QBUS6840 Lecture 5

Forecasting with Exponential Smoothing

Discipline of Business Analytics

The University of Sydney Business School

Recap...

In W3, we looked at Time Series Decomposition, mostly suitable for interpretation.

In W4, we have looked at forecasting with many time series:

- how to use linear regression for time series forecasting
- different covariates/predictors and how to select them
- several ways to check the modelling assumptions

In W5, we will study an important technique for forecasting (with a single time series): exponential smoothing.

Table of contents

Readings

Online textbook Sections 7.1-7.3, or BOK Sec 8.1-8.3

How this lecture is structured

We will provide the formulae for each forecast method

- these might look completely unintuitive at first!
- we will try to visualise each

We will next provide the underlying statistical model and assumptions for each method

Two equivalent forms will be discussed:

- Component form: easier to interpret and extend
- Error correction form: easier to derive the forecast intervals/uncertainty

Some slides will be dry and heavy with mathematical notations!

Outline

Simple Exponential Smoothing

Trend corrected exponential smoothing (TCES)

Exponential smoothing methods

- Exponential smoothing is the most basic, yet very successful forecasting method, developed in the 1950s. The idea of exponential smoothing has motivated the most successful forecasting methods being used nowadays.

- In simple terms, exponential smoothing forecasts are weighted averages of previous observations. The weights decay exponentially as we go further into the past.

Simple exponential smoothing (SES)

- Naïve Method

$$\widehat{y}_{T+1|1:T} = y_T$$

puts a weight of 1 to the latest observation

- Overall Average Method

$$\widehat{y}_{T+1|1:T} = \frac{1}{T}\sum_{t=1}^{T} y_t = \frac{1}{T}(y_T + y_{T-1} + \cdots + y_1)$$

puts the same weight of $1/T$ to ALL observations.

- Something between two extremes: $0 \leq \alpha \leq 1$

$$\widehat{y}_{T+1|1:T} = \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots + \alpha(1-\alpha)^{T-1} y_1$$

Simple exponential smoothing (SES)

- Naïve Method

$$\widehat{y}_{T+1|1:T} = y_T$$

  puts a weight of 1 to the latest observation
- Overall Average Method

$$\widehat{y}_{T+1|1:T} = \frac{1}{T}\sum_{t=1}^{T} y_t = \frac{1}{T}(y_T + y_{T-1} + \cdots + y_1)$$

  puts the same weight of $1/T$ to ALL observations.
- Something between two extremes: $0 \leq \alpha \leq 1$

$$\widehat{y}_{T+1|1:T} = \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots + \alpha(1-\alpha)^{T-1} y_1$$

Simple exponential smoothing (SES)

- Naïve Method

$$\widehat{y}_{T+1|1:T} = y_T$$

  puts a weight of 1 to the latest observation
- Overall Average Method

$$\widehat{y}_{T+1|1:T} = \frac{1}{T}\sum_{t=1}^{T} y_t = \frac{1}{T}(y_T + y_{T-1} + \cdots + y_1)$$

  puts the same weight of $1/T$ to ALL observations.
- Something between two extremes: $0 \le \alpha \le 1$

$$\begin{aligned}\widehat{y}_{T+1|1:T} &= \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots + \alpha(1-\alpha)^{T-1} y_1 \\ &= \alpha y_T + (1-\alpha)[\alpha y_{T-1} + \alpha(1-\alpha) y_{T-2} + \cdots + \alpha(1-\alpha)^{T-2} y_1]\end{aligned}$$

Simple exponential smoothing (SES)

- Naïve Method
$$\widehat{y}_{T+1|1:T} = y_T$$
puts a weight of 1 to the latest observation
- Overall Average Method
$$\widehat{y}_{T+1|1:T} = \frac{1}{T}\sum_{t=1}^{T} y_t = \frac{1}{T}(y_T + y_{T-1} + \cdots + y_1)$$
puts the same weight of $1/T$ to ALL observations.
- Something between two extremes: $0 \leq \alpha \leq 1$

$$\begin{aligned}
\widehat{y}_{T+1|1:T} &= \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots + \alpha(1-\alpha)^{T-1} y_1 \\
&= \alpha y_T + (1-\alpha)[\alpha y_{T-1} + \alpha(1-\alpha)y_{T-2} + \cdots + \alpha(1-\alpha)^{T-2} y_1] \\
&= \alpha y_T + (1-\alpha)\widehat{y}_{T|1:T-1}
\end{aligned}$$

Simple exponential smoothing (SES)

- SES for forecasting in the Weighted Average Form

$$\widehat{y}_{t+1|1:t} = \alpha y_t + (1-\alpha)\widehat{y}_{t|1:t-1}.$$

The forecast at time $t+1$ is equal to a weighted average between the most recent observation y_t and the most recent forecast $\widehat{y}_{t|t-1}$.

Two Alternative Forms of SES

- The Component Form

$$\begin{aligned} l_t &= \alpha y_t + (1-\alpha) l_{t-1}, \qquad 0 \le \alpha \le 1. \\ \widehat{y}_{t+1|1:t} &= l_t. \end{aligned}$$

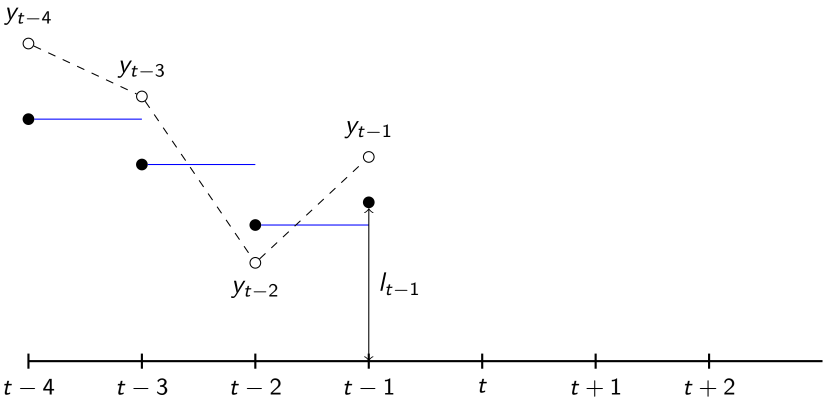
  l_t is called the level (or the smoothed value) of the series at time t. We first calculate the level l_t, then use it as the forecast $\widehat{y}_{t+1|1:t}$.

