DERACINATED QUANTILE REGRESSION: AN R VINAIGRETTE

ROGER KOENKER

1. Introduction

Suppose you have names for triangles and squares and now you need a name for a pentagon: should you call it a generalized triangle, or a generalized square? Having chosen, what would you now call an hexagon to distinguish from the pentagon? You see where this is going? Generalized isn't really very descriptive.[1]

What is "deracinated quantile regression" you may ask. It is my alternative term for "generalized quantile regression" as introduced in Powell (2020). Deracinated means uprooted, that sense of being torn away from your homeland by some involuntary force of man or nature. DQR isn't regression and it isn't even really about quantiles. What is regression? There are quite a few flavors, but like ice cream these flavors should have some common feature, and for regression that feature is that they should be methods for estimating *conditional* somethings – mostly, I would happily concede – conditional mean functions.[2] However, in the case of quantile regression conditional quantile functions are the objective, the parameter, if you will. Even instrumental variable quantile regression as introduced by Chernozhukov and Hansen (2004) is focused on estimating conditional quantile functions, as is the triangular system estimators of Chesher (2003) and Ma and Koenker (2006). What about "unconditional quantile regression" I hear you ask? This too is conditional, the model is

$$\mathbb{P}(Y < \xi | X = x) = \Lambda(x^\top \beta(\xi))$$

so conditional probabilities of the response falling below a fixed threshold ξ are modeled by a linear predictor transformed by a monotone link function. By varying ξ we trace out the complete family of the conditional distribution functions.[3]

Given either a family of conditional quantile functions or a family of conditional distribution functions, one can easily build other objects of interest. They are the bricks and if made with with the appropriate mixture of hypothetical straw one can estimate various counterfactuals: for example marginal distributions of the response by integrating out various marginal distributions of the covariates, as in Mata and Machado (2005) and Chernozhukov et al. (2013).

Version: October 8, 2020. A genre manifesto for R Vinaigrettes is available at `http://davoidofmeaning.blogspot.com/2016/12/r-vinaigrettes.html`.

[1] I confess that for most of 1975 Gib Bassett and I call regression quantiles "generalized sample quantiles for the linear model." But eventually we realized that this was foolish and we began using the term "regression quantiles," the term then stuck until Moshe Buchinsky and Gary Chamberlain began using the inverted term "quantile regression." At that point it became more than a statistical curiosity, and for me, for a time, became (almost) a way of life.

[2] For an elegant exposition of this viewpoint, see Goldberger (1968).

[3] For a very elementary comparison of quantile regression and distributional regression, see Koenker (2011).

DQR as imagined by Powell (2020) falls somewhere in a crevasse between the objectives of estimating conditionals and estimating marginals of this type. He posits two types of covariates, a policy manipulable sort, called D, and some others, called X. To illustrate this consider the first of Powell's simulation examples which looks like this:

$$Y_i = (1 - D_i) * (U_i + X_i), \quad X_i \sim U[0,1],\ D_i \sim U[0,1],\ U_i \sim [0,0.1].$$

The conditional quantile functions of the response look like this:

$$Q_{Y|D,X}(\tau|d,x) = x - xd + \tfrac{1}{10}(\tau - \tau d),$$

so there is an interaction term in a location shift and simple form of heteroscedasticity that implies that the distribution of the response is degenerate at $D = 1$. The quantile treatment effect is,

$$\frac{\partial Q_{Y|D,X}(\tau|d,x)}{\partial d} = -x - \tfrac{1}{10}\tau,$$

This is a rather strange model, but it can be estimated quite accurately.

```
# Powell Simulation Setting 1.
require(quantreg)
set.seed(1729)
n = 1000
D = runif(n)
X = runif(n)
U = runif(n,0,0.1)
Y = (X + U)*(1 - D)
f = rq(Y ~ X + D + I(X*D), tau = 1:9/10)
round(f$coef,3)
##             tau= 0.1 tau= 0.2 tau= 0.3 tau= 0.4 tau= 0.5 tau= 0.6 tau= 0.7
## (Intercept)    0.008    0.015    0.027    0.040    0.049    0.060    0.072
## X              1.002    1.007    1.003    1.002    1.002    1.003    0.999
## D             -0.007   -0.014   -0.027   -0.040   -0.050   -0.061   -0.072
## I(X * D)      -1.002   -1.008   -1.003   -1.002   -1.001   -1.002   -0.999
##             tau= 0.8 tau= 0.9
## (Intercept)    0.083    0.092
## X              0.996    0.996
## D             -0.083   -0.092
## I(X * D)      -0.996   -0.996
```

