{\mathrm{T}}, \ldots, \boldsymbol{B}_n^{\mathrm{T}}\big)^{\mathrm{T}}$, $\boldsymbol{X} = (\boldsymbol{X}_1^{\mathrm{T}}, \ldots, \boldsymbol{X}_n^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\Lambda}_{(k)} = \mathrm{diag}(\boldsymbol{\Lambda}_{1(k)}, \ldots, \boldsymbol{\Lambda}_{n(k)})$,

$$\boldsymbol{P}_{(k)} = \boldsymbol{B}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\Lambda}_{(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Lambda}_{(k)}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\Lambda}_{(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Lambda}_{(k)},$$

$$\tilde{\boldsymbol{X}}_{(k)} = (\boldsymbol{I} - \boldsymbol{P}_{(k)})\boldsymbol{X},$$

$$\boldsymbol{\Sigma} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\tilde{\boldsymbol{X}}_{i(k)}^{\mathrm{T}}\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Lambda}_{i(k)}\tilde{\boldsymbol{X}}_{i(k)},$$

$$\boldsymbol{\Xi} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{\boldsymbol{X}}_{i(k)}^{\mathrm{T}}\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{W}^{kk'}\boldsymbol{V}_{i(k')}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Lambda}_{i(k')}\tilde{\boldsymbol{X}}_{i(k')},$$

with $\boldsymbol{W}^{kk'} = \left(W_{jj'}^{kk'} \sqrt{\tau_k(1-\tau_k)}\sqrt{\tau_{k'}(1-\tau_{k'})} \right)_{j,j'=1}^{m}$ and $W_{jj'}^{kk'} = \mathrm{corr}(S_{ij(k)}^0, S_{ij'(k')}^0)$.

Theorem 1 *Under the regularity conditions (C1)–(C8) provided in the supplemental file, and if* $\kappa_n^{2a+1}/n^{1-2v} \to 0$, $\kappa_n \to \infty$ *and* $K_n \log K_n/n \to 0$ *as* $n \to \infty$, *we have the following results*

(a) $\sqrt{n}(\hat{\boldsymbol{\beta}}^o - \boldsymbol{\beta}_0) \to_d N(\mathbf{0}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Xi}\boldsymbol{\Sigma}^{-1})$,
(b) $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}(\hat{\alpha}_k^o(T_{ij}) - \alpha_{0k}(T_{ij}))^2 = O_p\left(\frac{K_n}{n} + K_n^{-2r}\right), k = 1, \ldots, q$,
(c) $\sqrt{n}(\hat{\boldsymbol{\beta}}^s - \boldsymbol{\beta}_0) \to_d N(\mathbf{0}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Xi}\boldsymbol{\Sigma}^{-1})$,
(d) $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}(\hat{\alpha}_k^s(T_{ij}) - \alpha_{0k}(T_{ij}))^2 = O_p\left(\frac{K_n}{n} + K_n^{-2r}\right), k = 1, \ldots, q$,

where v, a and r are defined in the regularity conditions. In particular, if $K_n = O(n^{1/(2r+1)})$, *then* $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}(\hat{\alpha}_k^o(T_{ij}) - \alpha_{0k}(T_{ij}))^2 = O_p\left(n^{\frac{-2r}{2r+1}}\right)$, $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}(\hat{\alpha}_k^s(T_{ij}) - \alpha_{0k}(T_{ij}))^2 = O_p\left(n^{\frac{-2r}{2r+1}}\right)$.

Theorem 1 gives the asymptotic distribution for the estimators of parametric components and the convergence rate for the estimators of nonparametric functions. Under the same smoothness assumptions (C6), we know that the convergence rate for nonparametric estimator is optimal. Theorem 1 also indicates that the smoothed estimating functions $\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})$ is equivalent to their original counterpart $\hat{\boldsymbol{U}}(\boldsymbol{\zeta})$.

3 Numerical experiment and real data analysis

In this section, we carry out simulation studies and real data analysis to investigate the finite sample performances of the proposed new method. In the CQR-based methods, we choose $K = 9$ as suggested by Zou and Yuan (2008).

