



University
of Glasgow

Neocleous, T. and Portnoy, S. (2008) On monotonicity of regression quantile functions. *Statistics and Probability Letters*.

<http://eprints.gla.ac.uk/3799/>

Deposited on: 2 November 2007

On Monotonicity of Regression Quantile Functions

Tereza Neocleous¹ and Stephen Portnoy²

October 11, 2007

Abstract

In the linear regression quantile model, the conditional quantile of the response, Y , given x is $Q_{Y|x}(\tau) \equiv x'\beta(\tau)$. Though $Q_{Y|x}(\tau)$ must be monotonically increasing, the Koenker-Bassett regression quantile estimator, $x'\hat{\beta}(\tau)$, is not monotonic outside a vanishingly small neighborhood of $x = \bar{x}$. Given a grid of mesh δ_n , let $x'\hat{\beta}^*(\tau)$ be the linear interpolation of the values of $x'\hat{\beta}(\tau)$ along the grid. We show here that for a range of rates, δ_n , $x'\hat{\beta}^*(\tau)$ will be strictly monotonic (with probability tending to one) and will be asymptotically equivalent to $x'\hat{\beta}(\tau)$ in the sense that $n^{1/2}$ times the difference tends to zero at a rate depending on δ_n .

¹Department of Statistics, University of Glasgow
15 University Gardens, Glasgow G12 8QQ, tereza@stats.gla.ac.uk

²Department of Statistics, University of Illinois at Urbana-Champaign
725 S. Wright St., Champaign IL 61801, sportnoy@uiuc.edu
Research partially supported by NSA Grant MDA904-02-1-0092

keywords: regression quantile, monotonicity, Bahadur Representation

1 Introduction.

From the earliest research on L_1 regression, it was recognized that the solutions were determined by fitting certain observations exactly. A general form of this was described by Gauss in 1809 (see Gauss, 1963), who also noted the subsequent invariance to perturbations not changing the residual signs. Gauss appears to use this as an argument against L_1 regression, though modern statisticians tend to take the invariance as a desirable robustness property, and do not find zero residuals to be problematic.

When general regression quantiles were introduced in Koenker and Bassett (1978, (see also Koenker, 2005), it was recognized that this exact-fit property forced the regression quantile fits to cross in general. Thus, the conditional quantile function must be non-monotonic in its argument, τ , at some values of the explanatory variable. Although this seems to have upset some Gauss-minded scholars, it can not be a serious statistical problem (under standard conditions for asymptotics) since the true (monotonic) conditional quantile function must lie within usual statistical accuracy. Specifically, if only a few τ -values are considered, standard asymptotic results show that the conditional quantile functions ($\hat{Q}_x(\tau)$) must be monotonic along these τ -values with high probability.

Nonetheless, it would be useful to have a general version of regression quantiles with strictly monotonic conditional quantile functions. Specifically, one application would be to the use of “direct” confidence intervals and confidence bands (see Zhou and Portnoy, 1996 and 1998). He (1997) provides

such a version when the model is of location-scale form. However, models for linear regression quantiles can be much more general and thus require a more general development. Here, based on Neocleous (2005), we show that by choosing an appropriate grid of τ -values and defining the quantile functions by linear interpolation between grid values, the resulting conditional quantile estimator is strictly monotonic with probability tending to one, and is asymptotically equivalent to the usual regression quantile estimator. Note that this asymptotic equivalence would be easy to establish if the grid mesh were $o(n^{-1/2})$, but then the interpolated grid estimator would not be monotonic. The crucial result proved here is that the equivalence holds if the mesh of the grid tends to zero strictly more slowly than $n^{-1/2}$ (so that monotonicity holds with probability tending to one) but more quickly than $n^{-1/4}$.

In fact, Koenker and Bassett (1978) noted that the empirical conditional quantile function is indeed monotonic at the mean, \bar{x} , of the explanatory variables, and by continuity in some interval about the mean. It may be of interest to note that this interval tends to be quite small, and in fact the length will generally tend to zero as the sample size increases. A rough argument for this is as follows: consider simple linear regression. If the observations are in general position, each regression quantile solution fits two points, one of which (say, x^*) remains in the basis for next solution (at the next largest τ value). Thus the solution pivots about x^* , so that if x^* is smaller than \bar{x} , the conditional quantile function must be decreasing everywhere to the left of x^* (and *vice versa* for $x^* > \bar{x}$). When x is random and under appropriate stronger assumptions, the argument in Portnoy (1991a)

can be used to show that the probability that each given basis element becomes a pivot is bounded below by a constant independent of sample size. Thus, if $k_n \rightarrow \infty$, of the k_n observations with x_i nearest to \bar{x} , at least one to the left of \bar{x} and one to the right of \bar{x} must be pivots at some τ -breakpoint (with probability tending to one). Therefore, for each x outside an interval about \bar{x} shrinking to zero, there must be some τ value at which the conditional quantile function is non-monotonic.

