Chapter 09

Time Series Analysis

Dr. Steffen Herbold
herbold@cs.uni-goettingen.de

Outline

- Overview
- Methods for Time Series Analysis
- Summary

Example of Time Series Analysis

420 in January

431 in February

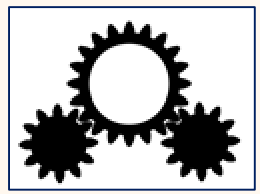
415 in March

⋮

509 in July

⋮

Data over time
- Sales
- Passengers
- Traffic
- …



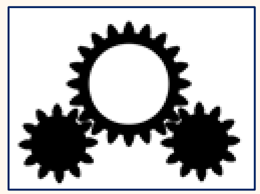
Time Series Analysis

Our sales in the next two months will be …

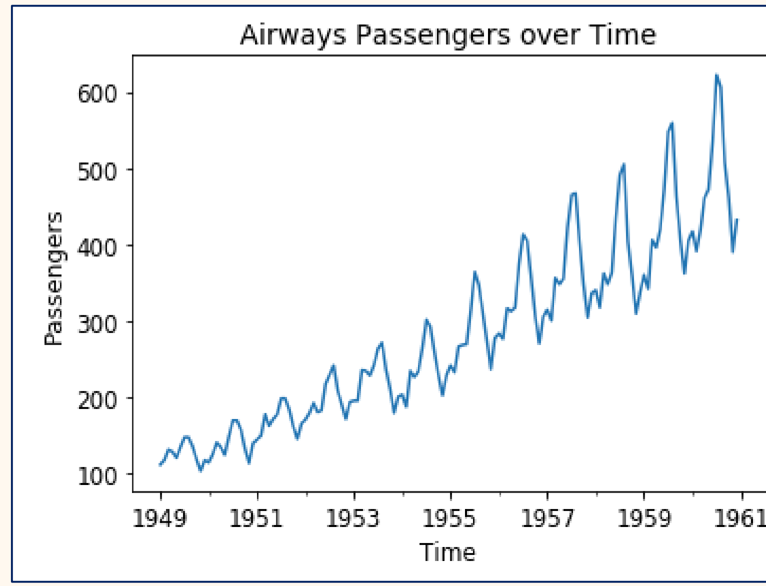
The General Problem

Data over time

- Value at time point 1
- Value at time point 2
- Value at time point 3
- Value at time point 4
- Value at time point 5
- Value at time point 6
- Value at time point 7
- Value at time point 8
- Value at time point 9
- …



Time Series Analysis



The Formal Problem

- Discrete values over time
  - $\{x_1, \dots, x_T\} = \{x_t\}_{t=1}^{T}$ with $x_t \in \mathbb{R}$
  - Can be seen as a series of random variables or a stochastic process
  - Time between t and $t+1$ must be equal for all $t = 1, \dots, T-1$
    - Minutes, hours, days, weeks, months, …

- Components of a time series
  - General trend of the time series T_t
  - Seasonal effects on the time series S_t
  - Autocorrelation between observations R_t
  - $x_t = T_t + S_t + R_t$

