QBUS6840 Lecture 6

Exponential Smoothing (Seasonal)

Professor Junbin Gao

The University of Sydney Business School

Outline

Exponential smoothing

- Holt-Winters smoothing
  - Exponential smoothing methods for seasonal data.
  - Additive seasonality.
  - Multiplicative seasonality.
- Damped Trend Exponential Smoothing
- Damped Trend Seasonal

Reading

- Online Textbook Sections 7.3-7.4 and 7.6: `https://otexts.org/fpp2/expsmooth.html` and/or
- BOK Sec 8.4-8.5

Introducing Additive Holt-Winters smoothing

- The ideal scenario

$$\overbrace{y_t = \omega_0 + \omega_1 t + S_t + \varepsilon_t}^{\text{trend}}$$

- Additive decomposition model: assuming ω_0, ω_0 and S_t (M different values) are fixed constants.
- Simple exponential method: modelling the case where $S_t = 0$, $\omega_1 = 0$ (or constant) and ω_0 changes with time
- Trend corrected exponential method: modelling the case where $S_t = 0$, both ω_1 and ω_0 are changing
- How to model the data if the level, the level growth rate (the trend), and seasonal patterns are changing?

Additive Holt-Winters smoothing

$$
\begin{aligned}
l_t &= \alpha(y_t - S_{t-M}) + (1-\alpha)(l_{t-1} + b_{t-1}), \qquad 0 \le \alpha \le 1 \\
b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}, \qquad 0 \le \beta \le 1 \\
S_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)S_{t-M}, \qquad 0 \le \gamma \le 1
\end{aligned}
$$

$$\widehat{y}_{t+1|1:t} = l_t + b_t + S_{t+1-M}.$$

Always read this as

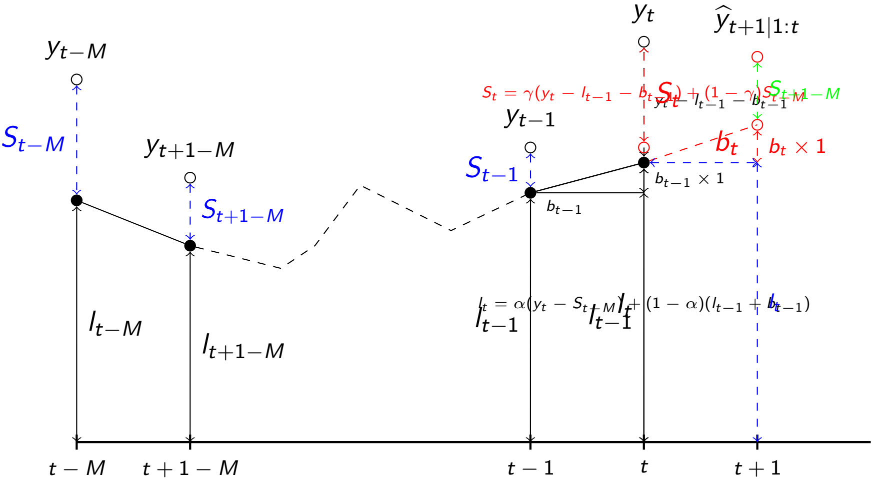
$$\boxed{\widehat{y}_{t+1|1:t} = l_t + b_t \times 1 + S_{t+1-M}.}$$