- The Error Correction Form

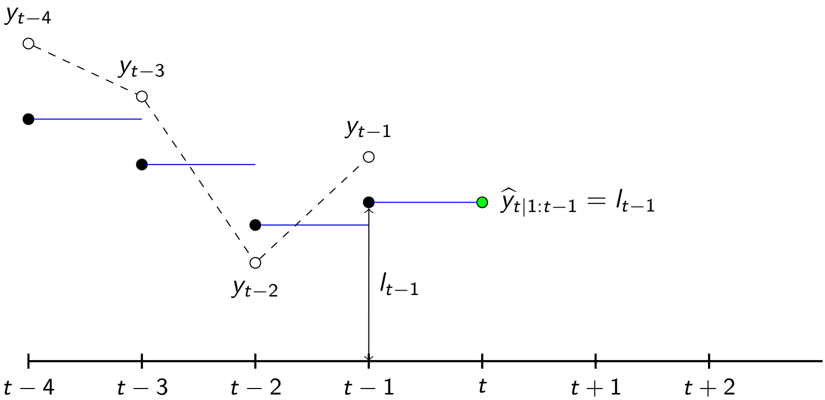
$$\begin{aligned} \widehat{y}_{t+1|1:t} &= l_t. \\ l_t &= \alpha y_t + (1-\alpha) l_{t-1} = l_{t-1} + \alpha (y_t - l_{t-1}) \\ &= l_{t-1} + \alpha \varepsilon_t \end{aligned}$$

  where $\varepsilon_t = y_t - l_{t-1} = y_t - \widehat{y}_{t|1:t-1}$ is the forecast error at time t.

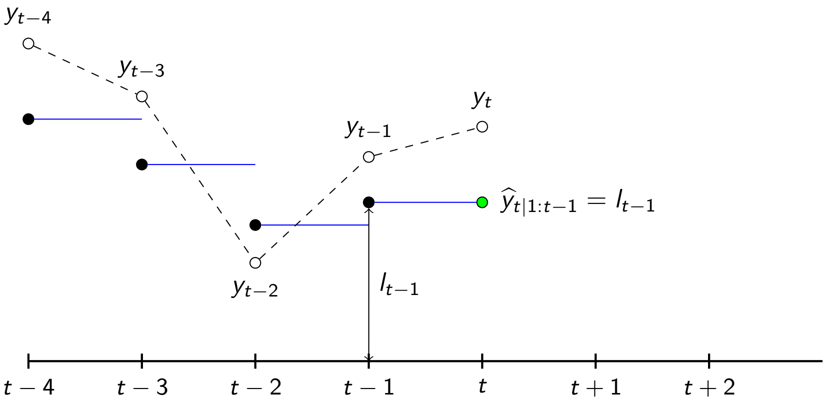
Explanation: Simple exponential smoothing



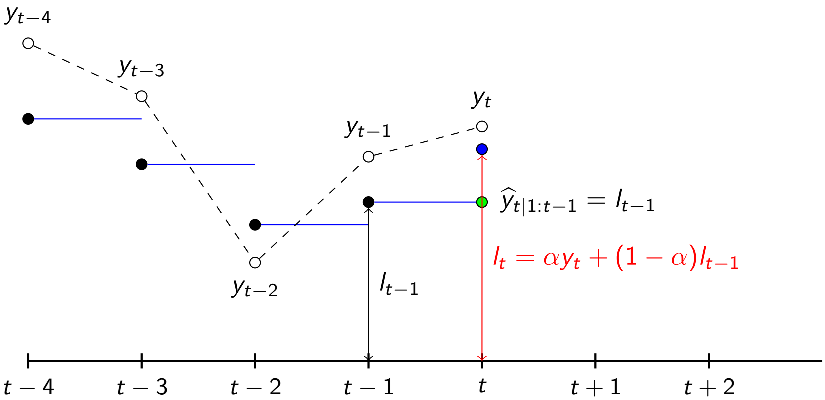
Explanation: Simple exponential smoothing



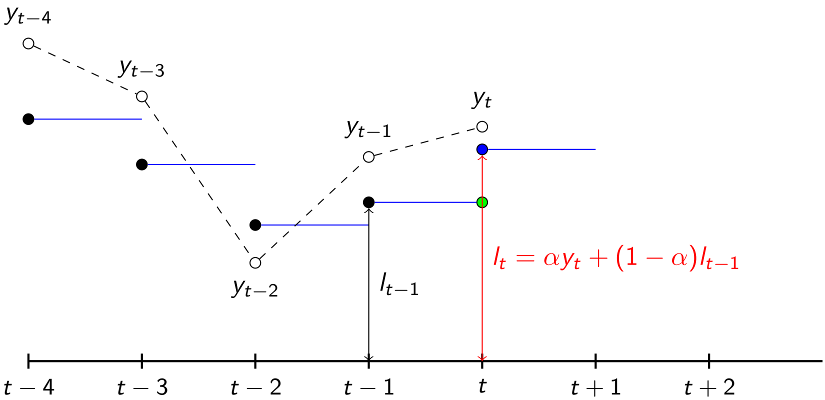
Explanation: Simple exponential smoothing



