

Simultaneous estimation of quantile curves using quantile sheets

Sabine K. Schnabel · Paul H. C. Eilers

Received: 5 December 2011 / Accepted: 21 June 2012 / Published online: 20 July 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract The results of quantile smoothing often show crossing curves, in particular, for small data sets. We define a surface, called a quantile sheet, on the domain of the independent variable and the probability. Any desired quantile curve is obtained by evaluating the sheet for a fixed probability. This sheet is modeled by P -splines in form of tensor products of B -splines with difference penalties on the array of coefficients. The amount of smoothing is optimized by cross-validation. An application for reference growth curves for children is presented.

Keywords P -splines · Quantiles · Smoothing · Tensor product

1 Introduction

In the wake of quantile regression (Koenker 2005; Koenker and Bassett 1978) quantile smoothing has become a popular research subject. It is also an effective tool to study how the conditional distribution of a response variable changes with a covariate like time or age. Many proposals have been published and a range of software implementations are available for the popular R system, e.g. `quantreg` (Koenker 2011), `cobs` (Ng and Maechler 2011) or `VGAM` (Yee 2011).

S. K. Schnabel (✉)
Biometris, Wageningen University and Research Centre,
Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands
e-mail: sabine.schnabel@wur.nl

P. H. C. Eilers
Department of Biostatistics, Erasmus Medical Center,
Dr Molewaterplein 50-60, 3015 GE Rotterdam, The Netherlands
e-mail: p.eilers@erasmusmc.nl

In theory, conditional quantile curves cannot cross, but in practice they do, especially for small data sets. In many cases this is only a visual annoyance, but it may also jeopardize further analysis, e.g. when studying conditional distributions at specific values of the independent variable.

In the statistical literature one can find several proposals to prevent crossing of quantile curves. Especially in recent years this problem has received considerable attention. Among recent publications on the topic are approaches using natural monotonicization (Chernozhukov et al. 2010), non-parametric techniques (Dette and Volgushev 2008; Shim et al. 2008; Takeuchi et al. 2006) as well as constraints enforcing non-crossing (Koenker and Ng 2005).

We propose an alternative approach. The basic idea is to introduce a surface on a two-dimensional domain. One axis is for the covariate x , the other is for the probability τ . The quantile curve for any probability is found by cutting the surface at that probability. This surface is called a *quantile sheet*. We prefer the name sheet over surface to avoid possible confusion with the generalization of quantile curves to multiple covariates. Effectively, all possible quantile curves are estimated at the same time and the crossing problem disappears completely if the sheet is monotonically increasing with τ for every x . We describe the tools to make this happen.

The quantile sheet is constructed as a sum of tensor products of B -splines. In the spirit of P -splines (Eilers and Marx 1996), a rather large number of tensor products is used that may generally lead to over-fitting and a quite wobbly sheet, but additional difference penalties on the model coefficients allow proper tuning of the smoothness. We use separate penalties along the x - and the τ -axes, because in general isotropic smoothing will be too restrictive.

The majority of quantile smoothing software is based on linear programming. It is an elegant approach, especially when interior point algorithms are being used. However, we fall back on the classic iteratively re-weighted least squares approach (Schlossmacher 1973) with a small modification. The reason for using Schlossmacher's proposal is that we wish to apply the fast array algorithm for multidimensional P -spline fitting (Currie et al. 2006; Eilers et al. 2006). It is not at all clear to us how to combine that with linear programming. Examples in Bassett and Koenker (1992) raise doubts about the convergence of Schlossmacher's approach, but this seems wrong. We analyzed the examples but found convergence to the right solution. More explanation and an example can be found in Appendix.

With moderate or large amounts of smoothing (in the direction of τ) the quantile sheet will be monotonically increasing. This is what we observed when we optimized the smoothing parameters by asymmetric cross-validation. But if one is not willing to trust that all will be well, additional asymmetric difference penalties can be adopted to enforce monotonicity as pioneered by Bollaerts et al. (2006) and Bondell et al. (2010).

The manuscript is structured in the following way: In Sect. 2, first we describe quantile smoothing using P -splines, in general. Then we present quantile sheets as well as the application of fast array computations and optimal smoothing in this context. Applications to empirical data from a longitudinal study monitoring growth of children can be found in Sect. 3. The manuscript closes with a conclusion and an overview of open questions and further research.

