THE UNIVERSITY OF TEXAS AT AUSTIN
McCOMBS SCHOOL OF BUSINESS

Time Series: Smoothing & Moving Average

Lecture 19

STA 371G

Trend & Seasonality in Time Series

- Although desired, most of the time series data are not stationary.

Trend & Seasonality in Time Series

- Although desired, most of the time series data are not stationary.
- Two main factors that make data non-stationary: Trend and seasonality.

Trend & Seasonality in Time Series

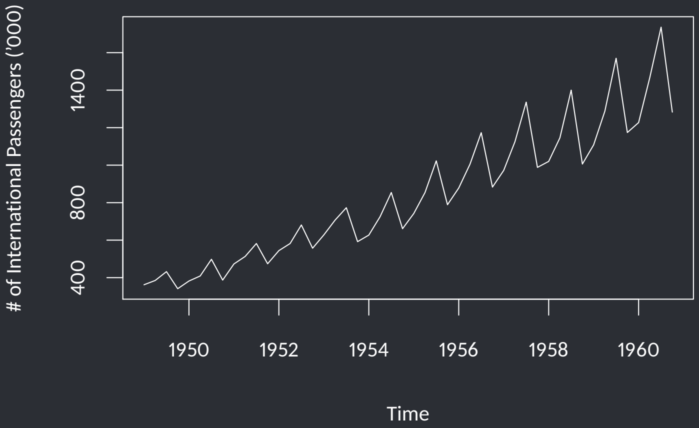
- Although desired, most of the time series data are not stationary.
- Two main factors that make data non-stationary: Trend and seasonality.
- Sales, economic, activity etc. data often show strong seasonality.

Trend & Seasonality in Time Series

- Although desired, most of the time series data are not stationary.
- Two main factors that make data non-stationary: Trend and seasonality.
- Sales, economic, activity etc. data often show strong seasonality.
- E.g. Quarterly totals of international airline passengers, 1949 to 1960.

Airline Passengers 1949-1960

```
# Convert data into time series, starting from 1st quarter of 1949.
air <- ts(air_passengers$number, start=c(1949,1), frequency=4)
# Frequency: 4 data points per year (this is quarterly data)
plot(air, ylab="# of International Passengers ('000)")
```



What is seasonality, exactly?

- Seasonality: The presence of variations that occur at specific regular intervals in a season.

What is seasonality, exactly?

- Seasonality: The presence of variations that occur at specific regular intervals in a season.
- In this case, a season is a year.

What is seasonality, exactly?

- Seasonality: The presence of variations that occur at specific regular intervals in a season.
- In this case, a season is a year.
- The frequency, the number of observations in a season, is 4.

What is seasonality, exactly?

- Seasonality: The presence of variations that occur at specific regular intervals in a season.
- In this case, a season is a year.
- The frequency, the number of observations in a season, is 4.
- The data behaves similarly in the same quarter of different years.

What is seasonality, exactly?

- Seasonality: The presence of variations that occur at specific regular intervals in a season.
- In this case, a season is a year.
- The frequency, the number of observations in a season, is 4.
- The data behaves similarly in the same quarter of different years.
- E.g. every year, the number of passengers peaks in the 3rd quarter. Similarly, it dips in the 4th quarter.

Exploring the data

- Overall, there is an upward trend in the number of airline passengers between 1949-1960.

Exploring the data

- Overall, there is an upward trend in the number of airline passengers between 1949-1960.
- The seasonality in the data makes it difficult to see how the trend changes over years.

Exploring the data

- Overall, there is an upward trend in the number of airline passengers between 1949-1960.
- The seasonality in the data makes it difficult to see how the trend changes over years.
- E.g. it is not clear wether the number is increasing at the same rate every year, or the rate varies in different years.

Exploring the data

- Overall, there is an upward trend in the number of airline passengers between 1949-1960.
- The seasonality in the data makes it difficult to see how the trend changes over years.
- E.g. it is not clear wether the number is increasing at the same rate every year, or the rate varies in different years.
- The trend in the data makes it difficult to see the impact of seasonality in the data.

Exploring the data

- Overall, there is an upward trend in the number of airline passengers between 1949-1960.
- The seasonality in the data makes it difficult to see how the trend changes over years.
- E.g. it is not clear wether the number is increasing at the same rate every year, or the rate varies in different years.
- The trend in the data makes it difficult to see the impact of seasonality in the data.
- E.g. it is not clear why the 2nd quarters are higher than the 1st quarters: Is it a more active traveling period of the year or is it just more people are using air travel as time goes by?