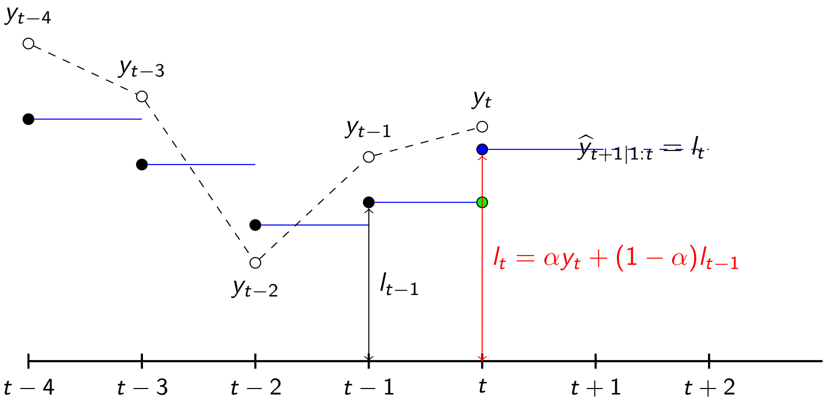
Explanation: Simple exponential smoothing



Explanation: Simple exponential smoothing

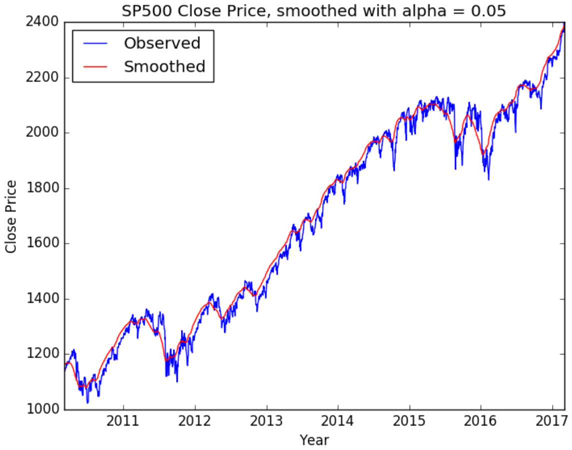


Explanation: Simple exponential smoothing



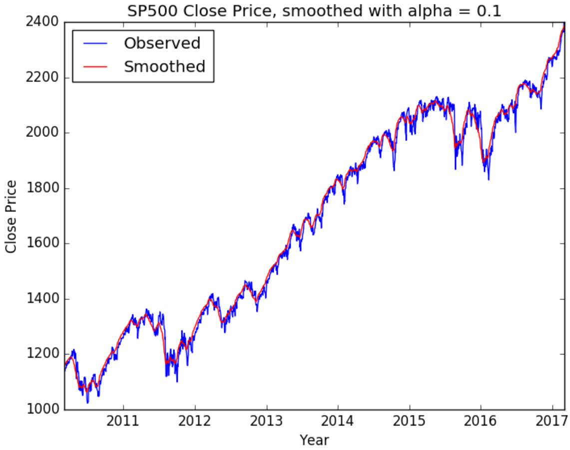
S&P 500 Closing Price (`Lecture05_Example01.py`)

Exponential smoothing with $\alpha = 0.05$



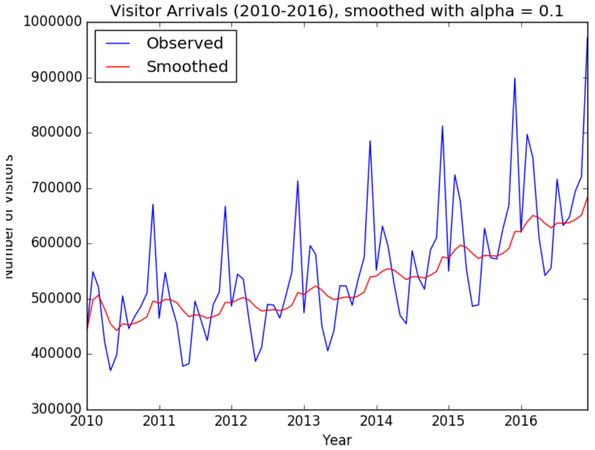
S&P 500 Closing Price (Lecture05_Example01.py)

Exponential smoothing with $\alpha = 0.1$



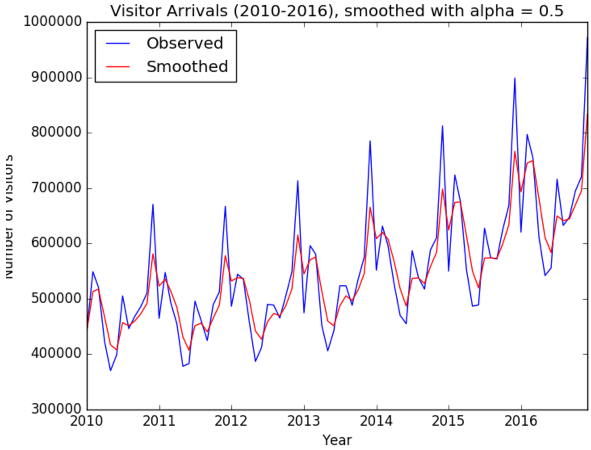
Visitor arrivals in Australia

Exponential smoothing with $\alpha = 0.1$



Visitor arrivals in Australia

Exponential smoothing with $\alpha = 0.5$



Simple exponential smoothing (SES)

SES as a Weighted Moving Average

Specify an initial value l_0 (an estimate or a guess, e.g., the average of y_1, y_2, y_3).

$$l_1 = \alpha y_1 + (1-\alpha) l_0$$

$$l_2 = \alpha y_2 + (1-\alpha) l_1 = \alpha y_2 + (1-\alpha)\alpha y_1 + (1-\alpha)^2 l_0$$

$$l_3 = \alpha y_3 + (1-\alpha) l_2 = \alpha y_3 + (1-\alpha)\alpha y_2 + (1-\alpha)^2 \alpha y_1 + (1-\alpha)^3 l_0$$

SES as a Weighted Moving Average

$$
\begin{aligned}
l_4 &= \alpha y_4 + (1-\alpha) l_3 \\
&= \alpha y_4 + (1-\alpha)\alpha y_3 + (1-\alpha)^2 \alpha y_2 + (1-\alpha)^3 \alpha y_1 + (1-\alpha)^4 l_0
\end{aligned}
$$

$$\vdots$$

$$
\begin{aligned}
l_t =& \alpha y_t + (1-\alpha) l_{t-1} \\
=& \alpha y_t + (1-\alpha)\alpha y_{t-1} + (1-\alpha)^2 \alpha y_{t-2} + \ldots + (1-\alpha)^{t-1} \alpha y_1 \\
& + (1-\alpha)^t l_0
\end{aligned}
$$

The smoothed/level values l_t are obtained from a WMA smoother $-->$

SES is sometimes called exponentially weighted moving average (EWMA).