Outline

- Overview
- **Methods for Time Series Analysis**
- Summary

Time Series Analysis with Box-Jenkins

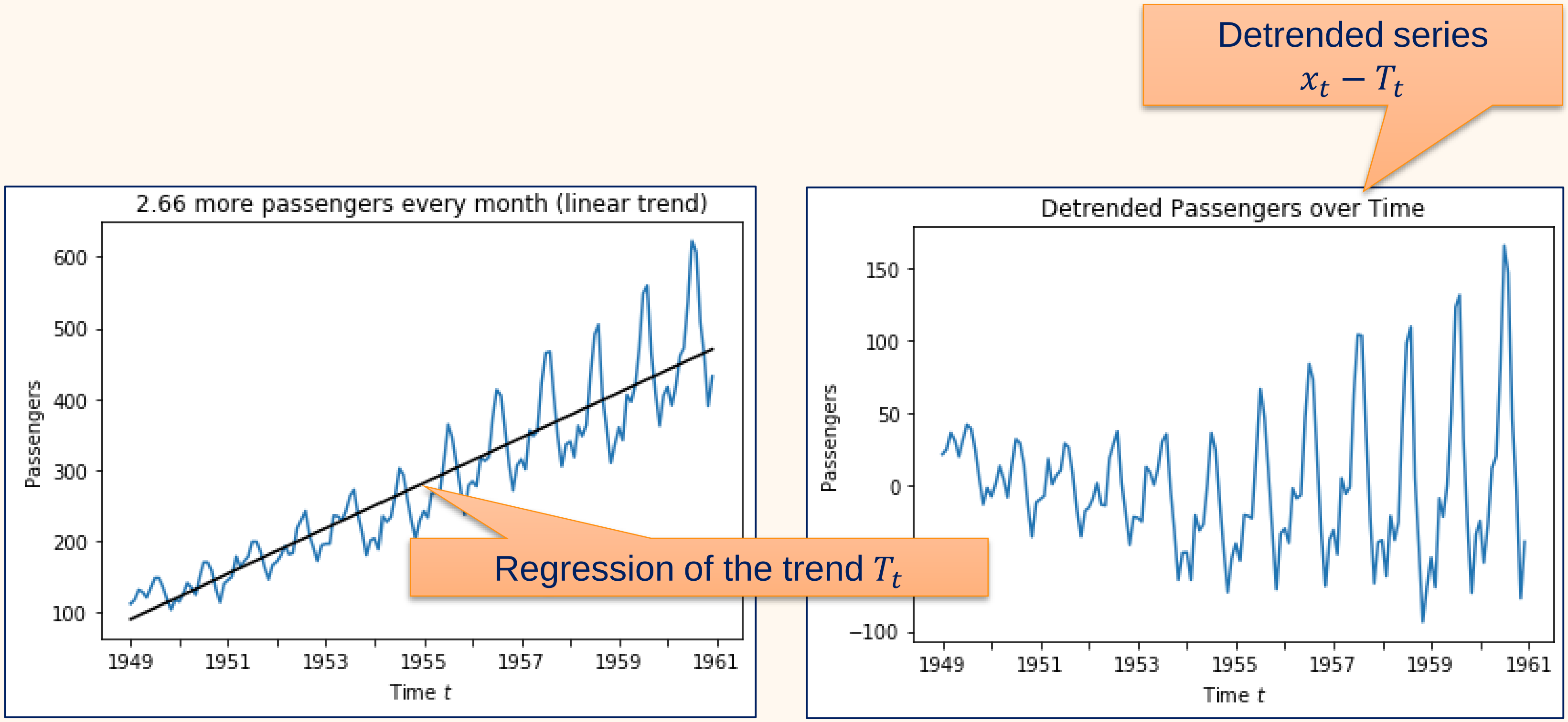
- For stationary data
  - Stationary means constant mean value and variance

  → Requires de-trending and seasonal adjustment

- Models autocorrelation as a stochastic process
  - Observations depend on past observation and a random component

