Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization

Wickramasuriya, Athanasopoulos & Hyndman, (2019)

Sven Pappert

06.02.2025

TU Dortmund University

pappert@statistik.tu-dortmund.de

Contents

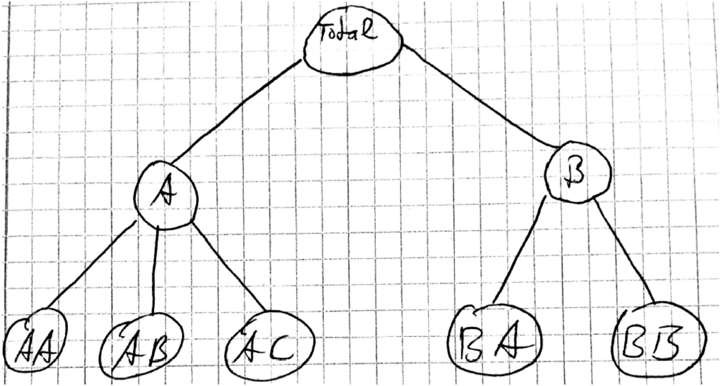
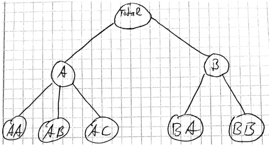


Figure: An easy example of a hierarchy. (Recreated from Wickramasuriya et al., 2017). AA, AB and AC have to sum up to A, BA and BB to B and B and A to Total.



- The time series $\{y_{it}\}_{i\in\{1,\dots,8\}}$ underlies a linear restriction.
- The restrictions can be comprised into the following equation:

$$y_t = \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ & & I_5 & & \end{pmatrix}}_{=:S} b_t, \tag{1}$$

- $b_t = (AA_t, AB_t, \ldots, BB_t)^T$ are called *bottom time series*,
- S is the *Summing Matrix*,
- In general: $y_t \in \mathbb{R}^n, b_t \in \mathbb{R}^m, S \in \{0,1\}^{n \times m}$.

- Now we (point-)forecast the time series h-steps ahead
- We obtain $\{\hat{y}_{it}(h)\}_{i\in\{1,\ldots,n\}}$.
- In general, these *base forecasts* are *incoherent*, meaning that they do not fulfill the hierarchical restrictions, i.e.

$$\hat{y}_t(h) \neq S\hat{b}_t(h). \tag{2}$$

- Coherency can be enforced by constructing the *reconciled forecasts* as

$$\tilde{y}_t(h) = S\tilde{b}_t(h) = SP\hat{y}(h), \quad S \in \mathbb{R}^{n\times m}, P \in \mathbb{R}^{m\times n} \tag{3}$$

- The question is: How to choose $\tilde{b}_t$, respectively P?

Reconciliation Strategies

Reconciliation strategies from Hyndman et al., 2011

1) Bottom-up reconciliation: $\tilde{b}_t(h) = \hat{b}_t(h)$, respectively $G = (0_{m\times(n-m)}|I_m)$.
2) Top-down reconciliation: $G = ((p_1, \ldots, p_m)^T|0_{m\times(n-1)})$, with $\sum_{i=1}^{m} p_i = 1, p_i > 0 \; \forall i \in \{1, \ldots, m\}$.

3) GLS reconciliation:

- Cast the relation between $\hat{y}_t(h)$ and $b_t\tilde{(h)}$ as regression problem,

$$\hat{y}_t = S\tilde{b}_t(h) + \varepsilon_t \tag{4}$$

- Find $\tilde{b}_t(h)$ such that MSE is minimized.
- The BLUE for $\tilde{b}_t(h)$ is the GLS estimator $(S^T\Sigma^{-1}S)^{-1}S^T\Sigma^{-1}\hat{y}_t(h)$, where $\Sigma = \text{Cov}(\varepsilon_t, \varepsilon_t^T)$.

$$\Rightarrow \tilde{y}_t(h) = S(S^T\Sigma^{-1}S)^{-1}S^T\Sigma^{-1}\hat{y}_t(h). \tag{5}$$

Unbiasedness

- Let $\hat{y}_t(h)$ be unbiased predictions,
- So $\mathbb{E}(\hat{y}_t(h)|\mathcal{F}_t) = \mathbb{E}(y_t(h)|\mathcal{F}_t) = S\mathbb{E}(b_t(h)|\mathcal{F}_t) =: S\beta_t(h)$ $(*)$
- We require that also $\tilde{y}_t(h) = SP\hat{y}_t(h)$ should be unbiased,

$$\mathbb{E}(\tilde{y}_t(h)|\mathcal{F}_t) = SP\mathbb{E}(\hat{y}_t(h)|\mathcal{F}_t) \overset{*}{=} SPS\beta \tag{6}$$

- Now Wickramasuriya et al., (2019) show that Σ is unidentifiable.
- To see this notice that from

$$\tilde{\varepsilon}_h = \hat{y}_{t+h} - \tilde{y}_{t+h} = (I - SG)\hat{y}_{t+h} \tag{7}$$

  we get

$$\text{Var}(\varepsilon_h | \mathcal{F}_t) = (I - SG)\Sigma_h (I - SG)^T \tag{8}$$

- $(I - SG)$ is not invertible, this follows from $SGS = S$, which is a consequence of requiring unbiasedness.
- GLS reconciliation is not possible!

MinT Reconciliation

MinT Reconciliation

Minimum Trace (MinT) reconciliation:

- Minimize tr(Var$(y_{t+h} - \tilde{y}_t(h))) =$ tr(Var$(y_{t+h} - SG\hat{y}_t(h)))$ subject to $SPS = S$.
- $\Rightarrow \tilde{y}_t(h) = S(S^T W_h^{-1} S)^{-1} S^T W^{-} \hat{y}_t(h)$
- Where W is the covariance matrix of base forecast errors, not coherency errors.
- The MinT approach is unbiased ($SPS = S$) and
- minimizes the variance of predictions.

- The MinT predictions fulfill:

$$
\begin{aligned}
&[y_{t+h} - \tilde{y}_t(h)]W_h^{-1}[y_{t+h} - \tilde{y}_t(h)] && (9)\\
\leq&[y_{t+h} - \hat{y}_t(h)]W_h^{-1}[y_{t+h} - \hat{y}_t(h)] && (10)
\end{aligned}
$$

- MinT Reconciliation is beneficial beyond coherency.

- The question remains on how to estimate W_h.
- Possible approaches from the paper ($k_h > 0$):
  - $W_h = k_h I$,
  - $W_h = k_h diag(\hat{W}_1)$,
  - $W_h = k_h,\ \Lambda = diag(S\mathbf{1})$,
  - $W_h = k_h \hat{W}_1$

Open Research

- (Point-)forecast reconciliation by $\tilde{y}_t(h) = S(g \circ \hat{y}_t(h))$.
- Probabilistic-forecast reconciliation,
- Temporal Hierarchies,
- Artificial Hierarchies.

Literature

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G. & Shang, H. L. (2011)
Optimal combination forecasts for hierarchical time series
Computational statistics & data analysis 55(9) 2579–2589.

Wickramasuriya, S. L., Athanasopoulos, G., & Hyndman, R. J. (2018).
Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization
Journal of the American Statistical Association.

Thanks for your Attention!