HTSeer

Elsevier[1]

Radarweg 29, Amsterdam

Elsevier Inc[a,b], *Global Customer Service*[b,∗]

[a]*1600 John F Kennedy Boulevard, Philadelphia*
[b]*360 Park Avenue South, New York*

Abstract

Keywords: Hierarchical Time Series, reconciliation methods.

---

1. Hierarchical Time Series forecasting

Hierarchical Time Series (HTS) has become a promising field of study, with many applications in the industry sector, government planning, trading and other topics. The idea is to generate forecasts that are coherent across an entire hierarchical structure with the best results on accuracy measures as possible.

In order to define a HTS, let $\mathbf{y_t}$ be a vector of size m containing values from all hierarchical levels at a given time t. In addition, consider a matrix $\mathbf{R}$ of dimension $m \times n$ such that we can write the following equation.

$$\mathbf{y_t} = \mathbf{R}\mathbf{y_t^g} \tag{1}$$

where $\mathbf{y_t^g}$ is a vector of size n that contains values in the most granular g-level of the hierarchical structure.

The idea of forecast reconciliation in HTS relies on a similar framework. Let $\hat{\mathbf{y}}_{t+h|t}$ be a vector of h-steps ahead forecasts with the same size of $\mathbf{y_t}$. For a given

[⋆]Fully documented templates are available in the elsarticle package on CTAN.

[∗]Corresponding author
Email address: support@elsevier.com (Global Customer Service)
URL: www.elsevier.com (Elsevier Inc)

[1]Since 1880.

matrix $\mathbf{G}$ of dimension $n \times m$ we have the construct the following equation.

$$\tilde{\mathbf{y}}_{t+h|t} = \mathbf{RG}\hat{\mathbf{y}}_{\mathbf{t+h|t}} \tag{2}$$

On the LHS we have $\tilde{\mathbf{y}}_{t+h|t}$ which are the reconciled forecasts. The $\mathbf{RG}$ matrices makes the reconciliation process, which transforms incoherent forecasts into coherent ones according to a aggregation constraint.

There are many strategies of time series reconciliation. Depending on the way we set $\mathbf{G}$, it is possible to adjust the equation (2) to different reconciliation approaches.

When making $\mathbf{G} = \left[\mathbf{0}_{\mathbf{n}\times(\mathbf{m}-\mathbf{n})}|\mathbf{I}_{\mathbf{n}}\right]$ where $0_{n\times(m-n)}$ is a null matrix, we have the *bottom-up* approach. Conversely, *top-down* forecasts can be achieved by making $\mathbf{G} = \left[\mathbf{p}|\mathbf{0}_{\mathbf{n}\times(\mathbf{m}-\mathbf{1})}\right]$, where $\mathbf{p}$ corresponds to proportions of forecasts[1]. There are different strategies to obtain these proportions. In this work, we considered the following ones:

- Average of historical proportions [2]:

$$p_j = \frac{1}{T}\sum_{t=1}^{T}\frac{y_{j,t}}{y_t} \tag{3}$$

  for $j = 1, \ldots, m$.

- Proportion of historical averages [2]:

$$p_j = \sum_{t=1}^{T}\frac{y_{j,t}}{T}\Bigg/\sum_{t=1}^{T}\frac{y_t}{T} \tag{4}$$

  for $j = 1, \ldots, m$.

- Proportion of forecasts [1].

$$p_j = \prod_{k=0}^{K-1}\frac{\hat{y}_{j,h}^{(k)}}{\hat{S}_{j,h}^{(k+1)}} \tag{5}$$

where k stands for the number of levels of the hierarchy, $\hat{y}_{j,h}^{(k)}$ is the h-steps ahead forecast of the series that corresponds to the node that is k levels above node j and $\hat{R}^{(k)}_{j,h}$ is the sum of h-steps forecasts ahead under the node which is k levels above node j, where $j = 1, \ldots, m$.

The optimal combination or reconciliation can be expressed according to the following regression model:

$$\hat{\mathbf{y}}_{t+h|t} = T\beta_{\mathbf{t+h|t}} + \epsilon_{\mathbf{t+h|t}} \tag{6}$$

where $\beta_{\mathbf{t+h|t}} = \mathbf{E}\left[\mathbf{y^b_{t+h}}|\mathcal{I}_\mathbf{t}\right]$, $\mathcal{I}_t = y_1, y_2, \ldots, y_t$ and $V\left(\epsilon_{t+h|t}|\mathcal{I}_t\right) = \Sigma_h$.

The Minimum Trace ($MinT$) reconciliation approach, which is an improvement of the optimal reconciliation, was introduced in [3]. This approach aims to find a matrix $\mathbf{G}$ that minimizes $tr(\mathbf{RGW_hG'R'})$ such that $\mathbf{RGR} = \mathbf{R}$, which corresponds to the unbiasedness condition.

Given the difficulty in estimating $\mathbf{W_h}$ matrix, there are different ways of doing it in an approximate fashion. In this work we considered the following ones:

- Ordinary least squares estimator (OLS): $\mathbf{W_h} = \mathbf{k_h I}$ where $k_h > 0$. This is the simplest hypothesis and means assuming that the $\mathbf{G}$ matrix is independent of the data([1]). This is basically the original formulation of the optimal combination.

- Weighted least squares estimator with variance scaling:$\mathbf{W_h} = \mathbf{k_h}\mathrm{diag}(\hat{\mathbf{W}}_\mathbf{1})$ where $k_h > 0$ and $\hat{\mathbf{W}}_1$ is the estimator of the covariance of one-step-ahead base forecast errors. In this approach, the off-diagonal elements are set to zero [4].

- Weighted least squares estimator with structural scaling: $\mathbf{W_h} = \mathbf{k_h\Lambda}$ where $k_h > 0$, $\mathbf{\Lambda} = \mathrm{diag}(\mathbf{R1})$ and $\mathbf{1}$ is a vector of ones.

The reconciliation strategies rely on producing a set of base forecasts and then combining then into coherent ones. In order to produce our base forecasts, we chose three well-known forecasting approaches: Exponential Smoothing, the ARIMA formulations and Prophet.

References

[1] G. Athanasopoulos, R. A. Ahmed, R. J. Hyndman, Hierarchical forecasts for australian domestic tourism, International Journal of Forecasting 25 (1) (2009) 146–166.

[2] C. W. Gross, J. E. Sohl, Disaggregation methods to expedite product line forecasting, Journal of Forecasting 9 (3) (1990) 233–254.

[3] S. L. Wickramasuriya, G. Athanasopoulos, R. J. Hyndman, Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization, Journal of the American Statistical Association 114 (526) (2019) 804–819.

[4] R. J. Hyndman, A. J. Lee, E. Wang, Fast computation of reconciled forecasts for hierarchical and grouped time series, Computational statistics & data analysis 97 (2016) 16–32.