- Tries to model time series with only few parameters
  - Goal are simple models

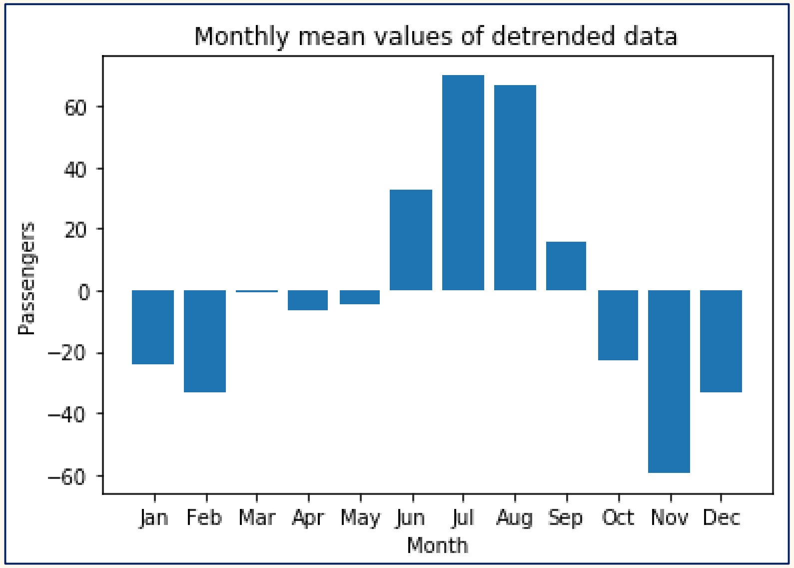
Detrending Through Regression

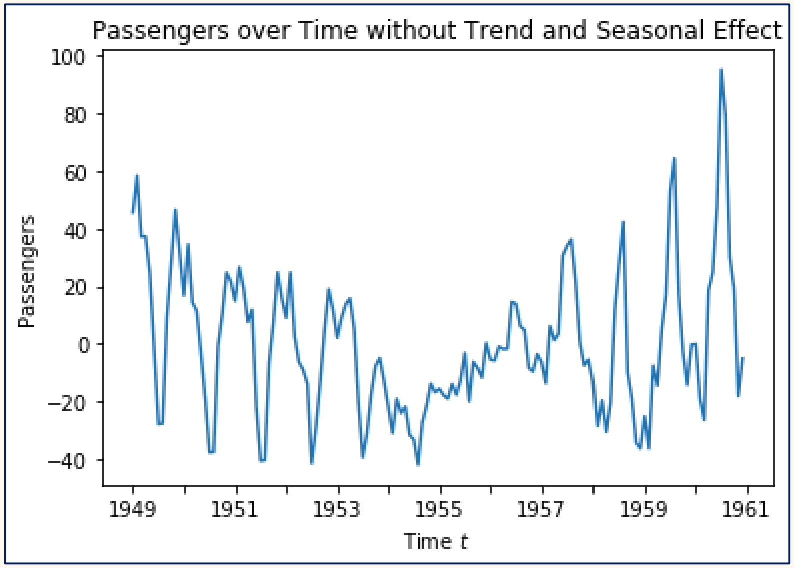


- Non-linear regression for non-linear trends

Seasonal Adjustment through the Mean

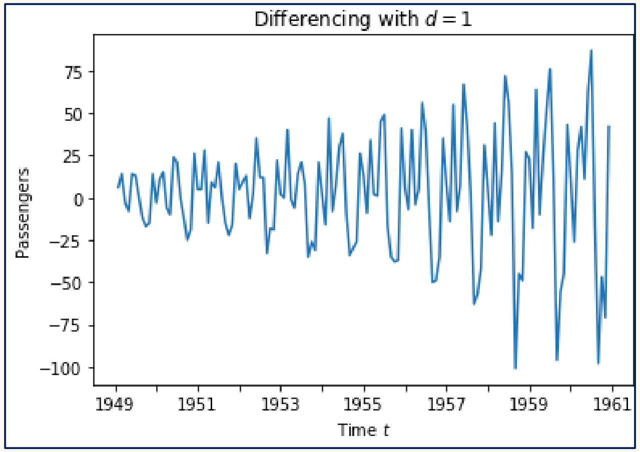
- Seasonal effect:
  - A regularly repeating pattern
  - Monthly, weekly, …
- Seasonal adjustment through the seasonal mean value





