0.1 Asymmetric CQR

0.1.1 Theory

This section proposes a first extension to the original CQR algorithm by relaxing the symmetry assumption. Instead of limiting ourselves to choose the *same* margin $Q_{1-\alpha}(E, I_2)$ on a *single* score vector E for adjusting the original lower and upper quantile predictions, we allow for individual and, thus, generally different margins $Q_{1-\alpha,low}(E_{low}, I_2)$ and $Q_{1-\alpha,high}(E_{high}, I_2)$ such that the post-processed prediction interval is given by

$$C(X_{n+1}) = \left[\hat{q}_{\alpha,low}(X_i) - Q_{1-\alpha,low}(E_{low}, I_2),\ \hat{q}_{\alpha,high}(X_i) + Q_{1-\alpha,high}(E_{high}, I_2)\right].$$

This asymmetric version additionally requires a change in the computation of the conformity scores. Instead of considering the elementwise maximum of the differences between observed values Y_i and original bounds, we simply compute two separate score vectors:

$$\begin{aligned} E_{i,low} &:= \hat{q}_{\alpha,low}(X_i) - Y_i \quad \forall\, i \in I_2 \\ E_{i,high} &:= Y_i - \hat{q}_{\alpha,high}(X_i) \quad \forall\, i \in I_2 \end{aligned}$$

0.1.2 Results

To avoid repetitions with respect to **??** we only briefly mention where asymmetric and vanilla CQR have similar effects and rather focus on the differences.

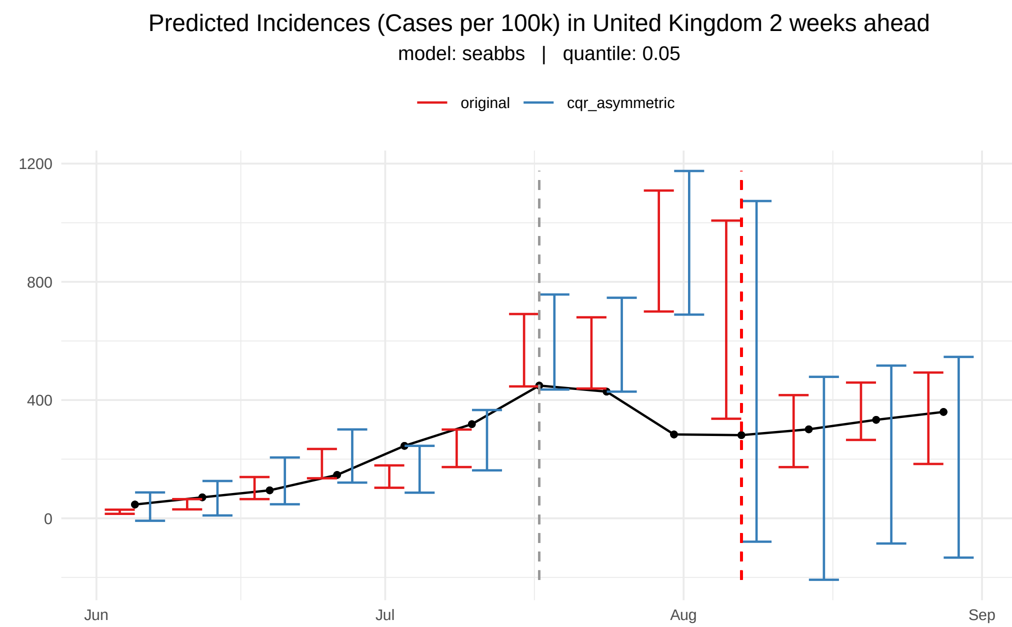

Figure 1: Illustration of CQR's slow reaction process

Figure 1 nicely demonstrates the key characteristics of asymmetric CQR: Adjustments of the lower and upper interval bounds are independent from each other. Considering the last interval on the far right the lower bound is adjusted downwards by a large amount whereas the upper bound is increased only slightly. This behaviour implies that, in contrast to traditional CQR, original and corrected prediction intervals are generally *not* centered around the same midpoint.