In the remainder of the paper, a general regression quantile model is introduced, and relatively mild conditions are imposed. A grid of τ values in $\epsilon \leq \tau \leq 1 - \epsilon$ is considered and the linearly interpolated regression quantile function, $\hat{\beta}^*(\tau)$ is defined. Two results are proven: (1) if the mesh of the grid tends to zero strictly slower than $n^{-1/2}$ (as $n \rightarrow \infty$), then $\hat{\beta}^*(\tau)$ is strictly monotonic (on $[\epsilon, 1 - \epsilon]$); and (2) if the mesh tends to zero strictly faster than $n^{-1/4}$, then $\hat{\beta}^*(\tau)$ has the same first order asymptotic Bahadur representation as the usual (Koenker-Bassett) regression quantile estimator.

2 Monotonicity.

Let $Y_i = x_i' \beta(\tau) + e_i$, $i = 1, \dots, n$, $x_i \in R^p$ with p fixed and the errors e_i having τ th quantile equal to zero. In quantile form, this is written as

$$Q_{Y_i}(\tau|x_i) = x_i' \beta(\tau).$$

The errors are assumed to be independent, but not necessarily identically distributed.

Assumptions.

(F) For D a bounded domain and $0 < \epsilon < 1$, there exist constants a, b, c with $a > 0$, $b < \infty$, and $c < \infty$ such that

$$a \leq f_{Y_i|x}(F_{Y_i|x}^{-1}(\tau)) \leq b \quad |f'_{Y_i|x}(F_{Y_i|x}^{-1}(\tau))| \leq c$$

uniformly for $x \in D$, $\epsilon \leq \tau \leq 1 - \epsilon$, and uniformly in $i = 1, \dots, n$.

(X) $\|x_i\| \leq d$ for some constant d uniformly in $i = 1, \dots, n$.

(XX) The matrix $\frac{1}{n} \sum_{i=1}^n x_i x_i'$ is positive definite.

Let $\hat{\beta}(\tau)$ be the Koenker-Bassett regression quantile estimator of $\beta(\tau)$ (see Koenker (2005) for the definition and basic properties). Consider a grid of τ -values covering $[\epsilon, 1 - \epsilon]$: let $\epsilon = t_1 < t_2 < \dots < t_M = 1 - \epsilon$ be the M gridpoints, with mesh $\delta_n = t_{k+1} - t_k$. At the j th gridpoint, t_j , the estimated regression coefficient is $\hat{\beta}(t_j) \in \mathbf{R}^p$. Starting with these estimates along the grid: $\hat{\beta} = (\hat{\beta}(t_1)', \dots, \hat{\beta}(t_M)')' \in \mathbf{R}^{Mp}$, let $\hat{\beta}_k \equiv \hat{\beta}(t_k)$, and define $\hat{\beta}^*(\tau)$ to be the estimator linearly interpolating $\{(t_k, \hat{\beta}_k) : k = 1, \dots, M\}$ (see equation (??) for a formal definition of the interpolating function).

Theorem 1. *Under the above conditions, let δ_n satisfy:*

$$\limsup_{n \rightarrow \infty} \delta_n n^\eta > 0 \quad \text{and} \quad \liminf_{n \rightarrow \infty} \delta_n n^{1/2} / \log n > 0. \quad (1)$$

for some $\eta > 0$. Then with probability tending to one, $x' \hat{\beta}^(\tau)$ is strictly monotonic uniformly on $\{\|x\| \leq d\}$ for $\epsilon \leq \tau \leq 1 - \epsilon$, and*

$$\sqrt{n}(\hat{\beta}^*(\tau) - \hat{\beta}(\tau)) = \mathcal{O}(\sqrt{n} \delta_n^2) + \mathcal{O}_p((\delta_n \log n)^{1/2}) + \mathcal{O}_p(n^{-1/4}(\log n)^{1/2})$$

uniformly in $\epsilon \leq \tau \leq 1 - \epsilon$.