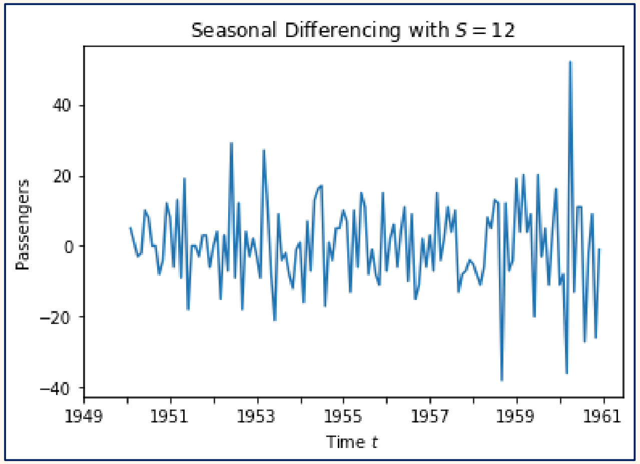
Differencing for Detrending

- Instead of regression / removal of mean seasonal effects
- Differencing for detrending of order d
  - First difference for moving mean values ($d = 1$)
    - Similar to linear trends
    - $\Delta^1 x_t = x_t - x_{t-1}$
  - Second difference for moving mean and the change in the movement (d =2)
    - Similar to quadratic trends
    - $\Delta^2 x_t = \Delta^1 x_t - \Delta^1 x_{t-1} = x_t - 2x_{t-1} + x_{t-2}$

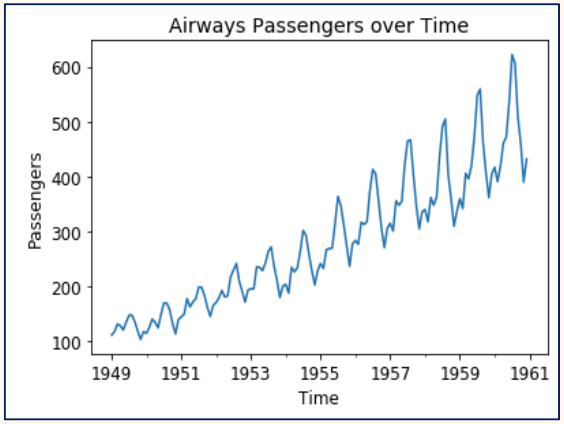


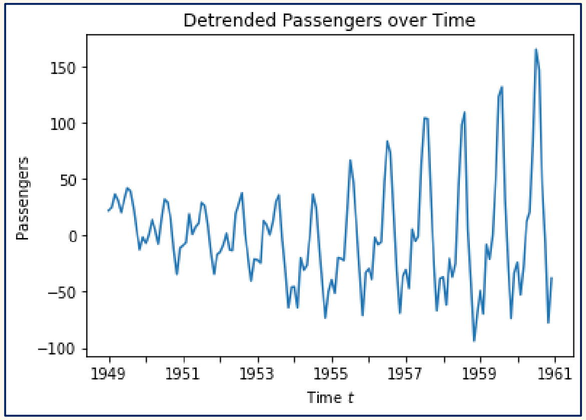
Differencing for Seasonal Adjustment

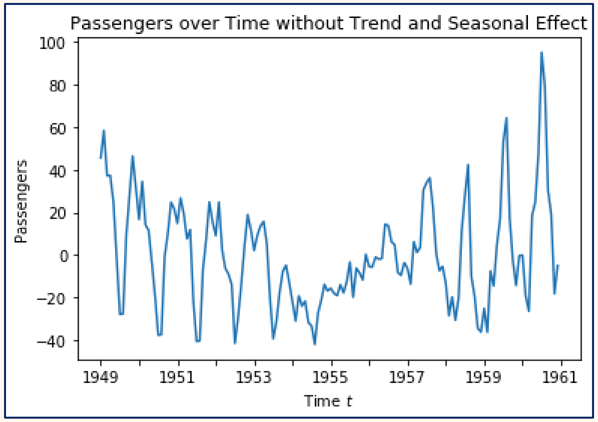
- Seasonal differencing for seasons of periodicity S
  - $\Delta_S \, x_t = x_t - x_{t-S}$
  - $\Delta_S^{12} x_t = x_t - x_{t-12}$ would be seasonal differencing for monthly data