Exploring the data

- Overall, there is an upward trend in the number of airline passengers between 1949-1960.
- The seasonality in the data makes it difficult to see how the trend changes over years.
- E.g. it is not clear wether the number is increasing at the same rate every year, or the rate varies in different years.
- The trend in the data makes it difficult to see the impact of seasonality in the data.
- E.g. it is not clear why the 2nd quarters are higher than the 1st quarters: Is it a more active traveling period of the year or is it just more people are using air travel as time goes by?
- Smoothing the data helps identifying the trend and seasonal effects in a clearer way.

Smoothing the data: one-sided moving average

(One-sided) Moving average is a simple average of all observations over the previous season (year).

y_t : Number of passengers traveled at time t

One-sided moving average at time t:

$$m_t = \frac{1}{4}y_{t-3} + \frac{1}{4}y_{t-2} + \frac{1}{4}y_{t-1} + \frac{1}{4}y_t$$

Each quarter has equal weight in the moving average.

Smoothing the data: one-sided moving average

	Y49Q1	Y49Q2	Y49Q3	Y49Q4	Y50Q1	Y50Q2	...
t	1	2	3	4	5	6	...
y_t	362	385	432	341	382	409	...
m_t	NA	NA	NA	380	385	391	...

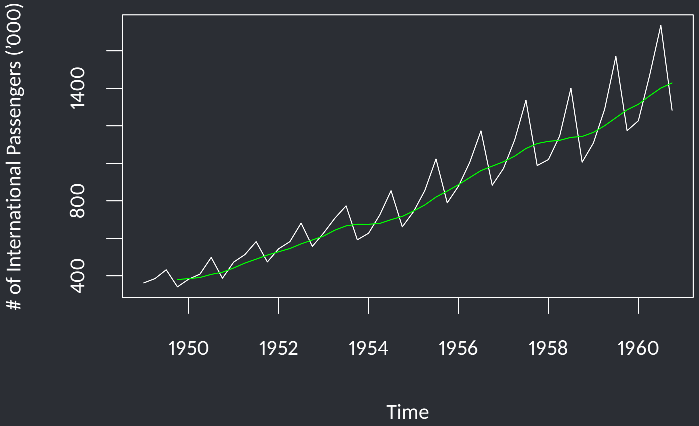
Smoothing the data: one-sided moving average

	Y49Q1	Y49Q2	Y49Q3	Y49Q4	Y50Q1	Y50Q2	...
t	1	2	3	4	5	6	...
y_t	362	385	432	341	382	409	...
m_t	NA	NA	NA	380	385	391	...

$$\begin{aligned} m_4 &= \frac{1}{4}y_1 + \frac{1}{4}y_2 + \frac{1}{4}y_3 + \frac{1}{4}y_4 \\ &= \frac{1}{4}362 + \frac{1}{4}385 + \frac{1}{4}432 + \frac{1}{4}341 \\ &= 380 \end{aligned}$$

Smoothing the data: one-sided moving average

```
# One sided moving average
air_ma_one <- filter(air, filter=c(1/4,1/4,1/4,1/4), sides=1)
plot(air, ylab="# of International Passengers ('000)")
lines(air_ma_one, col='green')
```



Smoothing the data

Smoothed data better shows the slow-downs in the trend around 1954 and 1959.

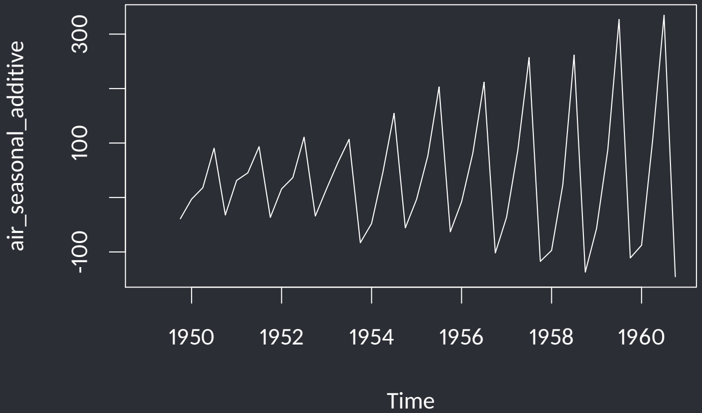
Let's also see the effect of the seasonality by eliminating the trend in the data. This is called detrending.

