Quantile Regression with Monotonicity Restrictions using P-splines and the L_1-norm

Kaatje Bollaerts

Center for Statistics, University Hasselt, Belgium

Paul H.C. Eilers

Leids University Medical Center, the Netherlands

Marc Aerts

Center for Statistics, University Hasselt, Belgium

Author notes: All correspondence concerning this paper should be sent to Kaatje Bollaerts, University Hasselt, Center for Statistics, Agoralaan 1 gebouw D, B-3590 Diepenbeek (Belgium).

E-mail: kaatje.bollaerts@uhasselt.be

Abstract

Quantile regression is an alternative to OLS-regression. In quantile regression, the sum of absolute deviations or the L_1-norm is minimized whereas the sum of squared deviations or the L_2-norm is minimized in OLS-regression. Quantile regression has the advantage over OLS-regression of being more robust to outlying observations. Furthermore, quantile regression provides information complementing the information provided by OLS-regression. In this study, a non-parametric approach to quantile regression is presented which constrains the estimated quantile function to be monotone increasing. In particular, P-splines with an additional asymmetric penalty enforcing monotonicity are used within a L_1-framework. This can be translated into a linear programming problem, which will be solved using an interior point algorithm. As an illustration, the presented approach will be applied to estimate quantile growth curves and to estimate quantile antibody levels as a function of age.

Keywords: Growth curves, Interior point, L_1-norm, Monotonicity, P-splines, Quantile regression

1 Introduction

In quantile regression as introduced by Koenker and Bassett (1978), conditional quantile functions are estimated by minimizing a (weighted) sum of absolute deviations or the L_1-norm. The latter can be considered as an alternative to classical least squares regression. Here, conditional mean functions are estimated by minimizing the sum of squared deviations or the L_2-norm. Compared to conditional mean functions, conditional quantile functions have the important advantage of being more robust to outlying observations in the response variable. Furthermore, a full range of different quantile functions can be fitted to the data, complementing the conditional mean function, and as such, providing a nuanced picture of the stochastic relationship between variables. Nevertheless and in spite of its multiple merits, quantile regression is far less used than classical least squares regression. Similarly, most non-parametric methods are adopted within a least squares framework. However, some authors, among others, Chaudhuri (1991), Welsh (1996) and Koenker (1994) investigated the applicability of non-parametric quantile regression. Eilers (2000) adopted L_1-estimation to P-splines regression models with the latter models being introduced within a L_2-setting in a target article by Eilers and Marx (1996). In the present paper, L_1-estimation of P-splines regression models with monotonicity restrictions is explored (for L_2-estimation, see Bollaerts et al., 2005).

The presented method can be naturally used in many application settings. For example, consider the estimation of pediatric growth curves, more specifically, the estimation of the lower and upper quantile reference growth curves for children's height and weight as a function of age. Note that, in this case, it is more plausible to assume that children's height (weight) is a monotone increasing function of age without presupposing any parametric relationship rather than assuming for instance a kwadratic or logarithmic function. As a second example, consider the estimation of longitudinal wage trends, which are commonly assumed to be monotonely increasing. In this respect, it is interesting to investigate whether wage discrimination is decreasing respectively increasing (that is, whether wage trends within different income categories are converging respectively diverging). Clearly, many applications of quantile regression with monotonicity restrictions exist.

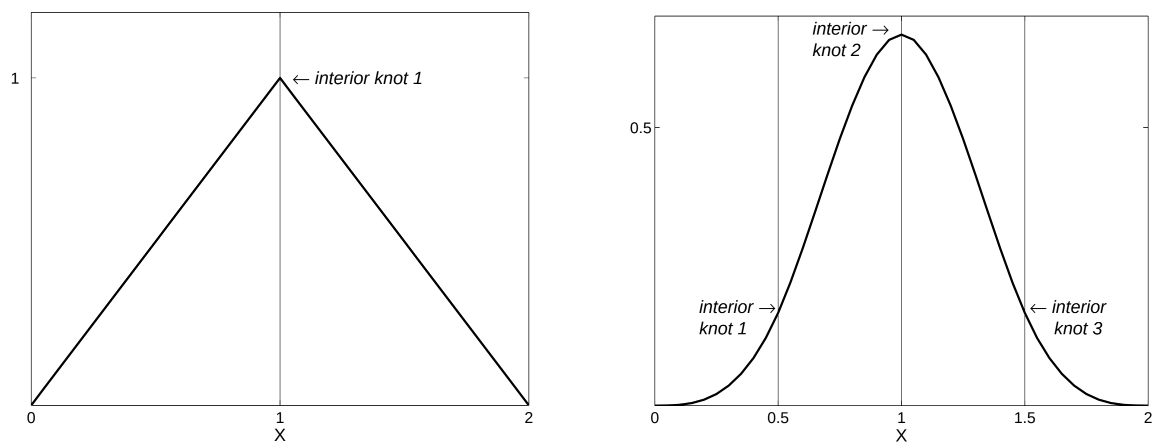
The remainder of this paper is organised as follows: In Section 2, unconstrained respectively monotonicity constrained quantile regression is introduced. In Section 3, computational issues are discussed. In Section 4, the presented method is applied to estimate growth curves regarding height and weight of boys. For this application, different number of basis functions are used and the corresponding estimated curves compared. In Section 5, an application on the estimation of Varicella-Zoster Virus antibody levels is given. For this application, a comparison is made between non-parametric quantile regression with and without monotonicity constraints. Finally, some concluding remarks and suggestions for further research are given in Section 6.

