Functional Data Analysis between Two Random Variables by Multilevel Monotone Splines

Daichi KITAHARA† Ke LENG† Yuji TEZUKA‡ Akira HIRABAYASHI†

† Graduate School of Information Science and Engineering, Ritsumeikan University

‡ Department of Diabetes and Endocrinology, Kusatsu General Hospital

Abstract As data analysis methods, hypothesis testing and regression analysis are famous. However, hypothesis testing can only detect significant differences between two groups divided by some threshold, and regression analysis can only construct an averaged model, whose information is limited. Quantile regression is a robust and flexible analysis method, and it can construct multilevel models. To make the most of quantile regression, we propose multilevel spline smoothing which considers the similarity between the adjacent quantile lines and can enforce non-crossing and monotone properties.

1 Introduction

Data analysis [1] is becoming more important in the big data era. In the simplest case, we analyze a pair of random variables from its observations. One famous analysis method is hypothesis testing [2]–[4]. In this approach, we first divide the observations into two groups by using some empirically determined threshold value on one random variable. Then we check whether the distributions of the other random variable are *significantly different between the two groups*. However, hypothesis testing cannot detect *small differences between the two groups* and *any differences among one group*.

Another famous method is regression analysis [1], [5]–[7]. In this approach, we construct a univariate continuous function which maps one random variable to the other one. Thus, we can analyze the continuous relation between the two random variables. This regression function is often constructed as a low-order polynomial having *the least square errors* due to its simplicity. In the spirit of *robust statistics* [4], [8]–[11], a polynomial having *the least absolute errors*, which leads to regression of the median, is also used for long-tailed data.

Quantile regression [12]–[15], which is the generalization of the above median regression, enables robust and flexible analysis. In this approach, by minimizing certain *asymmetric absolute errors*, we can construct multiple quantile lines, i.e., percentile lines. Hence, we can realize continuous and two-dimensional analysis using the multilevel regression results. Although quantile regression is an effective analysis method, there is a possibility that the polynomial regression model cannot approximate true quantile lines enough. To make the most of quantile regression, a spline regression model is used [16]–[22]. Splines are piecewise polynomials and have been widely used for construction of smooth functions, including regression analysis, due to their flexibility and optimality.

In this paper, we propose a multilevel spline smoothing technique for quantile regression. In the previous methods, a quantile line of each level is individually constructed even though all the quantiles are generated from one cumulative distribution function. We newly consider the smoothness of the cumulative distribution function, which makes the first derivatives of the adjacent quantiles similar. Moreover, we also enforce the non-crossing constraint [23]–[26], whom the adjacent quantiles should satisfy, and an optional monotone, i.e., *non-decreasing*, *non-increasing*, *convex* or *concave*, property [27]–[30], whom a linear or a quadratic polynomial has. Numerical experiments show that the proposed quantiles are more harmonious with each other than the previous ones.

2 Preliminaries

Let $\mathbb{R}$ and $\mathbb{N}$ denote the sets of all real numbers and non-negative integers, respectively. For any open interval (a, b) and $\rho \in \mathbb{N} \cup \{\infty\}$, $C^{\rho}(a, b)$ stands for the set of all ρ-times continuously differentiable real-valued functions on (a, b). For any $d \in \mathbb{N}$, $\mathbb{P}_d$ stands for the set of all univariate real polynomials of degree d at most, i.e., $\mathbb{P}_d := \{u : \mathbb{R} \to \mathbb{R} : x \mapsto \sum_{k=0}^{d} c_k x^k \mid c_k \in \mathbb{R}\}$. We write vectors and matrices with boldface small and boldface capital letters, respectively.