Smoothing the data

When trend and seasonality are additive

$$\text{Data} = \text{Trend} + \text{Seasonality} + \text{Randomness}$$

```
air_seasonal_additive <- air - air_ma_one
plot(air_seasonal_additive)
```

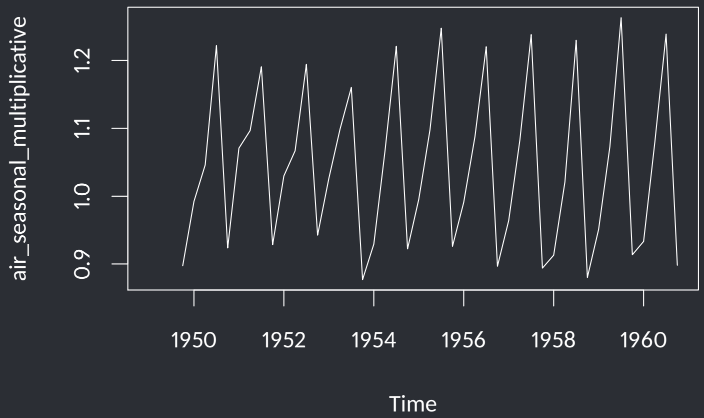


Smoothing the data

When trend and seasonality are multiplicative

$$\text{Data} = \text{Trend} \times \text{Seasonality} \times \text{Randomness}$$

```
air_seasonal_multiplicative <- air/ air_ma_one
plot(air_seasonal_multiplicative)
```



Forecasting using Simple Exponential Smoothing

So far, we have smoothed the data to better observe the trend and the seasonality.

In general, what we really want is to forecast the future numbers.

Simple exponential smoothing (SES) is one of the well known forecasting methods.

Simple Exponential Smoothing

In SES:

- Start with a forecast for the next step.

Simple Exponential Smoothing

In SES:

- Start with a forecast for the next step.
- Observe the actual number and the error in the forecast.

Simple Exponential Smoothing

In SES:

- Start with a forecast for the next step.
- Observe the actual number and the error in the forecast.
- Adjust your forecast based on the error.

Simple Exponential Smoothing

In SES:

- Start with a forecast for the next step.
- Observe the actual number and the error in the forecast.
- Adjust your forecast based on the error.
- Use your adjusted forecast for the subsequent step.

Simple Exponential Smoothing

$\hat{y}_t$: Forecast for time t
y_t: Actual number at time t
e_t: Error in time t. $e_t = y_t - \hat{y}_t$
α: Smoothing constant (to adjust for the error)

$$\hat{y}_{t+1} = \hat{y}_t + \alpha e_t$$

	t	y_t	$\hat{y}_t$	e_t	αe_t ($\alpha = 0.5$)
1949Q1	1	112	-	-	-
1949Q2	2	118	112	6	3
1949Q3	3	132	115	17	8.5
1949Q4	4	129	123.5	5.5	2.75
1950Q1	5	121	126.25	-5.25	-2.625
...	...	...	...	...	...

Simple Exponential Smoothing

$$\hat{y}_5 = \hat{y}_4 + \alpha e_4$$
$$\hat{y}_5 = \hat{y}_4 + \alpha(y_4 - \hat{y}_4)$$
$$\hat{y}_5 = \alpha y_4 + (1 - \alpha)\hat{y}_4$$

$$\hat{y}_4 = \hat{y}_3 + \alpha e_3$$
$$\hat{y}_4 = \hat{y}_3 + \alpha(y_3 - \hat{y}_3)$$
$$\hat{y}_4 = \alpha y_3 + (1 - \alpha)\hat{y}_3$$

Therefore,

$$\hat{y}_5 = \alpha y_4 + (1 - \alpha)(\alpha y_3 + (1 - \alpha)\hat{y}_3)$$
$$\hat{y}_5 = \alpha y_4 + \alpha(1 - \alpha)y_3 + (1 - \alpha)^2\hat{y}_3$$

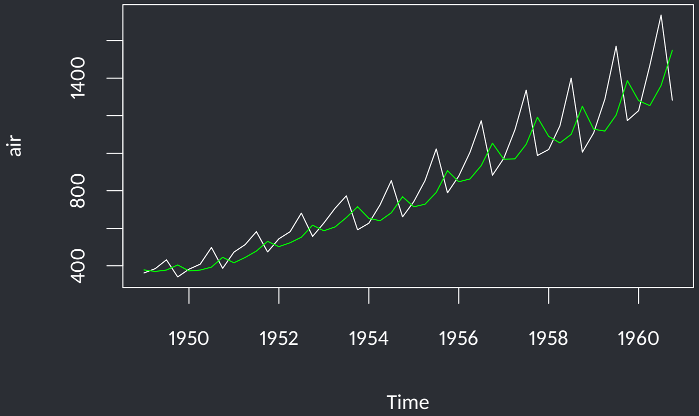
In general,

$$\hat{y}_{t+1} = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + (1 - \alpha)^2\hat{y}_{t-2} \ldots$$

$\hat{y}_{t+1}$ carries a portion of all past observations; the more recent the observation is, the more weight it has.