The plot also illustrates what we have already seen in **??**: Once the true value is not contained in the prediction interval and there is a large discrepancy towards the closest boundary, all CQR versions tend to *overcompensate* in the next time step. This jump can be observed from time step 9 to time step 10, where the latter is highlighted by the red dashed line. Even more problematic, the large correction margin only vanishes very gradually afterwards even if the observed Time Series has stabilized. In Figure 1 the lower quantile prediction of asymmetric CQR approaches the original lower quantile forecast very slowly after the jump in observed Cases.

The following paragraphs aim to explain this inflexibility in detail and draw the connection to the underlying statistical algorithm of **??**.

0.1.3 CQR Downsides

Going back to Section 0.1.1 asymmetric CQR computes two separate score vectors based on the original lower and upper quantile forecasts and the vector of observed values. To confirm our findings visually we now focus on the data subset of Figure 1.

Consider the intervals one step prior to the dashed red line. At this point in time the training set includes the first 9 elements of true values and predicted quantiles which are then used to compute a list of lower and upper scores:

```
scores_list <- compute_scores_asymmetric(
  true_values[1:9], quantiles_low[1:9], quantiles_high[1:9]
)
scores_list$scores_lower
## [1]  -31.443366  -40.808821  -29.765120  -11.289450 -141.757533 -145.173165
## [7]   -2.839344   10.514219  415.998372
```

The vector of lower scores E_{low} is given by $\hat{q}_{\alpha,low}(X) - Y$, i.e. elementwise differences of true values and predicted lower quantiles at each time step. Due to the jump from time point 9 to 10 the final element of the lower score vector has a large value of around 416.

Next, the (scalar) lower margin $Q_{1-\alpha,low}(E_{low}$ is computed:

```
margin <- compute_margin(scores_list$scores_lower, quantile)
margin
```

```
##     100%
## 415.9984
```

Due to the small sample size of 9 observations and the relatively small quantile level of 0.05 the margin is simply the *maximum* or 100% quantile of the lower scores. The *updated* lower quantile prediction for the 10th time point is simply $\hat{q}_{\alpha,low}(X_{10}) - Q_{1-\alpha,low}(E_{low})$, i.e. the original lower quantile prediction at time point 10 minus the margin:

```
quantiles_low[10] - margin
```

```
## [1] -79.18
```

which coincides with Figure 1.

The procedure now continues by consecutively adding the next elements to the vector of true values and original quantile predictions. Since the differences of observed incidences and predicted lower bounds are all much smaller for the remaining time steps, the *same* value 416 remains the maximum of the lower score vector until the end! Thus, if like in the case above, the margin always equaled the maximum score, the adjustments would stay that large independent of the future development of the time series.

In fact, the only difference from that scenario to *Step 4* of **??** is that the quantile of the score vector that determines the value of the margin depends on the *size* of the score vector. Since the size increases by one with each time step during the Cross Validation process, this quantile slowly declines. For instance, the margin which is responsible for adjusting forecasts at time point 11 is not simply the maximum anymore:

```
scores_list <- compute_scores_asymmetric(
  true_values[1:10], quantiles_low[1:10], quantiles_high[1:10]
)
```

Table 1: Performance of asymmetric CQR on Validation Set

method	uk_interval_score	hub_interval_score
cqr_asymmetric	63.97	34.37
original	65.74	29.84

```
margin <- compute_margin(scores_list$scores_lower, quantile)
margin
```

```
##      99%
## 383.5547
```

In this case the 99% quantile is an interpolation of the largest and second largest score, as implemented by the `stats::quantile()` function. Hence, even though the score outlier is not chosen directly, it strongly impacts the margins of future time steps.

The cycle proceeds in this way until the end. The conclusion of this brief case study is that all modifications of the traditional CQR algorithm suffer from a slow reaction time towards distribution shifts and particularly sudden jumps within observed values and original forecasts. This major downside of Conformalized Quantile Regression is an immediate consequence of the *margin* computation which finally determines the magnitude of forecast adjustments.