2 Model description

2.1 Quantile smoothing with P -splines

Quantile smoothing estimates, for a given probability τ , a smooth curve $\mu(x; \tau)$ that minimizes the weighted sum of absolute distances

$$S_1 = \sum_{i=1}^n w_i(\tau) |y_i - \mu(x_i; \tau)|, \tag{1}$$

where y is the response and $\mu(x; \tau)$ is the population quantile and with weights

$$w_i(\tau) = \begin{cases} \tau & \text{if } y_i > \mu(x_i; \tau) \\ 1 - \tau & \text{if } y_i \leq \mu(x_i; \tau) \end{cases}, \tag{2}$$

for probability τ with $0 < \tau < 1$, $\tau = \tau_j$, $j = 1, \dots, J$ and sample size n . This is equivalent to the classic description of quantile regression (see also [Koenker 2005](#) for a comprehensive description).

Most algorithms for quantile regression and smoothing use linear programming. We wish to avoid that, because when doing the two-dimensional smoothing (see Sect. 2.2) with tensor products of B -splines, we want to exploit the fast array algorithm GLAM ([Currie et al. 2006](#); [Eilers et al. 2006](#)). [Schlossmacher \(1973\)](#) showed how to approximate a sum of absolute values $S = \sum_i |u_i|$ as a sum of weighted squares $\tilde{S} = \sum_i u_i^2 / |\tilde{u}_i|$. Here \tilde{u}_i is an approximation to the solution. The idea is to start with $\tilde{u}_i \equiv 1$ and to repeatedly apply the approximation until convergence. In practice, it is safer to use the approximation $u_k^2 / \sqrt{\tilde{u}_k^2 + \beta^2}$, with β a small number of the order of 10^{-4} times the maximum absolute value of the elements of the solution \hat{u} for numerical stability.

Using the approximation described above, we iteratively minimize

$$S_2 = \sum_{i=1}^n v_i (y_i - \mu(x_i; \tau))^2, \tag{3}$$

where

$$v_i = w_i / \sqrt{\tilde{r}_i^2 + \beta^2} \quad \text{with } \tilde{r}_i = y_i - \tilde{\mu}(x_i; \tau). \tag{4}$$

with $w_i(\tau)$ given by (2). The model is fitted by alternating between weighted regression and recomputing weights until convergence. Equal weights ($\tau = 0.5$) give a convenient starting point. For notational convenience we will omit the dependence on x or τ of the quantile curve wherever suitable and unambiguous.

We model a quantile curve $\mu(x; \tau)$ with P -splines: $\mu(x; \tau) = Ba_\tau$, where B is a basis of B -splines and a_τ is the coefficient vector. In its general form for bivariate data P -splines ([Eilers and Marx 1996](#)) minimize

$$S = \|y - Ba_\tau\|^2 + \lambda \|D_d a_\tau\|^2, \tag{5}$$

with respect to the coefficients a_τ . The basis B contains a generous number of B -splines. The second term in (5) is a roughness penalty. The matrix D_d forms the d th differences. In practice, mainly second and third differences are used. Here, we apply second order differences. The smoothness of the estimated function is tuned by the smoothing parameter λ .

Our model includes a weight vector v which we can easily integrate into (5) and the solution can be expressed as

$$\hat{a}_\tau = (B^T V B + \lambda D_d^T D_d)^{-1} B^T V y \quad (6)$$

$$= (B^T V B + P)^{-1} B^T V y, \quad (7)$$

where $V = \text{diag}(v)$ is a diagonal matrix with weights v on the main diagonal as defined in (4).

2.2 Quantile sheets

When estimating smooth quantile curves as described above, we choose a handful of values for τ and separately compute a curve for each of them. Imagine a large number of values of τ and a corresponding set of *non-crossing* quantile curves. Taking this to the limit we have a surface, above a rectangular domain defined by the dimensions x and τ . If we invert this reasoning, we assume the existence of a surface $\mu(x; \tau)$, which we call a *quantile sheet* and we have to develop a procedure to estimate it for a given data set.