Simple Exponential Smoothing

```
library(forecast)
air_ses_05 <- ses(air, alpha = 0.5)
plot(air)
lines(air_ses_05$fitted, col='green')
```



Simple Exponential Smoothing

- α determines how aggresively the forecast will react to the observed error.

Simple Exponential Smoothing

- α determines how aggresively the forecast will react to the observed error.
- A larger α means a more aggressive adjustment (i.e. closer forecast to the observed value).

Simple Exponential Smoothing

- α determines how aggresively the forecast will react to the observed error.
- A larger α means a more aggressive adjustment (i.e. closer forecast to the observed value).
- The SES equation could also be written as:

Simple Exponential Smoothing

- α determines how aggresively the forecast will react to the observed error.
- A larger α means a more aggressive adjustment (i.e. closer forecast to the observed value).
- The SES equation could also be written as:

Simple Exponential Smoothing

- α determines how aggresively the forecast will react to the observed error.
- A larger α means a more aggressive adjustment (i.e. closer forecast to the observed value).
- The SES equation could also be written as:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

Simple Exponential Smoothing

- α determines how aggresively the forecast will react to the observed error.
- A larger α means a more aggressive adjustment (i.e. closer forecast to the observed value).
- The SES equation could also be written as:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

	$\hat{y}_t$				y_t
	100				200
$\hat{y}_{t+1}$	100	125	150	175	200
	$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$

Simple Exponential Smoothing

- α determines how aggresively the forecast will react to the observed error.
- A larger α means a more aggressive adjustment (i.e. closer forecast to the observed value).
- The SES equation could also be written as:

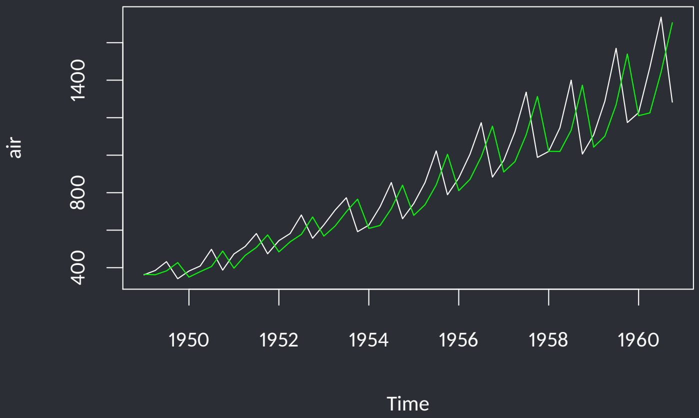
$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

	$\hat{y}_t$				y_t
	100				200
$\hat{y}_{t+1}$	100	125	150	175	200
	$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$

So, let's adjust the α to see if the forecasts get better.

Simple Exponential Smoothing

```
air_ses_09 <- ses(air, alpha = 0.9)
plot(air)
lines(air_ses_09$fitted, col='green')
```



Simple Exponential Smoothing

- Regardless of the size of the α parameter, there seems to be always a delay in the forecast.

Simple Exponential Smoothing

- Regardless of the size of the α parameter, there seems to be always a delay in the forecast.
- This is because SES method is for stationary data, i.e., when there is no trend or seasonality.

Simple Exponential Smoothing

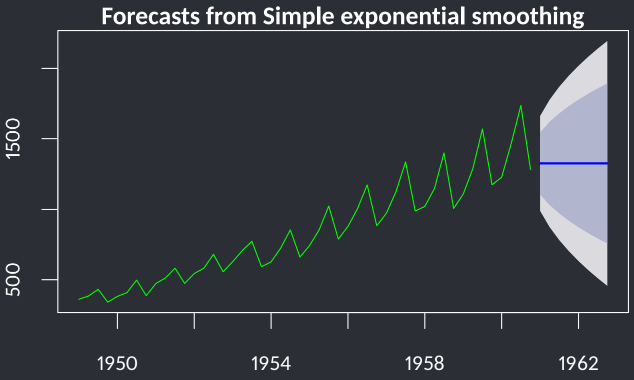
- Regardless of the size of the α parameter, there seems to be always a delay in the forecast.
- This is because SES method is for stationary data, i.e., when there is no trend or seasonality.
- When trend and seasonality exist, one needs to use lag operations to make it stationary.