2.1 Regression Analysis

Suppose that we have observations $\{(x_i, y_i)\}_{i=1}^{n}$ of a pair of random variables (X, Y) whose joint probability density function $f_{X,Y}(x, y)$ satisfies $\iint_{\mathbb{R}^2} f_{X,Y}(x, y)\,\mathrm{d}x\mathrm{d}y = 1$ and $f_{X,Y}(x, y) > 0$ for all $(x, y) \in \mathbb{R}^2$. Hence, the conditional probability density function of Y given X is $f_{Y|X}(y|x) := f_{X,Y}(x, y)/f_X(x) := f_{X,Y}(x, y)/\int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}y > 0$. When we analyze the continuous relation between the two random variables X and Y, the least squares regression

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \sum_{i=1}^{n} |y_i - r_{\boldsymbol{\theta}}(x_i)|^2 \tag{1}$$

is often used due to its low computational cost [1], [5]–[7], where $r_{\boldsymbol{\theta}}(x)$ is a certain regression model such as a polynomial and $\boldsymbol{\theta}$ stands for adjustable parameters to be optimized. As n approaches infinity, the optimal solution $r_{\boldsymbol{\theta}^*}(x)$ of (1) converges to *the conditional mean* $\mu_Y(x)$ *of* Y *given* $X = x$:

$$r_{\boldsymbol{\theta}^*}(x) \to \mu_Y(x) := E[Y|X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,\mathrm{d}y$$

under the assumption that $r_{\boldsymbol{\theta}}(x)$ can exactly express $\mu_Y(x)$ if we choose the appropriate $\boldsymbol{\theta}$ (see, e.g., [1] for proof).

It is well-known that the square error in (1) is sensitive to outliers and the reliability of the optimal solution $r_{\boldsymbol{\theta}^*}(x)$ of (1) significantly decreases for long-tailed data [4], [8]–[11]. In such situations, the least absolute deviations regression

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \sum_{i=1}^{n} |y_i - r_{\boldsymbol{\theta}}(x_i)| \tag{2}$$

is used [10], [11]. Since the absolute error in (2) does not over-evaluate the outliers differently from the square error, this regression is robust even if n is not so large. Moreover, as n approaches infinity, the optimal solution $r_{\boldsymbol{\theta}^*}(x)$ of (2) converges to *the conditional median $m_Y(x)$ of Y given $X = x$*:

$$r_{\boldsymbol{\theta}^*}(x) \to m_Y(x) \text{ satisfying } \int_{-\infty}^{m_Y(x)} f_{Y|X}(y|x)\,\mathrm{d}y = 0.5$$

under the assumption $r_{\boldsymbol{\theta}}(x)$ can express $m_Y(x)$ (see [10]).

2.2 Quantile Regression

By generalizing the fact that the least absolute deviations regression leads to the conditional median, we can estimate any *quantile line* as follows. First of all, we define the conditional cumulative distribution function of Y given $X = x$ by

$$F_{Y|x}(y) := \int_{-\infty}^{y} f_{Y|X}(t|x)\,\mathrm{d}t.$$

Since $F_{Y|x}(y)$ becomes a strictly increasing function from the positivity of $f_{Y|X}(y|x)$, the inverse function $F_{Y|x}^{-1}(p)$ is well-defined for $p \in (0,1)$. Actually, the conditional quantile function of Y given $X = x$ is equivalent to $F_{Y|x}^{-1}(p)$ [2], [31]:

$$Q_{Y|x}(p) := F_{Y|x}^{-1}(p).$$

The value of $q_{p,Y}(x) := Q_{Y|x}(p)$ is called *the p-th conditional quantile of Y given $X = x$*. Note that $q_{p,Y}(x)$ is called *the $100p$-th percentile* or *centile* in some papers [32]–[34]. If $p = 0.5$, then $q_{p,Y}(x)$ equals the conditional median $m_Y(x)$.

Define an *asymmetric absolute value function*

$$J_p(t) := \begin{cases} pt & \text{if } t \geq 0, \\ -(1-p)t & \text{if } t < 0, \end{cases}$$

and consider the following optimization problem [12]–[15]

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \sum_{i=1}^{n} J_p(y_i - r_{\boldsymbol{\theta}}(x_i)). \tag{3}$$

Then, as n approaches infinity, the optimal solution $r_{\boldsymbol{\theta}}(x)$ of (3) converges to the p-th conditional quantile $q_{p,Y}(x)$ of Y:

$$r_{\boldsymbol{\theta}^*}(x) \to q_{p,Y}(x) \text{ satisfying } \int_{-\infty}^{q_{p,Y}(x)} f_{Y|X}(y|x)\,\mathrm{d}y = p$$

under the assumption $r_{\boldsymbol{\theta}}(x)$ can express $q_{p,Y}(x)$ (see [14]). Any quantile line is easily obtained only by changing p in (3).

