MONASH University

Rob J Hyndman

Forecasting:

Principles and Practice

3. Exponential smoothing I

OTexts.com/fpp/7/

Outline

State space perspective

- Observed data: $y_1, \dots, y_T$.
- Unobserved state: $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T$.

- Forecast $\hat{y}_{T+h|T} = \mathrm{E}(y_{T+h} | \boldsymbol{x}_T)$.
- The "forecast variance" is $\mathrm{Var}(y_{T+h} | \boldsymbol{x}_T)$.
- A prediction interval or "interval forecast" is a range of values of y_{T+h} with high probability.

State space perspective

- Observed data: $y_1, \dots, y_T$.
- Unobserved state: $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T$.

- Forecast $\hat{y}_{T+h|T} = \text{E}(y_{T+h}|\boldsymbol{x}_T)$.
- The "forecast variance" is $\text{Var}(y_{T+h}|\boldsymbol{x}_T)$.
- A prediction interval or "interval forecast" is a range of values of y_{T+h} with high probability.

State space perspective

- Observed data: $y_1, \dots, y_T$.
- Unobserved state: $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T$.

- Forecast $\hat{y}_{T+h|T} = \mathrm{E}(y_{T+h}|\boldsymbol{x}_T)$.
- The "forecast variance" is $\mathrm{Var}(y_{T+h}|\boldsymbol{x}_T)$.
- A prediction interval or "interval forecast" is a range of values of y_{T+h} with high probability.

State space perspective

- Observed data: $y_1, \dots, y_T$.
- Unobserved state: $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T$.

- Forecast $\hat{y}_{T+h|T} = \text{E}(y_{T+h}|\boldsymbol{x}_T)$.
- The "forecast variance" is $\text{Var}(y_{T+h}|\boldsymbol{x}_T)$.
- A prediction interval or "interval forecast" is a range of values of y_{T+h} with high probability.

Outline

Simple Exponential Smoothing

Component form

Forecast equation $\hat{y}_{t+h|t} = \ell_t$

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

$\ell_1 = \alpha y_1 + (1-\alpha)\ell_0$

Simple Exponential Smoothing

Component form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

$\ell_1 = \alpha y_1 + (1-\alpha)\ell_0$

Simple Exponential Smoothing

Component form

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

$$\ell_1 = \alpha y_1 + (1-\alpha)\ell_0$$

$$\ell_2 = \alpha y_2 + (1-\alpha)\ell_1 = \alpha y_2 + \alpha(1-\alpha)y_1 + (1-\alpha)^2\ell_0$$

Simple Exponential Smoothing

Component form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

$$\ell_1 = \alpha y_1 + (1-\alpha)\ell_0$$

$$\ell_2 = \alpha y_2 + (1-\alpha)\ell_1 = \alpha y_2 + \alpha(1-\alpha)y_1 + (1-\alpha)^2\ell_0$$

$$\ell_3 = \alpha y_3 + (1-\alpha)\ell_2 = \sum_{j=0}^{2}\alpha(1-\alpha)^j y_{3-j} + (1-\alpha)^3\ell_0$$

Simple Exponential Smoothing

Component form

Forecast equation $\hat{y}_{t+h|t} = \ell_t$

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

$$\ell_1 = \alpha y_1 + (1 - \alpha)\ell_0$$

$$\ell_2 = \alpha y_2 + (1 - \alpha)\ell_1 = \alpha y_2 + \alpha(1 - \alpha)y_1 + (1 - \alpha)^2\ell_0$$

$$\ell_3 = \alpha y_3 + (1 - \alpha)\ell_2 = \sum_{j=0}^{2} \alpha(1 - \alpha)^j y_{3-j} + (1 - \alpha)^3\ell_0$$

$$\vdots$$

$$\ell_t = \sum_{j=0}^{t-1} \alpha(1 - \alpha)^j y_{t-j} + (1 - \alpha)^t\ell_0$$

Simple Exponential Smoothing

Forecast equation

$$\hat{y}_{t+h|t} = \sum_{j=1}^{t} \alpha(1-\alpha)^{t-j} y_j + (1-\alpha)^t \ell_0, \qquad (0 \le \alpha \le 1)$$

	Weights assigned to observations for:			
Observation	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$
y_t	0.2	0.4	0.6	0.8
y_{t-1}	0.16	0.24	0.24	0.16
y_{t-2}	0.128	0.144	0.096	0.032
y_{t-3}	0.1024	0.0864	0.0384	0.0064
y_{t-4}	$(0.2)(0.8)^4$	$(0.4)(0.6)^4$	$(0.6)(0.4)^4$	$(0.8)(0.2)^4$
y_{t-5}	$(0.2)(0.8)^5$	$(0.4)(0.6)^5$	$(0.6)(0.4)^5$	$(0.8)(0.2)^5$