Choice of weight α and initial level l_0

- We left l_0 unspecified above. How should we set it?
  - Use the average of very initial observations, i.e., y_1, y_2, y_3 etc., or even simply y_1
  - Take l_0 as a parameter, and use an algorithm to estimate it.
- How to set α?
  - Use expert's knowledge
  - Take α as a parameter, and use an algorithm to estimate it.
- We come back to this later.

Simple exponential smoothing (SES)

Some notes

- SES can be used for both smoothing and forecasting
- Useful for forecasting when the time series has no clear trend or seasonal patterns, hence
  - Should be used for seasonally adjusted and de-trended data

Statistical model

- SES considered so far can only produce point forecasts. How can we produce forecast intervals? We need a statistical model.
- The basic model:

$$l_t = \alpha y_t + (1-\alpha) l_{t-1}, \qquad 0 \le \alpha \le 1.$$

  and

$$y_{t+1} = l_t + \varepsilon_{t+1}, \text{ with } \varepsilon_{t+1} \sim N(0, \sigma^2).$$

- We assume all ε_t's are independent of each other
- The level is the hidden underlying mechanism, where the next observation is the current level corrupted with noise.
- What about when $\alpha = 1$?

Statistical model: Estimating Parameters

- Recall the basic model:

$$
\begin{aligned}
y_{t+1} &= l_t + \varepsilon_{t+1} \\
l_t &= \alpha y_t + (1-\alpha) l_{t-1}.
\end{aligned}
$$

- We can chose α (and l_0) by minimising

$$
\begin{aligned}
SSE &= \sum_{t=1}^{n} (y_t - l_{t-1})^2 \\
&= \sum_{t=1}^{n} (y_t - \alpha y_{t-1} - (1-\alpha) l_{t-2})^2.
\end{aligned}
$$

- You will try this in tutorial tasks.

A big picture

- The basic SES model:

$$\text{Measurement equation: } y_{t+1} = l_t + \varepsilon_{t+1}$$
$$\text{Transition equation: } l_t = \alpha y_t + (1-\alpha) l_{t-1}.$$

  The measurement equation represents the observation y_{t+1} as a function of a hidden/unobservable l_t. The transition equation describes how the hidden l_t evolves over time.

- Many state-of-the-art forecasting methods have this form: state space models, stochastic volatility models, recurrent neural network models

Error correction form

- The basic model for the level can be rewritten in terms of errors

$$
\begin{aligned}
l_t &= \alpha y_t + (1-\alpha) l_{t-1} \\
&= l_{t-1} + \alpha(y_t - l_{t-1}) \\
&= l_{t-1} + \alpha \varepsilon_t.
\end{aligned}
$$

- The next observation is

$$
\begin{aligned}
y_{t+1} &= l_t + \varepsilon_{t+1} = l_{t-1} + \alpha\varepsilon_t + \varepsilon_{t+1} = \cdots \\
&= l_0 + \alpha\varepsilon_1 + \cdots + \alpha\varepsilon_t + \varepsilon_{t+1}
\end{aligned}
$$

- Similarly

$$
\begin{aligned}
y_{t+2} &= l_{t+1} + \varepsilon_{t+2} = l_t + \alpha\varepsilon_{t+1} + \varepsilon_{t+2} \\
&= l_{t-1} + \alpha\varepsilon_t + \alpha\varepsilon_{t+1} + \varepsilon_{t+2} = \cdots \\
&= l_0 + \alpha\varepsilon_1 + \cdots + \alpha\varepsilon_t + \alpha\varepsilon_{t+1} + \varepsilon_{t+2}
\end{aligned}
$$

Forecast equations

- The forecast is defined as the average over all possible uncertainty

$$\widehat{y}_{t+h} := \mathbb{E}(y_{t+h}|y_{1:t}), \qquad \text{remember that } l_t = l_{t-1} + \alpha\varepsilon_t$$

- We have already observed up to time t, so l_t is certain. Uncertainty occurs after this time point, e.g., ε_{t+1}

$$\begin{aligned}\widehat{y}_{t+1} =& \mathbb{E}(l_t + \varepsilon_{t+1}|y_{1:t}) \\ =& \mathbb{E}(l_t) + \mathbb{E}(\varepsilon_{t+1}|y_{1:t}) = l_t + \mathbb{E}(\varepsilon_{t+1}) = l_t + 0 = l_t\end{aligned}$$

  where we have used the assumption $\varepsilon_{t+1} \sim N(0, \sigma^2)$, i.e., $\mathbb{E}(\varepsilon_{t+1}) = 0$.
- Similarly

$$\widehat{y}_{t+2} = \mathbb{E}(l_{t+1} + \varepsilon_{t+2}|y_{1:t}) = l_t \quad (\text{Can you prove this?})$$

$$\vdots$$

$$\widehat{y}_{t+h} = \mathbb{E}(l_t + \alpha\varepsilon_{t+1} + ... + \varepsilon_{t+h}|y_{1:t}) = l_t$$

Hence the forecast is always l_t after time t.