Simple Exponential Smoothing

- Regardless of the size of the α parameter, there seems to be always a delay in the forecast.
- This is because SES method is for stationary data, i.e., when there is no trend or seasonality.
- When trend and seasonality exist, one needs to use lag operations to make it stationary.
- Otherwise, SES predicts the same value for all future observations.

Simple Exponential Smoothing

```
air_ses_09 <- ses(air, alpha = 0.9, h=8)
plot(air_ses_09, col='green')
```



Simple Exponential Smoothing

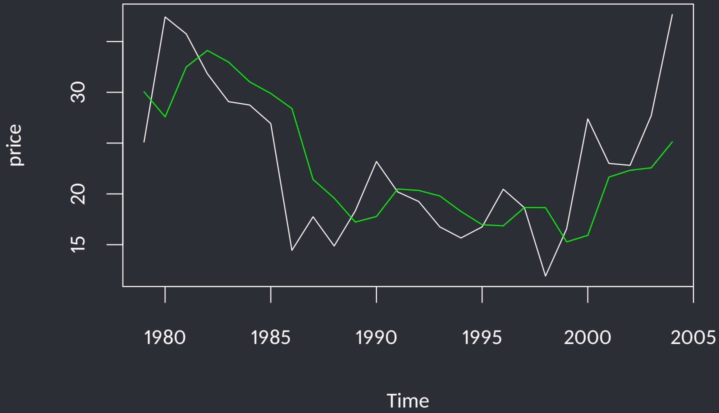
```
air_ses_09

        Point Forecast     Lo 80    Hi 80    Lo 95    Hi 95
1961 Q1       1325.377 1105.3915 1545.363 988.9380 1661.817
1961 Q2       1325.377 1029.4167 1621.338 872.7445 1778.010
1961 Q3       1325.377  969.2991 1681.456 780.8027 1869.952
1961 Q4       1325.377  917.9579 1732.797 702.2830 1948.472
1962 Q1       1325.377  872.3988 1778.356 632.6064 2018.148
1962 Q2       1325.377  831.0206 1819.734 569.3240 2081.431
1962 Q3       1325.377  792.8480 1857.907 510.9440 2139.811
1962 Q4       1325.377  757.2343 1893.520 456.4776 2194.277
```

Simple Exponential Smoothing for Oil Prices

Let’s try using SES to predict the oil pricess.

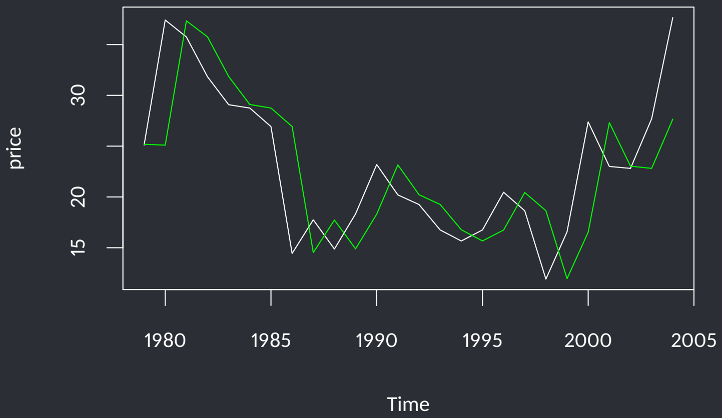
```
price <- ts(oil$price, start=1979, frequency=1)
price_ses_05 <- ses(price, alpha = 0.5)
plot(price)
lines(price_ses_05$fitted, col='green')
```



Simple Exponential Smoothing for Oil Prices

Let R decide on the optimal α.

```
price_ses <- ses(price)
plot(price)
lines(price_ses$fitted, col='green')
```



Simple Exponential Smoothing for Oil Prices

```
price_ses$model

Simple exponential smoothing

Call:
 ses(y = price)

  Smoothing parameters:
    alpha = 0.9932

  Initial states:
    l = 25.1832

  sigma:  5.3845

     AIC     AICc      BIC
178.2539 179.3448 182.0282
```

Simple Exponential Smoothing for Oil Prices

```
price_ses_h8 <- ses(price, h=8)
plot(price_ses_h8, col='green')
```