Proof. A basic ingredient of the proof is a Bahadur representation for $\hat{\beta}(\tau)$. A version for fixed τ appears in He and Shao (1996), though this result is for i.i.d. errors and does not include uniformity in τ . Portnoy (1991b) provides an appropriate uniform version under stronger conditions permitting the sample distributions to be dependent. The following result can be developed under the conditions here either by specializing the proof in Portnoy (1991b) for independence, or using the chaining argument there to extend the He-Shao (1996) result to apply to the nonstationary case and be uniform in τ :

$$\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) = n^{-1/2} A^{-1}(\tau) \sum x_i [\tau - I(Y_i \leq x_i' \beta(\tau))] + \mathcal{O}_p(r_n) \quad (2)$$

uniformly for $\epsilon \leq \tau \leq 1 - \epsilon$, where $r_n = \sqrt{\log n} n^{-1/4}$, and $A(\tau) = XV(\tau)X'$ with $V(\tau) = \text{Diag}(f_{Y_i|x_i}(F_{Y_i|x_i}^{-1}(\tau)))$.

Now, to show monotonicity, note that Condition (F) implies that

$$\frac{d}{d\tau} x' \beta(\tau) = \frac{1}{f_{Y|x}(F_{Y|x}^{-1}(\tau))} \geq \frac{1}{b}$$

so $x' \beta(\tau)$ must grow by at least $\frac{\delta_n}{b}$ in moving from t_k to t_{k+1} . From the Bahadur representation (??), $(\hat{\beta}(t_k) - \beta(t_k)) = \mathcal{O}_p(n^{-1/2})$ uniformly in k , and so $\{\hat{\beta}_k\}_{k=1}^M$ must also be strictly monotonic with probability tending to one as long as δ_n tends to zero strictly more slowly than $n^{-1/2}$.

To show asymptotic equivalence, note that the linearly interpolated quantile regression estimator $\hat{\beta}^*(\tau)$ for $t_k \leq \tau \leq t_{k+1}$ is

$$\alpha_n \hat{\beta}(t_k) + (1 - \alpha_n) \hat{\beta}(t_{k+1}) \quad (3)$$

where $\alpha_n = (t_{k+1} - \tau)/\delta_n$ (so that $\alpha_n t_k + (1 - \alpha_n)t_{k+1} = \tau$).

Then

$$\sqrt{n}(\hat{\beta}^*(\tau) - \beta(\tau)) = \sqrt{n}(\alpha_n \hat{\beta}(t_k) + (1 - \alpha_n) \hat{\beta}(t_{k+1}) - \beta(\tau)) \quad (4)$$

$$= \alpha_n \sqrt{n}(\hat{\beta}(t_k) - \beta(t_k)) \quad (5)$$

$$+ (1 - \alpha_n) \sqrt{n}(\hat{\beta}(t_{k+1}) - \beta(t_{k+1})) \quad (6)$$

$$- \sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) \quad (7)$$

$$+ \sqrt{n}(\alpha_n \beta(t_k) + (1 - \alpha_n) \beta(t_{k+1}) - \beta(\tau)) \quad (8)$$

Let $S \equiv (4)+(5)+(6)$. Then by (7), up to an error of order $\mathcal{O}_p(r_n)$ (with $r_n = n^{-1/4}(\log n)^{1/2}$) and uniformly in $\epsilon \leq \tau \leq 1 - \epsilon$.

$$\begin{aligned} S &= n^{-1/2} A^{-1}(\tau) \sum x_i \{ \alpha_n [I(Y_i \leq x'_i \beta_k) - I(Y_i \leq x'_i \beta(\tau))] \\ &\quad + (1 - \alpha_n) [I(Y_i \leq x'_i \beta_{k+1}) - I(Y_i \leq x'_i \beta(\tau))] \} . \end{aligned}$$

Note that S has mean zero:

$$ES = n^{-1/2} A^{-1}(\tau) \sum x_i \{ \alpha_n (t_k - \tau) + (1 - \alpha_n) (t_{k+1} - \tau) \} = 0 .$$