2.3 Spline Function

Let $\sqcup_b := \{I_i := (\xi_{i-1}, \xi_i)\}_{i=1}^{b}$ be a set of subintervals I_i on an open interval $I := (\xi_0, \xi_b)$ s.t. $\xi_i - \xi_{i-1} =: h_i > 0$ $(i = 1, 2, \ldots, b)$. For $\sqcup_b$ and $\rho, d \in \mathbb{N}$ s.t. $0 \leq \rho \leq d$, define

$$\mathcal{S}_d^\rho(\sqcup_b) := \{s \in C^\rho(\xi_0, \xi_b) \mid s = u_i \in \mathbb{P}_d \text{ on } I_i\}$$

as the set of all univariate spline functions of degree d and smoothness ρ on $\sqcup_b$. In what follows, we express a spline function $s \in \mathcal{S}_d^\rho(\sqcup_b)$ in the interval normalization form

$$s(x) := u_i(x) := \sum_{k=0}^{d} c_k^{\langle i \rangle} \left(\frac{x - \xi_{i-1}}{h_i}\right)^k \quad \text{for } x \in (\xi_{i-1}, \xi_i), \tag{4}$$

where $c_k^{\langle i \rangle} \in \mathbb{R}$ $(k = 0, 1, \ldots, d)$ are coefficients of $u_i \in \mathbb{P}_d$.

Spline functions are often used to construct smooth functions, e.g., for computer aided design and regression analysis [16]–[22], due to the optimality shown in Fact 1 below.

Fact 1 (Spline as the unique solution of a variational problem) *There is the unique solution of the following problem*

$$\underset{g \in C^2(-\infty,\infty)}{\text{minimize}} \sum_{i=1}^{n} |y_i - g(x_i)|^2 + \lambda \int_{-\infty}^{\infty} |g''(x)|^2\,\mathrm{d}x, \tag{5}$$

and it is a natural cubic spline, which is a kind of spline function of degree 3 and smoothness 2 [16]. In the problem of (5), the smoothing parameter $\lambda > 0$ controls the trade-off between the data fidelity and the smoothness of the solution.

By using the coefficients $c_k^{\langle i \rangle}$ in (4), we can easily evaluate the characteristics of spline functions as follows.

2.3.1 Quadratic Form of the Roughness Penalty Term

By restricting the domain of interest from $(-\infty, \infty)$ to $I = (\xi_0, \xi_b)$ $(\supseteq (x_{\min}, x_{\max}))$ and the function space from $C^2(-\infty, \infty)$ to $\mathcal{S}_d^\rho(\sqcup_b)$ $(2 \leq \rho \leq d)$ in Fact 1, the roughness penalty term used in (5) can be expressed as

$$\int_I |s''(x)|^2\,\mathrm{d}x = \sum_{i=1}^{b} \int_{I_i} |s''(x)|^2\,\mathrm{d}x. \tag{6}$$

By using the expression in (4), the roughness penalty on I_i is expressed as the following quadratic form

$$\int_{I_i} |s''(x)|^2\,\mathrm{d}x = \sum_{k=2}^{d} \sum_{j=2}^{d} \frac{k(k-1)j(j-1)}{h_i^3(k+j-3)} c_k^{\langle i \rangle} c_j^{\langle i \rangle}. \tag{7}$$

From (6) and (7), the roughness penalty on I is expressed as a quadratic form by $\int_I |s''(x)|^2\,\mathrm{d}x = \boldsymbol{c}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{c}$, where $\boldsymbol{c} := (c_d^{\langle 1 \rangle}, c_{d-1}^{\langle 1 \rangle}, \ldots, c_0^{\langle 1 \rangle}, c_d^{\langle 2 \rangle}, c_{d-1}^{\langle 2 \rangle}, \ldots, c_0^{\langle 2 \rangle}, \ldots, c_0^{\langle b \rangle}) \in \mathbb{R}^{b(d+1)}$ is the coefficient vector of $s(x)$ and $\boldsymbol{Q} \in \mathbb{R}^{b(d+1) \times b(d+1)}$ is a certain symmetric positive semidefinite matrix.