3.1 Numerical experiment

We consider the following model

$$Y_{ij} = \sum_{k=1}^{3} X_{ij,k}\beta_k + \sum_{k=1}^{4} Z_{ij,k}\alpha_k(T_{ij}) + \epsilon_{ij}, i = 1, \ldots, n, j = 1, \ldots, 5, \tag{14}$$

and generate 500 data sets from (14) with $n = 100$. In each replication, $T_{ij} \sim U[0,1]$, $\boldsymbol{\beta} = (1.5, 2.5, 3)^{\mathrm{T}}$, $\boldsymbol{Z}_{ij} \sim N_4(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}_{kl} = 0.5^{|k-l|}$, $X_{ij,1}, \ldots, X_{ij,3}$ are random realizations of Gaussian processes with zero mean and covariance structure $E[X_{ij,k_1}X_{ij,k_2}] = 4\exp(-|k_1 - k_2|)$. $\alpha_1(t) = 15 + 20\sin(0.5\pi t)$, $\alpha_2(t) = 2 - 3\cos(\frac{1}{3}\pi(6t-5))$, $\alpha_3(t) = 6 - 6t$ and $\alpha_4(t) = \cos(-3t - 1)$. For the random errors, the following three cases are considered.

Case 1: Random error $\boldsymbol{\epsilon}_i$ follows $N(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ has an AR(1) correlation structure with variance 1 and correlation coefficient $\delta = 0.5$.
Case 2: Random error $\boldsymbol{\epsilon}_i$ follows a multivariate t-distribution with the degree 3 and covariance matrix $\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is same as that in the Case 1.
Case 3: Random error $\boldsymbol{\epsilon}_i$ follows the same multivariate normal distribution as that in Case 1, but some outliers are included. We perturb the response Y_{ij} for two randomly chosen subjects to $Y_{ij} + 5$.

To evaluate the estimation accuracy, we use the mean squared error (MSE) for $\hat{\beta}_k, k = 1, 2, 3$, and the integrated mean squared error (IMSE) for $\hat{\alpha}_k(u), k = 1, 2, 3, 4$, defined as:

$$\text{IMSE}\{\hat{\alpha}_k(u)\} = \frac{1}{500}\sum_{i=1}^{500}\frac{1}{100}\sum_{j=1}^{100}\{\hat{\alpha}_{k,i}(t_j) - \alpha_k(t_j)\}^2,$$

where $\{t_j\}_{j=1}^{100}$ is a grid equally spaced on [0.02, 0.98], $\hat{\alpha}_{k,i}(\cdot), i = 1, \ldots, 500$ are the estimates of $\alpha_k(\cdot)$ in the ith replicate. Tables 1, 2 and 3 report the simulation results. For a clear illustration, we also compare our method with other methods, i.e., the conventional CQR estimator (Kai et al. 2011) without considering possible correlations (denoted as CQR); the Gaussian copula-based quantile regression method with $\tau = 0.5$ (denoted as CQR), which extend the method for linear regression model in Fu and Wang (2016) to the semiparametric model; the GEE method (Liang and Zeger 1986) and the robust GEE method (He et al. 2005, denoted as RGEE). For our copula and CQR-based method, we consider multivariate Gaussian copula and t-copula, they are denoted as "GCQR" and "TCQR," respectively, and we use the AR(1) and exchangeable working correlation structures, respectively.

Across all scenarios considered, the proposed new copula and CQR-based estimator shows higher efficiency than the working independence CQR estimator and $\text{GQR}_{0.5}$, and the efficiency gain is more obvious. Results also show that the proposed method is quite insensitive to the choice of copula function and correlation structures, the copula and CQR-based method performs well even under the misspecification of the copula function or correlation structure. Furthermore, by inheriting the robust and efficient properties of CQR and improving the efficiency by using the copula, the new estimators perform obviously better than the RGEE methods for the Cases 1–3. Even for the Case 1, our method performs equally well as the GEE since their IMSE and MSE have little difference.