We use tensor product P -splines. We have a n by K basis matrix $B = [b_{ik}]$ of the B -splines $B_k(x)$ on the domain of x . Let $\check{B} = [\check{b}_{jl}]$ be a similar, J by L , basis matrix on the domain $0 < \tau < 1$, containing the B -splines $\check{B}_l(\tau)$. For quantile curves we have a vector a_τ of B -spline coefficients for each separate τ . For the sheet we have a K by L matrix $A = [a_{kl}]$ of coefficients for the tensor products, in such a way that

$$\mu(x; \tau) = \sum_{k=1}^K \sum_{l=1}^L a_{kl} B_k(x) \check{B}_l(\tau). \quad (8)$$

To estimate the quantile sheet, we define an objective function containing two parts in (9). The first part defines an asymmetrically weighted sum of absolute values of residuals, equivalent to the one used for quantile curve estimation. The differences is that we sum not only over all observations but also over a set of probabilities. The second part of the objective function is a penalty on the coefficients in A , with different weights for the rows and the columns. Thus, we minimize

$$S = \sum_{i=1}^n \sum_{j=1}^J w_{ij} \left| y_i - \sum_{k=1}^K \sum_{l=1}^L a_{kl} b_{ik} \check{b}_{jl} \right| + P, \quad (9)$$

where P is the penalty part defined by

$$P = \lambda_x ||D_d A||_F + \lambda_\tau ||A \check{D}_d||_F, \tag{10}$$

where $||U||_F$ indicates the Frobenius norm of matrix U , i.e. the sum of squares of all its elements. The weight w_{ij} is the weight for observation i , as defined in (2) with $\tau = \tau_j$. The meaning of P is that we penalize sums of squares of differences for each column, weighted by λ_x and sums of squares of differences along each row, weighted by λ_τ .

2.3 Computation with array regression

There are two ways to perform the calculations for fitting the quantile sheet given the weights. One is to repeat the response vector y J -times in order to form a vector y^* of length nJ . Parallel to it the columns of the weight matrix V with $V = [v_{ij}]$, $i = 1, \dots, n$, $j = 1, \dots, J$ are put below each other to form the vector v^* . Here, v_{ij} is the weight v defined in (4) at x_i for $\tau = \tau_j$. Then penalized weighted regression is performed on the basis $\check{B} \otimes B$ where \check{B} is the B -spline basis over τ and B the B -spline basis over x . This approach has the charm of simplicity, but it consumes a lot of memory and computation time. The system to be solved in this case is

$$[\check{B} \otimes B]' V^* (\check{B} \otimes B) + Q \hat{a} = (\check{B} \otimes B)' V^* y, \tag{11}$$

where $V^* = \text{diag}(v^*)$. The solution \hat{a} gives the columns of the matrix \hat{A} of tensor product coefficients. The matrix Q follows from the combination of penalties on rows and columns of A :

$$Q = \lambda_x I_J \otimes D_d' D_d + \lambda_\tau \check{D}_d' \check{D}_d \otimes I_n, \tag{12}$$

where I_c is an identity matrix of size c .

More attractive is array regression as presented in Currie et al. (2006) and Eilers et al. (2006). Here, the construction and manipulation of the Kronecker product of the bases is avoided. The weight matrix V is kept as it is, and the n by J response matrix Y is formed by J -times repeating y . The details of the algorithm are complicated and will not be presented here, but we provide a short sketch of its essential features.

Let $C = B \square B$ be the row-wise Kronecker product of B with itself. This means that each row of C is the Kronecker product of the corresponding row of B with itself. Let $\check{C} = \check{B} \square \check{B}$. The system in (11) could also be written as

$$(G + Q) \hat{a} = r. \tag{13}$$

One can prove that

$$G^* = C' V \check{C} \quad \text{and} \quad R^* = B'(V \odot Y) \check{B}, \tag{14}$$

respectively, contain the elements of G and r but in a permuted order. In (14) the notation $V \odot Y$ stands for the element-wise product of the two matrices. In order to get r it is enough to vectorize R^* column-wise. To re-order the elements of G^* to

match with G , one has first to write it as a four-dimensional array, interchange second and third dimension and transform back to two dimensions. Because the row-wise tensor products are much smaller than the full Kronecker product $\check{B} \otimes B$, we obtain substantial savings in memory use. Additionally, depending on the size of the problem, computations are speeded up by at least an order of magnitude. Details can be found in [Currie et al. \(2006\)](#).

