ETS Time series model

Chapter 8. Forecasting: principles and practice (Hyndman & Athanasopoulos, 2018)

Yujin Kim 11-26-2022

Exponential smoothing (ETS) models

- Exponential smoothing models are based on a description of **the trend and seasonality (+ an error term)** in the data.
- ARIMA models aim to describe the autocorrelations in the data.
- Forecasts produced using exponential smoothing methods are ***weighted averages of past observations***, with the weights decaying exponentially as the observations get older.
    - the more recent the observation the higher the associated weight.

Simple exponential smoothing (SES)

The naïve method

- Suitable for forecasting data with no clear trend or seasonality
- all forecasts for the future are equal to the last observed value of the series

$$\hat{y}_{T+h|T} = y_T,$$

- all future forecasts are equal to a simple average of the observed data

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} y_t,$$

- However, we want to attach larger weights to more recent observations than to observations from the distant past.

Simple exponential smoothing (SES)

Weighted averages

- the weights decrease exponentially as observations come from further in the past — the smallest weights are associated with the oldest observations

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots,$$

- where $0 \le \alpha \le 1$ is the smoothing parameter

- For any α between 0 and 1, the weights attached to the observations decrease exponentially going back in time, indicates the name “exponential smoothing”.

- Component form:

$$\begin{aligned} \text{Forecast equation} \qquad \hat{y}_{t+h|t} &= \ell_t \\ \text{Smoothing equation} \qquad \ell_t &= \alpha y_t + (1-\alpha)\ell_{t-1}, \end{aligned}$$

Methods with trend

Holt's linear trend method

- A forecast equation and two smoothing equations

$$\begin{aligned} \text{Forecast equation} \quad & \hat{y}_{t+h|t} = \ell_t + h b_t \\ \text{Level equation} \quad & \ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\ \text{Trend equation} \quad & b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*) b_{t-1}, \end{aligned}$$

- α is the smoothing parameter for the level (0 ≤ α ≤ 1), a weighted average of observation y_t and the one-step-ahead training forecast for time t.
- β_* is the smoothing parameter for the trend (0 ≤ β ≤ 1), a weighted average of the estimated trend at time t, and the previous estimate of the trend.

Methods with trend

Damped trend methods

- The forecasts generated by Holt's linear method display a constant trend (increasing or decreasing) indefinitely into the future. => over-forecast
- Gardner & McKenzie (1985) introduced a parameter that "dampens" the trend to a flat line some time in the future. If $\phi = 1$, the method is identical to Holt's linear method. For values between 0 and 1, ϕ dampens the trend so that it approaches a constant some time in the future.

$$
\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + \underline{(\phi + \phi^2 + \cdots + \phi^h)} b_t \\
\ell_t &= \alpha y_t + (1-\alpha)(\ell_{t-1} + \phi b_{t-1}) \\
b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}.
\end{aligned}
$$

Methods with seasonality

Holt-Winters's extension

- Seasonal variations are roughly constant through the series — ***additive method*** (preferred)
- Seasonal variations are changing proportional to the level of the series — ***multiplicative method*** (preferred)

Methods with seasonality

Holt-Winters's additive term

- Component form:

$$
\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + hb_t + s_{t+h-m(k+1)} \\
\ell_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\
s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m},
\end{aligned}
$$

- *k* is the integral part of (h-1)/m, which ensures that the estimates of the seasonal indices used for forecasting come from the final year of the sample.

- The usual parameter restriction: $0 \leq \gamma \leq 1 - \alpha$

Methods with seasonality

Holt-Winters's multiplicative term

- Component form:

$$\hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$$

$$\ell_t = \alpha\frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$$

$$s_t = \gamma\frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1-\gamma)s_{t-m}.$$

A taxonomy of exponential smoothing methods

(T, S) — Trend and Seasonal components

- By considering variations in the combinations of the trend and seasonal components, nine exponential smoothing methods are possible.

Trend Component	Seasonal Component		
	N	**A**	**M**
	(None)	**(Additive)**	**(Multiplicative)**
N (None)	(N,N)	(N,A)	(N,M)
A (Additive)	(A,N)	(A,A)	(A,M)
A_d (Additive damped)	(A_d,N)	(A_d,A)	(A_d,M)

Short hand	Method
(N,N)	Simple exponential smoothing
(A,N)	Holt's linear method
(A_d,N)	Additive damped trend method
(A,A)	Additive Holt-Winters' method
(A,M)	Multiplicative Holt-Winters' method
(A_d,M)	Holt-Winters' damped method