2.3.2 Linear Equation for the ρ-Times Differentiability

For a spline function $s \in \mathcal{S}_d^\rho(\sqcup_b)$ in (4), to ensure the ρ-times continuous differentiability on (ξ_{i-1}, ξ_{i+1}), i.e., $s \in C^\rho(\xi_{i-1}, \xi_{i+1})$, the coefficients of the adjacent pieces $u_i(x)$ and $u_{i+1}(x)$ have to satisfy the following linear equation

$$s \in C^\rho(\xi_{i-1}, \xi_{i+1})$$
$$\Leftrightarrow \frac{1}{h_i^j} \sum_{k=j}^{d} \frac{k!}{(k-j)!} c_k^{\langle i \rangle} - \frac{j!}{h_{i+1}^j} c_j^{\langle i+1 \rangle} = 0 \quad (j = 0, 1, \ldots, \rho). \tag{8}$$

From (8), there is some matrix $\boldsymbol{H} \in \mathbb{R}^{(b-1)(\rho+1) \times b(d+1)}$ satisfying $s \in C^\rho(\xi_0, \xi_b) \Leftrightarrow \boldsymbol{H}\boldsymbol{c} = \boldsymbol{0}$.

2.3.3 Sufficient Condition for the Non-Negativity

In our previous works, we estimated probability density functions by splines [35], [36]. It is difficult to give a useful necessary and sufficient condition for the non-negativity of $s \in \mathcal{S}_d^\rho(\sqcup_b)$ over I_i. Instead, we used a sufficient condition

$$\sum_{k=0}^{j} \frac{(d-k)!}{(j-k)!(d-j)!} c_k^{\langle i \rangle} \geq 0 \quad (j = 0, 1, \ldots, d)$$
$$\Rightarrow s(x) \geq 0 \text{ for all } x \in (\xi_{i-1}, \xi_i) \tag{9}$$

in [37]. From (9), there is some matrix $\boldsymbol{G} \in \mathbb{R}^{b(d+1) \times b(d+1)}$ satisfying $\boldsymbol{G}\boldsymbol{c} \geq \boldsymbol{0} \Rightarrow s(x) \geq 0$ for all $x \in (\xi_0, \xi_b)$.

3 Data Analysis by Multilevel Splines

3.1 Quantile Regression via Spline Smoothing

In the problems of (1), (2), and (3), the most commonly used regression model is a polynomial $r_{\boldsymbol{\theta}}(x) = \sum_{k=0}^{d} c_k x^k$ of degree $d = 1$ or 2 [1], [5]–[7]. In this case, the adjustable parameters are coefficients $\boldsymbol{\theta} = (c_d, c_{d-1}, \ldots, c_0)^{\mathrm{T}} \in \mathbb{R}^{d+1}$. However, there is a high probability that such simple models cannot approximate the true quantile lines $q_{p,Y}(x)$ enough.

To deal with more complex quantile lines flexibly, we can employ a spline regression model $r_{\boldsymbol{\theta}}(x) = s(x) \in \mathcal{S}_d^{\rho}(\sqcup_b)$ as a generalization of the polynomial regression model. In this case, the adjustable parameters equal the coefficient vector $\boldsymbol{\theta} = \boldsymbol{c} \in \mathbb{R}^{b(d+1)}$ of $s(x)$ in Section 2.3.1. Although spline functions are very flexible, overfitting would be caused when the number n of observations is not so large. Therefore, by assuming the energy of local change of $q_{p,Y}(x)$ is small in the same manner as (5), we solve the following problem

$$\underset{s \in \mathcal{S}_d^{\rho}(\sqcup_b)}{\text{minimize}} \sum_{i=1}^{n} J_p(y_i - s(x_i)) + \lambda \int_I |s''(x)|^2 \, \mathrm{d}x, \tag{10}$$

instead of the problem of (3) [14], [15], [22]. By repeatedly solving the problem of (10) for $p = p_l$ ($p_1 < p_2 < \cdots < p_L$), we can construct quantile regression lines of L levels.