Sometimes we also write the seasonal update as

$$S_t = \gamma(y_t - l_t) + (1-\gamma)S_{t-M}, \qquad 0 \le \gamma \le 1$$

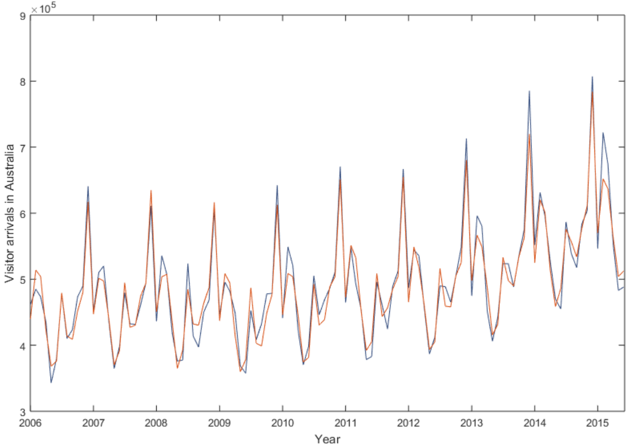
Explanation

Additive Holt-Winters smoothing



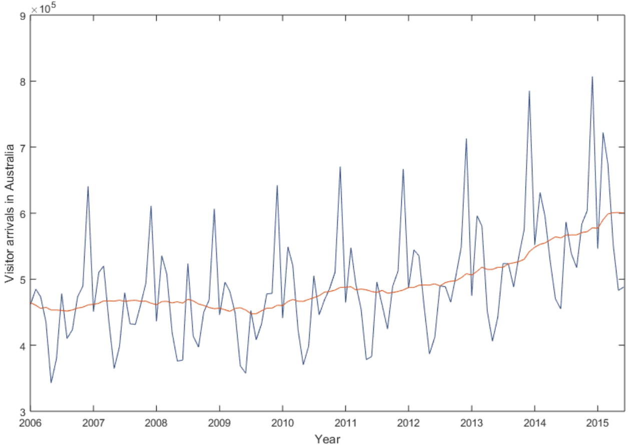
Visitor arrivals in Australia: `Lecture06_Example01.py`

Additive Holt-Winters method



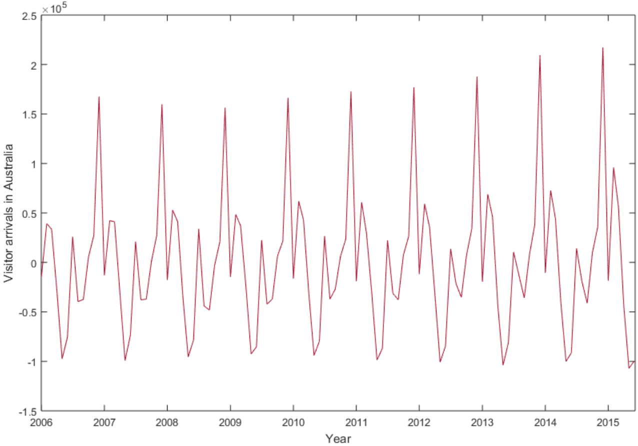
Visitor arrivals in Australia

Additive Holt-Winters level component estimate



Visitor arrivals in Australia

Additive Holt-Winters seasonal factors



How should we set the initial values l_0, b_0, s_0, s_{-1}, ..., s_{2-M}, s_{1-M}?

Suggested Method

1. Do a linear least square regression over the data $y_1, \ldots, y_T$ to find out
$$\widehat{y}_t = \widehat{\omega}_0 + \widehat{\omega}_1 t$$
2. Take $l_0 = \widehat{\omega}_0$ and $b_0 = \widehat{\omega}_1$
3. Find out $\widehat{s}_t = y_t - \widehat{y}_t$, then take the average of $\widehat{s}_t$ as one of s_0, s_{-1}, ..., s_{2-M}, s_{1-M} according to each season.

Additive Holt-Winters smoothing

Some notes

- Useful when level and/or trend and seasonal variation is changing
- Most useful when seasonal pattern changing in a cyclical or irregular fashion – but not too much!
- Choice of initial seasonal indices can be important.

Additive Holt-Winters smoothing

Model

$$
\begin{aligned}
l_t &= \alpha(y_t - S_{t-M}) + (1-\alpha)(l_{t-1} + b_{t-1}),\\
b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1},\\
S_t &= \gamma(y_t - l_t) + (1-\gamma)S_{t-M},
\end{aligned}
$$

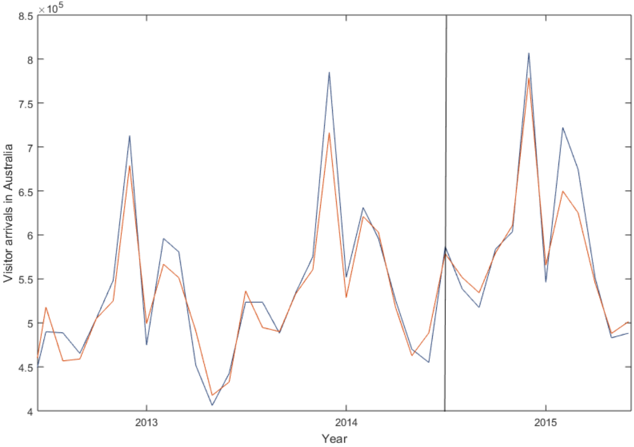
$$
y_{t+1} = l_t + b_t + S_{t+1-M} + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, \sigma^2).
$$

