Time-series Analytics

Giacomo Ziffer
Politecnico di Milano

POLITECNICO
MILANO 1863

Decomposition & Detrending

Let’s change perspective

Photo by Chokniti Khongchum from Pexels

Thinking …

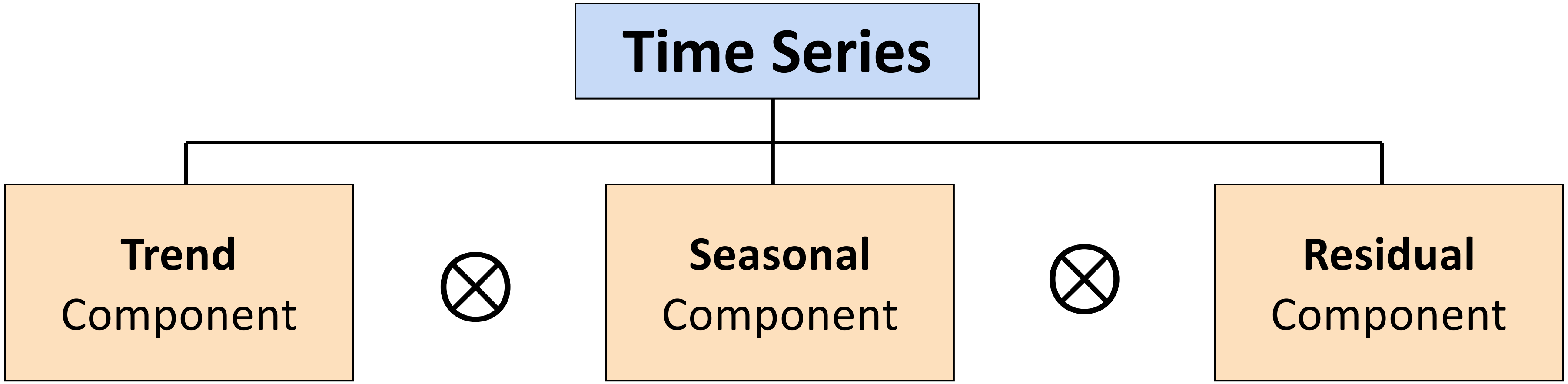
- If
  - stationary **implies** predictable
- and
  - changes in mean, variance
  - presence of seasonality
  - **imply** non-stationary
- **can we try to remove what causes of non-stationarity?**

Thinking …

- If
  - stationary **implies** predictable
- and
  - changes in mean, variance
  - presence of seasonality
  - **implies** non-stationary
- **can we try to remove what causes of non-stationarity?**
- **YES! Decomposing a time series**

Time Series Decomposition

Time-Series Components



Trend	Seasonal	Residual
Overall, persistent, long-term change	Regular periodic fluctuations	Residual fluctuations (hopefully stationary)

$\otimes$ this charather is a placeholder for various mathematical operation used to assemble the components

Models to decompose time series:

- **Additive** model

$$X_t = m_t + s_t + Y_t$$

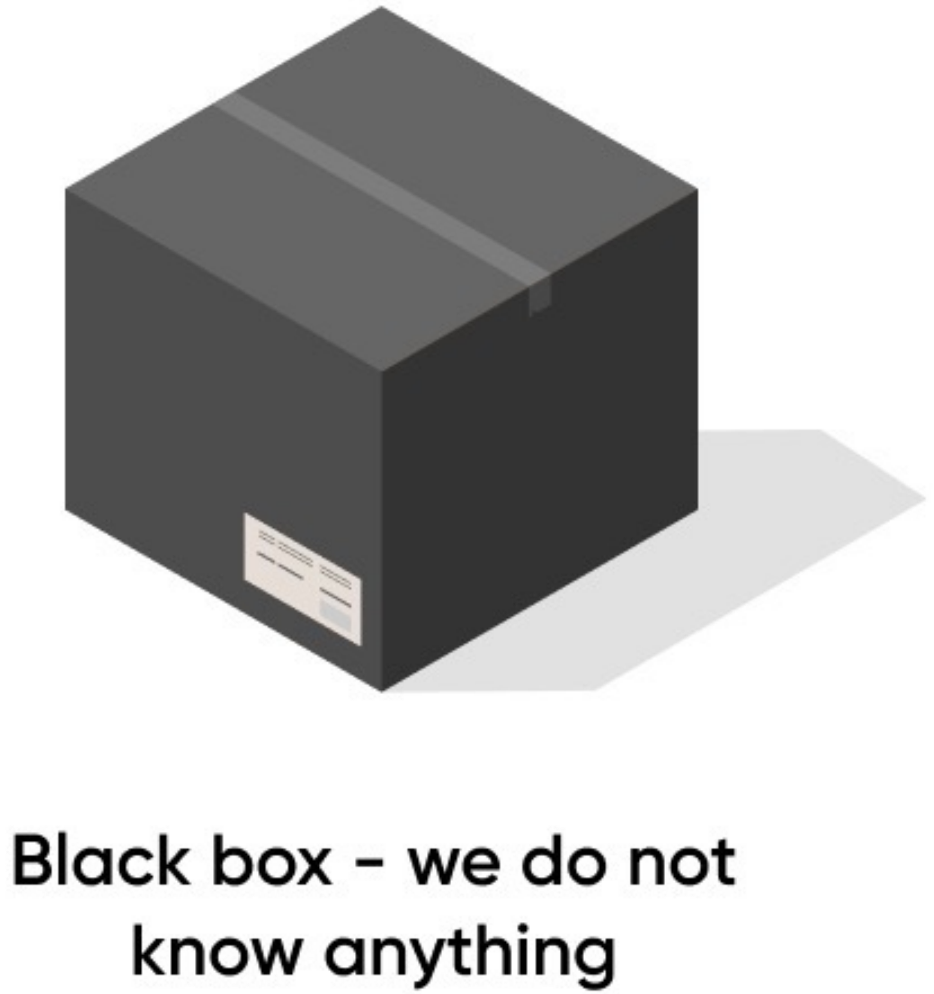
- **Multiplicative** model

$$X_t = m_t \, s_t \, Y_t$$

- Where
  - m_t is the trend component
  - s_t is the seasonal component
  - Y_t is the residual component

- **How to Choose** Between Additive and Multiplicative Decompositions
  - The **additive** model is useful when the **seasonal** variation is relatively **constant over time**.
  - The **multiplicative** model is useful when the **seasonal** variation **increases over time**.