3.2 Simultaneous Regression by Monotone Splines

In the above strategy [14], [15], [22], an important condition $\forall x\ q_{p_{l+1},Y}(x) > q_{p_l,Y}(x)$ ($l = 1, 2, \ldots, L-1$) is ignored. Some papers considered this condition and constructed *non-crossing* regression results [23]–[26]. Moreover, other papers enforced *non-decreasing*, *non-increasing*, *convex*, or *concave* property on each spline regression model $s_l(x)$ [27]–[30].

In this paper, in addition to the above properties, we newly utilize the smoothness of $f_{Y|X}(y|x)$, which makes the first derivatives of the adjacent quantiles similar. Hence, we solve

$$\begin{aligned} &\underset{(s_l \in \mathcal{S}_d^{\rho}(\sqcup_b))_{l=1}^{L}}{\text{minimize}} \sum_{l=1}^{L} \sum_{i=1}^{n} w_i J_{p_l}(y_i - s_l(x_i)) \\ &\quad + \lambda \sum_{l=1}^{L} \int_I |s_l''(x)|^2 \, \mathrm{d}x + \kappa \sum_{l=1}^{L-1} \int_I |s_{l+1}'(x) - s_l'(x)|^2 \, \mathrm{d}x \\ &\text{subject to} \quad \forall x\ s_{l+1}(x) \geq s_l(x) \quad (l = 1, 2, \ldots, L-1) \\ &\text{(and } \forall l\, \forall x\ s_l'(x) \geq 0,\ s_l'(x) \leq 0,\ s_l''(x) \geq 0, \text{ or } s_l''(x) \leq 0), \end{aligned} \tag{11}$$

where $w_i > 0$ and $\kappa > 0$. $s_l(x_i)$ is expressed as $\boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{c}_l$ with a certain vector $\boldsymbol{a}_i \in \mathbb{R}^{b(d+1)}$ and the coefficient vector $\boldsymbol{c}_l$ of s_l. The second term is expressed as $\sum_{l=1}^{L} \boldsymbol{c}_l^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{c}_l = \bar{\boldsymbol{c}}^{\mathrm{T}} \boldsymbol{Q}_1 \bar{\boldsymbol{c}}$ with $\bar{\boldsymbol{c}} := (\boldsymbol{c}_1^{\mathrm{T}}, \boldsymbol{c}_2^{\mathrm{T}}, \ldots, \boldsymbol{c}_L^{\mathrm{T}}) \in \mathbb{R}^{b(d+1)L}$. From (6) and

$$\begin{aligned} \int_{I_i} |s_{l+1}'(x) - s_l'(x)|^2 \, \mathrm{d}x &= \sum_{k=1}^{d} \sum_{j=1}^{d} \frac{kj}{h_i(k+j-1)} \\ &\cdot \left(c_{l+1,k}^{\langle i \rangle} c_{l+1,j}^{\langle i \rangle} - c_{l+1,k}^{\langle i \rangle} c_{l,j}^{\langle i \rangle} - c_{l,k}^{\langle i \rangle} c_{l+1,j}^{\langle i \rangle} + c_{l,k}^{\langle i \rangle} c_{l,j}^{\langle i \rangle} \right), \end{aligned} \tag{12}$$

the third term is expressed as $\bar{\boldsymbol{c}}^{\mathrm{T}} \boldsymbol{Q}_2 \bar{\boldsymbol{c}}$. From (9), a sufficient condition for the non-crossing constraint can be given by

$$\begin{aligned} &\sum_{k=0}^{j} \frac{(d-k)!}{(j-k)!(d-j)!} (c_{l+1,k}^{\langle i \rangle} - c_{l,k}^{\langle i \rangle}) \geq 0 \quad (j = 0, 1, \ldots, d) \\ &\qquad \Rightarrow s_{l+1}(x) \geq s_l(x) \text{ for all } x \in (\xi_{i-1}, \xi_i), \end{aligned} \tag{13}$$

and that for the non-decreasing property can be given by

$$\begin{aligned} &\sum_{k=0}^{j} \frac{(d-k-1)!(k+1)}{(j-k)!(d-j-1)!} c_{l,k+1}^{\langle i \rangle} \geq 0 \quad (j = 0, 1, \ldots, d-1) \\ &\qquad \Rightarrow s_l'(x) \geq 0 \text{ for all } x \in (\xi_{i-1}, \xi_i). \end{aligned} \tag{14}$$