We can chose the parameters α, β and γ by minimising

$$
SSE = \sum_{t=1}^{n} (y_t - l_{t-1} - b_{t-1} - S_{t-M})^2
$$

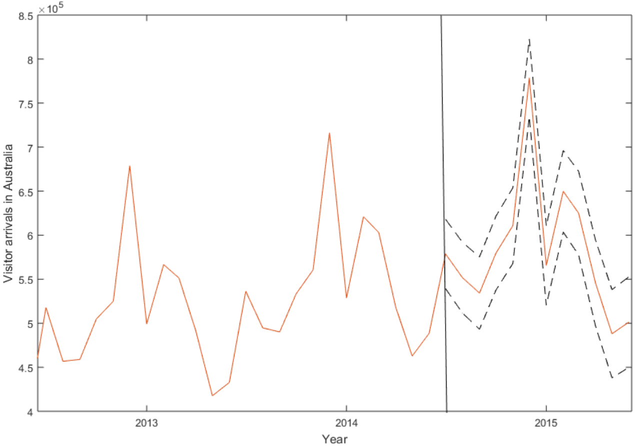
Visitor arrivals in Australia

Additive Holt-Winters forecast



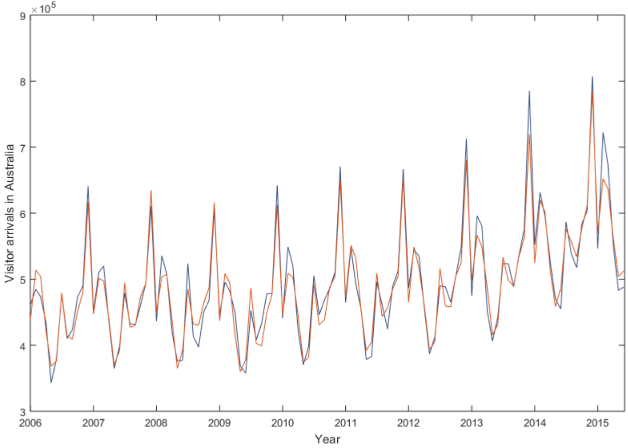
Visitor arrivals in Australia

Additive Holt-Winters forecast



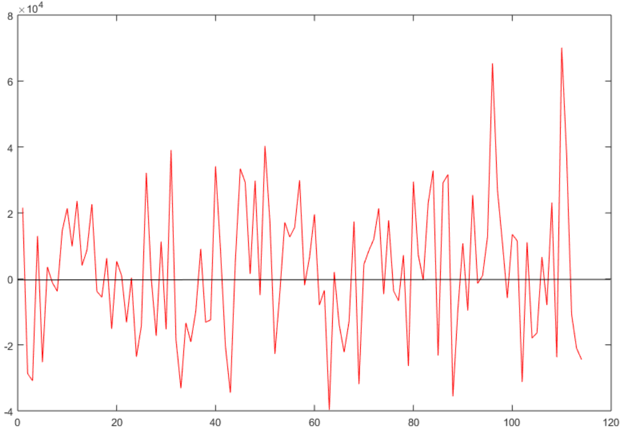
Visitor arrivals in Australia

Additive Holt-Winters fit



Visitor arrivals in Australia

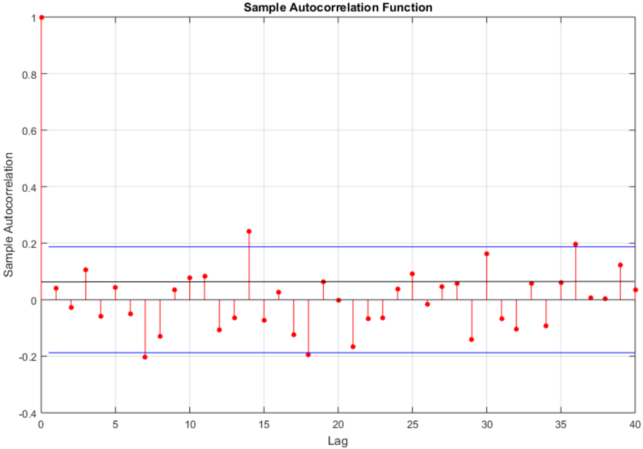
Additive Holt-Winters residuals



Visitor arrivals in Australia

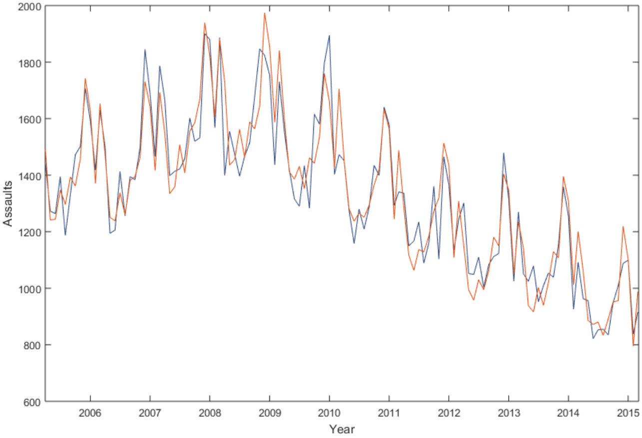
Additive Holt-Winters residual autocorrelations

We will talk about this in Week 7



Alcohol related assaults in NSW

Additive Holt-Winters fit



Additive Holt-Winters smoothing

Error correction formulation

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

Additive Holt-Winters smoothing

Error correction formulation

First from the first model, we have

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

Hence

$$
l_t - l_{t-1} - b_{t-1} = \alpha(y_t - S_{t-M} - l_{t-1} - b_{t-1})
$$

Hence,

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

Additive Holt-Winters smoothing

Error correction formulation

First from the first model, we

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

Hence