Let’s first go
black-box



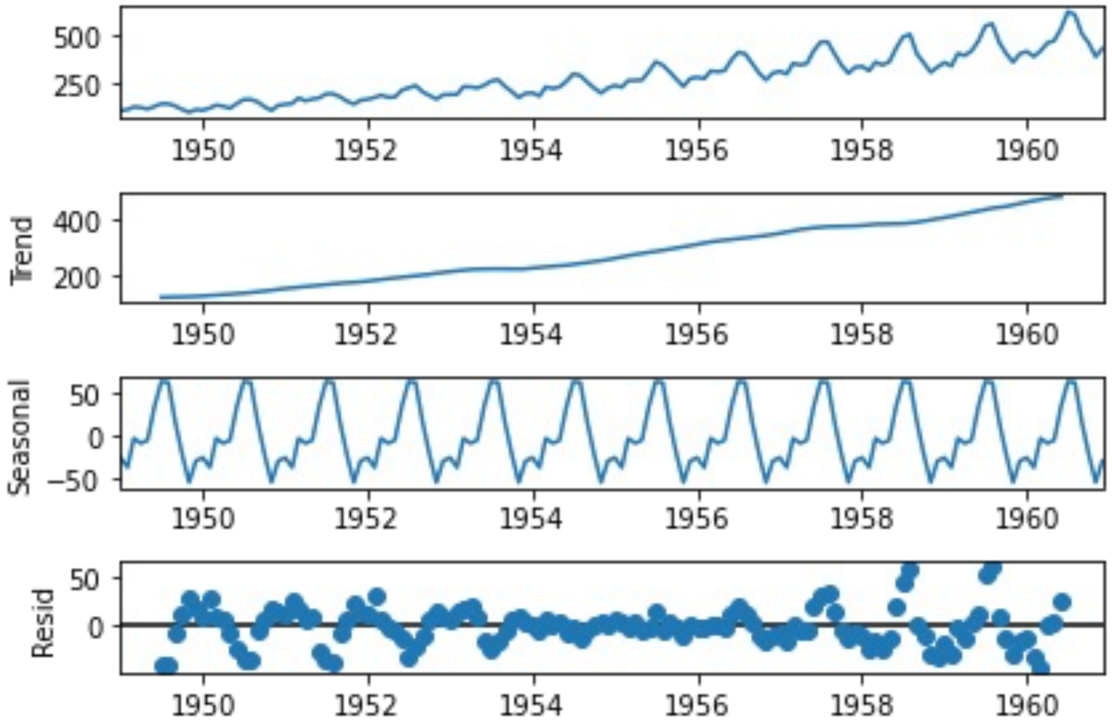
statsmodels

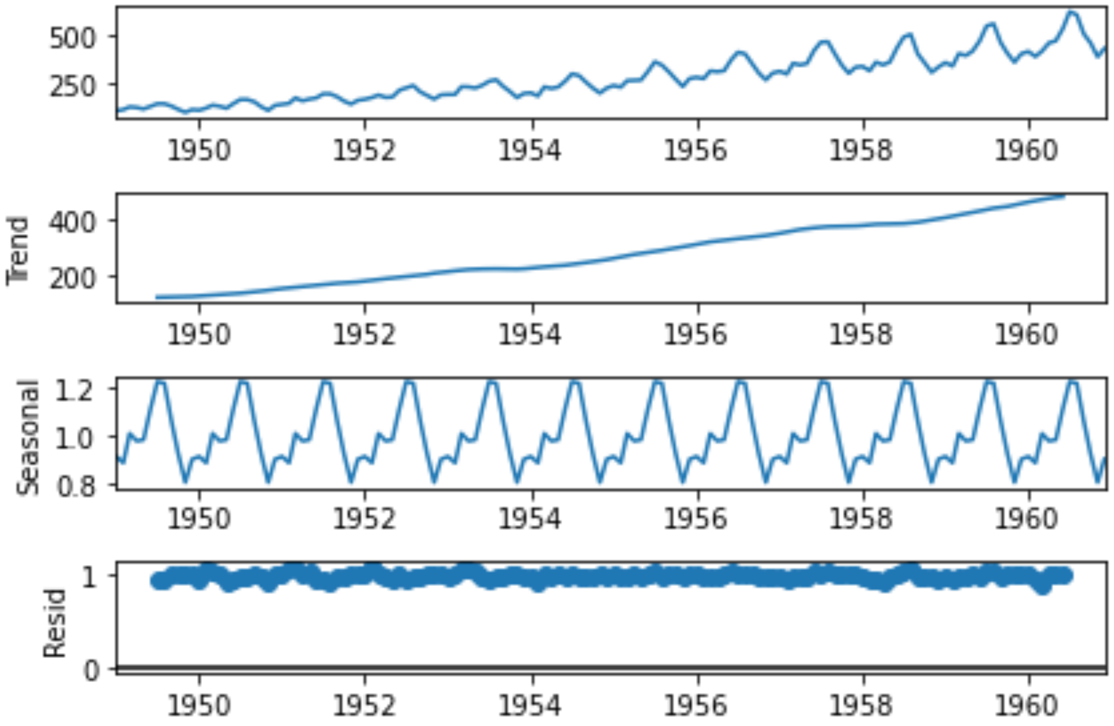
- Use statsmodels as a *black-box* able to decompose a time-series
- Try both methods
  - Additive
  - Multiplicative