2 Method

2.1 Quantile regression using P-splines

In a target article, Eilers and Marx (1996) introduced P-splines regression for one-dimensional data within a least squares framework. This is essentially least squares regression with an excessive number of univariate B-splines (De Boor, 1978; Dierckx, 1993) and a additional discrete penalty to correct for overfitting. In the following, univariate B-splines are introduced. Then, we discuss regression with B-splines first before discussing regression with P-splines. Each time, both least squares and quantile regression are discussed.

Univariate B-splines are piecewise polynomial functions with local support. A B-spline of degree q consists of $q+1$ polynomial pieces of degree q joined smoothly (i.e., differentiable



(a) first degree

(b) third degree

Figure 1: Single B-splines of first and third degree.

in order $q-1$) at q points t_i (called interior knots) between boundaries t_{min} and t_{max} (called exterior knots) and with a positive value between these boundaries and a value of zero outside these boundaries. In Figure 1a, an example of a B-spline of the first degree is given; it is clear that this B-spline consists of two linear pieces joined at one interior knot. Figure 1b gives an example of a B-spline of the third degree (which is the most commonly used degree in B-splines regression); this B-spline consists of four cubic pieces joined smoothly at three interior knots.

In order to use B-splines for non-parametric regression, a basis of r overlapping B-splines is constructed, which is such that

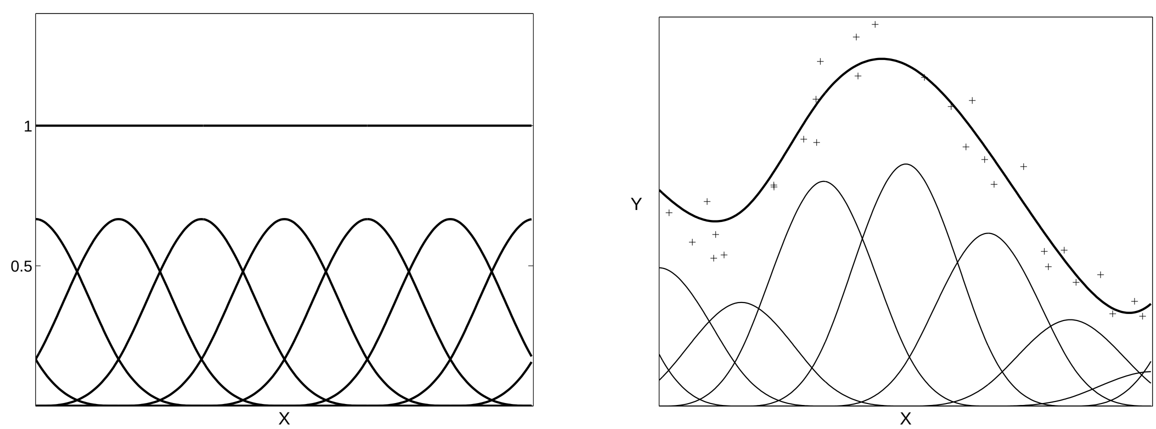
$$\forall x : \sum_{j=1}^{r} B_j(x, q) = 1 \tag{2.1}$$

with $B_j(x, q)$ denoting a B-spline of degree q with left most knot j. In Figure 2a, one can see an example of a basis of B-splines of the third degree. Then, the B-splines of a B-splines basis act as predictors in spline regression. With, for m observations (x_i, y_i),

$$\widehat{y}_{(\alpha)_i} = \sum_{j=1}^{r} \alpha_j B_j(x_i, q), \text{ i=1,... m} \tag{2.2}$$

and with α_j being the coefficient of the corresponding B-spline. The vector $\boldsymbol{\alpha}$ is commonly estimated using the L_2-norm or

$$S_2 = \sum_{i=1}^{m} (y_i - \widehat{y}_{(\alpha)_i})^2 \tag{2.3}$$



(a) B-splines basis

(b) spline regression

Figure 2: Spline regression with B-splines of third degree.

