Nonparametric Quantile Regression with Applications to Stock Returns

Siegfried Heiler

Klaus Abberger

University of Konstanz, Department of Mathematics and Statistics
Universitätsstr. 10
78457 Konstanz, Germany
Siegfried.Heiler@uni-konstanz.de

For a random variable Y with covariate (vector) X and conditional cumulative distribution function F(y|x):=F(y|X=x) various possibilities for a nonparametric estimation of the conditional quantile function $q_\alpha(x) = F^{-1}(\alpha|x)$, $0 < \alpha < 1$, are studied and applied to some series of stock returns. Two approaches for nonparametric estimation of q_α are considered in the literature. One starts with estimating F(y|x) first where then in a second step from the empirical conditional c.d.f. $F_n(y|x)$ the empirical quantile is derived. The other one tries to estimate q_α directly by a local polynomial expansion of q_α which then is estimated by local least squares. Other approaches like orthonormal series approximation (based, e.g. on a cosine system or a Haar system) or spline smoothing (with appropriate smoothness penalty) are also considered, but will not be discussed here.

A local neighborhood in the X-space is selected by using a kernel function K (i.e. a symmetrical density) to generate local neighborhood weights K $\left(\frac{X_i - x}{h}\right)$ for an estimation at x. The range of the local neighborhood is driven by the bandwidth h, which has to fulfil $h \to 0$ and $nh \to \infty$ with sample size $n \to \infty$.

The estimation of F(y|x) in a first step by

$$F_n(y|x) = \sum_{i=1}^{n} \mathbf{1}_{(-\infty,y]}(Y_i) K\left(\frac{X_i - x}{h}\right) / \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right) \qquad (1)$$

was considered by *Horvath and Yandell (1988)* and later extended by *Abberger (1997)* to stationary time series and applied to some series of stock returns.

From (1) the estimated quantile function is derived by

$$q_{n,a}(x) = \inf \{ y \in \Re \mid F_n(y|x) \geq a, \quad 0 < a < 1. \qquad (2)$$

Quantile regression for the parametric case was introduced by *Koenker and Basset (1978).* They suggested the use of the loss function $r_a(u) = a\mathbf{1}_{[0,\infty)}(u)u + (a-1)\mathbf{1}_{(-\infty,o)}(u)u.$

With this, a local polynomial approximation for q_α - motivated by Taylor series expansion arguments – can be estimated (for scalar X) directly by solving the problem

$$\arg\min \left\{ \sum_{i=1}^{n} r_a\left(Y_i - \sum_{j=0}^{r} (X_i - x)^j b_j(x)\right) K\left(\frac{X_i - x}{h}\right) \right\} \qquad (3)$$

and putting $\hat{q}_{\alpha}(x) = \hat{b}_0(x)$. With $\hat{q}_{\alpha}^{(j)}(x) = j!\,\hat{b}_j(x)$ this procedure also yields estimates for derivatives. The case r=0 (local constant approximation) was considered by *Jones and Hall (1990)* and the local linear case in *Yu and Jones (1998).* In the latter paper also a double kernel approach for a local linear estimation of F(y|x) with a second kernel K_2 and corresponding c.d.f. F_2 is discussed, based on the approximation

$$E\left[F_2\left(\frac{y-Y}{h_2}\right)\middle| X = x_0\right] \cong F(y \mid x_0) + \dot{F}(y \mid x)(x - x_{0)} = b_0(x_0) + b_1(x_0)(x - x_0)\,, \qquad (4)$$

where $\dot{F}$ denotes the derivative with respect to x.

The discussed estimates react very sensitively on the choice of bandwidths.

In *Jones and Hall (1990)* and in *Yu and Jones (1998)* procedures for bandwidth selection based on minimization of the mean squared error are discussed. *Abberger (1997)* uses a cross validation criterion based on the loss function r_{α} of *Koenker and Basset*. Bandwidth selection procedures with good asymptotic properties and satisfactory practical performance are still looked for.

If in the case of time series the covariate X is just the time index, then we arrive at quantile smoothing. In the case of past Y-values this leads to quantile autoregression. In application to stock returns and compared with GARCH models, the quantile procedures turn out to be very robust, to yield smoother courses of quantile curves and to be able to catch and to reflect asymmetries in the distributions of the data.

REFERENCE

Abberger, K. (1997). Quantile Smoothing in Financial Time Series. Statistical Papers, 38, 125-148.

Heiler, S. (2001). Nonparamtric Time Series Analysis: Nonparametric Regression, Locally weighted regression, Autoregression, and Quantile Regression. Chapter 12 in (Peña, D., Tiao, G.C. and R.S. Tsay editors): A Course in Time Series Analysis. J. Wily & Sons, New York.

Horvath, L. and Yandell, B.S. (1988). Asymptotics for Conditional Empirical Processes. J. Multivar. Anal. 26, 184-206.

Jones, M.C. and Hall, P. (1990) Mean Squared Error Properties of Kenel Estimates of Regression Quantiles. Statistics & Probability Letters 10, 283 –289.

Koenker, R. and Basset, G. (1978). Regression Quantiles Econometric 46, 33-50.

Yu, K. and Jones, M.C. (1998). Local Linear Quantile Regression. J. Amer. Statist. Assoc. 98, 228-237.

RESUME

Pour une variable aléatoire Y avec vecteurs de covariables X diverses méthodes d'estimation non-paramétrique pour fonctions de quantile d'ordre α sont discutées et on fait un rapport sur quelques application aux séries temporelles provenant de marchés financiers.