1 Conformalized Quantile Regression

This chapter introduces *Conformalized Quantile Regression (CQR)* as the first of two main post-processing procedures that we implemented in the `postforecasts` package.

Section 1.1 explains the original Conformalized Quantile Regression algorithm as proposed in Romano, Patterson, and Candès (2019). There, we highlight potential limitations of the traditional implementation that could potentially be fixed by more flexible modifications, which are discussed in Section 1.2 and Section 1.3.

1.1 Traditional CQR

All derivations in this sections are taken from the original paper (Romano, Patterson, and Candès 2019). The authors motivate Conformalized Quantile Regression by stating two criteria that an ideal procedure for generating prediction intervals should satisfy:

- It should provide valid coverage in finite samples without making strong distributional assumptions
- The resulting intervals should be as short as possible at each point in the input space

According to the authors, CQR performs well on both criteria while being *distribution-free* and adaptive to *heteroscedasticity*.

1.1.1 Statistical Validity

The algorithm that CQR is build upon is statistically supported by the following Theorem:

Theorem 1.1. *If $(X_i, Y_i), i = 1, \ldots, n+1$ are exchangeable, then the $(1-\alpha) \cdot 100\%$ prediction interval $C(X_{n+1})$ constructed by the CQR algorithm satisfies*

$$P\left(Y_{n+1} \in C(X_{n+1})\right) \geq 1 - \alpha.$$

Moreover, if the conformity scores E_i are almost surely distinct, then the prediction interval is nearly perfectly calibrated:

$$P\left(Y_{n+1} \in C(X_{n+1})\right) \leq 1 - \alpha + \frac{1}{|I_2| + 1},$$

where I_2 denotes the calibration set.

Thus, the first statement of Theorem 1.1 provides a coverage *guarantee* in the sense that the adjusted prediction interval is *lower-bounded* by the desired coverage level. The second statement adds an *upper-bound* to the coverage probability which gets tighter with increasing sample size and asymptotically converges to the desired coverage level $1 - \alpha$ such that lower bound and upper bound are asymptotically identical.

1.1.2 Algorithm

The CQR algorithm is best described as a multiple step procedure.

1.2 Asymmetric CQR

1.3 Multiplicative CQR

text

Romano, Yaniv, Evan Patterson, and Emmanuel J. Candès. 2019. "Conformalized Quantile Regression." *arXiv:1905.03222 [Stat]*, May. http://arxiv.org/abs/1905.03222.