0.1.4 Global Results

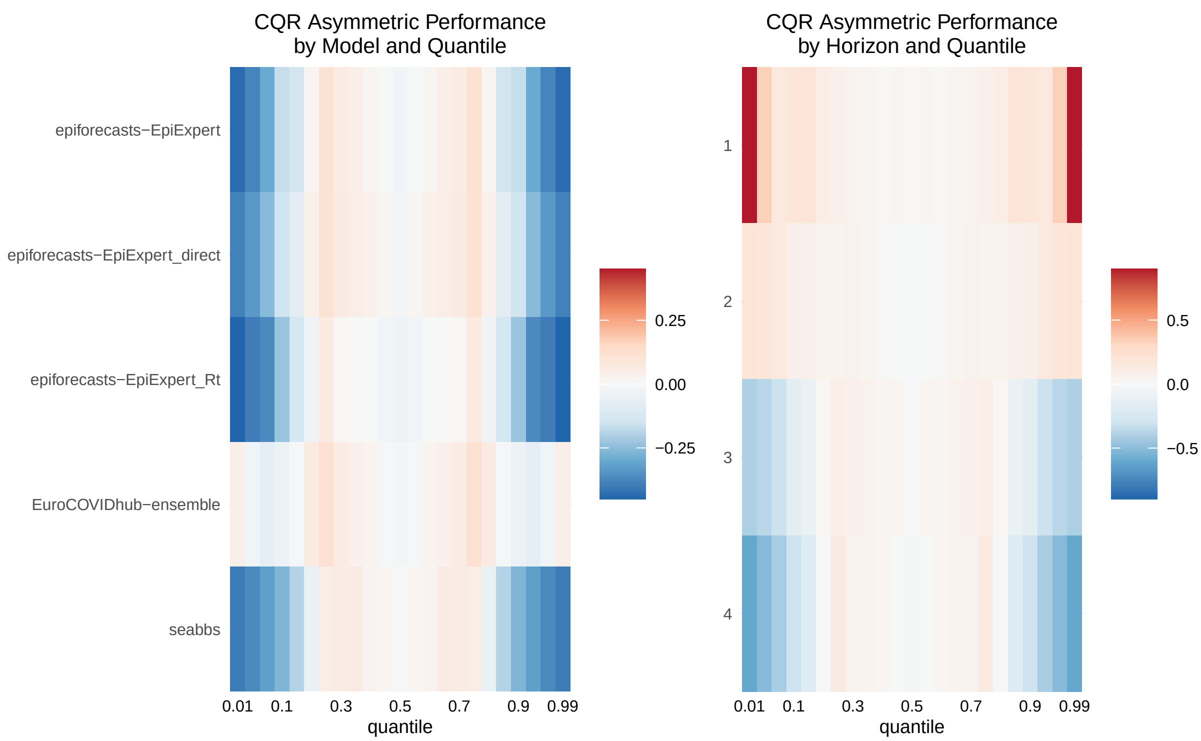

Figure 2: Mixed Results of asymmetric CQR on UK data

In summary, asymmetric Conformalized Quantile Regression can lead to improved prediction intervals as it is the case for the UK data set. However, the vast majority of countries in the European Forecast Hub did not benefit from this first CQR modification.

