1.4 Exponential smoothing
Outline

1. Simple exponential smoothing
2. Trend methods
3. Lab session 7
4. Seasonal methods
5. Lab session 8
6. Taxonomy of exponential smoothing methods
### Simple methods

**Time series** $y_1, y_2, \ldots, y_T$.

**Random walk forecasts**

$$\hat{y}_{T+h|T} = y_T$$
### Simple methods

**Time series** $y_1, y_2, \ldots, y_T$.

**Random walk forecasts**

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

**Average forecasts**

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

Want something in between that weights most recent data more highly. Simple exponential smoothing uses a weighted moving average with weights that decrease exponentially.
Simple methods

Time series $y_1, y_2, \ldots, y_T$.

Random walk forecasts

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

Average forecasts

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

- Want something in between that weights most recent data more highly.
- Simple exponential smoothing uses a weighted moving average with weights that decrease exponentially.
Simple Exponential Smoothing

Forecast equation

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

where \(0 \leq \alpha \leq 1\).
Simple Exponential Smoothing

Forecast equation

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

where \(0 \leq \alpha \leq 1\).

<table>
<thead>
<tr>
<th>Observation</th>
<th>(\alpha = 0.2)</th>
<th>(\alpha = 0.4)</th>
<th>(\alpha = 0.6)</th>
<th>(\alpha = 0.8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(y_T)</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>(y_{T-1})</td>
<td>0.16</td>
<td>0.24</td>
<td>0.24</td>
<td>0.16</td>
</tr>
<tr>
<td>(y_{T-2})</td>
<td>0.128</td>
<td>0.144</td>
<td>0.096</td>
<td>0.032</td>
</tr>
<tr>
<td>(y_{T-3})</td>
<td>0.1024</td>
<td>0.0864</td>
<td>0.0384</td>
<td>0.0064</td>
</tr>
<tr>
<td>(y_{T-4})</td>
<td>(0.2)(0.8)^4</td>
<td>(0.4)(0.6)^4</td>
<td>(0.6)(0.4)^4</td>
<td>(0.8)(0.2)^4</td>
</tr>
<tr>
<td>(y_{T-5})</td>
<td>(0.2)(0.8)^5</td>
<td>(0.4)(0.6)^5</td>
<td>(0.6)(0.4)^5</td>
<td>(0.8)(0.2)^5</td>
</tr>
</tbody>
</table>
### Simple Exponential Smoothing

#### Component form

- **Forecast equation**
  \[ \hat{y}_{t+h|t} = \ell_t \]

- **Smoothing equation**
  \[ \ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1} \]

- \( \ell_t \) is the level (or the smoothed value) of the series at time \( t \).

- \( \hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1} \)
  
  Iterate to get exponentially weighted moving average form.

#### Weighted average form

\[
\hat{y}_{T+1|T} = \sum_{j=0}^{T-1} \alpha (1 - \alpha)^j y_{T-j} + (1 - \alpha)^T \ell_0
\]
Need to choose value for $\alpha$ and $\ell_0$

Similarly to regression — we choose $\alpha$ and $\ell_0$ by minimising SSE:

$$SSE = \sum_{t=1}^{T} (y_t - \hat{y}_{t|t-1})^2.$$ 

Unlike regression there is no closed form solution — use numerical optimization.
```r
fc <- ses(oil, h=5)
summary(fc[["model"]])
```

```r
## Simple exponential smoothing
##
## Call:
## ses(y = oil, h = 5)
##
## Smoothing parameters:
## alpha = 0.9999
##
## Initial states:
## l = 110.8832
##
## sigma: 49.05
##
## AIC  AICc   BIC
## 576.2 576.7 581.8
```
Example: Oil production

```r
oil %>% ses(PI=FALSE) %>% autoplot
```

Forecasts from Simple exponential smoothing
1. Simple exponential smoothing
2. Trend methods
3. Lab session 7
4. Seasonal methods
5. Lab session 8
6. Taxonomy of exponential smoothing methods
## Holt’s linear trend

### Component form

<table>
<thead>
<tr>
<th>Component</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forecast</strong></td>
<td>( \hat{y}_{t+h</td>
</tr>
<tr>
<td><strong>Level</strong></td>
<td>( \ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) )</td>
</tr>
<tr>
<td><strong>Trend</strong></td>
<td>( b_t = \beta^<em>(\ell_t - \ell_{t-1}) + (1 - \beta^</em>)b_{t-1}, )</td>
</tr>
</tbody>
</table>