NOTE: for more information check out https://www.statsmodels.org/stable/examples/notebooks/generated/statespace_seasonal.html

Time series decomposition
A comparison

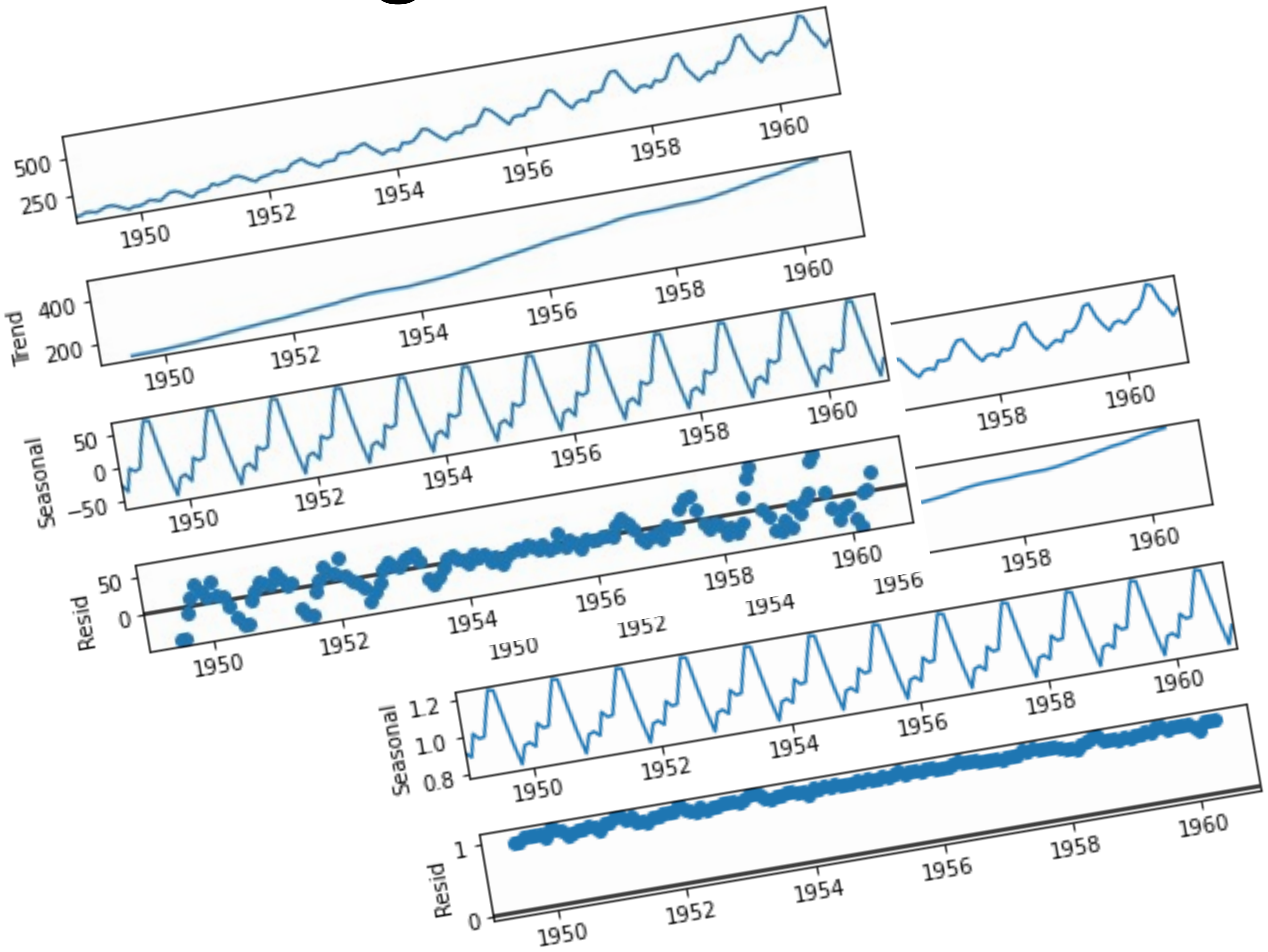
Additive vs. Multiplicative





Time series decomposition
Let’s see if you are following …

- Q: Is additive better than multiplicative?
- Q: Why?



Let’s now go **white-box**

White box - we know everything

Two simplified time series models

Non-seasonal Decomposition Model with Trend

$$X_t = m_t \otimes Y_t$$

Time-series trend residual

components

Decomposition Model **with** Trend and **Seasonal** Components

$$X_t = m_t \otimes s_t \otimes Y_t$$

Time-series trend seasonal residual

components

Non-seasonal Decomposition Models with Trend

There are three basic methods for estimating/removing trend:

- Method 1: *Trend elimination by differencing*
  - Differencing one or more times removes the trend

