ETF3231/5231 Week 5 - Exponential smoothing, Part 1

Sao Yang Hew

Friday - 30 March 2023 (12:30pm - 2:00pm)

2023

Table of Contents

Intuitively

Intuitively, exponential smoothing allows us to represent time series y_t and predictions $\hat{y}_{t+1|t}, \hat{y}_{t+2|t}, \ldots$ as a **weighted average** of past time series, these can be:

- Past time series of itself, $y_t, y_{t-1}, \ldots$
- Or past time series of **latent, unobservable components** we need to derive from y_t. These include:
  - Smoothed values of y_t, which we refer to as l_t
  - Trend values of y_t, which we refer to as b_t
  - Seasonal values of y_t, which we refer to as s_t
- We need to note that l_t, b_t and s_t, are all components that depend on y_t and their past values. They are all functions of y_t
- Essentially, the exponential smoothing method just takes weighted averages of the past to predict the future! though how the weighted averages are constructed is complex and clever

Exponential smoothing in component form

We first begin with writing out the component form of a simple exponential smoothing method (no trend, no seasonality):

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

Notice how the **smoothing equation** is a weighted average of **past time series** y_t and **unobservable components** b_t

Exponential smoothing in component form

$$
\begin{aligned}
\hat{y}_{t+h|t} &= l_t \\
l_t &= \underbrace{\alpha y_t}_{\text{weighted average of past value}} + \underbrace{(1-\alpha)l_{t-1}}_{\text{weighted average of unobservable component}} \\
&= \alpha y_t + (1-\alpha)[\alpha y_{t-1} + (1-\alpha)l_{t-2}] \\
&= \alpha y_t + (1-\alpha)\alpha y_{t-1} + (1-\alpha)^2 l_{t-2} \\
&= \alpha y_t + (1-\alpha)\alpha y_{t-1} + (1-\alpha)^2[\alpha y_{t-2} + (1-\alpha)l_{t-3}] \\
&\ldots \\
l_t &= \underbrace{f(y_t, y_{t-1}, y_{t-2}, \ldots|\alpha, l_0)}_{\text{weighted average of past values}} \\
\therefore \hat{y}_{t+h|t} &= f(y_t, y_{t-1}, y_{t-2}, \ldots|\alpha, l_0)
\end{aligned}
$$

l_0 and α needs to be estimated, e.g., using the maximum-likelihood estimation, or Bayesian methods (more on this next week).

Coding exercise: quick demo

In this class we will also run-through a quick demo - on how the simple exponential smoothing method can be coded up as a function in R.

- This is so we could give you a taste of what it's like to work on R code on the back-end. In the unit, we only need to use these functions to produce results, rather than programming from the ground up.
- How are `tsibble`, `fable` and the other packages developed?
- No need to understand this in detail - they are **not assessed**.
- For more information on how to code functions, please have a look at the Advanced R textbook recommended on Week 1 (adv-r.hadley.nz)

Adding seasonality and trend

If we add a trends and seasonality components (assuming undamped, and with additive seasonality):

$$
\begin{aligned}
\text{Forecast equation} \quad & \hat{y}_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)} \\
\text{Level equation} \quad & l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\
\text{Trend equation} \quad & b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\
\text{Season equation} \quad & s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}
\end{aligned}
$$

Notice how these equations are similar to the simple method before? They are weighted averages of **past values** and **unobservable components**.

Adding seasonality and trend

$$
\begin{aligned}
\hat{y}_{t+h|t} &= l_t + hb_t + s_{t+h-m(k+1)} \\
l_t &= \underbrace{\alpha(y_t - s_{t-m})}_{\text{past and unobservable component}} + \underbrace{(1-\alpha)(l_{t-1} + b_{t-1})}_{\text{unobservable component}} \\
b_t &= \underbrace{\beta(l_t - l_{t-1})}_{\text{unobservable component}} + \underbrace{(1-\beta)b_{t-1}}_{\text{unobservable component}} \\
s_t &= \underbrace{\gamma(y_t - l_{t-1} - b_{t-1})}_{\text{unobservable component}} + \underbrace{(1-\gamma)s_{t-m}}_{\text{unobservable component}}
\end{aligned}
$$

Notice how every equation is essentially a weighted average?

Adding seasonality and trend

$$
\begin{aligned}
\hat{y}_{t+h|t} &= l_t + hb_t + s_{t+h-m(k+1)} \\
l_t &= g(y_t, y_{t-1}, y_{t-2}, ...|\alpha, l_0, s_0, b_0) \\
b_t &= h(y_t, y_{t-1}, y_{t-2}, ...|\beta, l_0, b_0) \\
s_t &= k(y_t, y_{t-1}, y_{t-2}, ...|\gamma, l_0, b_0, s_0)
\end{aligned}
$$

And in a nutshell:

$$
\hat{y}_{t+h|t} = f(\underbrace{y_t, y_{t-1}, y_{t-2}, ...}_{\text{past observations}} | \underbrace{\alpha, \beta, \gamma, l_0, b_0, s_0}_{\text{parameters to estimate}})
$$

With the damped trend method though:

$$
\hat{y}_{t+h|t} = f(\underbrace{y_t, y_{t-1}, y_{t-2}, ...}_{\text{past observations}} | \underbrace{\alpha, \beta, \phi, \gamma, l_0, b_0, s_0}_{\text{parameters to estimate}})
$$

Revisit intuition from before

- Exponential smoothing allows us to represent y_t and predictions $\hat{y}_{t+1|t}, \hat{y}_{t+2|t}, ...$ as a **weighted average** of past time series.
- These **weights** come in the form of up to 7 **parameters** $\alpha, \beta, \phi, \gamma, l_0, s_0, b_0$
- These **parameters** help capture the **trend-cycle component** and **seasonal component** of our time series
- All of our forecasts would then just be a **function of past values** and these **parameters**:

$$\hat{y}_{t+h|t} = f(y_t, y_{t-1}, y_{t-2}, ...|\alpha, \beta, \phi, \gamma, l_0, b_0, s_0)$$

Weeks 6 and 7 will teach

- We have learnt about the **component form** equations of exponential smoothing
- We are going to learn about representing exponential smoothing equations in **state-space form** (it's the same intuition, but just represented a bit differently)
- We are also going to learn about using the `ETS()` model in modelling and forecasting, also **revise and remember**:
    - The benchmark models (e.g., naive), because we can compare these simple models to the more complex, ETS model.
    - How to `model()`, `forecast()` and `autoplot()` forecasts.
    - How do we assess residual diagnostics (`gg_tsresiduals()`).
    - How do we evaluate our forecasts using **train-test split** and **time series cross validation**