with the conditional mean function being the minimand. However, following Koenker and Bassett (1978), one might also consider estimating $\boldsymbol{\alpha}$ by minimizing the asymmetric L_1-norm defined as

$$S_1 = \sum_{i=1}^{m} \rho_\theta(y_i - \widehat{y}_{(\alpha)_i}) \tag{2.4}$$

with ρ_θ being the check function

$$\rho_\theta(\tau) = \begin{cases} \theta\tau & \text{if } \tau \geq 0 \\ (\theta - 1)\tau & \text{otherwise} \end{cases}$$

This loss function has the conditional quantile θ as minimand.

A major problem in B-splines regression, irrespective of the norm used, is the choice of the optimal number of B-splines. An insufficient number of B-splines leads to underfitting, whereas too many B-splines leads to overfitting. To regularize smoothness, Eilers and Marx (1996) propose to use an excessive number of equally spaced B-splines with, in order to correct for overfitting, a smoothness penalty based on differences of the coefficients of adjacent B-splines. They call this approach P-splines regression. The corresponding loss functions based on the L_2-norm and the asymmetric L_1-norm are respectively equal to

$$S_2 = \sum_{i=1}^{m} (y_i - \widehat{y}_{(\alpha)_i})^2 + \lambda \sum_{j=d+1}^{r} (\triangle^d \alpha_j)^2 \tag{2.5}$$

and

$$S_1 = \sum_{i=1}^{m} \rho_\theta(y_i - \widehat{y}_{(\alpha)_i}) + \lambda \sum_{j=d+1}^{r} |\triangle^d \alpha_j| \tag{2.6}$$

with $\triangle^d \alpha_j$ being the d^{th} order differences, that is $\triangle^d \alpha_j = \triangle^1(\triangle^{d-1}\alpha_j)$ with $\triangle^1 \alpha_j = \alpha_j - \alpha_{j-1}$ and with λ being a user-defined smoothness parameter, which can be optimized using for instance cross-validation. Mostly, a penalty on second order differences is used. However, lower or higher order penalties can be used equally well with a penalty on first order differences yielding a piecewise linear fit whereas penalties on higher order differences yield more smooth fits. In general, if $\lambda \to \text{inf}$, then, for a regression with a smoothness penalty on d^{th} order differences, the fitted function will approach a polynomial of degree $n - 1$.

2.2 Quantile regression with monotonicity restrictions using P-splines

The P-splines approach discussed so far can be easily adapted in order to impose monotonicity. This can be done by adding an additional asymmetric penalty on the first order differences. Indeed, the first order derivative of a B-splines function with equally spaced knots equals

$$f'(\alpha, x) = \frac{\partial f(\alpha, x)}{\partial x} = \frac{\partial}{\partial x} \sum_{j=1}^{r} \alpha_j B_j(x, q) = h^{-1} \sum_{j=1}^{r+1} \triangle^1 \alpha_j B_j(x, q-1) \tag{2.7}$$

with h denoting the distance between two adjacent knots (De Boor, 1978). Then, since q, h and $B_j(x, q-1)$ are all positive by definition, restricting $\triangle^1 \alpha_j$ to be positive (resp. negative) is a sufficient condition for $f'(\alpha, x)$ to be positive (resp. negative). Hence, the L_2 loss function imposing monotonicity reads as

$$S_2 = \sum_{i=1}^{m} (y_i - \widehat{y}_{(\alpha)_i})^2 + \lambda \sum_{j=3}^{r} (\triangle^2 \alpha_j)^2 + \kappa \sum_{j=2}^{r} w_{(\alpha_j)} (\triangle^1 \alpha_j)^2 \tag{2.8}$$

with

$$w_{(\alpha_j)} = \begin{cases} 0 & \text{if } \triangle^1 \alpha_j \geq 0 \text{ (resp. } \triangle^1 \alpha_j \leq 0) \\ 1 & \text{otherwise} \end{cases}$$