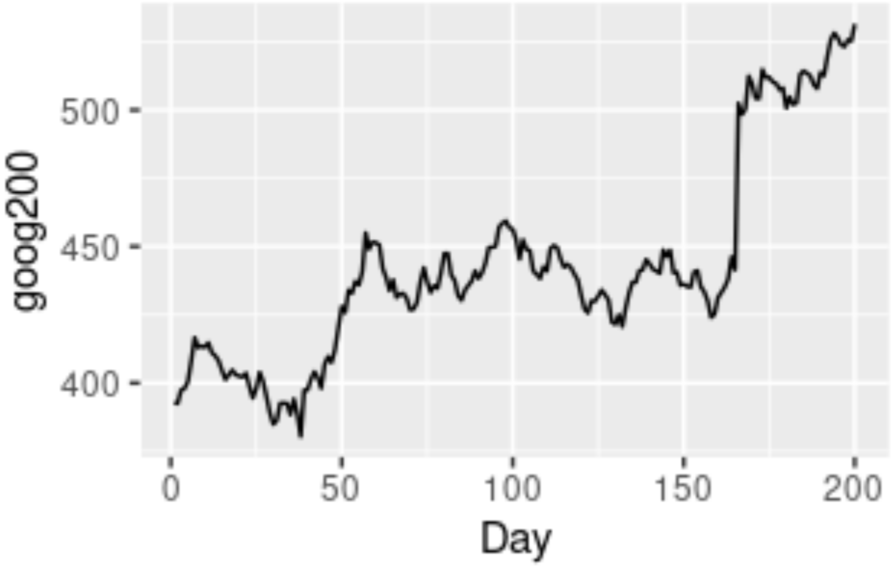
- Method 2: *Trend estimation by model fitting & removal*
  - We estimate the trend fitting a model and then we remove it

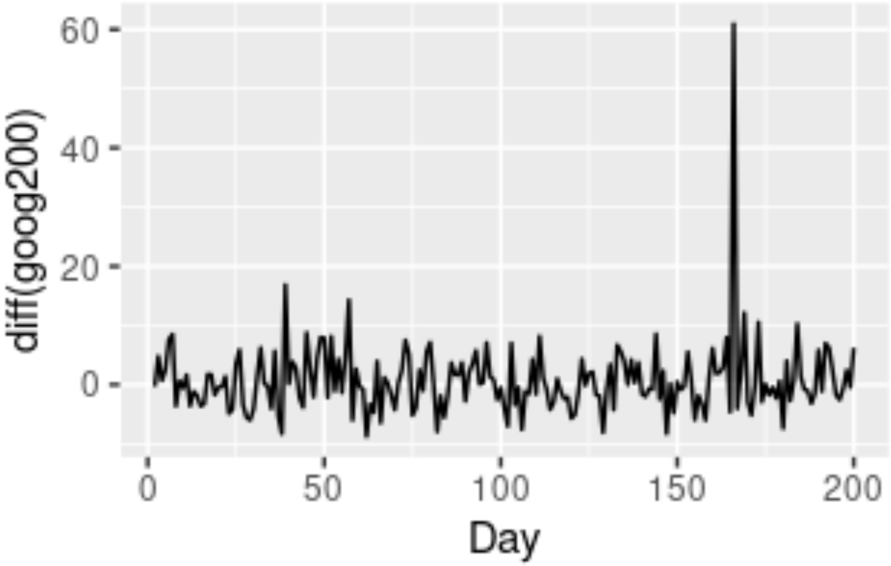
- Method 3: *Trend estimation by using "centered" moving averages*
  - We estimate the trend using a "centered" moving average and we remove it

Non-seasonal Decomposition Models with Trend - Method 1
Trend elimination by differencing

- Differencing of a time series $\{X_t\}$ in discrete time *t* is the transformation of the series to a new time series $\{D_t\}$ where the values are the differences between consecutive values of $\{X_t\}$.