- Limiting cases: $\alpha \to 1, \quad \alpha \to 0.$

Simple Exponential Smoothing

Forecast equation

$$\hat{y}_{t+h|t} = \sum_{j=1}^{t} \alpha(1-\alpha)^{t-j} y_j + (1-\alpha)^t \ell_0, \qquad (0 \le \alpha \le 1)$$

	Weights assigned to observations for:			
Observation	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$
y_t	0.2	0.4	0.6	0.8
y_{t-1}	0.16	0.24	0.24	0.16
y_{t-2}	0.128	0.144	0.096	0.032
y_{t-3}	0.1024	0.0864	0.0384	0.0064
y_{t-4}	$(0.2)(0.8)^4$	$(0.4)(0.6)^4$	$(0.6)(0.4)^4$	$(0.8)(0.2)^4$
y_{t-5}	$(0.2)(0.8)^5$	$(0.4)(0.6)^5$	$(0.6)(0.4)^5$	$(0.8)(0.2)^5$

- Limiting cases: $\alpha \to 1$, $\quad \alpha \to 0$.

Simple Exponential Smoothing

Forecast equation

$$\hat{y}_{t+h|t} = \sum_{j=1}^{t} \alpha(1-\alpha)^{t-j} y_j + (1-\alpha)^t \ell_0, \qquad (0 \le \alpha \le 1)$$

	Weights assigned to observations for:			
Observation	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$
y_t	0.2	0.4	0.6	0.8
y_{t-1}	0.16	0.24	0.24	0.16
y_{t-2}	0.128	0.144	0.096	0.032
y_{t-3}	0.1024	0.0864	0.0384	0.0064
y_{t-4}	$(0.2)(0.8)^4$	$(0.4)(0.6)^4$	$(0.6)(0.4)^4$	$(0.8)(0.2)^4$
y_{t-5}	$(0.2)(0.8)^5$	$(0.4)(0.6)^5$	$(0.6)(0.4)^5$	$(0.8)(0.2)^5$

- Limiting cases: $\alpha \to 1$, $\quad \alpha \to 0$.

Simple Exponential Smoothing

Component form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

State space form

Observation equation	$y_t = \ell_{t-1} + e_t$
State equation	$\ell_t = \ell_{t-1} + \alpha e_t$

- $e_t = y_t - \ell_{t-1} = y_t - \hat{y}_{t|t-1}$ for $t = 1, \dots, T$, the one-step within-sample forecast error at time t.
- ℓ_t is an unobserved "state".
- Need to estimate α and ℓ_0.

Simple Exponential Smoothing

Component form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

State space form

Observation equation	$y_t = \ell_{t-1} + e_t$
State equation	$\ell_t = \ell_{t-1} + \alpha e_t$

- $e_t = y_t - \ell_{t-1} = y_t - \hat{y}_{t|t-1}$ for $t = 1, \dots, T$, the one-step within-sample forecast error at time t.
- ℓ_t is an unobserved "state".
- Need to estimate α and ℓ_0.

Simple Exponential Smoothing

Component form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

State space form

Observation equation	$y_t = \ell_{t-1} + e_t$
State equation	$\ell_t = \ell_{t-1} + \alpha e_t$

- $e_t = y_t - \ell_{t-1} = y_t - \hat{y}_{t|t-1}$ for $t = 1, \dots, T$, the one-step within-sample forecast error at time t.
- ℓ_t is an unobserved "state".
- Need to estimate α and ℓ_0.

Simple Exponential Smoothing

Component form

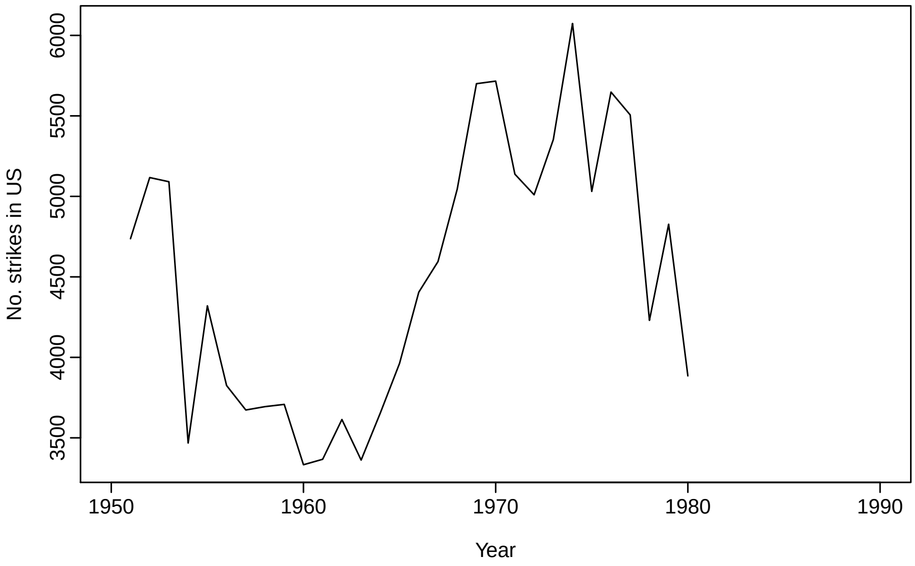
Forecast equation	$\hat{y}_{t+h\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