$$\begin{aligned} S_t &= \gamma(y_t - l_t) + (1-\gamma)S_{t-M} \\ &= S_{t-M} + \gamma(y_t - l_t - S_{t-M}) \\ &= S_{t-M} + \gamma(1-\alpha)\varepsilon_t \end{aligned}$$

Additive Holt-Winters smoothing

Error correction formulation

$$
\begin{aligned}
l_t &= l_{t-1} + b_{t-1} + \alpha\varepsilon_t \\
b_t &= b_{t-1} + \alpha\beta\varepsilon_t \\
S_t &= S_{t-M} + \gamma(1-\alpha)\varepsilon_t \\
y_t &= l_{t-1} + b_{t-1} + S_{t+1-M} + \varepsilon_t
\end{aligned}
$$

$$
e.g. \quad y_{t+1} = l_{t-1} + 2b_{t-1} + S_{t-M+1} + \alpha(1+\beta)\varepsilon_t + \varepsilon_{t+1}
$$

Additive Holt-Winters smoothing

Forecasting equations

$$
\begin{aligned}
\widehat{y}_{t+1|1:t} &= E(l_t + b_t + S_{t-M+1} + \varepsilon_{t+1}|y_{1:t}) \\
&= l_t + b_t + S_{t-M+1} \\
&(= l_{t-1} + 2b_{t-1} + S_{t-M+1} + \alpha(1+\beta)\varepsilon_t) \\
\\
\widehat{y}_{t+2|1:t} &= E(l_{t+1} + b_{t+1} + S_{t-M+2} + \varepsilon_{t+2}|y_{1:t}) \\
&= E(l_t + 2b_t + S_{t-M+2} + \alpha(1+\beta)\varepsilon_{t+1} + \varepsilon_{t+2}|y_{1:t}) \\
&= l_t + 2b_t + S_{t-M+2} \\
&\vdots \\
\widehat{y}_{t+h|1:t} &= l_t + hb_t + S_{t-M+(h \bmod M)}
\end{aligned}
$$

Additive Holt-Winters smoothing

Variance for interval forecasts

$$y_{t+1} = l_{t-1} + 2b_{t-1} + S_{t-M+1} + \alpha(1+\beta)\varepsilon_t + \varepsilon_{t+1}$$

$$\begin{aligned}
\mathrm{Var}(y_{t+1}|y_{1:t}) &= \mathrm{Var}(l_t + b_t + S_{t-M+1} + \varepsilon_{t+1}|y_{1:t}) \\
&= \sigma^2
\end{aligned}$$