$$d_t = x_t - x_{t-1}$$

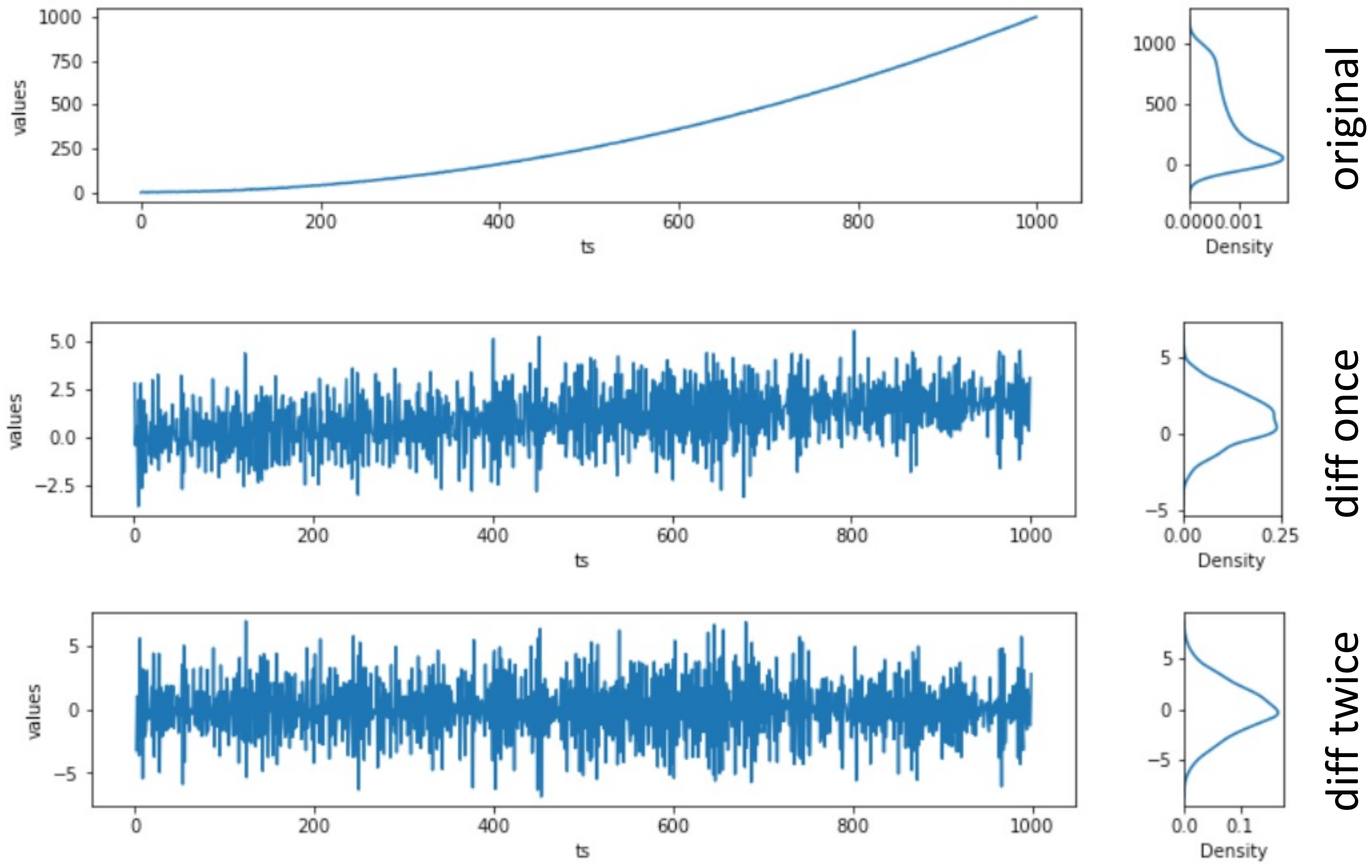




Non-seasonal Decomposition Models with Trend - Method 1

Trend elimination by differencing (cont.)

- If a trend is linear differencing once is sufficient to remove it
- If a trend is quadratic, you need two difference twice
- If a trend can be model with a polynomial of order *n*, then you need to difference *n* times



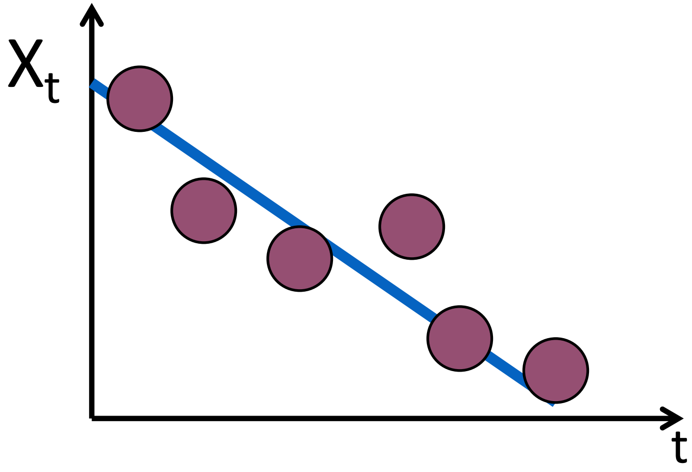
Non-seasonal Decomposition Models with Trend - Method 1

Trend elimination by differencing

1. Generate synthetic data
2. Differencing one to remove a linear trend
3. Differencing twice to remove a quadratic trend

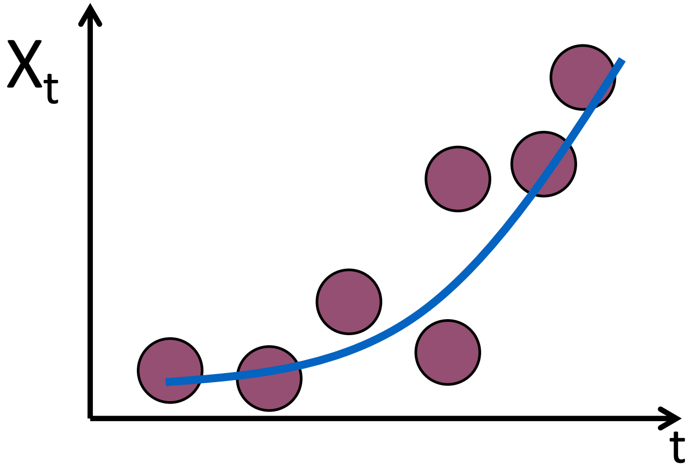
Non-seasonal Decomposition Models with Trend - Method 2
Trend estimation by model fitting & removal

$$X_t = b_0 + b_1 t$$



Downward linear trend

$$X_t = b_0 + b_1 t + b_2 t^2$$



Upward quadratic trend

Non-seasonal Decomposition Models with Trend - Method 2
Trend estimation by model fitting & removal

1. Generate synthetic data
2. Fit a linear regression
3. Fit a quadratic regression

Non-seasonal Decomposition Models with Trend

Combining Method 1 and 2

1. Generate synthetic data with quadratic trend
2. Differentiate
3. Observe the time-series still shows a trend
4. Detrend fitting a linear Regression

Decomposition Models with Trend and Seasonality

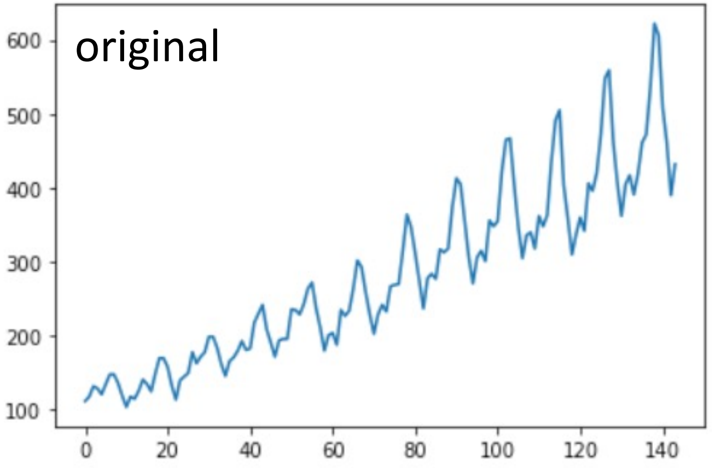
There are three basic methods for estimating/removing the trend and seasonal components:

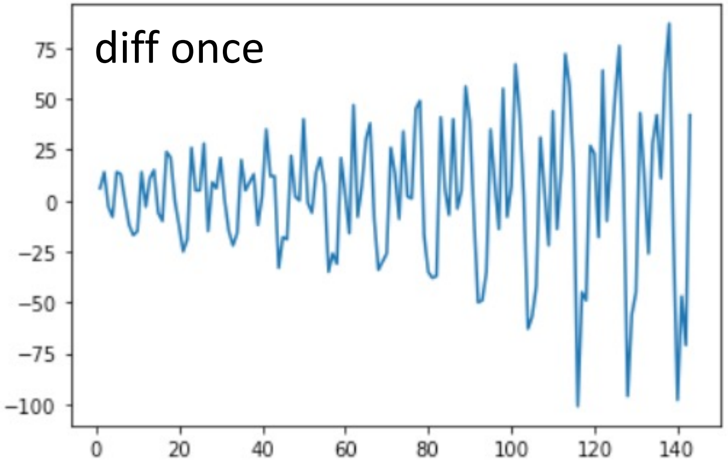
- Method 1 – Differencing
  - First we difference one or mote time to remove the trend
  - Then we perform "**seasonal differencing**" to directly removes the season
- Method 2 - Filtering
  - First we estimate and remove the trend using a **"centered" moving average**
  - then we estimate and remove the seasonal component using "**periodic averages**"
- Method 3 - Joint-fit method
  - fitting a combined polynomial and **dynamic harmonic regression**

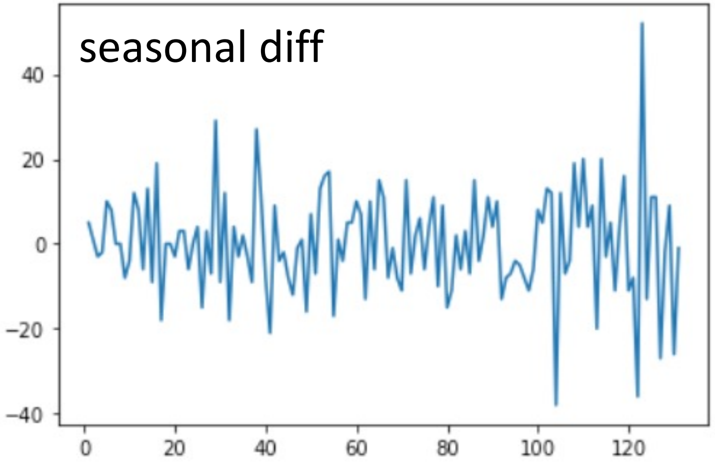
Decomposition Models with Trend and Seasonality - Method 1
Seasonal differencing

- Seasonal differencing of a time series $\{X_t\}$ in discrete time *t* given the seasonality's period *d* is the transformation of the series to a new time series $\{S_t\}$ where the values are the differences between the value of $\{X_t\}$ at time *t* and the the value of $\{X_t\}$ a period *d* before.

$$s_t = x_t - x_{t-d}$$







Decomposition Models with Trend and Seasonality - Method 1

Seasonal differencing

1. Generate synthetic data using a sine form
2. Apply ”seasonal differencing”
3. Observe you removed the seasonality

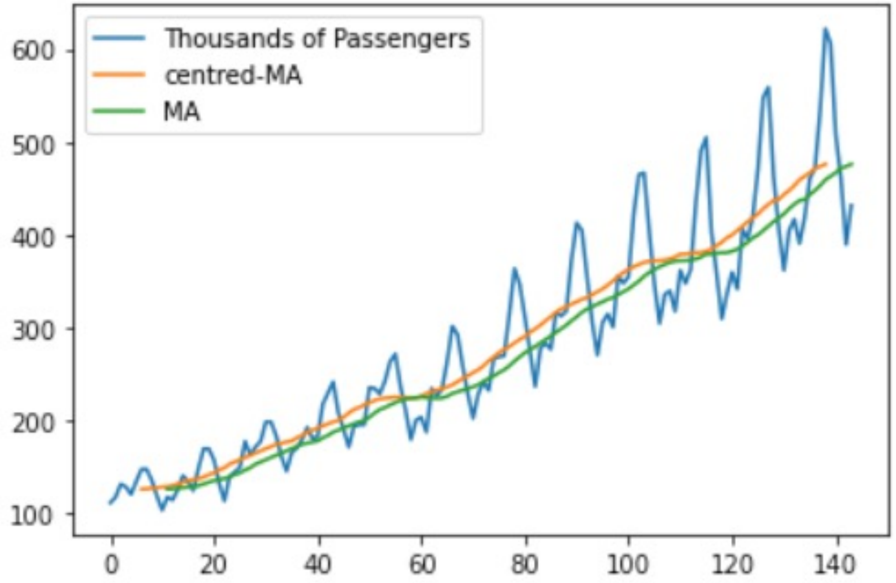
Decomposition Models with Trend and Seasonality - Method 2
Estimate the trend using "centered" moving average

- Given the seasonality's period *d*
  - If *d* is even, the «centered» moving average is defined as

$$\widehat{m_t} = \left(0.5x_{t-q} + x_{t-q+1} + \cdots + x_{t+q-1} + 0.5x_{t+q}\right)/d$$

  - If *d* is odd, the «centered» moving average is defined as

$$\widehat{m_t} = \left(x_{t-q} + x_{t-q+1} + \cdots + x_{t+q-1} + x_{t+q}\right)/d$$



- Notes:
  - there are no values for either the first *q* or the last *q* data points, because we do not have enough observations on either side to define the moving average for those values of *t*.
  - This «centered» moving average is different from the «normal» moving average

Decomposition Models with Trend and Seasonality - Method 2

Estimate the seasonal component

- To estimate the seasonal component using "**periodic averages**"
  1. Divide the detrended value in seasons of length *d*
  2. Compute the seasonal component values w_k by averaging each of the *d* points of the season (i.e., $k = 1, \ldots, d$)
  3. Compute adjusted the seasonal component values s_k to ensure that they add to zero