State space form

Observation equation	$y_t = \ell_{t-1} + e_t$
State equation	$\ell_t = \ell_{t-1} + \alpha e_t$

- $e_t = y_t - \ell_{t-1} = y_t - \hat{y}_{t|t-1}$ for $t = 1, \dots, T$, the one-step within-sample forecast error at time t.
- ℓ_t is an unobserved "state".
- Need to estimate α and ℓ_0.

Simple exponential smoothing



Simple exponential smoothing

Optimisation

- Need to choose value for α and ℓ_0
- Similarly to regression — we choose α and ℓ_0 by minimising MSE:

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

- Unlike regression there is no closed form solution — use numerical optimization.

Optimisation

- Need to choose value for α and ℓ_0
- Similarly to regression — we choose α and ℓ_0 by minimising MSE:

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

- Unlike regression there is no closed form solution — use numerical optimization.

Optimisation

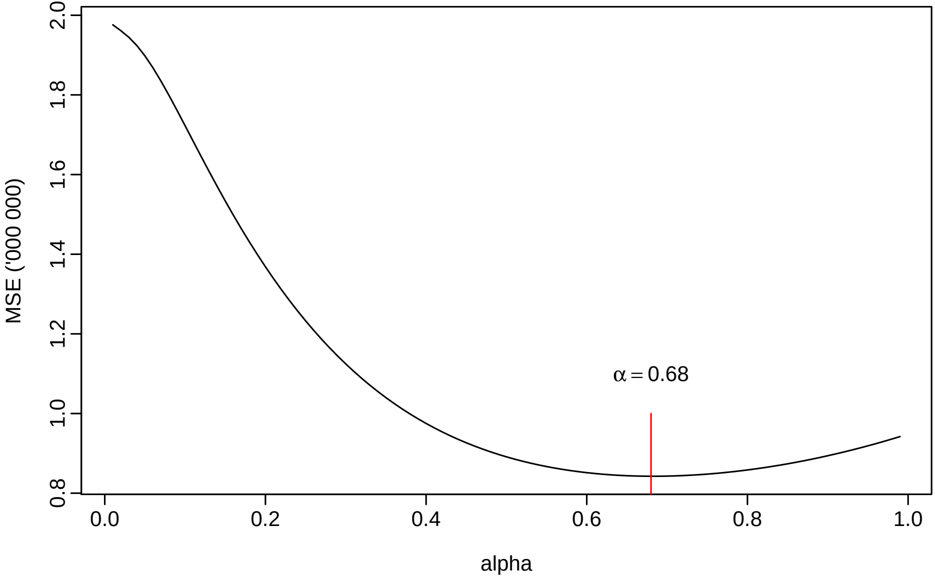
- Need to choose value for α and ℓ_0
- Similarly to regression — we choose α and ℓ_0 by minimising MSE:

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

- Unlike regression there is no closed form solution — use numerical optimization.

Simple exponential smoothing

Simple exponential smoothing



Simple exponential smoothing

Multi-step forecasts

$$\hat{y}_{T+h|T} = \hat{y}_{T+1|T}, \qquad h = 2, 3, \ldots$$

- A "flat" forecast function.
- Remember, a forecast is an estimated mean of a future value.
- So with no trend, no seasonality, and no other patterns, the forecasts are constant.

Simple exponential smoothing

Multi-step forecasts

$$\hat{y}_{T+h|T} = \hat{y}_{T+1|T}, \qquad h = 2, 3, \ldots$$

- A “flat” forecast function.
- Remember, a forecast is an estimated mean of a future value.
- So with no trend, no seasonality, and no other patterns, the forecasts are constant.

Simple exponential smoothing

Multi-step forecasts

$$\hat{y}_{T+h|T} = \hat{y}_{T+1|T}, \qquad h = 2, 3, \dots$$

- A “flat” forecast function.
- Remember, a forecast is an estimated mean of a future value.
- So with no trend, no seasonality, and no other patterns, the forecasts are constant.