Each of the terms in S is the difference of a weighted empirical process between two values of the argument. Thus, we need a bound on the modulus of continuity for the weighted empirical process. To obtain a result in the form needed here, let $D \equiv n^{-1/2} \sum x_i [I(Y_i \leq u) - I(Y_i \leq u + \delta_n)]$. Then, since the standard deviation of each summand is bounded by $\sqrt{\delta_n}$, one can use the chaining argument in Portnoy (1991b) (for example) to obtain:

$$D - ED = \mathcal{O}_p((\delta_n \log n)^{1/2}) + \mathcal{O}_p(n^{-1/4} (\log n)^{1/2})$$

uniformly in u , as long as δ_n satisfies the condition (??) in the hypotheses of the Theorem. Inserting this in S, it follows that

$$S = S - ES = \mathcal{O}_p\left((\delta_n \log n)^{1/2}\right) + \mathcal{O}_p\left(n^{-1/4}(\log n)^{1/2}\right)$$

Finally, for (??), note that

$$\left|\frac{d^2}{d\tau^2}x'\beta(\tau)\right| = \left|-\frac{f'_{Y|x}(F_{Y|x}^{-1}(\tau))}{f_{Y|x}^3(F_{Y|x}^{-1}(\tau))}\right| < \frac{c}{a^3}$$

Thus, by a second-order Taylor expansion for $x'_i\beta_k$ and $x'_i\beta_{k+1}$ and using $\alpha_n t_k + (1 - \alpha_n)t_{k+1} = \tau$,

$$|x'_i(\alpha_n\beta(t_k) + (1 - \alpha_n)\beta(t_{k+1}) - \beta(\tau))| < \frac{c\delta_n^2}{a^3},$$

and the Theorem follows. \parallel

Remark. If $\delta_n = n^{-1/2}(\log n)^\epsilon$, then $x'\hat{\beta}^*(\tau)$ is strictly monotonic (with probability tending to 1) and

$$\sqrt{n}(\hat{\beta}^*(\tau) - \hat{\beta}(\tau)) = \mathcal{O}_p\left(n^{-1/4}(\log n)^{1/2+\epsilon/2}\right)$$

which is (essentially) the best the Bahadur representation will permit. In fact, choosing δ_n larger will make the probability of monotonicity larger, perhaps moderately so for moderately small n . However, this comes at the cost of increasing the difference between $\hat{\beta}^*(\tau)$ and $\hat{\beta}(\tau)$. For example, if $\delta_n = n^{-1/3}$, the two largest error terms have the same exponent, providing $\sqrt{n}(\hat{\beta}^*(\tau) - \hat{\beta}(\tau)) = \mathcal{O}_p\left(n^{-1/6}(\log n)^{1/2}\right)$.

References

- [1] Gauss, K. (1963). *Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*, a translation of *Theoria Motus* by Charles Henry Davis, Section 186, Dover Publications, Inc. New York.
- [2] He, X. (1997). Quantile curves without crossing, *Amer. Statistician*, 51, 186-192.
- [3] He, X. and Shao, Q.-M. (1996). A general Bahadur representation of M-estimators and its application to linear regression with nonstochastic designs, *Ann. Statist.*, 24, 2608-2630.
- [4] Koenker, R. (2005). *Quantile Regression*, Cambridge University Press, Cambridge.
- [5] Koenker, R. and Bassett, G. (1978). Regression Quantiles, *Econometrica*, 46, 33-50.
- [6] Neocleous, T. (2005). *A Partly Linear Model for Censored Regression Quantiles*. Ph. D. Thesis, Department of Statistics, University of Illinois.
- [7] Portnoy, S. (1991a). Asymptotic behavior of the number of regression quantile breakpoints, *SIAM J. Sci. Stat. Comp.*, 12, 867-883.
- [8] Portnoy, S. (1991b). Asymptotic behavior of regression quantiles in non-stationary, dependent cases, *J. Multivariate Analysis*, 38, 100-113.
- [9] Portnoy, S. (2003). Censored Regression Quantiles, *J. Amer. Statist. Assoc.*, 98, 1001-1012.
- [10] Zhou, K. and Portnoy S (1996). Direct Use of Regression Quantiles to Construct Confidence Sets in Linear Models, *Ann. Statist.*, 24, 287-306.
- [11] Zhou, K. and Portnoy S (1998). Statistical Inference on Heteroscedastic Models Based on Regression Quantiles, *J. Nonpar. Statist.*, 9, 239-260.