Variance of the forecasts

- Recall the model again:

$$y_{t+1} = l_t + \varepsilon_{t+1}, \qquad l_t = l_{t-1} + \alpha\varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma^2).$$

- Consider the variance of the new observation

$$\begin{aligned}
\mathrm{Var}(y_{t+1}|y_{1:t}) &= \mathrm{Var}(l_t + \varepsilon_{t+1}|y_{1:t}) \\
&= \mathrm{Var}(l_t) + \mathrm{Var}(\varepsilon_{t+1}|y_{1:t}) \\
&= 0 + \sigma^2 = \sigma^2. \quad \text{Why?}
\end{aligned}$$

- Similarly

$$\begin{aligned}
\mathrm{Var}(y_{t+2}|y_{1:t}) &= \mathrm{Var}(l_{t+1} + \varepsilon_{t+2}|y_{1:t}) = \mathrm{Var}(l_t + \alpha\varepsilon_{t+1} + \varepsilon_{t+2}|y_{1:t}) \\
&= \mathrm{Var}(l_t) + \mathrm{Var}(\alpha\varepsilon_{t+1}) + \mathrm{Var}(\varepsilon_{t+2}|y_{1:t}) \\
&= 0 + \alpha^2\mathrm{Var}(\varepsilon_{t+1}) + \mathrm{Var}(\varepsilon_{t+2}|y_{1:t}) \\
&= \alpha^2\sigma^2 + \sigma^2 = \sigma^2(1 + \alpha^2)
\end{aligned}$$

Variance of the forecasts

$$
\begin{aligned}
\mathrm{Var}(y_{t+3}|y_{1:t}) &= \mathrm{Var}(l_{t+2} + \varepsilon_{t+3}|y_{1:t}) \\
&= \mathrm{Var}(l_{t+1} + \alpha\varepsilon_{t+2} + \varepsilon_{t+3}|y_{1:t}) \\
&= \mathrm{Var}(l_t + \alpha\varepsilon_{t+1} + \alpha\varepsilon_{t+2} + \varepsilon_{t+3}|y_{1:t}) \\
&= \sigma^2(1 + 2\alpha^2)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(y_{t+h}|y_{1:t}) &= \mathrm{Var}(l_{t+h-1} + \varepsilon_{t+h}|y_{1:t}) \\
&= \mathrm{Var}(l_{t+h-2} + \alpha\varepsilon_{t+h-1} + \varepsilon_{t+h}|y_{1:t}) \\
&= \mathrm{Var}(l_t + \sum_{i=1}^{h-1} \alpha\varepsilon_{t+h-i} + \varepsilon_{t+h}|y_{1:t}) \\
&= \sigma^2(1 + (h-1)\alpha^2)
\end{aligned}
$$

Forecasting: collecting the results

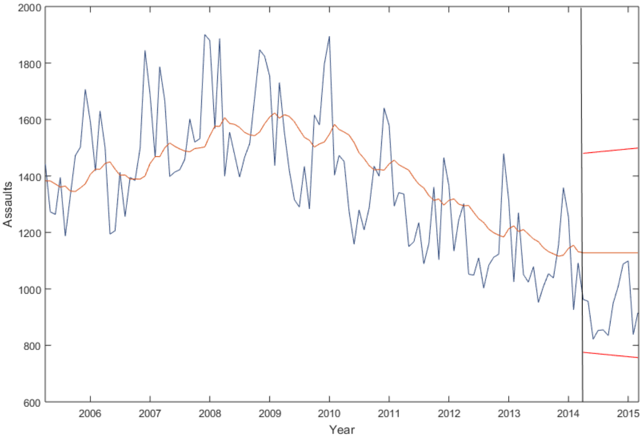
$$\widehat{y}_{t+h|1:t} = \mathbb{E}(y_{t+h}|y_{1:t}) = l_t$$

$$\mathrm{Var}(y_{t+h}|y_{1:t}) = \sigma^2(1 + (h-1)\alpha^2)$$

What happens as h increases?

Alcohol related assaults in NSW

Forecasting the seasonally adjusted series (last 12 months)



Outline

Simple Exponential Smoothing

Trend corrected exponential smoothing (TCES)

Trend corrected exponential smoothing (TCES)

Recall the component form of SES:

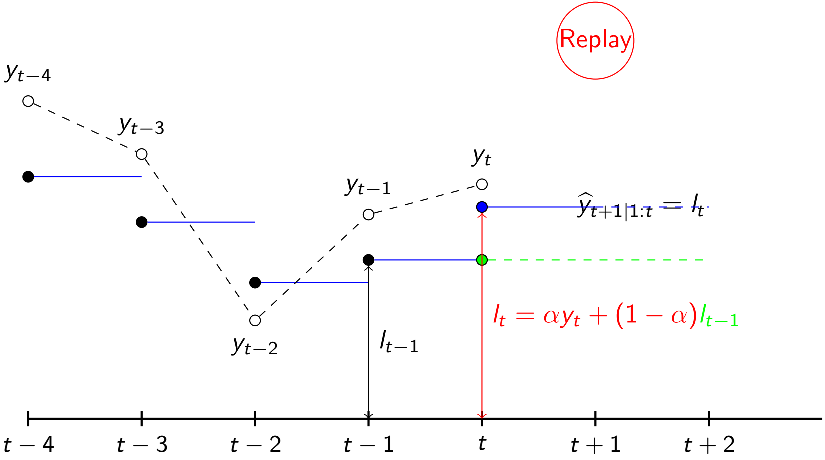
$$
\begin{aligned}
l_t &= \alpha y_t + (1-\alpha)\widehat{y}_t = \alpha y_t + (1-\alpha) l_{t-1}, \qquad 0 \le \alpha \le 1 \\
\widehat{y}_{t+1} &= l_t
\end{aligned}
$$

SES doesn't take trend into account.