3.2 Real data analysis

In this subsection, we illustrate our method by analyzing a subset of data from the Multi-Center AIDS Cohort study. The data set reports the human immunodeficiency virus (HIV) status of 283 homosexual men that were infected with HIV during the follow-up period between 1984 and 1991. Details of the study design, methods, and medical implications have been given by Kaslow et al. (1987). In our analysis, we take x_1 to be the smoking status: (1 for a smoker and 0 for a nonsmoker), x_2 to be the

Table 1 Simulation results for the Case 1

	IMSE $\times 10^2$				MSE $\times 10^3$		
	$\alpha_1(t)$	$\alpha_2(t)$	$\alpha_3(t)$	$\alpha_4(t)$	β_1	β_2	β_3
GCQR(ex)	3.012	3.532	2.775	2.972	4.161	4.619	3.237
GCQR(ar1)	3.118	3.577	2.797	3.013	4.212	4.597	3.189
TCQR(ex)	3.198	3.519	2.801	2.989	4.263	4.713	3.327
TCQR(ar1)	3.203	3.602	2.818	3.097	4.272	4.737	3.176
CQR	4.032	5.101	3.115	3.365	5.203	6.411	5.603
$GQR_{0.5}$(ex)	4.155	5.236	3.209	3.414	5.411	6.489	5.711
$GQR_{0.5}$(ar1)	4.207	5.262	3.311	3.398	5.397	6.545	6.009
RGEE(ex)	3.997	5.003	3.069	3.252	5.111	6.350	5.420
RGEE(ar1)	4.032	5.101	3.115	3.365	5.322	6.411	5.603
GEE(ex)	2.898	4.631	2.572	2.967	4.558	5.762	4.996
GEE(ar1)	3.197	4.689	2.616	3.018	4.621	5.883	5.133

$\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{i5})^{\mathrm{T}}$ follows $N(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ has an AR(1) correlation structure with variance 1 and correlation $\delta = 0.5$. IMSE: the integrated mean squared error, MSE: the mean squared error

Table 2 Simulation results for the Case 2

	IMSE $\times 10^2$				MSE $\times 10^3$		
	$\alpha_1(t)$	$\alpha_2(t)$	$\alpha_3(t)$	$\alpha_4(t)$	β_1	β_2	β_3
GCQR(ex)	3.271	3.738	3.119	3.198	4.793	5.323	3.407
GCQR(ar1)	3.302	3.762	3.201	3.203	4.815	5.339	3.411
TCQR(ex)	3.262	3.665	3.035	3.115	4.776	5.268	3.353
TCQR(ar1)	3.277	3.712	3.113	3.176	4.801	5.317	3.376
CQR	9.129	7.819	9.758	9.616	9.979	9.765	9.892
$GQR_{0.5}$(ex)	9.167	7.916	10.007	9.732	10.117	9.868	10.011
$GQR_{0.5}$(ar1)	9.213	8.026	10.131	9.739	10.205	9.785	10.105
RGEE(ex)	9.553	9.759	9.967	9.135	10.529	9.723	10.167
RGEE(ar1)	10.011	9.838	11.016	9.258	10.727	9.919	10.663
GEE(ex)	24.986	26.577	28.701	25.086	30.121	27.927	26.008
GEE(ar1)	25.196	27.467	29.258	25.173	30.353	28.115	26.168

$\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{i5})^{\mathrm{T}}$ follows a multivariate t-distribution with the degree 3 and covariance matrix $\mathbf{\Sigma}$, where $\mathbf{\Sigma}$ is same as that in the Case 1. IMSE: the integrated mean squared error, MSE: the mean squared error

standardized variable for age, x_3 to be the standardized variable for PreCD4, and set the standardized individual CD4 percentage as response variable y.

Several researchers have studied the same data set by fitting nonparametric models (Huang et al. 2002; Fan and Zhang 2000) or semi-parametric models (Wang and Lin 2015; Wang et al. 2019). Fan and Li (2004) showed that there are quadratic and interaction effects. So the quadratic terms of x_2 and x_3 and the interactions of the three covariates are also considered, and we consider the following PLVC model

Table 3 Simulation results for the Case 3

	IMSE $\times 10^2$				MSE $\times 10^3$		
	$\alpha_1(t)$	$\alpha_2(t)$	$\alpha_3(t)$	$\alpha_4(t)$	β_1	β_2	β_3
GCQR(ex)	3.511	3.818	3.218	3.231	4.872	5.587	3.612
GCQR(ar1)	3.507	3.876	3.267	3.307	4.907	5.603	3.637
TCQR(ex)	3.515	3.912	3.232	3.322	4.917	5.615	3.688
TCQR(ar1)	3.522	3.898	3.229	3.335	4.879	5.621	3.657
CQR	9.013	8.979	9.937	9.609	9.632	9.786	8.397
$GQR_{0.5}$(ex)	9.158	9.093	10.122	9.858	10.117	10.015	9.711
$GQR_{0.5}$(ar1)	9.173	9.115	10.087	9.797	10.205	10.078	9.616
RGEE(ex)	9.179	9.961	10.216	9.317	9.901	10.128	8.839
RGEE(ar1)	9.236	10.121	10.977	9.502	10.114	10.357	8.918
GEE(ex)	27.736	29.612	30.878	27.257	29.971	30.258	29.369
GEE(ar1)	27.892	29.893	31.408	27.416	30.732	30.397	30.532