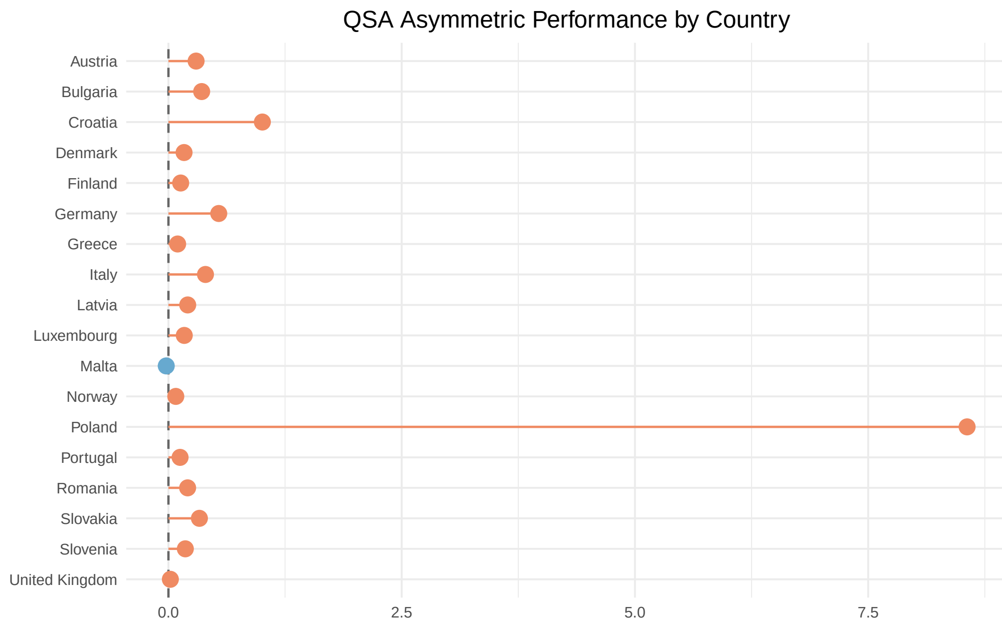

Figure 3: Asymmetric CQR has negative effects on almost all countries

Table 2: Performance of asymmetric CQR for Covid-19 Cases and Deaths

method	target_type	uk_interval_score	hub_interval_score
cqr_asymmetric	Cases	127.78	70.59
original	Cases	131.29	62.69
cqr_asymmetric	Deaths	0.16	0.50
original	Deaths	0.18	0.32

Compared to the traditional CQR algorithm, relaxing the symmetry assumption had a negative effect on the Weighted Interval Score across both data sets. It is worth noting that allowing for separate margins did *not* cause significant overfitting as one might assume, the original CQR algorithm outperforms the asymmetric version even on the training set. The magnitude of the performance differences between the two methods is analyzed in **??**.

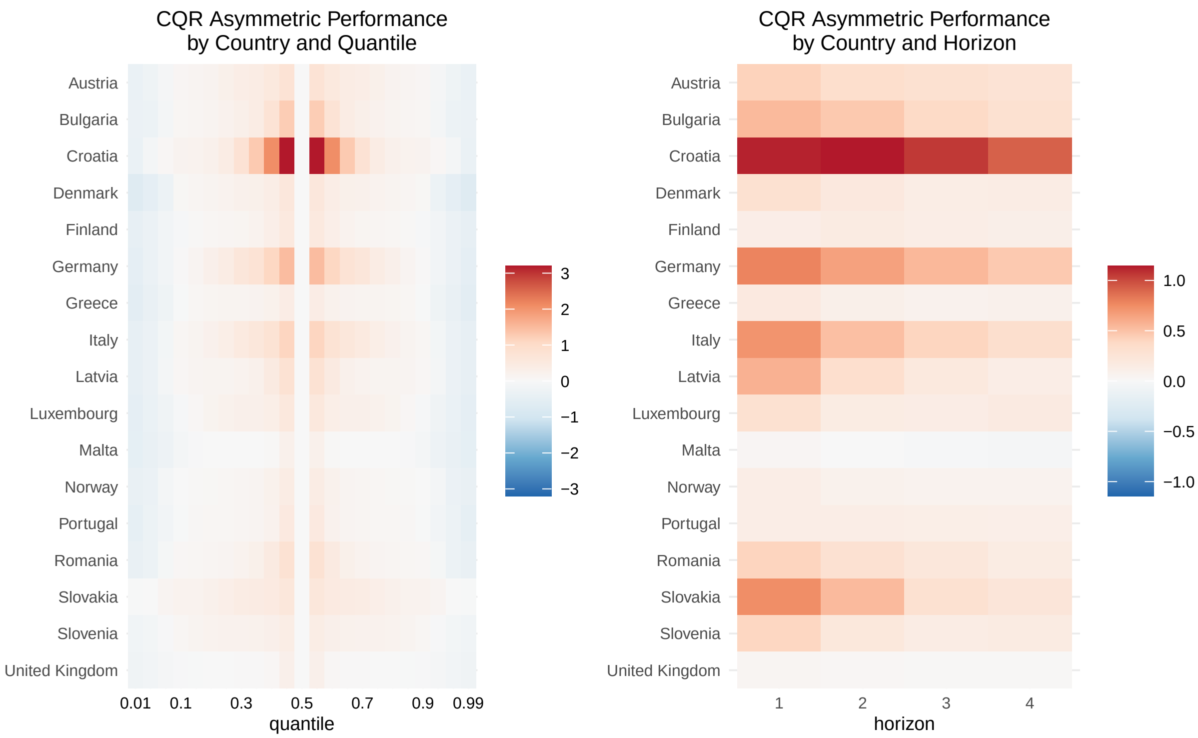

Figure 4: Performance of asymmetric CQR on European Forecast Hub data