being asymmetric weights and with κ being a user-defined constraint parameter by which the strength of the constraint can be fine-tuned. In particular, for $\kappa = 0$, monotonicity is not imposed whereas for $\kappa \to \infty$ violations against monotonicity are negligible. The corresponding asymmetric L_1 loss function imposing monotonicity equals

$$S_1 = \sum_{i=1}^{m} \rho_\theta(y_i - \widehat{y}_{(\alpha)_i}) + \lambda \sum_{j=d+1}^{r} |\triangle^d \alpha_j| + \kappa \sum_{j=2}^{r} \rho_\omega(\triangle^1 \alpha_j) \tag{2.9}$$

with ρ_ω being defined as

$$\rho_\omega(\tau) = \begin{cases} \omega\tau & \text{if } \tau \geq 0 \text{ (resp. } \tau \leq 0) \\ (\omega - 1)\tau & \text{otherwise} \end{cases}$$

and with $\omega \to 1$.

3 Computation

Solving an L_1-regression problem relies on reformulating the corresponding loss function as a linear programming problem. Consider the loss function regarding quantile regression with B-splines as defined in (2.4). Reformulating this loss function as a linear programming problem gives

$$\begin{aligned} \text{min!} \quad & \theta \sum_{i=1}^{m} u_i + (1-\theta) \sum_{i=1}^{m} v_i \\ \text{s.t.} \quad & \sum_{i=1}^{m} \widehat{y}_{(\alpha)_i} + u_i - v_i = y_i \end{aligned} \tag{3.1}$$

with $u_i \geq 0$ and $v_i \geq 0$. For unconstrained quantile regression with B-splines in combination with a smoothness penalty (shortly, P-splines), the loss function is defined in (2.9). In this case, the linear programming problem equals

$$\begin{aligned} \text{min!} \quad & \theta \sum_{i=1}^{m} u_i + (1-\theta) \sum_{i=1}^{m} v_i + \lambda \sum_{j=d+1}^{r} s_j + \lambda \sum_{j=d+1}^{r} t_j \\ \text{s.t.} \quad & \sum_{i=1}^{m} \widehat{y}_{(\alpha)_i} + u_i - v_i = y_i \\ \text{and} \quad & \sum_{j=d+1}^{r} (\triangle^d \alpha_j) + s_j - t_j = 0 \end{aligned} \tag{3.2}$$

with $u_i \geq 0$, $v_i \geq 0$, $s_j \geq 0$ and $t_j \geq 0$ (Eilers, 2000). Further extensions are needed for quantile regression using B-splines, a smoothness penalty and a penalty enforcing monotoncity, of which de loss function is defined in (6). In this case, the linear programming problem is

$$\begin{aligned} \text{min!} \quad & \theta \sum_{i=1}^{m} u_i + (1-\theta) \sum_{i=1}^{m} v_i + \lambda \sum_{j=d+1}^{r} s_j + \lambda \sum_{j=d+1}^{r} t_j + \kappa\omega \sum_{j=2}^{r} p_j + \kappa(1-\omega) \sum_{j=2}^{r} q_j \\ \text{s.t.} \quad & \sum_{i=1}^{m} \widehat{y}_{(\alpha)_i} + u_i - v_i = y_i \\ \text{and} \quad & \sum_{j=d+1}^{r} (\triangle^d \alpha_j) + s_j - t_j = 0 \\ \text{and} \quad & \sum_{j=2}^{r} (\triangle^1 \alpha_j) + p_j - q_j = 0 \end{aligned} \tag{3.3}$$

with $u_i \geq 0$, $v_i \geq 0$, $s_j \geq 0$, $t_j \geq 0$, $p_j \geq 0$ and $q_j \geq 0$. In order to solve these L_1-related linear programming problems, we adopt the approach proposed by Portnoy and Koenker (1997).

They proposed, as an alternative to simplex based methods, interior point optimization in combination with statistical preprocessing for L_1-type of problems. The latter approach has the advantage over simplex based methods of being computationally less demanding, especially in large datasets. Portnoy and Koenker (1997) even showed that their approach is comparably as fast as classical L_2-methods of estimation, irrespective of the probem size. Matlab code to solve L_1-type of problems using Portnoy and Koenker's approach as well as quantile regression sofware in R can be found on Koenker's home-page at the University of Illinois (http://www.econ.uiuc.edu/ roger).

