Quantile regression in varying coefficient models: non-crossingness and heteroscedasticity

Y. Andriyana[1], I. Gijbels[2] and **A. Verhasselt**[3*]

[1] *Statistics Department, Universitas Padjadjaran, Bandung, Indonesia; yudhie.andriyana@unpad.ac.id.*
[2] *Department of Mathematics and Leuven Statistics Research Center (LStat), KU Leuven, Leuven, Belgium; irene.gijbels@wis.kuleuven.be*
[3] *Censtat, Interuniversity Institute for Biostatistics and statistical Bioinformatics, Universiteit Hasselt, Belgium; anneleen.verhasselt@uhasselt.be.*
[*] *Presenting author*

Keywords. *Crossing quantile curves; Heteroscedasticity; P-splines; Quantile regression; Varying coefficient models.*

Quantile regression is an important tool for describing the characteristics of conditional distributions. In real applications, the impact of explanatory variables on a variable of interest, leads to the study of conditional quantile functions or regression quantiles. In practice, the conditional quantile functions are estimated, from data, for various fixed values of the order of the quantile τ. Conditional quantile functions are by definition, for any given fixed values of the covariates, an increasing function in the argument τ. Unfortunately estimated regression quantile curves often violate this non-crossingness property, which can be very annoying for interpretation and further analysis. There is thus an interest to prevent this crossing to happen in finite-samples.

To describe accurately complex data, one often considers regression models that are on the one hand flexible enough to capture this complexity, but on the other hand still allow for estimation methods with good practical performance. We focus on varying coefficient models, which naturally extend linear regression models by allowing the regression coefficients to change with another covariate.

We consider flexible varying coefficient models, and develop methods (based on P-splines) for quantile regression that ensure that the estimated quantile curves do not cross. A second aim is to allow for some heteroscedasticity in the error modelling, and to also estimate the associated scaling/variability function. We investigate the finite-sample performance of the discussed methods via simulation studies. Some applications to real data illustrate the use of the methods in practical settings.