2.4 Optimal smoothing

In the presented model we have to choose two smoothing parameters $(\lambda_x, \lambda_\tau)$ for smoothing in direction of the independent variable x and the probability τ , respectively. Cross-validation is a classic criterion for choosing an optimal model. We will adapt the definition of the score to our situation with weights. This can be used to perform n -fold cross-validation. First, we introduce g_i as an indicator function whether or not an observation i is included in the training set. This vector will be multiplied element-wise with the weights v to form the effective weights \tilde{v} :

$$\tilde{v}_{ij} = v_{ij} g_i, \quad (15)$$

where v_{ij} is the weight at x_i for $\tau = \tau_j$. The effective weights are then used in the fitting procedure and the additional weight vector reflects the included and excluded observations. We define the weighted cross-validation score as

$$WCV = \sum_i \sum_j (1 - g_i) \tilde{v}_{ij} |y_i - \hat{\mu}(x_i; \tau_j)|. \quad (16)$$

The introduction of effective weights simplifies the calculation for n -fold cross-validations as the fitted values and weights for the excluded observations are automatically determined in one step. A comparison of smooth quantile sheets chosen by weighted cross-validation and quantile curves estimated by constrained B -splines COBS and its automatic smoothing will be shown in Sect. 3.2.

3 Application

3.1 Examples

As an example we analyze growth data of Dutch children ([Van Buuren and Fredriks 2001](#)) from the R package `expectreg` ([Sobotka et al. 2012](#)). The relation between age and weight—especially for babies—is often illustrated by means of percentiles, e.g. by the World Health Organization ([WHO Multicentre Growth Reference Study Group 2009](#)) issuing child growth standards. Standards and comparisons for these measures are commonly expressed in terms of quantiles. Therefore, it is of special interest to provide good and reliable methods for quantile curve estimation. We computed quantile sheets for these data. In Fig. 1 the resulting quantile curves of the relation between age (in months) and weight (in kg) for Dutch boys from birth to the age of 2 years are shown. The sample consists of more than 1,700 observations.

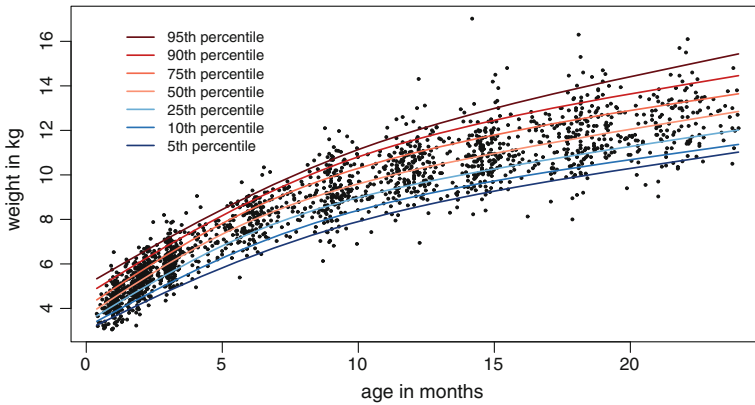


Fig. 1 Data: Dutch boys aged 0 months to 2 years. Estimated quantile curves using quantile sheets for the relation of age and weight (in kg) ($\tau = 0.05, \dots, 0.95$; bottom to top, blue to red) (colour figure online)

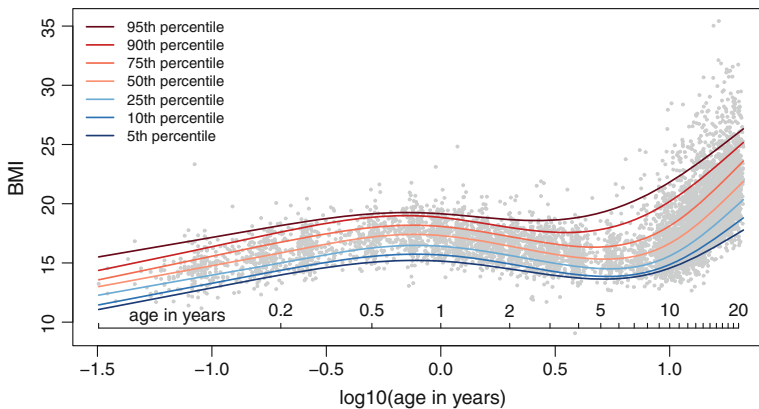


Fig. 2 Data: Dutch boys aged 0 months to 20 years ($n > 6,500$). Estimated quantile curves using quantile sheets for the relation of $\log_{10}(\text{age})$ and body mass index (BMI) ($\tau = 0.05, \dots, 0.95$; bottom to top, blue to red) (colour figure online)

This type of chart is also often used in practice by pediatricians informing parents about the developmental status of their small children and to judge whether or not the child is developing normally in terms of weight.