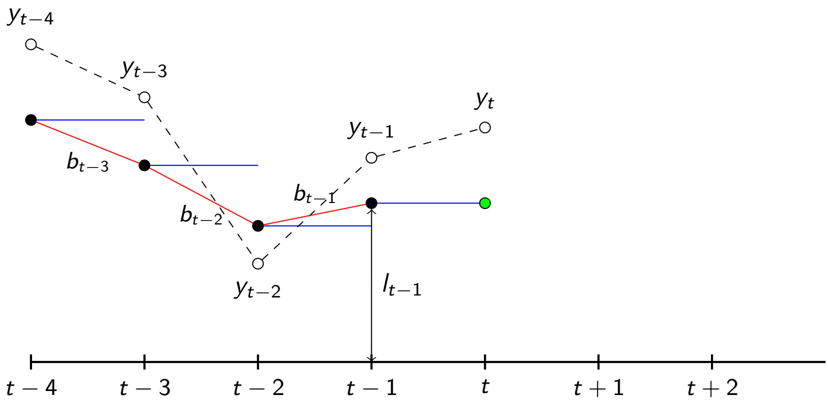
Trend corrected exponential smoothing:

$$
\begin{aligned}
l_t &= \alpha y_t + (1-\alpha)\widehat{y}_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}), \qquad 0 \le \alpha \le 1 \\
b_t &= \gamma(l_t - l_{t-1}) + (1-\gamma) b_{t-1}, \qquad 0 \le \gamma \le 1 \\
\widehat{y}_{t+1} &= l_t + b_t.
\end{aligned}
$$

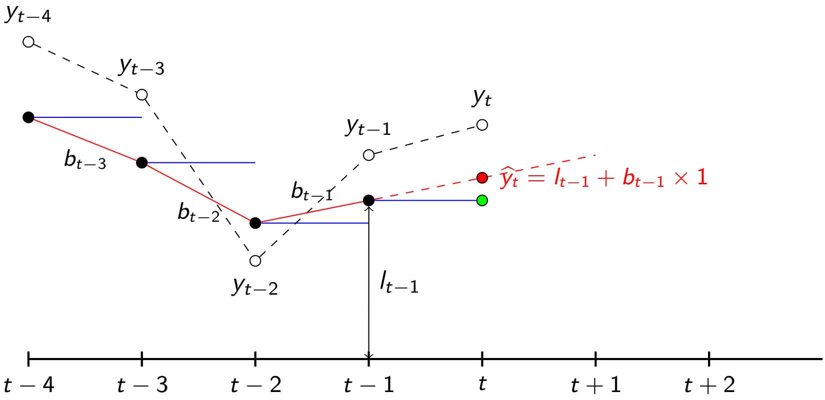
Explanation: Including Trend Information



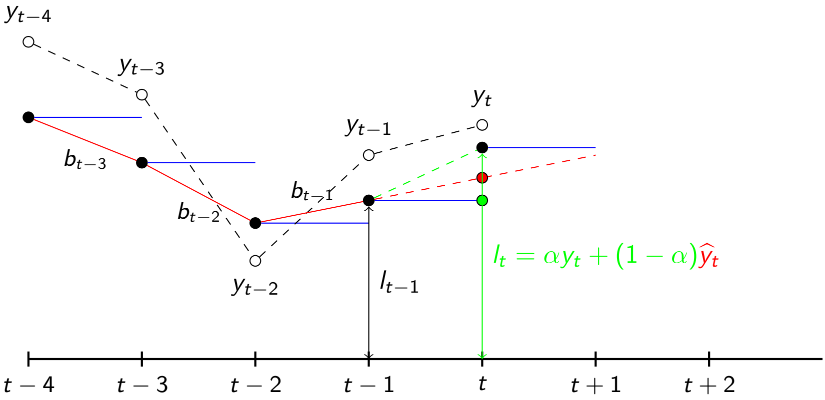
Explanation: Including Trend Information



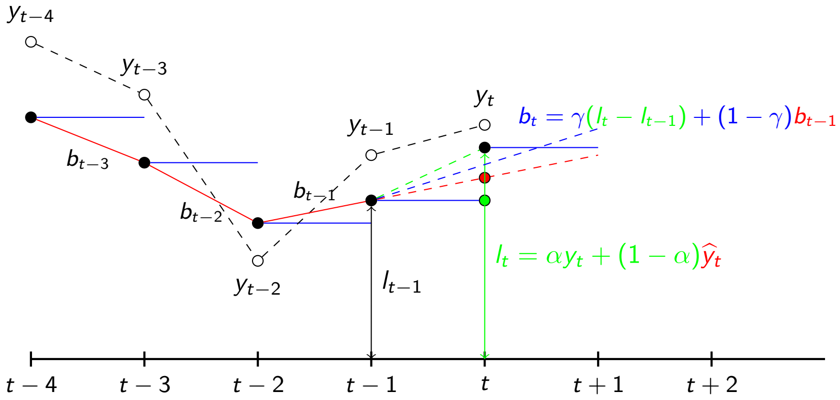
Explanation: Including Trend Information



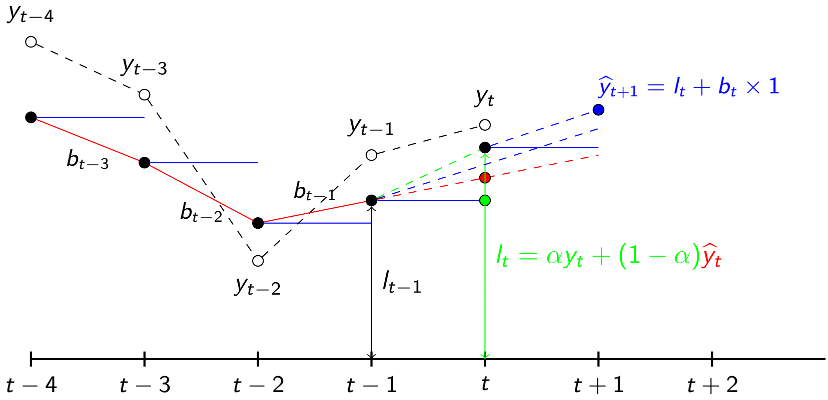
Explanation: Including Trend Information