Simple exponential smoothing

Multi-step forecasts

$$\hat{y}_{T+h|T} = \hat{y}_{T+1|T}, \qquad h = 2, 3, \ldots$$

- A "flat" forecast function.
- Remember, a forecast is an estimated mean of a future value.
- So with no trend, no seasonality, and no other patterns, the forecasts are constant.

SES in R

```
library(fpp)

fit <- ses(oil, h=3)

plot(fit)

summary(fit)
```

Outline

Holt's local trend method

- Holt (1957) extended SES to allow forecasting of data with trends.
- Two smoothing parameters: α and β^* (with values between 0 and 1).

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

Holt’s local trend method

- Holt (1957) extended SES to allow forecasting of data with trends.
- Two smoothing parameters: α and β^* (with values between 0 and 1).

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

- ℓ_t denotes an estimate of the level of the series at time t

Holt's local trend method

- Holt (1957) extended SES to allow forecasting of data with trends.
- Two smoothing parameters: α and β^* (with values between 0 and 1).

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

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the slope of the series at time t.

Holt's local trend method

- Holt (1957) extended SES to allow forecasting of data with trends.
- Two smoothing parameters: α and β^* (with values between 0 and 1).

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

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the slope of the series at time t.

Holt's local trend method

- Holt (1957) extended SES to allow forecasting of data with trends.
- Two smoothing parameters: α and β^* (with values between 0 and 1).

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

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the slope of the series at time t.

Holt's local trend method

- Holt (1957) extended SES to allow forecasting of data with trends.
- Two smoothing parameters: α and β^* (with values between 0 and 1).

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

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the slope of the series at time t.

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

State space form

Observation equation	$y_t = \ell_{t-1} + b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + b_{t-1} + \alpha e_t$
	$b_t = b_{t-1} + \beta e_t$

- $\beta = \alpha\beta^*$
- $e_t = y_t - (\ell_{t-1} + b_{t-1}) = y_t - \hat{y}_{t|t-1}$
- Need to estimate $\alpha, \beta, \ell_0, b_0$.

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

State space form

Observation equation	$y_t = \ell_{t-1} + b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + b_{t-1} + \alpha e_t$
	$b_t = b_{t-1} + \beta e_t$

- $\beta = \alpha\beta^*$
- $e_t = y_t - (\ell_{t-1} + b_{t-1}) = y_t - \hat{y}_{t|t-1}$
- Need to estimate $\alpha, \beta, \ell_0, b_0$.

Holt's linear trend

Component form

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

State space form

$$
\begin{aligned}
\text{Observation equation} \qquad y_t &= \ell_{t-1} + b_{t-1} + e_t \\
\text{State equations} \qquad \ell_t &= \ell_{t-1} + b_{t-1} + \alpha e_t \\
b_t &= b_{t-1} + \beta e_t
\end{aligned}
$$

- $\beta = \alpha\beta^*$
- $e_t = y_t - (\ell_{t-1} + b_{t-1}) = y_t - \hat{y}_{t|t-1}$
- Need to estimate $\alpha, \beta, \ell_0, b_0$.

Holt's linear trend

Component form

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

State space form

$$\begin{aligned} \text{Observation equation} \qquad y_t &= \ell_{t-1} + b_{t-1} + e_t \\ \text{State equations} \qquad \ell_t &= \ell_{t-1} + b_{t-1} + \alpha e_t \\ b_t &= b_{t-1} + \beta e_t \end{aligned}$$

- $\beta = \alpha\beta^*$
- $e_t = y_t - (\ell_{t-1} + b_{t-1}) = y_t - \hat{y}_{t|t-1}$
- Need to estimate $\alpha, \beta, \ell_0, b_0$.

Holt’s method in R

```
fit2 <- holt(ausair, h=5)

plot(fit2)

summary(fit2)
```

Holt's method in R

```r
fit1 <- holt(strikes)
plot(fit1$model)
plot(fit1, plot.conf=FALSE)
lines(fitted(fit1), col="red")
fit1$model

fit2 <- ses(strikes)
plot(fit2$model)
plot(fit2, plot.conf=FALSE)
lines(fit1$mean, col="red")

accuracy(fit1)
accuracy(fit2)
```

Comparing Holt and SES

- Holt's method will almost always have better in-sample RMSE because it is optimized over one additional parameter.
- It may not be better on other measures.
- You need to compare out-of-sample RMSE (using a test set) for the comparison to be useful.
- But we don't have enough data.
- A better method for comparison will be in the next session!

Comparing Holt and SES

- Holt's method will almost always have better in-sample RMSE because it is optimized over one additional parameter.
- It may not be better on other measures.
- You need to compare out-of-sample RMSE (using a test set) for the comparison to be useful.
- But we don't have enough data.
- A better method for comparison will be in the next session!