$\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{i5})^{\mathrm{T}}$ follows the same multivariate normal distribution as that in Case 1, but we perturb the response Y_{ij} for two randomly chosen subjects to $Y_{ij} + 5$. IMSE: the integrated mean squared error, MSE: the mean squared error

$$y_{ij} = \alpha_0(t_{ij}) + \alpha_1(t_{ij})x_{i1} + \alpha_2(t_{ij})x_{i2} + \alpha_3(t_{ij})x_{i3} + \beta_1 x_{i2}^2 + \beta_2 x_{i3}^2 + \beta_3 x_{i1}x_{i2} + \beta_4 x_{i1}x_{i3} + \beta_5 x_{i2}x_{i3} + \epsilon_{ij}. \tag{15}$$

Wang et al. (2009) indicated that there are outliers in this data set. This motivates us to fit the model (15) using our robust CQR estimating equations method. The first 15 subjects are used to evaluate the predictive ability of the estimated model, and the remaining subjects are used as a training data set to fit the model. Furthermore, the median absolute prediction error (MAPE), which is defined as the median of $\{|y_{ij} - \hat{y}_{ij}|, i = 1, \ldots, 15, j = 1, \ldots, m_i\}$, is used to measure the prediction performance.

We report the analysis results for the Gaussian copula and t-copula with exchangeable and AR(1) correlation structures, respectively. Figure 1a–d shows the estimated varying coefficient functions, the estimation results for the linear part and the MAPE are reported in Table 4. Smoking and Age tend to have a negative interaction, that is, the elder smokers tend to have lower median CD4 counts, and this agrees with the findings in Fan and Li (2004). In addition, our method suggests that the quadratic terms other two interactions are positive. What is more, from Table 4 and Fig. 1, we also find that our method is quite insensitive to the choice of copula function and correlation structure.

Table 4 Estimation results for the parametric components in model (15)

Variable	Gaussian copula		t-copula	
	Exchangeable	AR(1)	Exchangeable	AR(1)
Age^2	0.014	0.023	0.027	0.026
$PreCD4^2$	0.009	0.017	0.010	0.004
Smoking*Age	− 0.132	− 0.138	− 0.148	− 0.151
Smoking*PreCD4	0.139	0.102	0.064	0.104
Age*PreCD4	0.033	0.018	0.007	0.008
MAPE	7.056	7.211	7.155	7.023