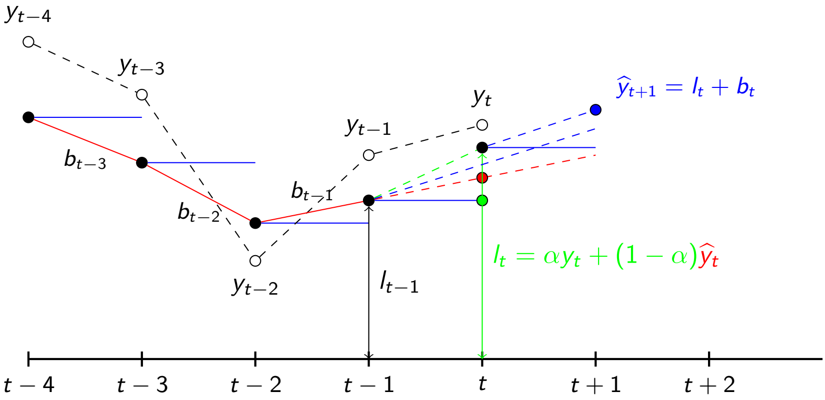
Explanation: Including Trend Information



Explanation: Including Trend Information



Explanation: Including Trend Information



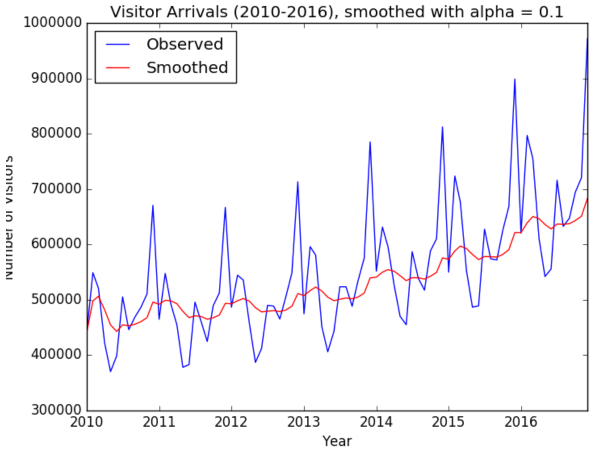
Statistical Model

$$
\begin{aligned}
y_{t+1} &= l_t + b_t + \varepsilon_{t+1} \\
l_t &= \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \\
b_t &= \gamma(l_t - l_{t-1}) + (1-\gamma) b_{t-1} \\
\varepsilon_{t+1} &\sim N(0, \sigma^2)
\end{aligned}
$$

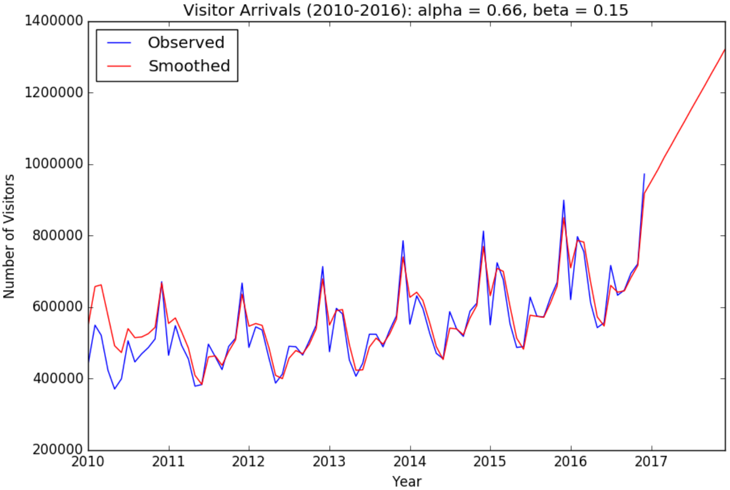
We can choose α and γ by minimising

$$\text{SSE}(\alpha, \gamma) = \sum_{t=2}^{n} (y_t - l_{t-1} - b_{t-1})^2$$

Visitor arrivals in Australia: `Lecture05_Example02.py`



Visitor arrivals in Australia: TCES



TCES: Error correction form

The basic model for the trend corrected exponential smoothing can be written in many ways. We can express all the components in terms of errors:

$$
\begin{aligned}
l_t &= \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \\
&= l_{t-1} + b_{t-1} + \alpha(y_t - l_{t-1} - b_{t-1}) \\
&= l_{t-1} + b_{t-1} + \alpha\varepsilon_t
\end{aligned}
$$

$$
\begin{aligned}
y_{t+1} &= l_t + b_t + \varepsilon_{t+1} \\
&= l_{t-1} + b_{t-1} + b_t + \alpha\varepsilon_t + \varepsilon_{t+1}
\end{aligned}
$$

TCES: Error correction form

$$\begin{aligned}
b_t &= \gamma(l_t - l_{t-1}) + (1-\gamma)b_{t-1} \\
&= b_{t-1} + \gamma(l_t - l_{t-1} - b_{t-1}) \\
&= b_{t-1} + \gamma\alpha\varepsilon_t \ \ (\text{from } l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t) \\
&(= b_{t-1} + \gamma\alpha(y_t - l_{t-1} - b_{t-1}))
\end{aligned}$$

TCES: Error correction form

$$
\begin{aligned}
l_t &= l_{t-1} + b_{t-1} + \alpha\varepsilon_t \\
b_t &= b_{t-1} + \gamma\alpha\varepsilon_t \\
y_{t+1} &= l_t + b_t + \varepsilon_{t+1} \\
\epsilon &\sim N(0, \sigma^2)
\end{aligned}
$$

Forecasting equations

$$\begin{aligned}
\widehat{y}_{t+1} &:= \mathbb{E}(y_{t+1}|y_{1:t}) \\
&= \mathbb{E}(l_t + b_t + \varepsilon_{t+1}|y_{1:t}) \\
&= l_t + b_t
\end{aligned}$$

$$\begin{aligned}
\widehat{y}_{t+2} &= \mathbb{E}(l_{t+1} + b_{t+1} + \varepsilon_{t+2}|y_{1:t}) \\
&= \mathbb{E}(l_t + 2b_t + \alpha(1+\gamma)\varepsilon_{t+1} + \varepsilon_{t+2}|y_{1:t}) \\
&= l_t + 2b_t
\end{aligned}$$

Trick: We iteratively expand the formula until we arrive at the time point where all are known.