Table 1 *Optimal values for smoothness parameter λ based on Generalized Cross-Validation.*

	height			weight		
quantile	B = 5	B = 15	B = 25	B = 5	B = 15	B = 25
0.05	0.07	0.21	18.39	0.07	1.93	21.80
0.50	2.82	0.61	27.50	0.02	9.60	68.75
0.95	0.36	4.30	18.42	0.24	7.22	28.26

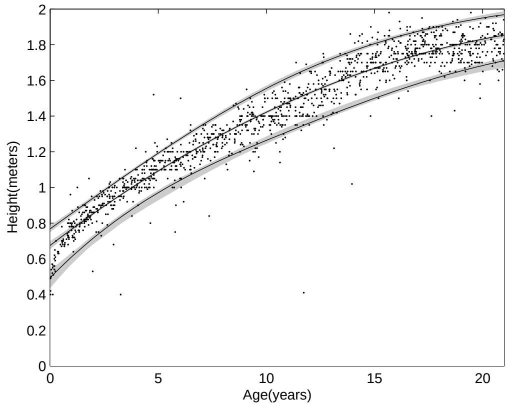
4 Application: growth curves

In this section, quantiles of growth curves are estimated using P-splines regression with monotonicity constraints to avoid the estimation of decreasing trends in growth but still allow maximal flexibility. The data come from the Belgian Health Interview Survey, conducted by the National Institute for Statistics (NIS) in 2001. For this application, only data referring to male respondents aged between 0 and 21 years are considered. The number of observations equals $N = 1295$. In order to provide a nuanced picture of growth, isotone P-splines regression will be used to estimate quantiles of boys' height and weight as a function of age. In particular, for both height and weight the conditional quantile functions 0.05, 0.50 and 0.95 are estimated using a basis of 5, 15 and 25 B-splines of third degree. The optimal values for the smoothness parameters λ are determined using Generalized Cross Validation. The values of the monotonicity parameter κ are chosen as high as 10^{12} to ensure that violations against monotonicity are negligible. The variability of the estimated quantile functions $f(\hat{\alpha}, x)$ is assessed by means of pointwise bootstrap confidence intervals. For each quantile function, $B = 1000$ bootstrap samples are generated by resampling the original data (x_i, y_i) with replacement. For each bootstrap sample, which contains $N = 1295$ observations (x_i^*, y_i^*), the corresponding quantile function is estimated. This leads to $B = 1000$ different estimates $f^*(\hat{\alpha}, x)$. In order to estimate the $100(1-2\alpha)\%$ pointwise confidence interval for $f(\hat{\alpha}, x)$, percentile intervals are calculated conditional on $X = x$. The latter are defined as $[f^*(\hat{\alpha}, x)_{[(B+1)\alpha]}; f^*(\hat{\alpha}, x)_{[(B+1)(1-\alpha)]}]$ with $f^*(\hat{\alpha}, x)_{[(B+1)\alpha]}$ being the $[(B+1)\alpha]^{th}$ order statistic of $f^*(\hat{\alpha}, x)$.

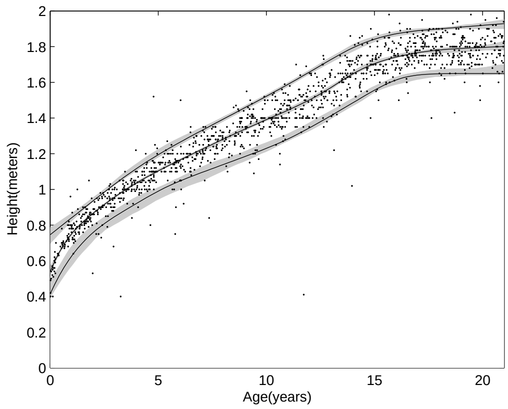
The optimal values for the corresponding smoothness parameters are given in Table 1.

As can be seen, the larger the number of B-splines used, the higher the optimal value of the smoothness parameter. Only one exception is observed when estimating the median of height using a basis of respectively 5 and 15 B-splines which might be due to the fact that using only 5 B-splines is far from enough. However, it is a general result that the number of basis functions used (in this case, B-splines) influences smoothness with less basis functions yielding a smoother fit. Hence, if the number of basis functions used is small, a small weight of the smoothness penalty λ suffices to yield a smooth result. Second, it can be seen that, for the data at hand, the optimal value of the smoothness parameter for the median is larger compared to the ones for quantile 0.05 and 0.95. This might follow from the fact that, for this data, there is much more variability at quantile 0.05 and 0.95 compared to the median.