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

Additive Holt-Winters smoothing

Variance for interval forecasts

$$y_{t+1} = l_{t-1} + 2b_{t-1} + S_{t-M+1} + \alpha(1+\beta)\varepsilon_t + \varepsilon_{t+1}$$

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

$$\begin{aligned}
\mathrm{Var}(y_{t+h}|y_{1:t}) &= \mathrm{Var}\left(l_t + hb_t + S_{t-M+h} + \alpha\sum_{i=1}^{h-1}(1+i\beta)\varepsilon_{t+i} + \varepsilon_{t+h}|y_{1:t}\right) \\
&= \sigma^2\left(1 + \alpha^2\sum_{i=1}^{h-1}(1+i\beta)^2\right), \quad \text{for } h \leq M \text{ only.}
\end{aligned}$$

Additive Holt-Winters smoothing

Variance for interval forecasts

$$y_{t+1} = l_{t-1} + 2b_{t-1} + S_{t-M+1} + \alpha(1+\beta)\varepsilon_t + \varepsilon_{t+1}$$

For $h > M$,

$$\begin{aligned}
\mathrm{Var}(y_{t+h}|y_{1:t}) &= \mathrm{Var}\left(l_t + hb_t + S_{t-M+h} + \alpha\sum_{i=1}^{h-1}(1+i\beta)\varepsilon_{t+i} + \varepsilon_{t+h}|y_{1:t}\right) \\
&= \mathrm{Var}\Bigg(l_t + hb_t + S_{t-2M+h} + \gamma(1-\alpha)\varepsilon_{t-M+h} \\
&\quad + \alpha\sum_{i=1}^{h-1}(1+i\beta)\varepsilon_{t+i} + \varepsilon_{t+h}|y_{1:t}\Bigg) \\
&= \sigma^2\left(1 + \sum_{i=1}^{h-1}\left[\alpha(1+i\beta) + I_{i,M}\gamma(1-\alpha)\right]^2\right),
\end{aligned}$$

where $I_{i,M} = 1$ if i is an integer multiple of M and 0 otherwise.

Additive Holt-Winters smoothing

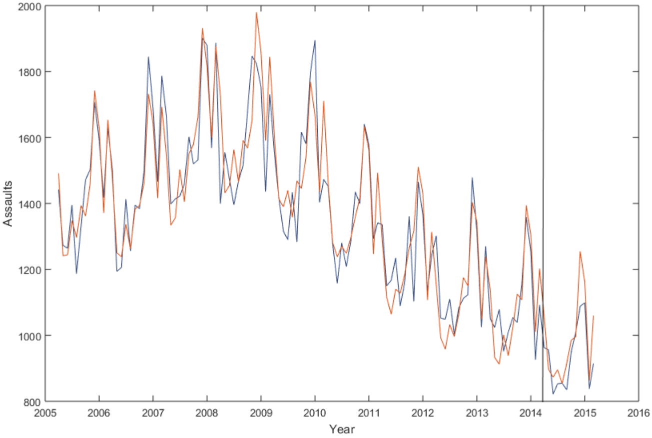
Forecasting: collecting the results

$$\widehat{y}_{t+h|1:t} = \widehat{l}_t + h\widehat{b}_t + S_{t-M+(h \bmod M)}.$$

$$\mathrm{Var}(y_{t+h}|y_{1:t}) = \sigma^2 \left(1 + \sum_{i=1}^{h-1} \left[\alpha(1+i\beta) + I_{i,M}\gamma(1-\alpha)\right]^2\right).$$

Alcohol related assaults in NSW

Additive Holt-Winters forecast



Alcohol related assaults in NSW

Additive Holt-Winters forecast

Multiplicative Holt-Winters smoothing

Most useful when the seasonal pattern changes in a strong pattern and is proportional to the level of the series.