So as expected the main effect of X and the interaction term have coefficients of 1 and -1, respectively, while the main effect of the treatment variable, D, fluctuates around its correct value, $\tau/10$. We can contrast this with what happens when we fail to condition on the covariate X. Now,

$$Q_{Y|D}(\tau|d) = Q_V(\tau)(1 - d)$$

where $V = X + U$. The quantile function of the convolution, V, because U is supported on the interval $[0, 0.1]$ is almost the same as the quantile function of the standard uniform, that is to say, the identity function. See Figure 1. Thus, $Q_{Y|D}(\tau|d) \approx \tau(1-d)$, and the QTE from this perspective is simply, $-\tau$. Powell comments that, "quantile regression with covariates [i.e. conditioning on X] is not estimating the quantile function of interest since controlling for additional covariates alters the quantile function." It is not at all clear in any general sense

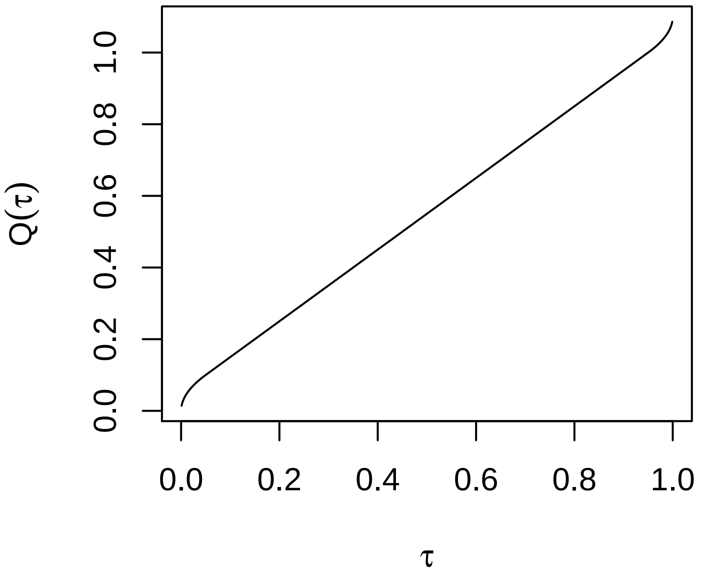

Figure 1. Quantile Function of V

what he means by the "quantile function of interest," but when D is randomly assigned, i.e. independent of X, then apparently $Q_{Y|D}$ is his intended target. But why? Wouldn't we like to know that the treatment acts quite differently on patients having different values of X? Indeed, most of the variability of the response in this, admittedly somewhat bizarre, circumstance is due to X and not to U, so ignoring X seems quite imprudent.

2. Conclusion

There are a variety of other troubling aspects of the Powell paper. Aside from the ambiguity of what might be meant by the "quantile function of interest," there are several places where assumptions are casually introduced that drastically alter the nature of the model under consideration. In particular there seems to be a implicit assumption that the instrumental variables, denoted by Z are valid thoughout the full range of the QTE. Such assumptions strike me as dangerous since in practice it is often plausible that they are only effective in some narrow range of τ. This is a point very effectively made in the fundemental paper of Chesher (2003). Computational implementation of the proposed methods by the generalized method of moments[4] is also a travesty given the non-smooth nature of the underlying problem.

[4] What is needed in such circumstances is a generalized method of mollifiers; neither of these methods should be confused with the generalized method of emoluments that seems to be the dominant feature of modern political economy.

References

Chernozhukov, V., Fernández-Val, I. and Melly, B. (2013), 'Inference on counterfactual distributions', *Econometrica* **81**(6), 2205–2268.

Chernozhukov, V. and Hansen, C. (2004), 'An IV model of quantile treatment effects', *Econometrica* **73**, 245–261.

Chesher, A. (2003), 'Identification in nonseparable models', *Econometrica* **71**, 1405–1441.

Goldberger, A. (1968), *Topics in Regression Analysis*, Collier Macmillan.

Koenker, R. (2011), 'On distributional vs. quantile regression'. Available from `http://www.econ.uiuc.edu/~roger/research/vinaigrettes/dreg.pdf`.

Ma, L. and Koenker, R. (2006), 'Quantile regression methods for recursive structural equation models', *Journal of Econometrics* **134**, 471 – 506.

Mata, J. and Machado, J. A. F. (2005), 'Counterfactual decomposition of changes in wage distributions using quantile regression', *Journal of Applied Econometrics* **20**, 445–465.

Powell, D. (2020), 'Quantile treatment effects in the presence of covariates', *Review of Economic and Statistics* **??**, ??–??