Sufficient conditions for the other properties can be obtained in similar manners. From (6), (7), (8), (12), (13) and (14), the problem of (11) is reduced to the following problem

$$\begin{aligned} &\underset{\bar{\boldsymbol{c}} \in \mathbb{R}^{b(d+1)L}}{\text{minimize}} \sum_{l=1}^{L} \sum_{i=1}^{n} w_i J_{p_l}(y_i - \boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{c}_l) + \lambda \bar{\boldsymbol{c}}^{\mathrm{T}} \boldsymbol{Q}_1 \bar{\boldsymbol{c}} + \kappa \bar{\boldsymbol{c}}^{\mathrm{T}} \boldsymbol{Q}_2 \bar{\boldsymbol{c}} \\ &\text{subject to} \qquad \forall l\ \boldsymbol{H} \boldsymbol{c}_l = \boldsymbol{0} \ \text{ and } \ \boldsymbol{G}_1 \bar{\boldsymbol{c}} \geq \boldsymbol{0} \\ &\text{(and} \quad \forall l\ \boldsymbol{G}_2 \boldsymbol{c}_l \geq \boldsymbol{0},\ \boldsymbol{G}_2 \boldsymbol{c}_l \leq \boldsymbol{0},\ \boldsymbol{G}_3 \boldsymbol{c}_l \geq \boldsymbol{0}, \text{ or } \boldsymbol{G}_3 \boldsymbol{c}_l \leq \boldsymbol{0}). \end{aligned}$$

The optimal solution of this problem is obtained, e.g., by the alternating direction method of multipliers (ADMM) [38].

4 Numerical Experiments

We estimate the $p_l = \frac{l}{4}$-th ($l = 1, 2, 3$) quantiles $q_{p_l,Y}(x)$ of

$$f_{Y|X}(y|x) := \frac{1}{\sqrt{2\pi} \hat{\sigma}(x) y} e^{-\frac{(\log y - \hat{\mu}(x))^2}{2(\hat{\sigma}(x))^2}} \qquad (y > 0)$$

by the method in (10) and the proposed one in (11), where

$$\begin{cases} \hat{\mu}(x) := \begin{cases} -0.5(x+0.15)^2 + 1 & \text{if } x \in (-\infty, -0.15], \\ 1 & \text{if } x \in (-0.15, 0.15], \\ 0.75(x-0.15)^2 + 1 & \text{if } x \in (0.15, \infty), \end{cases} \\ \hat{\sigma}(x) := \begin{cases} 0.5 & \text{if } x \in (-\infty, 0.15], \\ 0.5(x-0.15)^2 + 0.5 & \text{if } x \in (0.15, \infty), \end{cases} \end{cases}$$

and we define the probability density function of X by

$$f_X(x) := \frac{0.3}{\sqrt{2\pi} \sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \frac{0.7}{\sqrt{2\pi} \sigma_2} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$$

with $(\mu_1, \sigma_1) := (-0.4, 0.2)$ and $(\mu_2, \sigma_2) := (0.2, 0.25)$. We set $n = 1000$, $d = 5$, $\rho = 2$, $b = 40$, $\xi_0 = -1$, $\xi_{40} = 1$, and $h_i = \frac{1}{20}$. Figures 1 and 2 show the results by (10) with $\lambda = \frac{1}{40}$ and by (11) with $w_i = \sqrt[4]{f_X(x_i)}$, $(\lambda, \kappa) = (\frac{1}{20}, \frac{1}{100})$ and the non-decreasing property. Black lines are the true monotone quantiles, blue circles are observations, green lines are the results by (10), and red ones are the results by (11). We find that, differently from (10), the proposed method in (11) can reconstruct the monotone and harmonious quantiles.

5 Conclusion

In this paper, we have proposed a novel spline smoothing technique for quantile regression. Differently from the other methods, we constructed multilevel quantile lines simultaneously by utilizing the similarity between the adjacent quantile lines. We also considered the non-crossing constraint and an optional (non-decreasing/non-increasing/convex/concave) property. Numerical experiments showed that the proposed method can construct harmonious multilevel quantile lines.