As a second illustration in Fig. 2 we include quantile curves for the relation between age and the body mass index (BMI) for the whole sample in this growth study. Comprehensive reference curves for children’s BMI have been first published in Cole et al. (1995) for the United Kingdom. This relationship is not monotone and it is still convincingly captured by quantile sheets.

3.2 A comparison with COBS

For a comparison between quantile sheets and quantile curves estimated with COBS we used a sample of size 500 from the data used in Fig. 2. Since the data show a

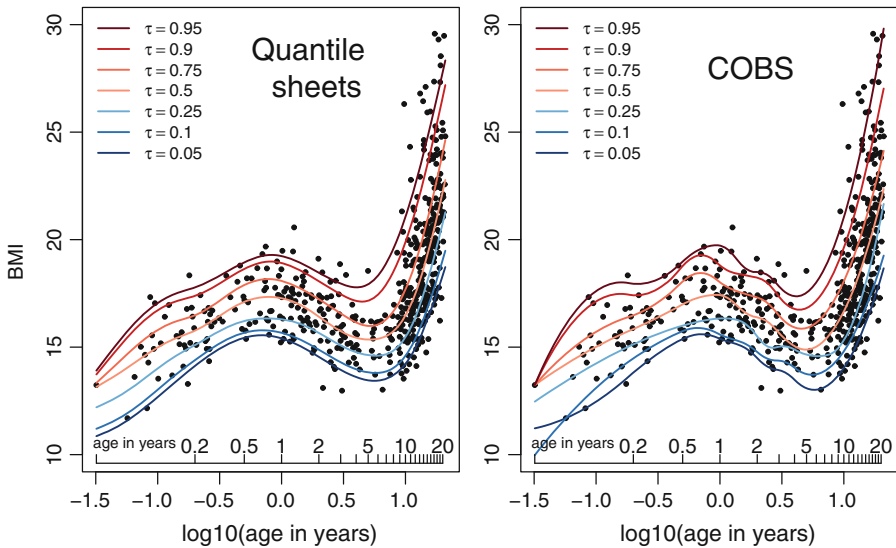


Fig. 3 Data: Dutch boys aged 0 months to 20 years. Relation between $\log_{10}(\text{age})$ and BMI. Subsample of 500 observations. Selected quantile curves estimated with quantile sheets (*left*) and COBS (*right*). Smoothing parameters for the quantile sheet: $\lambda_x = 1, \lambda_p = 10$, settings for coBS: `coBS(x, y, lambda = -20, tau = tau, lambda.length = 50, method = "quantile", nknots = 20, lambda.lo = 1e-2, lambda.hi = 1e5)`

complicated trend, a flexible model is needed. In Fig. 3, we show the results for these estimations. The smoothing parameter for the quantile sheet was chosen manually. The fit using COBS involved a considerable amount of fine-tuning and yet the resulting curves still show few crossings and undersmoothing. Each curve is modeled separately while the quantile sheet (shown in the left panel) fits all curves simultaneously without additional fine-tuning. By opting for an automatic selection of the smoothing parameter in COBS the quantile curves exhibit even a considerable number of overcrossings as well as more undersmoothing. This is especially true for quantile curves away from the median.

4 Conclusion

We have introduced a new approach to the estimation of smooth non-crossing quantile curves. Each curve is interpreted as a level curve of a sheet (a surface) above the (x, τ) plane. Tensor product P -splines are used to estimate the sheet. Regular difference penalties allow tuning of the smoothness and additional asymmetric difference penalties can enforce monotonicity (in direction of probability τ). Application to a real data showed convincing results.