Multiplicative Holt-Winters smoothing

Model

$$
\begin{aligned}
l_t &= \alpha(y_t/S_{t-M}) + (1-\alpha)(l_{t-1} + b_{t-1}),\\
b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1},\\
S_t &= \gamma(y_t/l_t) + (1-\gamma)S_{t-M},
\end{aligned}
$$

$$
y_{t+1} = (l_t + b_t) \times S_{t+1-M} + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, \sigma^2).
$$

We can chose the parameters α, β and γ by minimising

$$
SSE = \sum_{t=1}^{n} (y_t - (l_{t-1} + b_{t-1})S_{t-M})^2
$$

Multiplicative Holt-Winters smoothing

Error correction formulation

$$
\begin{aligned}
l_t &= \alpha(y_t/S_{t-M}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\
&= l_{t-1} + b_{t-1} + \alpha(y_t/S_{t-M} - l_{t-1} - b_{t-1}) \\
&= l_{t-1} + b_{t-1} + \alpha\left(\frac{y_t - (l_{t-1} + b_{t-1})S_{t-M}}{S_{t-M}}\right) \\
&= l_{t-1} + b_{t-1} + \alpha\frac{\varepsilon_t}{S_{t-M}}
\end{aligned}
$$

Multiplicative Holt-Winters smoothing

Error correction formulation

$$
\begin{aligned}
b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\
&= b_{t-1} + \beta\alpha\left(\frac{y_t - (l_{t-1} + b_{t-1})S_{t-M}}{S_{t-M}}\right) \quad \text{see previous slide} \\
&= b_{t-1} + \alpha\beta\frac{\varepsilon_t}{S_{t-M}}
\end{aligned}
$$

Multiplicative Holt-Winters smoothing

Error correction formulation

$$S_t = \gamma(y_t/l_t) + (1-\gamma)S_{t-M} = S_{t-M} + \gamma\frac{y_t - l_t S_{t-M}}{l_t}$$

From the first model we have

$$l_t S_{t-M} = (l_{t-1} + b_{t-1})S_{t-M} + \alpha(y_t - (l_{t-1} + b_{t-1})S_{t-M})$$

Hence

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

Hence

$$S_t = S_{t-M} + \gamma(1-\alpha)\frac{\varepsilon_t}{l_t}$$

Multiplicative Holt-Winters smoothing

Error correction formulation

$$
\begin{aligned}
l_t &= l_{t-1} + b_{t-1} + \alpha \frac{\varepsilon_t}{S_{t-M}} \\
b_t &= b_{t-1} + \alpha\beta \frac{\varepsilon_t}{S_{t-M}} \\
S_t &= S_{t-M} + \gamma(1-\alpha)\frac{\varepsilon_t}{l_t} \\
y_t &= (l_{t-1} + b_{t-1}) \times S_{t-M} + \varepsilon_t
\end{aligned}
$$

Multiplicative Holt-Winters smoothing

Forecasting equations

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

$$
\begin{aligned}
\widehat{y}_{t+2|1:t} &= E((l_{t+1} + b_{t+1})S_{t-M+2} + \varepsilon_{t+2}|y_{1:t}) \\
&= E\left(\left[l_t + 2b_t + \alpha(1+\beta)\frac{\varepsilon_{t+1}}{S_{t-M+1}}\right] S_{t-M+2} + \varepsilon_{t+2} \,\middle|\, y_{1:t}\right) \\
&= (l_t + 2b_t)S_{t-M+2} \\
&\vdots \\
\widehat{y}_{t+h|1:t} &= (l_t + hb_t)S_{t-M+(h \bmod M)}
\end{aligned}
$$

Multiplicative Holt-Winters smoothing
Variance for interval forecasts

$$y_{t+1} = \left[l_{t-1} + 2b_{t-1} + \alpha(1+\beta)\frac{\varepsilon_t}{S_{t-M}} \right] S_{t-M+1} + \varepsilon_{t+1}$$

$$\begin{aligned} \mathrm{Var}(y_{t+1}|y_{1:t}) &= \mathrm{Var}((l_t + b_t)S_{t-M+1} + \varepsilon_{t+1}|y_{1:t}) \\ &= \sigma^2 \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(y_{t+2}|y_{1:t}) &= \mathrm{Var}((l_{t+1} + b_{t+1})S_{t-M+2} + \varepsilon_{t+2}|y_{1:t}) \\ &= \mathrm{Var}\left(\left[l_t + 2b_t + \alpha(1+\beta)\frac{\varepsilon_{t+1}}{S_{t-M+1}} \right] S_{t-M+2} + \varepsilon_{t+2}|y_{1:t} \right) \\ &= \sigma^2(1 + \alpha^2(1+\beta)^2(S^2_{t-M+2}/S^2_{t-M+1})) \end{aligned}$$