Comparing Holt and SES

- Holt's method will almost always have better in-sample RMSE because it is optimized over one additional parameter.
- It may not be better on other measures.
- You need to compare out-of-sample RMSE (using a test set) for the comparison to be useful.
- But we don't have enough data.
- A better method for comparison will be in the next session!

Comparing Holt and SES

- Holt's method will almost always have better in-sample RMSE because it is optimized over one additional parameter.
- It may not be better on other measures.
- You need to compare out-of-sample RMSE (using a test set) for the comparison to be useful.
- But we don't have enough data.
- A better method for comparison will be in the next session!

Comparing Holt and SES

- Holt's method will almost always have better in-sample RMSE because it is optimized over one additional parameter.
- It may not be better on other measures.
- You need to compare out-of-sample RMSE (using a test set) for the comparison to be useful.
- But we don't have enough data.
- A better method for comparison will be in the next session!

Exponential trend method

Multiplicative version of Holt's method

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t b_t^h$
Observation equation	$y_t = (\ell_{t-1} b_{t-1}) + e_t$
State equations	$\ell_t = \ell_{t-1} b_{t-1} + \alpha e_t$
	$b_t = b_{t-1} + \beta e_t / \ell_{t-1}$

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the relative growth of the series at time t.
- In R: holt(x, exponential=TRUE)

Exponential trend method

Multiplicative version of Holt's method

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t b_t^h$
Observation equation	$y_t = (\ell_{t-1} b_{t-1}) + e_t$
State equations	$\ell_t = \ell_{t-1} b_{t-1} + \alpha e_t$
	$b_t = b_{t-1} + \beta e_t / \ell_{t-1}$

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the relative growth of the series at time t.
- In R: `holt(x, exponential=TRUE)`

Exponential trend method

Multiplicative version of Holt's method

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t b_t^h$
Observation equation	$y_t = (\ell_{t-1} b_{t-1}) + e_t$
State equations	$\ell_t = \ell_{t-1} b_{t-1} + \alpha e_t$
	$b_t = b_{t-1} + \beta e_t / \ell_{t-1}$

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the relative growth of the series at time t.
- In R: `holt(x, exponential=TRUE)`

Exponential trend method

Multiplicative version of Holt's method

State space form

$$
\begin{aligned}
\text{Forecast equation} \quad & \hat{y}_{t+h|t} = \ell_t b_t^h \\
\text{Observation equation} \quad & y_t = (\ell_{t-1} b_{t-1}) + e_t \\
\text{State equations} \quad & \ell_t = \ell_{t-1} b_{t-1} + \alpha e_t \\
& b_t = b_{t-1} + \beta e_t / \ell_{t-1}
\end{aligned}
$$

- ℓ_t denotes an estimate of the level of the series at time t
- b_t denotes an estimate of the relative growth of the series at time t.
- In R: `holt(x, exponential=TRUE)`

Additive damped trend

- Gardner and McKenzie (1985) suggested that the trends should be "damped" to be more conservative for longer forecast horizons.
- Damping parameter $0 < \phi < 1$.

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t$
Observation equation	$y_t = \ell_{t-1} + \phi b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha e_t$
	$b_t = \phi b_{t-1} + \beta e_t$

Additive damped trend

- Gardner and McKenzie (1985) suggested that the trends should be “damped” to be more conservative for longer forecast horizons.
- Damping parameter $0 < \phi < 1$.

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t$
Observation equation	$y_t = \ell_{t-1} + \phi b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha e_t$
	$b_t = \phi b_{t-1} + \beta e_t$

- If $\phi = 1$, identical to Holt's linear trend.

Additive damped trend

- Gardner and McKenzie (1985) suggested that the trends should be "damped" to be more conservative for longer forecast horizons.
- Damping parameter $0 < \phi < 1$.

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t$
Observation equation	$y_t = \ell_{t-1} + \phi b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha e_t$
	$b_t = \phi b_{t-1} + \beta e_t$

- If $\phi = 1$, identical to Holt's linear trend.
- As $h \to \infty$, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T/(1-\phi)$.
- Short-run forecasts trended, long-run forecasts constant.

Additive damped trend

- Gardner and McKenzie (1985) suggested that the trends should be "damped" to be more conservative for longer forecast horizons.
- Damping parameter $0 < \phi < 1$.

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t$
Observation equation	$y_t = \ell_{t-1} + \phi b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha e_t$
	$b_t = \phi b_{t-1} + \beta e_t$

- If $\phi = 1$, identical to Holt's linear trend.
- As $h \to \infty$, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T/(1-\phi)$.
- Short-run forecasts trended, long-run forecasts constant.

Additive damped trend

- Gardner and McKenzie (1985) suggested that the trends should be "damped" to be more conservative for longer forecast horizons.
- Damping parameter $0 < \phi < 1$.