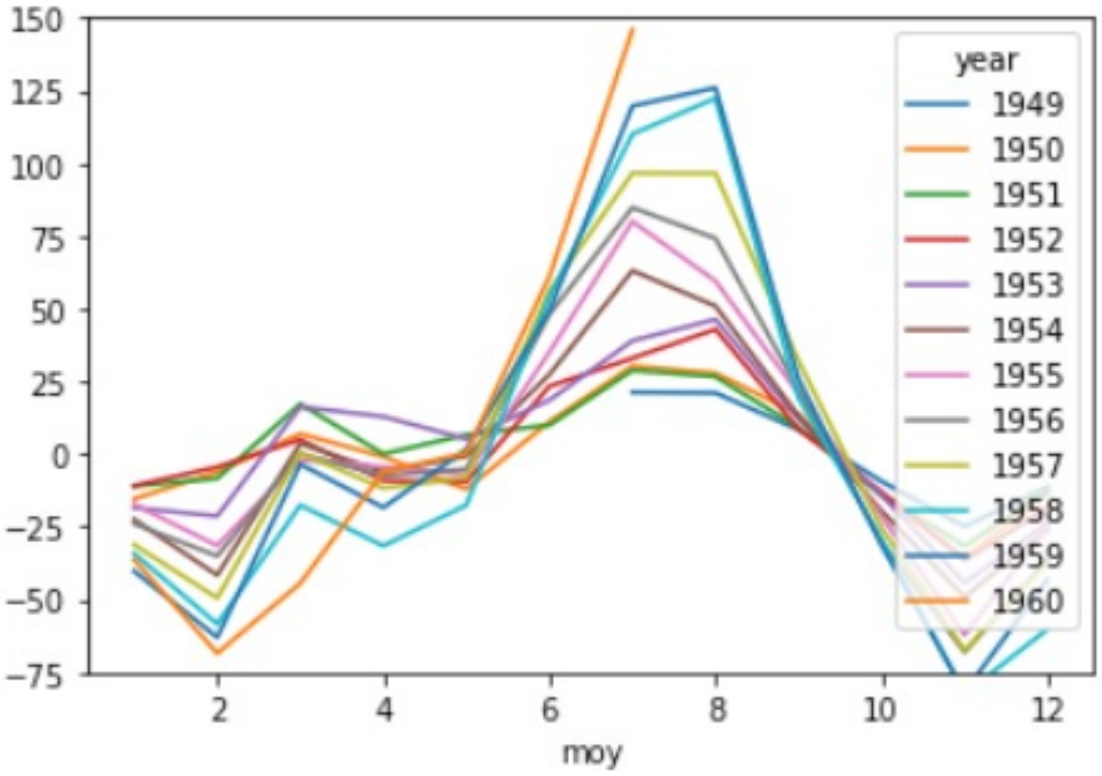
$$\hat{s}_k = w_k - d^{-1} \sum_{j=1}^{d} w_j, \ \ k = 1, \cdots, d$$

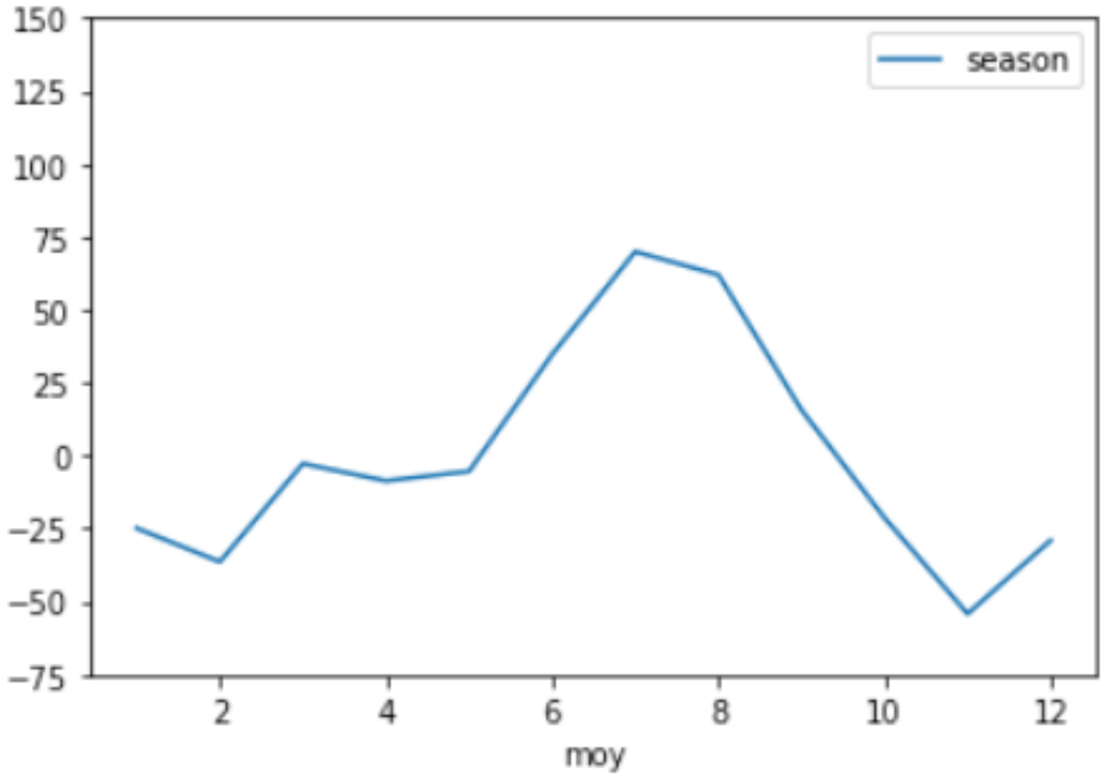
  4. String together the adjusted seasonal component values in a sequence
  5. Replace the sequence for each season