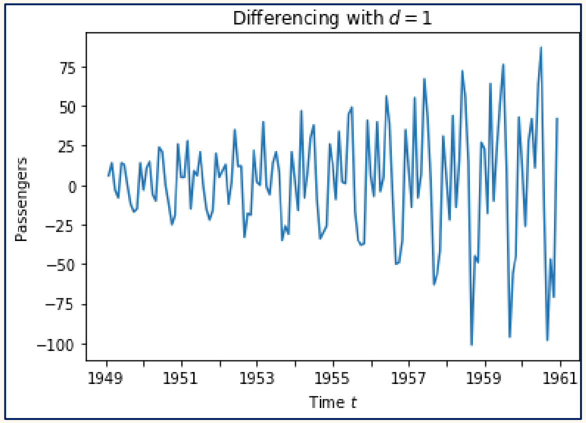


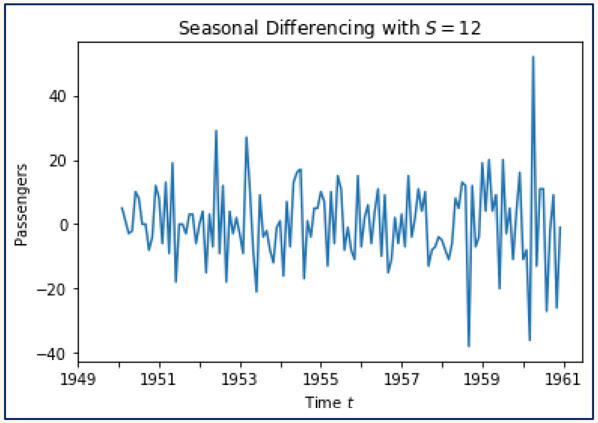
Comparison of Adjustments





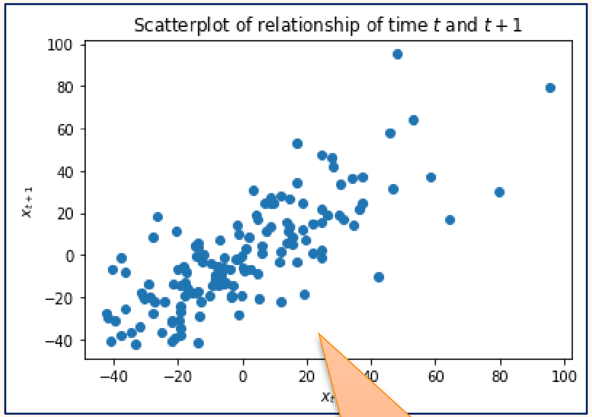


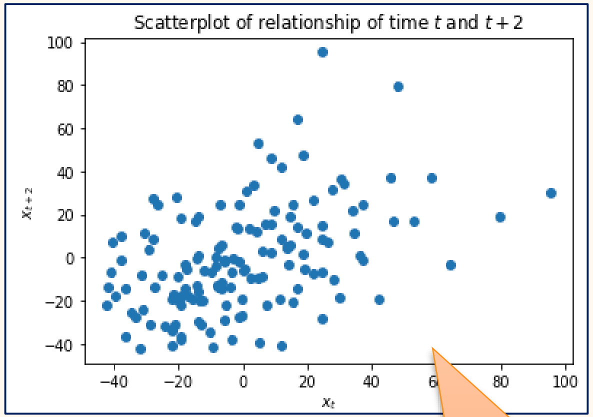