Acknowledgment

This work was supported in part by JSPS Grants-in-Aid for Early-Career Scientists (JP19K20361). The authors are grateful to Dr. Atsunori Kashiwagi and Dr. Osamu Sekine in Kusatsu General Hospital for very fruitful discussion.

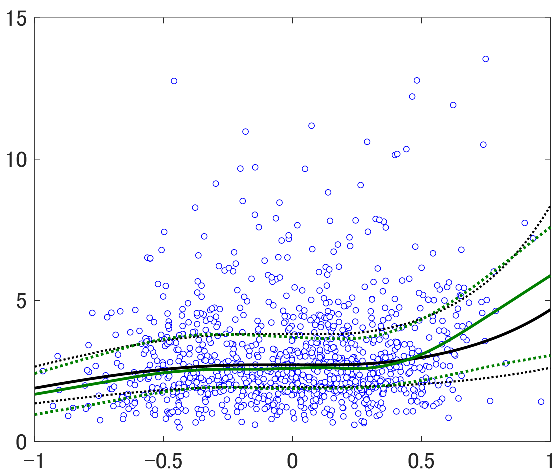

Figure 1: Quantile Regression Results by (10)

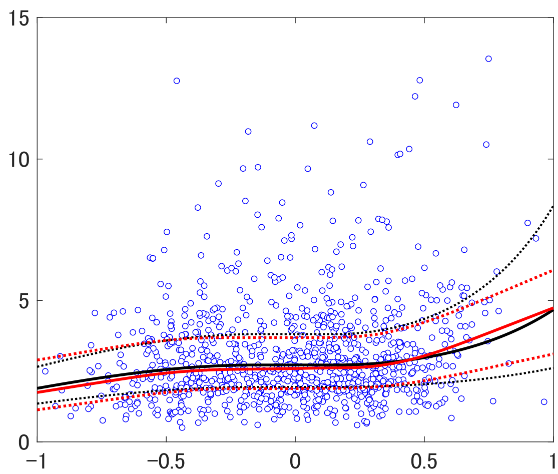

Figure 2: Quantile Regression Results by (11)

References

[1] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., ser. Springer Series in Statistics. New York, NY, USA: Springer, 2009.

[2] J. Shao, *Mathematical Statistics*, 2nd ed., ser. Springer Texts in Statistics. New York, NY, USA: Springer, 2003.

[3] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed., ser. Springer Texts in Statistics. New York, NY, USA: Springer, 2008.

[4] R. R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing*, 4th ed., ser. Statistical Modeling and Decision Science. London, UK: Academic Press, 2016.

[5] B. Abraham and J. Ledolter, *Introduction to Regression Modeling*, ser. Duxbury Advanced Series. Belmont, CA, USA: Duxbury, 2005.

[6] J. A. Rice, *Mathematical Statistics and Data Analysis*, 3rd ed., ser. Duxbury Advanced Series. Belmont, CA, USA: Duxbury, 2006.

[7] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 5th ed., ser. Wiley Series in Probability and Statistics. New York, NY, USA: Wiley, 2013.

[8] P. J. Huber and E. M. Ronchetti, *Robust Statistics*, 2nd ed., ser. Wiley Series in Probability and Statistics. New York, NY, USA: Wiley, 2009.

[9] R. A. Maronna, R. D. Martin, V. J. Yohai, and M. Salibián-Barrera, *Robust Statistics: Theory and Methods (with R)*, 2nd ed., ser. Wiley Series in Probability and Statistics. New York, NY, USA: Wiley, 2019.

[10] G. Bassett Jr. and R. Koenker, "Asymptotic theory of least absolute error regression," *Journal of the American Statistical Association*, vol. 73, no. 363, pp. 618–622, 1978.

[11] P. Bloomfield and W. L. Steiger, *Least Absolute Deviations: Theory, Applications, and Algorithms*, ser. Progress in Probability and Statistics, P. Huber and M. Rosenblatt Eds. Boston, MA, USA: Birkhäuser, 1983, vol. 6.

[12] R. Koenker and G. Bassett Jr., "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.