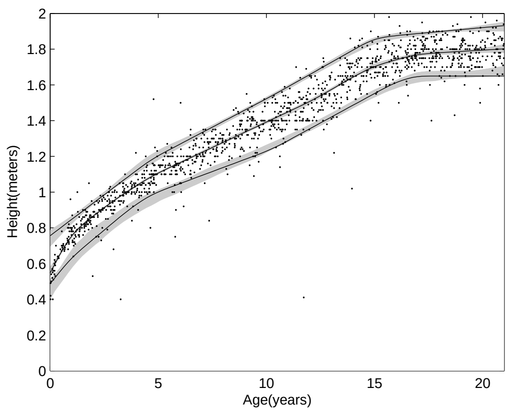
In Figure 3 (resp. Figure 4), the isotone quantile functions 0.05, 0.50 and 0.95 of height (resp. weight) are shown together with their 95% bootstrap confidence intervals for a basis of 5, 15 and 25 B-splines of third degree. Clearly, using a basis of only 5 B-splines is not enough since large and systematic deviations between the data and the fitted curve are still observed. The fitted curves for a basis of 15 respectively 25 B-splines, which are very similar, describe the data well. The latter result illustrates the rationale behind P-splines: take an excessive number of B-splines and correct for overfitting using the smoothness penalty, for which the weight can be optimally chosen using for instance Generalized Cross-Validation. Hence, whenever more than enough B-splines are used, the smoothness penalty will 'automatically' correct for overfitting, regardless of the number of B-splines used. Second, as indicated by the width of the confidence intervals, there is more variability in estimating the quantile function 0.95 and 0.05 compared to the median for both height and weight. This result is in accordance with the previous finding that the optimal value of the smoothness parameter for the median is larger compared to the ones for quantile 0.05 and 0.95. Finally, comparing Figure 3 and 4, it can be seen that the curves indicate that height is symmetrically distributed conditional on age whereas weight is clearly not. Indeed, for weight, the absolute deviation between the quantile functions 0.95 and 0.50 is larger than the absolute deviation between the quantile functions 0.05 and 0.50.



(a) B = 5

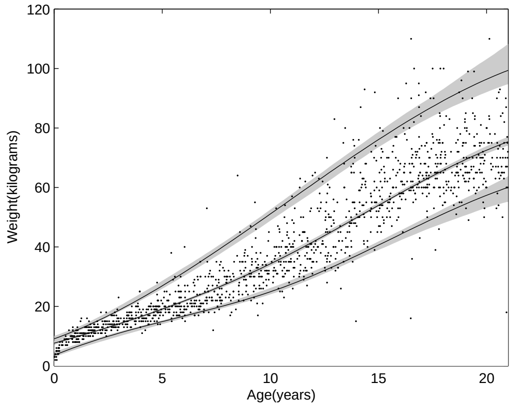


(b) B = 15

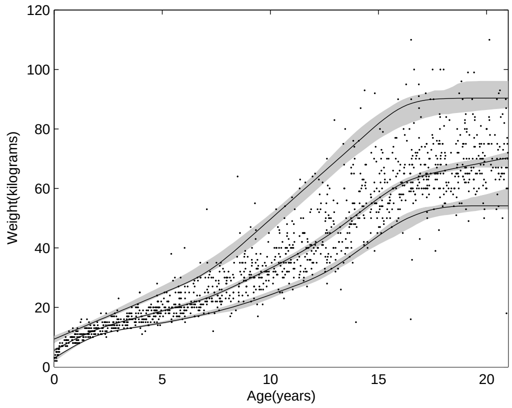


(c) B = 25

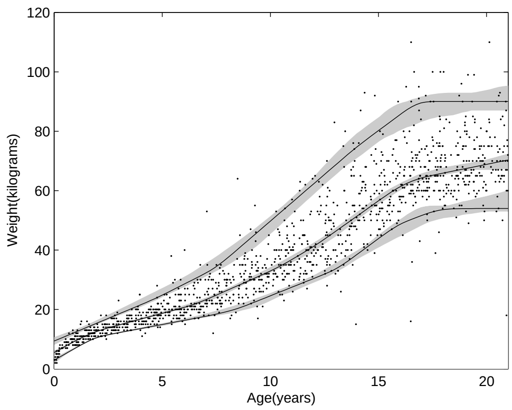
Figure 3: Height: isotone quantiles 0.05, 0.50 and 0.95 as function of age.



(a) B = 5



(b) B = 15



(c) B = 25

Figure 4: Weight: isotone quantiles 0.05, 0.50 and 0.95 as function of age.