Forecasting equations

$$
\begin{aligned}
\widehat{y}_{t+3} &= \mathbb{E}(l_{t+2} + b_{t+2} + \varepsilon_{t+3} | y_{1:t}) \\
&= \mathbb{E}((l_{t+1} + b_{t+1} + \alpha\varepsilon_{t+2}) + (b_{t+1} + \gamma\alpha\varepsilon_{t+2}) + \varepsilon_{t+3}) \\
&= \mathbb{E}(l_{t+1} + 2b_{t+1} + \alpha(1+\gamma)\varepsilon_{t+2} + \varepsilon_{t+3}) \\
&= \mathbb{E}((l_t + b_t + \alpha\varepsilon_{t+1}) + 2(b_t + \gamma\alpha\varepsilon_{t+1}) + \alpha(1+\gamma)\varepsilon_{t+2} + \varepsilon_{t+3}) \\
&= \mathbb{E}(l_t + 3b_t + \alpha(1+2\gamma)\varepsilon_{t+1} + \alpha(1+\gamma)\varepsilon_{t+2} + \varepsilon_{t+3}) \\
&= l_t + 3b_t \\
&\vdots \\
\widehat{y}_{t+h} &= l_t + hb_t
\end{aligned}
$$

Variance for forecasts

$$
\begin{aligned}
\mathrm{Var}(y_{t+1}|y_{1:t}) &= \mathrm{Var}(l_t + b_t + \varepsilon_{t+1}|y_{1:t}) \\
&= \sigma^2 \\
\mathrm{Var}(y_{t+2}|y_{1:t}) &= \mathrm{Var}(l_{t+1} + b_{t+1} + \varepsilon_{t+2}|y_{1:t}) \\
&= \mathrm{Var}(l_t + 2b_t + \alpha(1+\gamma)\varepsilon_{t+1} + \varepsilon_{t+2}|y_{1:t}) \\
&= \sigma^2(1 + \alpha^2(1+\gamma)^2)
\end{aligned}
$$

Variance for forecasts

$$
\begin{aligned}
\text{Var}(y_{t+3}|y_{1:t}) &= \text{Var}(l_{t+2} + b_{t+2} + \varepsilon_{t+3}|y_{1:t}) \\
&= \text{Var}(l_{t+1} + 2b_{t+1} + \alpha(1+\gamma)\varepsilon_{t+2} + \varepsilon_{t+3}|y_{1:t}) \\
&= \text{Var}(l_t + 3b_t + \alpha(1+2\gamma)\varepsilon_{t+1} + \alpha(1+\gamma)\varepsilon_{t+2} + \varepsilon_{t+3}|y_{1:t}) \\
&= \text{Var}(\alpha(1+2\gamma)\varepsilon_{t+1} + \alpha(1+\gamma)\varepsilon_{t+2} + \varepsilon_{t+3}|y_{1:t}) \\
&= \alpha^2(1+2\gamma)^2\sigma^2 + \alpha^2(1+\gamma)^2\sigma^2 + \sigma^2 \\
&= \sigma^2(1 + \alpha^2(1+\gamma)^2 + \alpha^2(1+2\gamma)^2)
\end{aligned}
$$

$$
\begin{aligned}
\text{Var}(y_{t+h}|y_{1:t}) &= \text{Var}\left(l_t + hb_t + \alpha\sum_{i=1}^{h-1}(1+i\gamma)\varepsilon_{t+i} + \varepsilon_{t+h}|y_{1:t}\right) \\
&= \sigma^2\left(1 + \alpha^2\sum_{i=1}^{h-1}(1+i\gamma)^2\right) \\
&= \sigma^2\left(1 + \alpha^2\left(\frac{\gamma^2}{3}h(h-1)(h-2) + (\gamma+1)h - 1\right)\right)
\end{aligned}
$$

(I used the formula for the sum of an arithmetic progression to get to the last step)

Forecasting: collecting the results

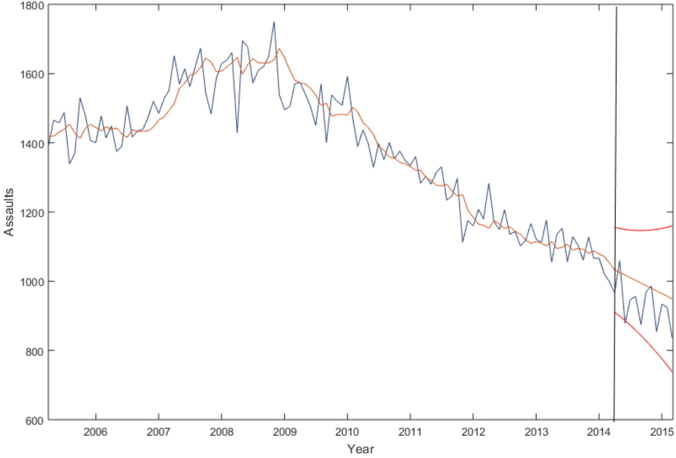
$$\widehat{y}_{t+h} = l_t + hb_t$$

$$\text{Var}(y_{t+h}|y_{1:t}) = \sigma^2 \left(1 + \alpha^2 \left(\frac{\gamma^2}{3} h(h-1)(h-2) + (\gamma+1)h - 1\right)\right)$$

What happens as h increases?

Alcohol related assaults in NSW

Forecasting the seasonally adjusted series (last 12 months)



Alcohol related assaults in NSW

Forecasting the seasonally adjusted series (last 12 months)

SES and TCES, which one is better?

Table: **One month ahead forecasts**

	SES	TCES
RMSE	70.9	63.5
MAE	56.5	55.8
MAPE	6.2	6.0

Recap

We have looked at

- simple exponential smoothing as weighted averaging
- the underlying statistical model and equivalent forms of the updates
- the error correction form for getting the prediction intervals
- extending the model to handle trend

Next lecture: still in the exponential smoothing land, we will look at how to handle seasonality and how to dampen the trend for long-term predictions

Thank you and see you next week!