[13] R. Koenker and K. F. Hallock, "Quantile regression," *Journal of Economic Perspectives*, vol. 15, no. 4, pp. 143–156, 2001.

[14] R. Koenker, *Quantile regression*. New York, NY, USA: Cambridge University Press, 2005.

[15] R. Koenker, V. Chernozhukov, X. He, and L. Peng, *Handbook of Quantile Regression*, ser. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. New York, NY, USA: Chapman & Hall, 2017.

[16] C. de Boor, "Best approximation properties of spline functions of odd degree," *Journal of Mathematics and Mechanics*, vol. 12, no. 5, pp. 747–749, 1963.

[17] S. Wold, "Spline functions in data analysis," *Technometrics*, vol. 16, no. 1, pp. 87–111, 1974.

[18] B. W. Silverman, "Some aspects of the spline smoothing approach to non-parametric regression curve fitting," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 47, no. 1, pp. 1–52, 1985.

[19] G. Wahba, *Spline Models for Observational Data*, ser. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia, PA, USA: SIAM, 1990, vol. 59.

[20] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, 1999.

[21] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, 2nd ed. New York, NY, USA: Springer, 2005.

[22] R. J. Bosch, Y. Ye, and G. G Woodworth, "A convergent algorithm for quantile regression with smoothing splines," *Computational Statistics & Data Analysis*, vol. 19, no. 6, pp. 613–630, 1995.

[23] X. He, "Quantile curves without crossing," *The American Statistician*, vol. 51, no. 2, pp. 186–192, 1997.

[24] I. Takeuchi, Q. V. Le, T. Sears, and A. J. Smola, "Nonparametric quantile regression," *Journal of Machine Learning Research*, vol. 7, pp. 1231–1264, 2006.

[25] V. M. R. Muggeo, M. Sciandra, A. Tomasello, and S. Calvo, "Estimating growth charts via nonparametric quantile regression: a practical framework with application in ecology," *Environmental and Ecological Statistics*, vol. 20, no. 4, pp. 519–531, 2013.

[26] Y. Yuan, N. Chen, and S. Zhou, "Modeling regression quantile process using monotone B-splines," *Technometrics*, vol. 59, no. 3, pp. 338–350, 2017.

[27] X. He and P. Shi, "Monotone B-spline smoothing," *Journal of the American Statistical Association*, vol. 93, no. 442, pp. 643–650, 1998.

[28] R. Koenker and P. Ng, "Inequality constrained quantile regression," *The Indian Journal of Statistics*, vol. 67, no. 2, pp. 418–440, 2005.

[29] K. Bollaerts, P. H. C. Eilers, and M. Aerts, "Quantile regression with monotonicity restrictions using P-splines and the L_1-norm," *Statistical Modelling*, vol. 6, pp. 189–207, 2006.

[30] T. Neocleous and S. Portnoy, "On monotonicity of regression quantile functions," *Statistics & Probability Letters*, vol. 78, no. 10, pp. 1226–1229, 2008.

[31] G. R. Shorack, *Probability for Statistics*, 2nd ed., ser. Springer Texts in Statistics. New York, NY, USA: Springer, 2017.

[32] R. J. Casady and J. D. Cryer, "Monotone percentile regression," *The Annals of Statistics*, vol. 4, no. 3, pp. 532–541, 1976.

[33] D. Griffiths and M. Willcox, "Percentile regression: A parametric approach," *Journal of the American Statistical Association*, vol. 73, no. 363, pp. 496–498, 1978.

[34] T. J. Cole, "Fitting smoothed centile curves to reference data," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 151, no. 3, pp. 385–418, 1988.

[35] D. Kitahara and I. Yamada, "Probability density function estimation by positive quartic C^2-spline functions," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, Australia, 2015, pp. 3556–3560.

[36] D. Kitahara and I. Yamada, "Two-dimensional positive spline smoothing and its application to probability density estimation," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, 2016, pp. 4219–4223.

[37] W. Heß and J. W. Schmidt, "Positive quartic, monotone quintic C^2-spline interpolation in one and two dimensions," *Journal of Computational and Applied Mathematics*, vol. 55, no. 1, pp. 51–67, 1994.

[38] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.