If we take a wider perspective, quantile sheets are an interesting new statistical formalism, extending beyond a tool to avoid crossing curves. In fact, the estimation of quantile curves can be seen as a far from perfect way to estimate the more fundamental quantile sheet. Also, the smooth quantile sheet we obtain, as a sum of tensor products

of (cubic) splines, can easily be differentiated and/or inverted to obtain estimates of the joint density or conditional cumulative distributions.

Quantile sheets result in smooth surfaces. Derivatives of our quantile sheets are smooth, too, and piecewise quadratic. The iteratively reweighting approach allows the use of squares of differences in the penalties. The results compare favorably with the piecewise linear derivatives obtained with the triogram smoothing (Koenker 2005) where the penalty is based on total variation (sums of absolute values of differences).

Another advantage of our technique is that it can be extended to the case of two or more covariates resulting in “quantile volumes”. If we write it as $\mu(x, t; \tau)$ in two dimensions, the quantile surface above the (x, t) plane, for a chosen probability τ^* is obtained by evaluating $\mu(x, t; \tau^*)$ at a chosen grid for x and t .

Instead of asymmetrically weighted absolute values of residuals, one can use asymmetrically weighted squares, resulting in an *expectile sheet*. Expectiles have been introduced in Newey and Powell (1987) as a least squares alternative to quantile curves. Optimal smoothing for expectile curves has been studied in Schnabel and Eilers (2009). We have obtained very promising results for expectile sheets (Schnabel and Eilers 2011).

One issue needs further study. To estimate the quantile sheet, we introduced a grid of values for the probability τ . The objective function we minimize is a sum of the asymmetric objective functions we know from the estimation of individual smooth quantile curves. The result does depend on the choice of grid for τ . In the limit, with a large uniform grid, we are approximating an integral over the domain of τ . A non-uniform grid implies a weighted integral. An interesting question is whether a non-uniform grid is better, and if so, what the optimal grid should be.

In addition, there is the challenge to combine array regression with the interior point algorithm. In principle, this looks feasible, because the core of the interior point algorithm boils down to weighted regression. We hope to work on this in the future.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

5 Appendix: Schlossmacher’s (1973) method and the Bassett and Koenker (1992) counterexamples

Bassett and Koenker (1992) raised doubts about the proper convergence of Schlossmacher’s algorithm. Actually they do not attack it directly, but mention it in passing, as having the same problems that they document for the methods of Soliman et al. (1988, 1991) But if we apply Schlossmacher directly to their counterexamples, no problems occur. Our R code is presented below. To keep it simple we run many iterations instead of checking for convergence.

The core of Bassett and Koenker’s paper is an attack on methods that use least squares residuals to select subsets of observations. It is well known that an n -dimensional linear model that minimizes the sum of absolute values of residuals exactly fits n observations. If these observations can be found by any other method than linear programming, one can simply compute the model parameters by computing

an exact fit to them. Schlossmacher's method does not try such a prior selection. In fact it keeps all the observations but with every iteration increases the weights of the ones that finally will give an (almost) exact fit.

```
library(quantreg)

# Location example
v <- 1
y0 <- v * c(0, 0, 0, 0, 6, 23, 132)

# Regression example
x <- c(1, 2, 2, 3, 3, 4)
y1 <- c(0, 1, -1, 2, -2, 0)

# Regression quantiles
eqn <- y1 ~ x # Use eqn = y0 ~ v for location example
mod1 <- rq(eqn)

# Schlossmacher
mod2 <- lm(eqn)
beta <- 1e-10
for (it in 1:100) {
  r <- mod2$residuals
  w <- 1 / sqrt(r ^ 2 + beta ^ 2)
  mod2 <- lm(eqn, weights = w)
}

print(mod1)

#Call:
#rq(formula = eqn)
#Coefficients:
#(Intercept)          x
#           0           0
#Degrees of freedom: 6 total; 4 residual

print(mod2)

#Call:
#lm(formula = eqn, weights = w)
#Coefficients:
#(Intercept)          x
# 1.308e-21  -5.234e-22
```