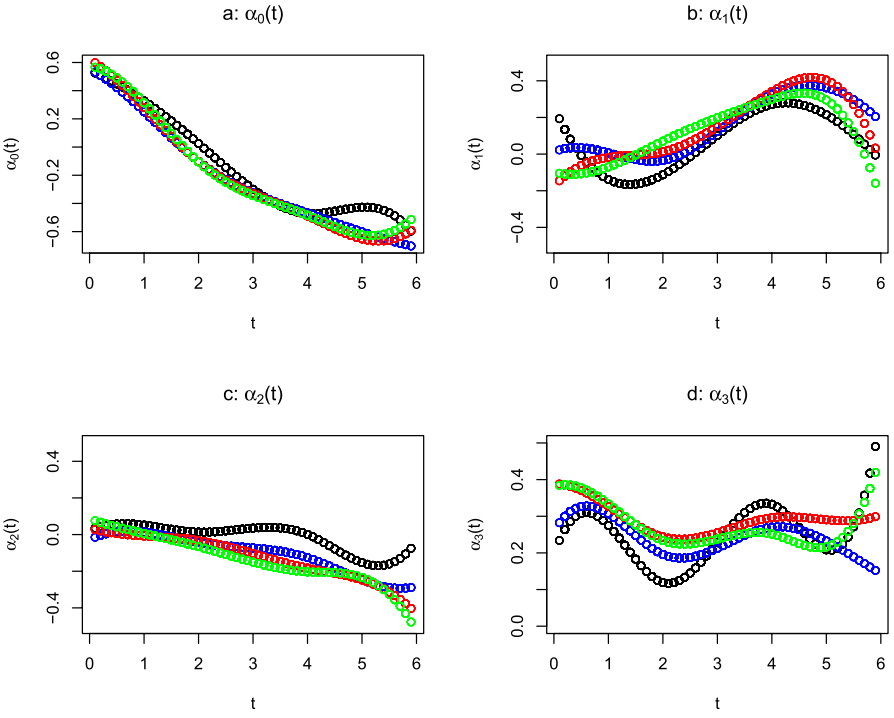

Fig. 1 **a**–**d** show the estimated $\alpha_0(t)$, $\alpha_1(t)$, $\alpha_2(t)$ and $\alpha_3(t)$ for Gaussian copula with exchangeable correlation structure (black circle), Gaussian copula with AR(1) correlation structure (blue circle), t-copula with exchangeable correlation structure (red circle) and t-copula with AR(1) correlation structure (green circle)

Acknowledgements The research was supported by NNSF Project (11901356), wealth management Project (2019ZBKY047) of Shandong Technology and Business University.

References

Bai, Y., Kang, J., Song, P. (2014). Efficient pairwise composite likelihood estimation for spatial clustered data. *Biometrics*, *70*, 661–670.

Brown, B., Wang, Y. (2005). Standard errors and covariance matrices for smoothed rank estimators. *Biometrika*, *92*, 149–158.

Fan, J., Li, R. (2004). New estimation and model selection procedures for semiparametric modeling in longitudinal data analysis. *Journal of the American Statistical Association*, *99*, 710–723.

Fan, J., Zhang, J. (2000). Two-step estimation of functional linear models with applications to longitudinal data. *Journal of the Royal Statistical Society*: *Series B*, *62*, 303–322.

Fan, Y., Qin, G., Zhu, Z. (2012). Variable selection in robust regression models for longitudinal data. *Journal of Multivariate Analysis*, *109*, 156–167.

Fan, Y., Härdle, W., Wang, W., Zhu, L. (2018). Single-index-based CoVaR with very high-dimensional covariates. *Journal of Business and Economic Statistics*, *36*, 212–226.

Fu, L., Wang, Y. (2016). Efficient parameter estimation via Gaussian copulas for quantile regression with longitudinal data. *Journal of Multivariate Analysis*, *143*, 492–502.

Haff, I., Aas, K., Frigessi, A. (2010). On the simplified pair-copula construction-simply useful or too simplistic? *Journal of Multivariate Analysis*, *101*, 1296–1310.

Hall, P., Sheather, S. (1988). On the distribution of a studentized quantile. *Journal of the Royal Statistical Society*: *Series B*, *50*, 381–391.

He, X., Fung, W., Zhu, Z. (2005). Robust estimation in generalized partial linear models for clustered data. *Journal of the American Statistical Association*, *100*, 1176–1184.

Hendricks, W., Koenker, R. (1992). Hierarchical spline models for conditional quantiles and the demand for electricity. *Journal of the American Statistical Association*, *87*, 58–68.

Huang, J., Wu, C., Zhou, L. (2002). Varying-coefficient models and basis function approximation for the analysis of repeated measurements. *Biometrika*, *89*, 111–128.

Jiang, X., Jiang, J., Song, X. (2012). Oracle model selection for nonlinear models based on weighted composite quantile regression. *Statistica Sinica*, *22*, 1479–1506.

Jung, S. (1996). Quasi-likelihood for median regression models. *Journal of the American Statistical Association*, *91*, 251–257.

Kai, B., Li, R., Zou, H. (2010). Local composite quantile regression smoothing: An efficient and safe alternative to local polynomial regression. *Journal of the Royal Statistical Society: Series B*, *72*, 49–69.

Kai, B., Li, R., Zou, H. (2011). New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. *The Annals of Statistics*, *39*, 399–411.

Kaslow, R., Ostrow, D., Detels, R., Phair, J., Polk, B., Rinaldo, C. (1987). The multicenter AIDS cohort study: Rationale, organization and selected characteristics of the participants. *American Journal of Epidemiology*, *126*, 310–318.