Two smoothing parameters \( \alpha \) and \( \beta^* \) (0 ≤ \( \alpha, \beta^* \) ≤ 1). \( \ell_0 \) and \( b_0 \) are chosen to minimize SSE.
Holt’s linear trend

**Component form**

| Forecast | \( \hat{y}_{t+h|t} = \ell_t + hb_t \) |
|----------|----------------------------------|
| Level    | \( \ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \) |
| Trend    | \( b_t = \beta^* (\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \) |

- Two smoothing parameters \( \alpha \) and \( \beta^* \) \((0 \leq \alpha, \beta^* \leq 1)\).
- \( \ell_t \) level: weighted average between \( y_t \) and one-step ahead forecast for time \( t \), \((\ell_{t-1} + b_{t-1} = \hat{y}_{t|t-1})\)
- \( b_t \) slope: weighted average of \((\ell_t - \ell_{t-1})\) and \( b_{t-1} \), current and previous estimate of slope.
Holt’s method in R

```
window(ausair, start=1990, end=2004) %>%
holt(h=5, PI=FALSE) %>%
autoplot()
```

Forecasts from Holt's method
Damped trend method

Component form

\[ \hat{y}_{t+h|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t \]

\[ \ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1}) \]

\[ b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}. \]
Damped trend method

**Component form**

\[ \hat{y}_{t+h|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t \]

\[ \ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1}) \]

\[ b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}. \]

- Damping parameter \(0 < \phi < 1\).
- If \(\phi = 1\), identical to Holt’s linear trend.
- As \(h \rightarrow \infty\), \(\hat{y}_{T+h|T} \rightarrow \ell_T + \phi b_T/(1 - \phi)\).
- Short-run forecasts trended, long-run forecasts constant.
Example: Sheep in Asia

```r
livestock2 <- window(livestock, start=1970, end=2000)
fc1 <- ses(livestock2)
f22 <- holt(livestock2)
f33 <- holt(livestock2, damped = TRUE)

accuracy(fc1, livestock)
accuracy(fc2, livestock)
accuracy(fc3, livestock)
```

<table>
<thead>
<tr>
<th></th>
<th>SES</th>
<th>Linear trend</th>
<th>Damped trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test RMSE</td>
<td>25.46</td>
<td>11.88</td>
<td>14.73</td>
</tr>
<tr>
<td>Test MAE</td>
<td>20.38</td>
<td>10.71</td>
<td>13.30</td>
</tr>
<tr>
<td>Test MAPE</td>
<td>4.60</td>
<td>2.54</td>
<td>3.07</td>
</tr>
<tr>
<td>Test MASE</td>
<td>2.26</td>
<td>1.19</td>
<td>1.48</td>
</tr>
</tbody>
</table>
Example: Sheep in Asia

```r
autoplot(window(livestock, start=1970)) +
  autolayer(fc1, series="SES", PI=FALSE) +
  autolayer(fc2, series="Linear trend", PI=FALSE) +
  autolayer(fc3, series="Damped trend", PI=FALSE) +
  ylab("Livestock, sheep in Asia (millions)")
```
Outline

1. Simple exponential smoothing
2. Trend methods
3. Lab session 7
4. Seasonal methods
5. Lab session 8
6. Taxonomy of exponential smoothing methods
Lab Session 7
1. Simple exponential smoothing
2. Trend methods
3. Lab session 7
4. Seasonal methods
5. Lab session 8
6. Taxonomy of exponential smoothing methods
Holt-Winters additive method

Holt and Winters extended Holt’s method to capture seasonality.

Component form

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

- \( \ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \)
- \( b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \)
- \( s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \)

- \( k = \) integer part of \((h - 1)/m\). Ensures estimates from the final year are used for forecasting.
- Parameters: \( 0 \leq \alpha \leq 1, \ 0 \leq \beta^* \leq 1, \ 0 \leq \gamma \leq 1 - \alpha \) and \( m = \) period of seasonality (e.g. \( m = 4 \) for quarterly data).
Holt-Winters multiplicative method

For when seasonal variations are changing proportional to the level of the series.