Multiplicative Holt-Winters smoothing

Forecasting formula

$$\widehat{y}_{t+h|1:t} = (\widehat{l}_t + h\widehat{b}_t) \times \widehat{S}_{t+h-M}.$$

Dampened trend ES

无限期地推断未来的趋势可能会有问题。

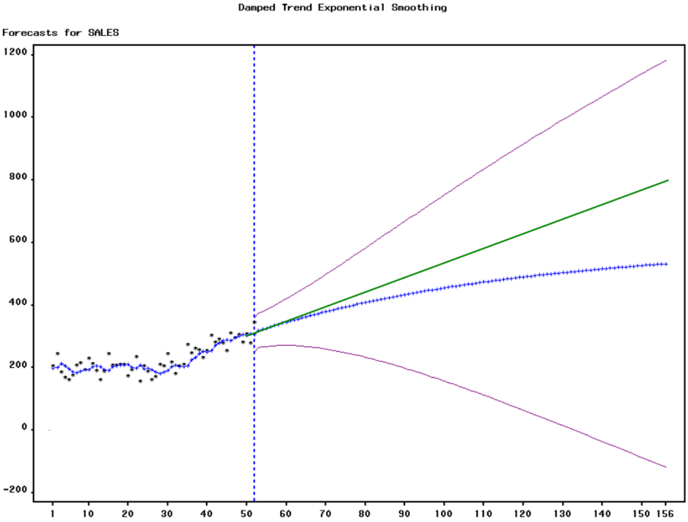
Extrapolating trends indefinitely into the future can be problematic.

Dampened trend exponential smoothing aims to deal with this problem.

阻尼趋势指数平滑旨在解决这个问题。

Dampened trend ES

Illustration



Dampened trend ES

Model

$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + \phi b_{t-1}),$$
$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)\phi b_{t-1},$$

$$y_{t+1} = l_t + \phi b_t + \varepsilon_{t+1},$$

where ϕ is the dampening factor, with $0 \leq \phi \leq 1$.

Dampened trend ES

Forecasting and variance equations

$$y_{t+1} = l_t + \phi b_t + \varepsilon_{t+1}$$

$$\widehat{y}_{t+1|1:t} = l_t + \phi b_t$$

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

Dampened trend ES

Forecasting and variance equations

$$
\begin{aligned}
y_{t+2} &= l_{t+1} + \phi b_{t+1} + \varepsilon_{t+2} \\
&= l_t + \phi b_t + \phi^2 b_t + \alpha(1 + \phi\beta)\varepsilon_{t+1} + \varepsilon_{t+2}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{y}_{t+2|1:t} &= l_t + b_t(\phi_t + \phi^2) \\
\mathrm{Var}(y_{t+1}|y_{1:t}) &= \sigma^2(1 + \alpha^2(1 + \phi\beta)^2)
\end{aligned}
$$

Dampened trend ES

Forecasting formula

$$\widehat{y}_{t+h|1:t} = l_t + \phi b_t + \phi^2 b_t + \phi^3 b_t + \ldots + \phi^h b_t$$

Compared with the forecast of the trend correct exponential method

$$\widehat{y}_{t+h|1:t} = l_t + h \times b_t$$

What happens as h gets larger?
For the dampened forecast $\widehat{y}_{t+h|1:t} \rightarrow l_t + \frac{\phi}{1-\phi} b_t$
For the trend corrected forecast

$$\widehat{y}_{t+h|1:t} \rightarrow \infty$$

Dampened trend seasonal

Model

$$\begin{aligned}
l_t &= \alpha(y_t - S_{t-M}) + (1-\alpha)(l_{t-1} + \phi b_{t-1}), \\
b_t &= \beta(l_t - l_{t-1}) + (1-\beta)\phi b_{t-1}, \\
S_t &= \gamma(y_t - l_t) + (1-\gamma)S_{t-M},
\end{aligned}$$

$$y_{t+1} = l_t + \phi b_t + S_{t-M+1} + \varepsilon_{t+1},$$

where ϕ is the dampening factor, with $0 \leq \phi \leq 1$.

Dampened trend seasonal
Forecasting formula

$$\widehat{y}_{t+h|1:t} = l_t + \phi b_t + \phi^2 b_t + \phi^3 b_t + \ldots + \phi^h b_t + S_{t+h-M}$$