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t$
Observation equation	$y_t = \ell_{t-1} + \phi b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha e_t$
	$b_t = \phi b_{t-1} + \beta e_t$

- If $\phi = 1$, identical to Holt's linear trend.
- As $h \to \infty$, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T/(1-\phi)$.
- Short-run forecasts trended, long-run forecasts constant.

Additive damped trend

- Gardner and McKenzie (1985) suggested that the trends should be "damped" to be more conservative for longer forecast horizons.
- Damping parameter $0 < \phi < 1$.

State space form

Forecast equation	$\hat{y}_{t+h\|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t$
Observation equation	$y_t = \ell_{t-1} + \phi b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha e_t$
	$b_t = \phi b_{t-1} + \beta e_t$

- If $\phi = 1$, identical to Holt's linear trend.
- As $h \to \infty$, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T/(1 - \phi)$.
- Short-run forecasts trended, long-run forecasts constant.

Additive damped trend

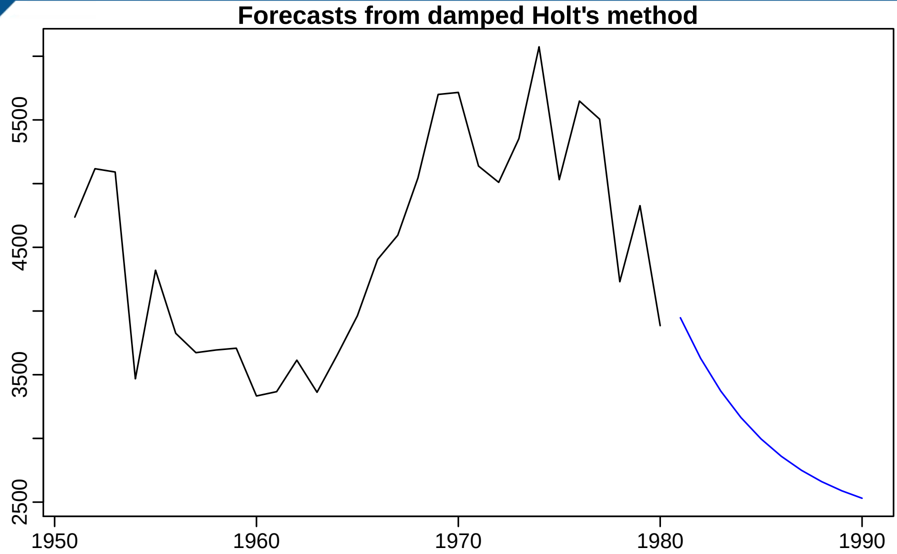
- Gardner and McKenzie (1985) suggested that the trends should be "damped" to be more conservative for longer forecast horizons.
- Damping parameter $0 < \phi < 1$.

State space form

Forecast equation	$\hat{y}_{t+h\vert t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h) b_t$
Observation equation	$y_t = \ell_{t-1} + \phi b_{t-1} + e_t$
State equations	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha e_t$
	$b_t = \phi b_{t-1} + \beta e_t$

- If $\phi = 1$, identical to Holt's linear trend.
- As $h \to \infty$, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T/(1 - \phi)$.
- Short-run forecasts trended, long-run forecasts constant.

Damped trend method



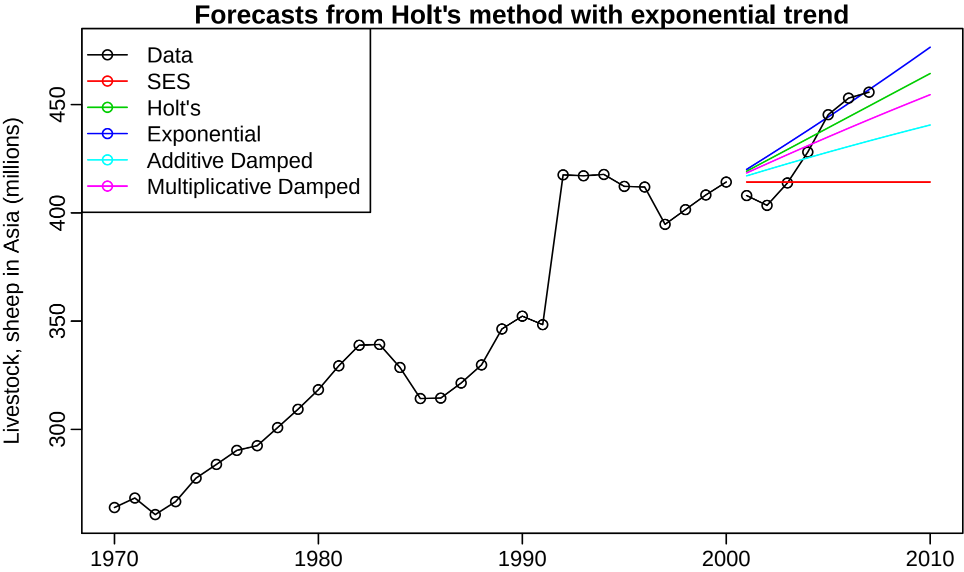
Trend methods in R

```
fit4 <- holt(air, h=5, damped=TRUE)

plot(fit4)

summary(fit4)
```

Example: Sheep in Asia



Multiplicative damped trend method

Taylor (2003) introduced multiplicative damping.