**Component form**

\[
\hat{y}_{t+h|t} = (l_t + hb_t)s_{t+h-m(k+1)}.
\]

\[
l_t = \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1})
\]

\[
b_t = \beta^* (l_t - l_{t-1}) + (1 - \beta^*)b_{t-1}
\]

\[
s_t = \gamma \frac{y_t}{(l_{t-1} + b_{t-1})} + (1 - \gamma)s_{t-m}
\]

- \(k\) is integer part of \((h - 1)/m\).

- **Additive**: \(s_t\) in absolute terms: within each year \(\sum_i s_i \approx 0\).

- **Multiplicative**: \(s_t\) in relative terms: within each year \(\sum_i s_i \approx m\).
Example: Visitor Nights

```r
aust <- window(austourists,start=2005)
fcl <- hw(aust,seasonal="additive")
fcl2 <- hw(aust,seasonal="multiplicative")
```
Estimated components

Additive states

Multiplicative states
Holt-Winters damped method

Often the single most accurate forecasting method for seasonal data:

\[
\hat{y}_{t+h|t} = [\ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t]s_{t+h-m(k+1)}
\]

\[
\ell_t = \alpha \left( \frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})
\]

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

\[
st = \gamma \left( \frac{y_t}{(\ell_{t-1} + \phi b_{t-1})} \right) + (1 - \gamma)s_{t-m}
\]
Outline

1. Simple exponential smoothing
2. Trend methods
3. Lab session 7
4. Seasonal methods
5. Lab session 8
6. Taxonomy of exponential smoothing methods
Lab Session 8
Outline

1. Simple exponential smoothing
2. Trend methods
3. Lab session 7
4. Seasonal methods
5. Lab session 8
6. Taxonomy of exponential smoothing methods
# Exponential smoothing methods

<table>
<thead>
<tr>
<th>Trend Component</th>
<th>Seasonal Component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (None)</td>
</tr>
<tr>
<td>N (None)</td>
<td>(N,N)</td>
</tr>
<tr>
<td>A (Additive)</td>
<td>(A,N)</td>
</tr>
<tr>
<td>A&lt;sub&gt;d&lt;/sub&gt; (Additive damped)</td>
<td>(A&lt;sub&gt;d&lt;/sub&gt;,N)</td>
</tr>
</tbody>
</table>

- **(N,N):** Simple exponential smoothing
- **(A,N):** Holt’s linear method
- **(A<sub>d</sub>,N):** Additive damped trend method
- **(A,A):** Additive Holt-Winters’ method
- **(A,M):** Multiplicative Holt-Winters’ method
- **(A<sub>d</sub>,M):** Damped multiplicative Holt-Winters’ method
### Recursive formulae

<table>
<thead>
<tr>
<th>Trend</th>
<th>N</th>
<th>Seasonal</th>
<th>A</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{y}_{t+h</td>
<td>t} = \ell_t )</td>
<td>( \ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1} )</td>
<td>( \hat{y}_{t+h</td>
<td>t} = \ell_t + s_{t-m+h_m^+} )</td>
</tr>
<tr>
<td>( \ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)\ell_{t-1} )</td>
<td>( b_t = \beta^<em>(\ell_t - \ell_{t-1}) + (1 - \beta^</em>)b_{t-1} )</td>
<td>( s_t = \gamma(y_t - \ell_{t-1}) + (1 - \gamma)s_{t-m} )</td>
<td>( \ell_t = \alpha(y_t/s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) )</td>
<td>( b_t = \beta^<em>(\ell_t - \ell_{t-1}) + (1 - \beta^</em>)b_{t-1} )</td>
</tr>
<tr>
<td>( \ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) )</td>
<td>( b_t = \beta^<em>(\ell_t - \ell_{t-1}) + (1 - \beta^</em>)b_{t-1} )</td>
<td>( s_t = \gamma(y_t - s_{t-m} - b_{t-1}) + (1 - \gamma)s_{t-m} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
R functions

- Simple exponential smoothing: no trend.
  \texttt{ses(y)}

- Holt’s method: linear trend.
  \texttt{holt(y)}

- Damped trend method.
  \texttt{holt(y, damped=TRUE)}

- Holt-Winters methods
  \texttt{hw(y, damped=TRUE, seasonal="additive" )}
  \texttt{hw(y, damped=FALSE, seasonal="additive" )}
  \texttt{hw(y, damped=TRUE, seasonal="multiplicatively" )}
  \texttt{hw(y, damped=FALSE, seasonal="multiplicatively" )}

- Combination of no trend with seasonality not possible using these functions.