Lai, P., Wang, Q., Lian, H. (2012). Bias-corrected GEE estimation and smooth-threshold GEE variable selection for single-index models with clustered data. *Journal of Multivariate Analysis*, *105*, 422–432.

Li, G., Lian, H., Feng, S., Zhu, L. (2013). Automatic variable selection for longitudinal generalized linear models. *Computational Statistics and Data Analysis*, *61*, 174–186.

Lian, H., Liang, H., Wang, L. (2014). Generalized additive partial linear models for clustered data with diverging number of covariates using GEE. *Statistica Sinica*, *23*, 173–196.

Liang, K., Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*, 13–22.

Lv, J., Yang, H., Guo, C. (2015). An efficient and robust variable selection method for longitudinal generalized linear models. *Computational Statistics and Data Analysis*, *82*, 74–88.

Noh, H., Ghouch, A., Van Keilegom, I. (2015). Semiparametric conditional quantile estimation through copula-based multivariate models. *Journal of Business and Economic Statistics*, *33*, 167–178.

Qin, G., Bai, Y., Zhu, Z. (2012). Robust empirical likelihood inference for generalized partial linear models with longitudinal data. *Journal of Multivariate Analysis*, *105*, 32–44.

Qin, G., Zhu, Z. (2007). Robust estimation in generalized semiparametric mixed models for longitudinal data. *Journal of Multivariate Analysis*, *98*, 1658–1683.

Qin, G., Zhu, Z., Fung, W. (2009). Robust estimation of covariance parameters in partial linear model for longitudinal data. *Journal of Statistical Planning and Inference*, *139*, 558–570.

Schumaker, L. (1981). *Spline functions*: *Basic theory*. New York: Wiley.
Shi, P., Frees, E. (2010). Long-tail longitudinal modeling of insurance company expenses. *Insurance*: *Mathematics and Economics*, *47*, 303–314.
Smith, M., Min, A., Almeida, C., Czado, C. (2010). Modeling longitudinal data using a pair-copula decomposition of serial dependence. *Journal of the American Statistical Association*, *105*, 1467–1479.
Sun, J., Frees, E., Rosenberg, M. (2008). Heavy-tailed longitudinal data modeling using copulas. *Insurance: Mathematics and Economics*, *42*, 817–830.
Sun, J., Gai, Y., Lin, L. (2013). Weighted local linear composite quantile estimation for the case of general error distributions. *Journal of Statistical Planning and Inference*, *143*, 1049–1063.
Song, P. (2000). Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics*, *27*, 305–320.
Tian, R., Xue, L., Hu, Y. (2015). Smooth-threshold GEE variable selection for varying coefficient partially linear models with longitudinal data. *Journal of the Korean Statistical Society*, *44*, 419–431.
Wang, H., Feng, X., Dong, C. (2018). Copula-based quantile regression for longitudinal data. *Statistica Sinica*. https://doi.org/10.5705/ss.202016.0135.
Wang, K., Li, S., Sun, X., Lin, L. (2019). Modal regression statistical inference for longitudinal data semivarying coefficient models: Generalized estimating equations, empirical likelihood and variable selection. *Computational Statistics and Data Analysis*, *133*, 257–276.
Wang, K., Lin, L. (2015). Variable selection in semiparametric quantile modeling for longitudinal data. *Communications in Statistics-Theory and Methods*, *44*, 2243–2266.
Wang, K., Sun, X. (2017). Efficient parameter estimation and variable selection in partial linear varying coefficient quantile regression model with longitudinal data. *Statistical Papers*. https://doi.org/10.1007/s00362-017-0970-0.
Wang, L., Zhou, J., Qu, A. (2012). Penalized generalized estimating equations for high-dimensional longitudinal data analysis. *Biometrics*, *68*, 353–360.
Wang, H., Zhu, Z., Zhou, J. (2009). Quantile regression in partially linear varying coefficient models. *The Annals of Statistics*, *37*, 3841–3866.
Zhao, P., Li, G. (2013). Modified SEE variable selection for varying coefficient instrumental variable models. *Statistical Methodology*, *12*, 60–70.
Zhao, W., Lian, H., Song, X. (2017). Composite quantile regression for correlated data. *Computational Statistics and Data Analysis*, *109*, 15–33.
Zou, H., Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. *The Annals of Statistics*, *36*, 1108–1126.