

PAVA-Reordering Horizon Monotonicity Violations

Rike Becker

October 23, 2023

This is an outline for how the PAVA-Algorithm would apply in our case. Main reference is De Leeuw et al. [2009], which contains the most accessible explanation of the PAVA-Algorithm that I could find.

We want to obtain central prediction intervals of level $\alpha \in \{0.5, 0.8\}$, based on quantile forecasts $q_{t,h}^\tau$ for a quantity y_t . $\tau \in (0, 1)$ is the quantile level, $t \in \{1, \dots, T\}$ is target year, $h \in \{2, 8, 14, 20\}$ indexes the forecast horizon (monthly units). The length of the corresponding central prediction intervals should be monotonic with respect to h , that is,

$$c_{t,h}^\alpha \leq c_{t,h+1}^\alpha, \quad \text{for all } t, h; \quad \text{where} \quad (1)$$

$$c_{t,h}^\alpha := q_{t,h}^{1-\alpha/2} - q_{t,h}^{\alpha/2} \quad (2)$$

Assuming we have a minimal information set \mathcal{A}_0 consisting only of past forecast-observation pairs $(\hat{y}_{t,h}, y_t)_{t=1}^T$ and further assuming (some sort of) stationarity in the series of forecast errors $\hat{e} = y_t - \hat{y}_{t,h}$, the optimal τ -quantile forecast in terms of the weighted interval score should be the respective τ -quantile in the set of past forecast errors

$$\mathcal{E}_{t,h} = \{\hat{e}_{t^*,h} | t - R \leq t^* < t\}, \quad (3)$$

where we denote the value of the quantile with $q_{t,h}^{\tau, \mathcal{E}}$.¹

Of course, while there is a natural tendency for the quantiles obtained from (3) to fulfill constraint (1), there is no guarantee for this and we practically observe a considerable proportion of violations in our setting.

I would propose that applying the pool-adjacent-violators algorithm as detailed in section 3.1. in De Leeuw et al. [2009] to this problem would look like the following:²

The initial solution is simply $q_h^{(0)} := q_h^\mathcal{E}$. The index for the blocks is $r = 1, \dots, B$ where at the initial step we have $B := 4$. Then:

1. Merge $q^{(l)}$ -values from blocks r and $r + 1$ into block r if $c_{r+1}^{(l)} < c_r^{(l)}$
2. Choose $q_r^{(l+1)}$ such that our loss function is optimized.³
3. Go back to 1. in case of any further violations.

Basically, we would simply merge horizons if they don't uphold the natural ordering for a given α and set the respective quantiles to the same (average) value.

Some things to note/discuss:

1. To my (albeit relatively new) knowledge of the algorithm, it is unusual that the violators are not registered on the same level that they are treated. That is, here we check for violators on the level of the central interval lengths, but we treat them directly on the quantiles.

¹If we set $R = 21$, we can extract quantiles directly from the order statistics, simplifying the question of which interpolation method would be optimal.

²For a given t and α - respective indices dropped for simplicity.

³I would suspect that in terms of the WIS, this would amount to taking an average of the respective quantiles.

2. This is inherently a "univariate" method with respect to the horizon monotonicity constraint. We of course have a second monotonicity constraint with respect to the quantile levels, that is, we require $q^\tau \leq q^{\tau'}$ for $\tau < \tau'$ (note that this constraint is again formulated for the quantiles, not for the length of the central intervals). This could simply be accounted for with a step 3., where we either (a) automatically pool all other quantiles in block r or (b) do so in cases where quantile ordering is not upheld across that block.
3. Note that this document was formulated for the *directional* error method. This is due to the fact that I suspect that the quantiles of (3) are the optimal solution in that case only. In the case of absolute error treatment, choosing the new pooled value in step 2. of the algorithm could perhaps not be theoretically motivated in terms of a loss function, at least not in a way that is immediately obvious to me.⁴ At least, in the case of absolute errors, we would diagnose and treat violations on the same level (that is, length of the central intervals).

References

- J. De Leeuw, K. Hornik, and P. Mair. Isotone optimization in r: pool-adjacent-violators algorithm (pava) and active set methods. *Journal of statistical software*, 32(5):1–24, 2009.

⁴I would still argue in favor of absolute error treatment however, as it empirically shows much better coverage levels.