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

- $\phi = 1$ gives exponential trend method
- Forecasts converge to $\ell_T + b_T^{\phi/(1-\phi)}$ as $h \to \infty$.

Multiplicative damped trend method

Taylor (2003) introduced multiplicative damping.

$$\hat{y}_{t+h|t} = \ell_t b_t^{(\phi+\phi^2+\dots+\phi^h)}$$
$$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} b_{t-1}^{\phi})$$
$$b_t = \beta^*(\ell_t/\ell_{t-1}) + (1-\beta^*) b_{t-1}^{\phi}$$

- $\phi = 1$ gives exponential trend method
- Forecasts converge to $\ell_T + b_T^{\phi/(1-\phi)}$ as $h \rightarrow \infty$.

Multiplicative damped trend method

Taylor (2003) introduced multiplicative damping.

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

- $\phi = 1$ gives exponential trend method
- Forecasts converge to $\ell_T + b_T^{\phi/(1-\phi)}$ as $h \to \infty$.

Outline

Holt-Winters additive method

- Holt and Winters extended Holt's method to capture seasonality.
- Three smoothing equations—one for the level, one for trend, and one for seasonality.
- Parameters: $0 \le \alpha \le 1$, $0 \le \beta^* \le 1$, $0 \le \gamma \le 1-\alpha$ and $m =$ period of seasonality.

State space form

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+}$$

$$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + e_t$$

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha e_t$$

$$b_t = b_{t-1} + \beta e_t$$

$$s_t = s_{t-m} + \gamma e_t.$$

Holt-Winters additive method

- Holt and Winters extended Holt's method to capture seasonality.
- Three smoothing equations—one for the level, one for trend, and one for seasonality.
- Parameters: $0 \le \alpha \le 1$, $0 \le \beta^* \le 1$, $0 \le \gamma \le 1 - \alpha$ and $m =$ period of seasonality.

State space form

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+}$$

$$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + e_t$$

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha e_t$$

$$b_t = b_{t-1} + \beta e_t$$

$$s_t = s_{t-m} + \gamma e_t.$$

Holt-Winters additive method

- Holt and Winters extended Holt's method to capture seasonality.
- Three smoothing equations—one for the level, one for trend, and one for seasonality.
- Parameters: $0 \le \alpha \le 1$, $0 \le \beta^* \le 1$, $0 \le \gamma \le 1 - \alpha$ and $m =$ period of seasonality.

State space form

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+}$$

$$\begin{aligned}
y_t &= \ell_{t-1} + b_{t-1} + s_{t-m} + e_t \\
\ell_t &= \ell_{t-1} + b_{t-1} + \alpha e_t \\
b_t &= b_{t-1} + \beta e_t \\
s_t &= s_{t-m} + \gamma e_t.
\end{aligned}$$

Holt-Winters additive method

- Holt and Winters extended Holt's method to capture seasonality.
- Three smoothing equations—one for the level, one for trend, and one for seasonality.
- Parameters: $0 \le \alpha \le 1$, $0 \le \beta^* \le 1$, $0 \le \gamma \le 1-\alpha$ and $m =$ period of seasonality.

State space form

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+}$$

$$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + e_t$$

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha e_t$$

$$b_t = b_{t-1} + \beta e_t$$

$$s_t = s_{t-m} + \gamma e_t.$$

Holt-Winters additive method

- Holt and Winters extended Holt's method to capture seasonality.
- Three smoothing equations—one for the level, one for trend, and one for seasonality.
- Parameters: $0 \le \alpha \le 1$, $0 \le \beta^* \le 1$, $0 \le \gamma \le 1 - \alpha$ and $m =$ period of seasonality.

State space form

$$\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + hb_t + s_{t-m+h_m^+} \\
y_t &= \ell_{t-1} + b_{t-1} + s_{t-m} + e_t \\
\ell_t &= \ell_{t-1} + b_{t-1} + \alpha e_t \\
b_t &= b_{t-1} + \beta e_t \\
s_t &= s_{t-m} + \gamma e_t.
\end{aligned}$$

Holt-Winters additive method

- Holt and Winters extended Holt's method to capture seasonality.
- Three smoothing equations—one for the level, one for trend, and one for seasonality.
- Parameters: $0 \le \alpha \le 1$, $0 \le \beta^* \le 1$, $0 \le \gamma \le 1 - \alpha$ and $m =$ period of seasonality.