Decomposition Models with Trend and Seasonality - Method 2

Estimate the seasonal component

- The case of air travel time-series





Decomposition Models with Trend and Seasonality - Method 2

Filtering

1. Load the time series
2. estimate the trend by the "centered" moving average
3. Remove the trend from the time series
4. Estimate the seasonal component using "**periodic averages**"
5. Remove it from the detrended series

Decomposition Models with Trend and Seasonality - Method 3
Joint-fit method

- we can fit a combined polynomial linear regression and harmonic functions to estimate and then remove the trend and seasonal component simultaneously as the following

$$\begin{aligned} X_t &= m_t + s_t + Y_t \\ &= (\beta_0 + \beta_1 t + \beta_2 t^2) + \left[\sum_{j=1}^{k} \left(\alpha_j \cos(\lambda_j t) + \gamma_j \sin(\lambda_j t)\right)\right] + Y_t \end{aligned}$$

Decomposition Models with Trend and Seasonality - Method 3

Anatomy of the harmonic function

Add k harmonics

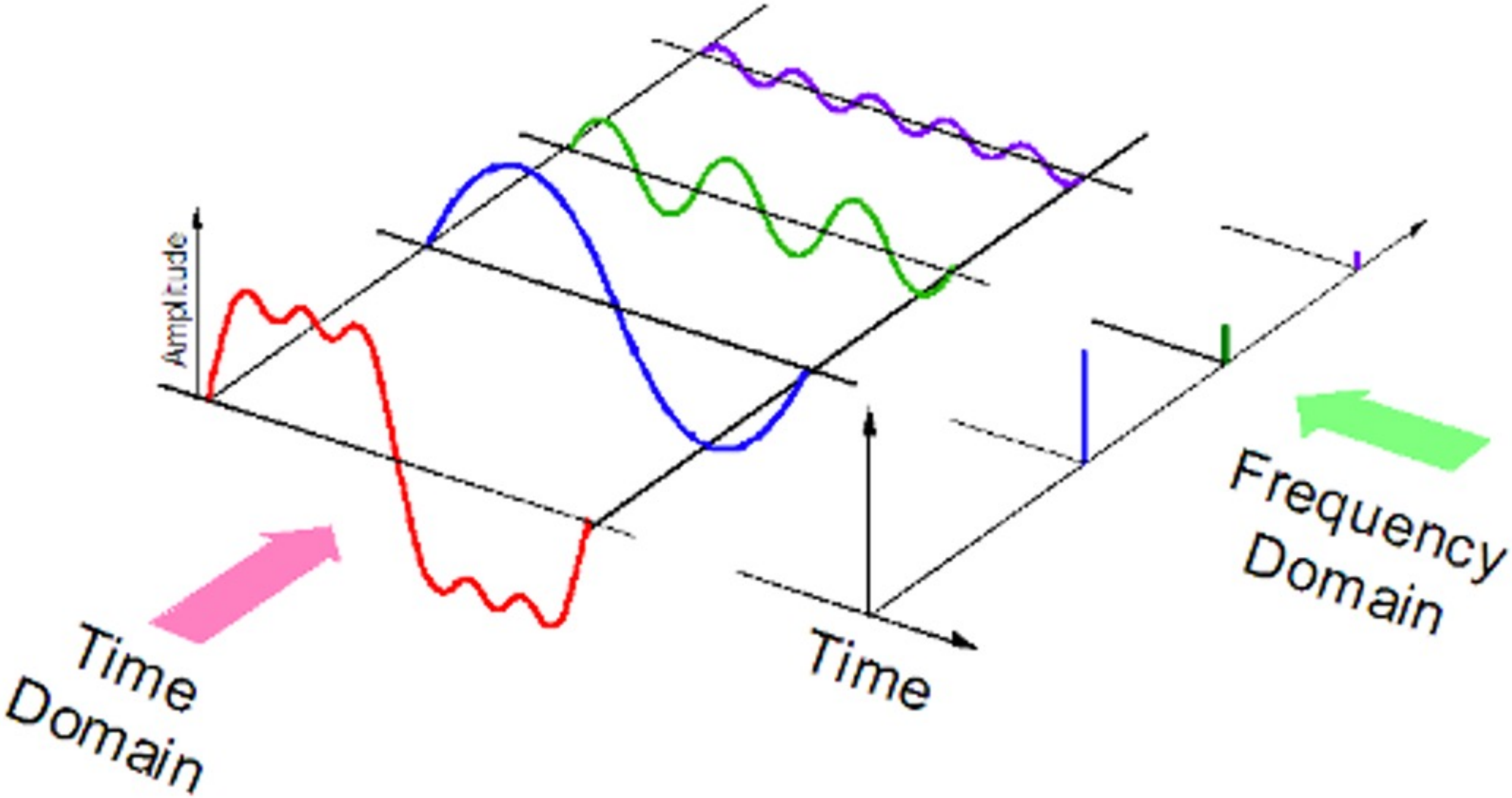
sin and cosin deals with phases of the partials

$$\left[\sum_{j=1}^{k}\left(\alpha_j \cos(\lambda_j t) + \gamma_j \sin(\lambda_j t)\right)\right] + Y_t$$

Coefficients gives energies to the partial

Number of the armonic

Decomposition Models with Trend and Seasonality - Method 3
Intuitively ….



[src: https://kinder-chen.medium.com/denoising-data-with-fast-fourier-transform-a81d9f38cc4c]

Decomposition & Detrending

Quiz

Time-series Analytics

Giacomo Ziffer
Politecnico di Milano

POLITECNICO
MILANO 1863