5 Application: antibody levels

In this section, quantile functions of Varicella-Zoster virus (VZV) antibody levels are estimated. The data, collected at the Centre for the Evaluation of Vaccination, Antwerp University, contain the antibody levels (October 1999- April 2000) of 1673 sera from a sample of the Flemish (Belgian) population. (Thiry et al, 2002)

Given the assumption of lifelong immunity, it is reasonable to assume a monotone increasing trend of the mean/median antibody level as a function of age. In particular, the quantiles 0.05, 0.25, 0.50, 0.75 and 0.95 are estimated as a function of age. For this application, the performance of unconstrained and monotonicity constrained quantile regression using P-splines will be compared graphically. Each time, a B-splines basis of 15 B-splines is used. Again, the optimal values for the smoothness parameters λ are determined using Generalized Cross Validation and the values of the monotonicity parameter κ are chosen as high as 10^{12}. The variability of the estimated quantile functions $f(\hat{\alpha}, x)$ is assessed by means of pointwise bootstrap confidence intervals as described in Section 4.

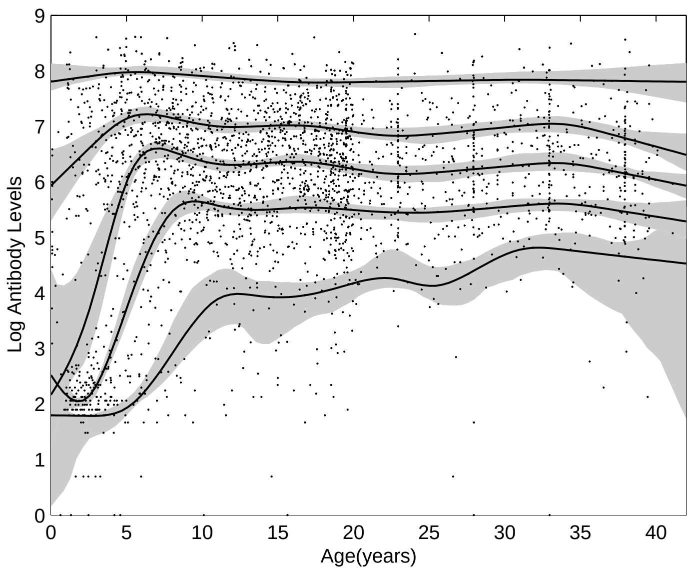
The optimal values for the smoothness parameters λ are shown in Table 2. As can be seen, the optimal value of the smoothness parameter is smaller when imposing the constraint of monotonicity compared to not imposing the constraint. This makes perfect sense since imposing constraints by adding an additional penalty is also a way of smoothing.

In Figure 5, the estimated unconstrained and monotonicity constrained quantile functions are shown together with their corresponding 95% pointwise confidence intervals. Furthermore, when looking at the pointwise confidence intervals, it can be seen that the confidence

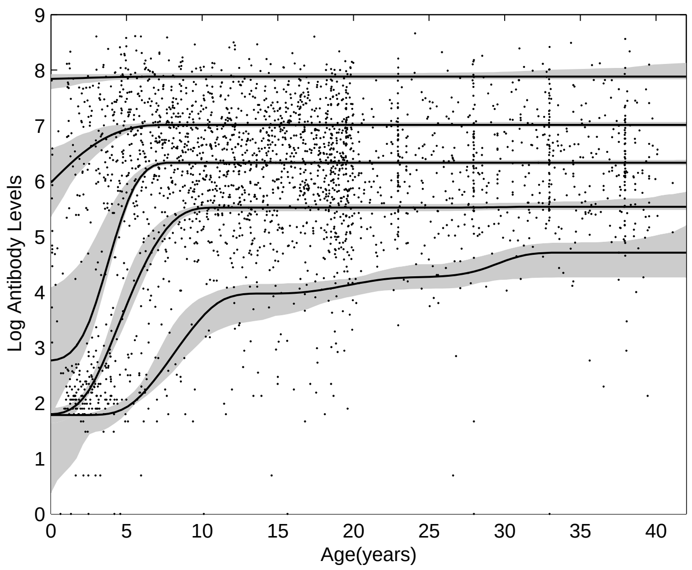
Table 2 *Optimal values for smoothness parameter λ based on Generalized Cross-Validation.*

quantile	unconstrained	constrained
0.05	0.37	0.33
0.25	10.04	0.32
0.50	20.66	6.59
0.75	10.99	2.8
0.95	37.89	10.27

intervals are wider for the unconstrained quantiles compared to their corresponding constrained quantiles. This is as expected, since as said before, adding an additional penalty enforcing monotonicity is a way of smoothing. Finally, it can be seen that the confidence intervals are wider at the boundaries of the range of data. This is as expected as well because less information is available to estimate the amplitudes of the basis functions located at the boundaries. This results from the fact that the latter basis functions substantially exceed the actual range of data. In addition, for the data at hand, there are only few observations for the very young and very old ages, yielding wide confidence intervals as well.