References

- Bassett, G.W., Koenker, R.W.: A note on recent proposals for computing l_1 estimates. *Comput. Stat. Data Anal.* **14**, 207–211 (1992)
- Bollaerts, K., Eilers, P.H.C., van Mechelen, I.: Simple and multiple P-splines regression with shape constraints. *Br. J. Math. Stat. Psychol.* **59**, 451–469 (2006)
- Bondell, H.D., Reich, B.J., Wang, H.: Noncrossing quantile regression curve estimation. *Biometrika* **97**(4), 825–838 (2010)
- Chernozhukov, V., Fernandez-Val, I., Galichon, A.: Quantile and probability curves without crossing. *Econometrica* **78**(3), 1093–1125 (2010)
- Cole, T.J., Freeman, J.V., Preece, M.A.: Body mass index reference curves for the UK, 1990. *Arch. Dis. Child.* **73**, 25–29 (1995)
- Currie, I.D., Durban, M., Eilers, P.H.C.: Generalized linear array models with applications to multidimensional smoothing. *J. Roy. Stat. Soc. Ser. B* **68**(2), 259–280 (2006)
- Dette, H., Volgushev, S.: Non-crossing non-parametric estimates of quantile curves. *J. Roy. Stat. Soc. Ser. B* **70**(3), 609–627 (2008)
- Eilers, P.H., Currie, I.D., Durban, M.: Fast and compact smoothing on large multidimensional grids. *Comput. Stat. Data Anal.* **50**, 61–76 (2006)
- Eilers, P.H.C., Marx, B.D.: Flexible smoothing with B-splines and penalties. *Stat. Sci.* **11**(2), 89–121 (1996)
- Koenker, R.: *Quantile Regression*. Cambridge University Press, Cambridge (2005)
- Koenker, R.: (2011) *quantreg*: Quantile regression. R package version 4.71. <http://CRAN.R-project.org/package=quantreg>
- Koenker, R., Bassett, G.W.: Regression quantiles. *Econometrica* **46**(1), 33–50 (1978)
- Koenker, R., Ng, P.: (2005) Inequality constrained quantile regression. Special issue on quantile regression and related models. *Sankhyā. Indian J. Stat.* **67**(2):418–440
- Newey, W.K., Powell, J.L.: Asymmetric least squares estimation and testing. *Econometrica* **55**(4), 819–847 (1987)
- Ng, P.T., Maechler, M.: (2011) *cobs*: COBS-Constrained B-splines (Sparse matrix based). R package version 1.2-2. <http://CRAN.R-project.org/package=cobs>
- Schlossmacher, E.J.: An iterative technique for absolute deviations curve fitting. *J. Am. Stat. Assoc.* **68**(344), 857–859 (1973)
- Schnabel, S.K., Eilers, P.H.C.: Optimal expectile smoothing. *Comput. Stat. Data Anal.* **53**, 4168–4177 (2009)
- Schnabel, S.K., Eilers, P.H.C.: (2011) Expectile sheets for joint estimation of expectile curves (submitted)
- Shim, J., Hwang, C., Seok, K.H.: Non-crossing quantile regression via doubly penalized kernel machine. *Comput. Stat.* **23**, 83–94 (2009)
- Sobotka, F., Kneib, T., Schnabel, S., Eilers, P.: (2012) *expectreg*: Expectile Regression. R package version 0.30. <http://CRAN.R-project.org/package=expectreg>
- Soliman, S.A., Christensen, G.S., Rouhi, A.: A new technique for curve fitting based on minimum absolute deviations. *Comput. Stat. Data Anal.* **6**(4), 341–351 (1988)
- Soliman, S.A., Christensen, G.S., Rouhi, A.H.: A new algorithm for nonlinear l_1 -norm minimization with nonlinear equality constraints. *Comput. Stat. Data Anal.* **11**(1), 97–109 (1991)
- Takeuchi, I., Le, Q.V., Sears, T.D., Smola, A.J.: Nonparametric quantile estimation. *J. Mach. Learn. Res.* **7**, 1231–1264 (2006)
- Van Buuren, S., Fredriks, A.M.: Worm plot: a simple diagnostic device for modeling growth reference curves. *Stat. Med.* **20**, 1259–1277 (2001)
- WHO Multicentre Growth Reference Study Group: WHO child growth standards: growth velocity based on weight, length and head circumference: methods and developments, Tech. rep. World Health Organization, Geneva (2009)
- Yee, T.W. (2011) *VGAM*: vector generalized linear and additive models. R package version 0.8-3. <http://CRAN.R-project.org/package=VGAM>