State space form

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+} \qquad h_m^+ = \lfloor (h-1) \mod m \rfloor + 1$$

$$\begin{aligned}
y_t &= \ell_{t-1} + b_{t-1} + s_{t-m} + e_t \\
\ell_t &= \ell_{t-1} + b_{t-1} + \alpha e_t \\
b_t &= b_{t-1} + \beta e_t \\
s_t &= s_{t-m} + \gamma e_t.
\end{aligned}$$

Holt-Winters multiplicative method

Holt-Winters multiplicative method

$$\hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t-m+h_m^+}$$
$$y_t = (\ell_{t-1} + b_{t-1})s_{t-m} + e_t$$
$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha e_t / s_{t-m}$$
$$b_t = b_{t-1} + \beta e_t / s_{t-m}$$
$$s_t = s_{t-m} + \gamma e_t / (\ell_{t-1} + b_{t-1}).$$

- Most textbooks use $s_t = \gamma(y_t/\ell_t) + (1-\gamma)s_{t-m}$
- We optimize for α, β^*, γ, ℓ_0, b_0, s_0, $s_{-1}, \dots, s_{1-m}$.

Holt-Winters multiplicative method

Holt-Winters multiplicative method

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

- Most textbooks use $s_t = \gamma(y_t/\ell_t) + (1-\gamma)s_{t-m}$
- We optimize for α, β^*, γ, ℓ_0, b_0, $s_0, s_{-1}, \dots, s_{1-m}$.

Seasonal methods in R

```r
aus1 <- hw(austourists)
aus2 <- hw(austourists, seasonal="mult")

plot(aus1)
plot(aus2)

summary(aus1)
summary(aus2)
```

Holt-Winters damped method

Often the single most accurate forecasting method for seasonal data:

State space form

$$
\begin{aligned}
y_t &= (\ell_{t-1} + \phi b_{t-1}) s_{t-m} + e_t \\
\ell_t &= \ell_{t-1} + \phi b_{t-1} + \alpha e_t / s_{t-m} \\
b_t &= \phi b_{t-1} + \beta e_t / s_{t-m} \\
s_t &= s_{t-m} + \gamma e_t / (\ell_{t-1} + \phi b_{t-1}).
\end{aligned}
$$

Seasonal methods in R

```
aus3 <- hw(austourists, seasonal="mult",
    damped=TRUE)

summary(aus3)

plot(aus3)
```

Outline

Exponential smoothing methods

- Simple exponential smoothing: no trend.
  `ses(x)`
- Holt's method: linear trend.
  `holt(x)`
- Exponential trend method.
  `holt(x, exponential=TRUE)`
- Damped trend method.
  `holt(x, damped=TRUE)`
- Holt-Winters methods
  `hw(x, damped=TRUE, exponential=TRUE, seasonal="additive")`

Exponential smoothing methods

- Simple exponential smoothing: no trend.
  `ses(x)`
- Holt's method: linear trend.
  `holt(x)`
- Exponential trend method.
  `holt(x, exponential=TRUE)`
- Damped trend method.
  `holt(x, damped=TRUE)`
- Holt-Winters methods
  `hw(x, damped=TRUE, exponential=TRUE, seasonal="additive")`

Exponential smoothing methods

- Simple exponential smoothing: no trend.
  `ses(x)`
- Holt's method: linear trend.
  `holt(x)`
- Exponential trend method.
  `holt(x, exponential=TRUE)`
- Damped trend method.
  `holt(x, damped=TRUE)`
- Holt-Winters methods
  `hw(x, damped=TRUE, exponential=TRUE, seasonal="additive")`

Exponential smoothing methods

- Simple exponential smoothing: no trend.
  `ses(x)`
- Holt's method: linear trend.
  `holt(x)`
- Exponential trend method.
  `holt(x, exponential=TRUE)`
- Damped trend method.
  `holt(x, damped=TRUE)`
- Holt-Winters methods
  `hw(x, damped=TRUE, exponential=TRUE, seasonal="additive")`

Exponential smoothing methods

- Simple exponential smoothing: no trend.
  `ses(x)`
- Holt's method: linear trend.
  `holt(x)`
- Exponential trend method.
  `holt(x, exponential=TRUE)`
- Damped trend method.
  `holt(x, damped=TRUE)`
- Holt-Winters methods
  ```
  hw(x, damped=TRUE, exponential=TRUE,
    seasonal="additive")
  ```