(a) unconstrained



(b) monotonicity constrained

Figure 5: VZV antibody levels: unconstrained and monotonicity constrained quantiles 0.05, 0.25, 0.50, 0.75 and 0.95 as function of age.

6 Discussion

In this paper, monotonicity constrained quantile regression using P-splines is presented. This non-parametric method allows flexible estimation of quantile curves subject to the constraint of monotonicity. This method is suited for the estimation of, among others, growth curves. Indeed, adopting a fully parametric approach would be too restrictive regarding the shape of the fitted quantile curve. On the other hand, adopting an unconstrained non-parametric approach would be too flexible in the sense that the latter approaches allow the 'unnatural' estimation of decreasing trends in growth. The method proposed in this paper can be extended in order to (1) impose the constraint of convexity/concavity and to (2) fit data in two (or more) dimensions. The first extension is achieved by using second order differences in the constraint penalty. Remember that restricting the first order differences of the coefficients of a B-splines function to be positive (resp. negative) is a sufficient condition for the first order derivative of the B-splines function to be positive (resp. negative). By induction, one can prove that restricting the second order differences to be positive (resp. negative) is a sufficient condition for the second order derivative of the B-splines function to be positive (resp. negative), with the latter implying convexity (resp. concavity). The second extension is achieved by using n-variate B-splines (with a n-variate B-spline being the tensorproduct of an univariate and a $(n-1)$-variate B-spline) in combination with penalties for each dimension. An extensive discussion of P-splines models for two-dimensional data within a L_2-framework can be found in Bollaerts et al. (2005). Matlab code to fit P-splines models for two-dimensional data using the L_1-norm has been developed as well and can be obtained from the authors of the current paper.

Acknowledgements

We thank Pierre Van Damme and Nancy Thiry, Center for the Evaluation of Vaccination, Faculty of Medicine, University of Antwerp, for making the Varicella-Zoster antibody level data available to us. We would also like to mention the Scientific Institute of Public Health - Louis Pasteur in Brussels for access to the Belgian Health Interview Survey 1997 (http://www.iph.fgov.be). Data can be requested from the corresponding institutes. The first and third author gratefully acknowledge support from the IAP research network nr P5/24 of the Belgian Government (Belgian Science Policy).

References

Bollaerts, K., Eilers, P.H.C. & Van Mechelen, I. (2005). Simple and multiple P-splines regression with shape constraints. *Manuscript submitted for publication.*

Chaudhuri, P. (1991). Global nonparametric estimation of conditional quantile functions and their derivatives. *Journal of multivariate statistics, 39,* 246-269.

De Boor, C. (1978). *A practical guide to splines.* Berlin: Springer.

Dierckx, P. (1993). *Curve and surface fitting with splines.* Oxford: Clarendon.

Eilers, P.H.C. & Marx, B.D. (1996). Flexible smoothing using B-splines and penalized likelihood (with comments and rejoinder). *Statistical Science, 11(2)*,89-121.

Eilers, P.H.C. (2000). Robust and quantile smoothing with P-splines and the L_1-norm. *Proceedings of the 15th International Workshop on Statistical Modelling, Barcelona*, Nuñez-Anton V. & Ferreira, E., *eds.*

Koenker, R. & Bassett, G. (1978). Regression quantiles. *Econometrica, 46,* 33-50.

Koenker, R., Ng, P. & Portnoy, S. (1994). Quantile smoothing splines. *Biometrika, 81,* 676-680.

Portnoy, S. & Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: computability of squared-error vs. absolute-error estimators (with discussion). *Statistical Science, 12,* 279-296.

Thiry, N., Beutels, P., Shkedy, Z., Vranckx, R., Vandermeulen, C., Wielen, M.V. & Damme, P.V. (2002). The seroepidemiology of primary varicella-zoster virus infection in Flanders (Belgium). *European Journal of Pediatrics, 161(11),* 588-593.

Welsh, A.H. (1996). Robust estimation of smooth regression and spread functions and their derivatives. *Statistica Sinica, 6,* 347-366.