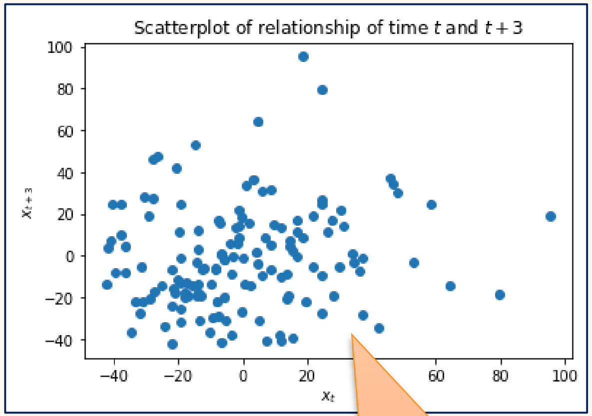


Autocorrelation

- Relationship between time series values with other time series values





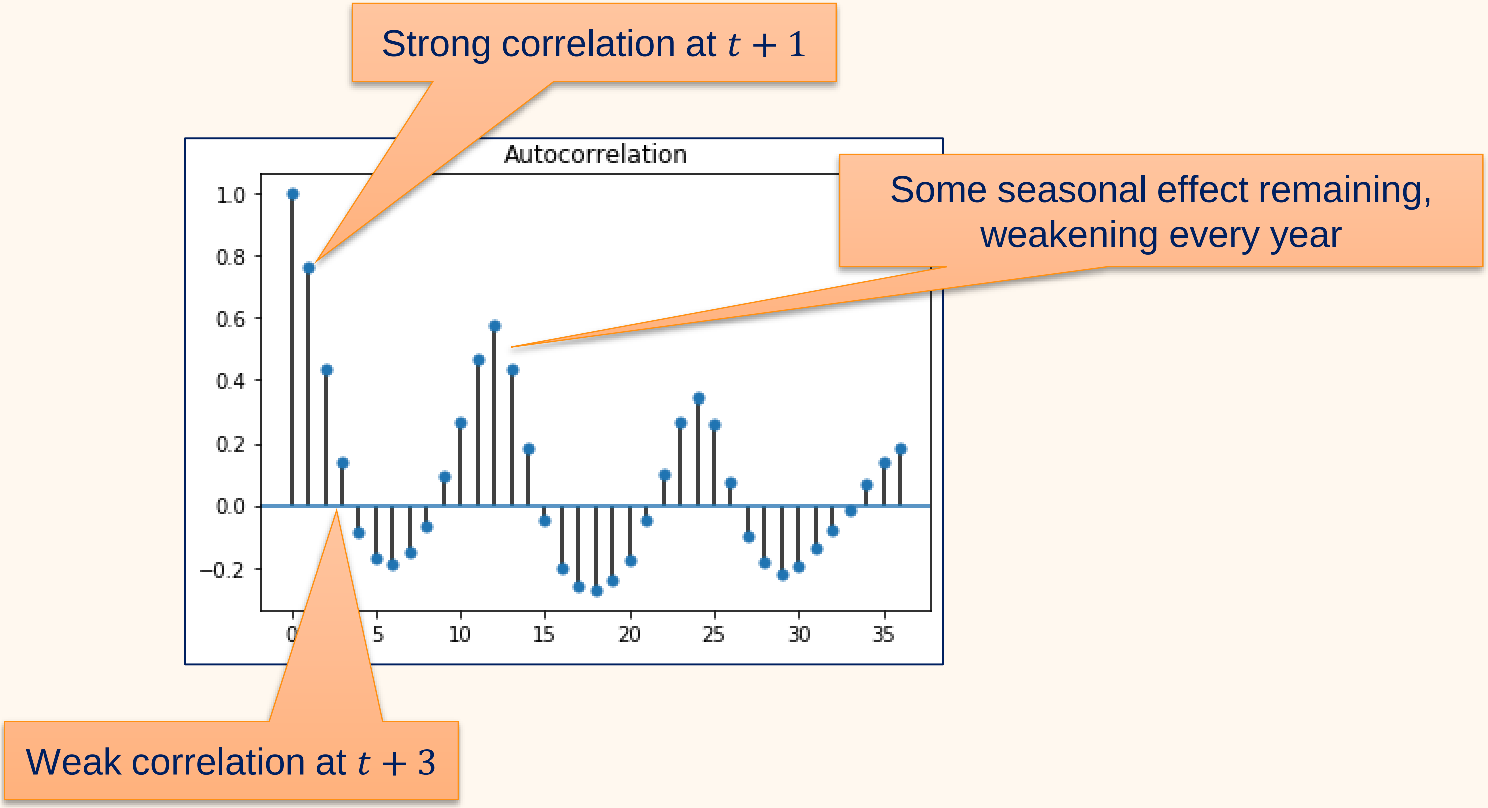


Fairly linear relationship

Spread increasing

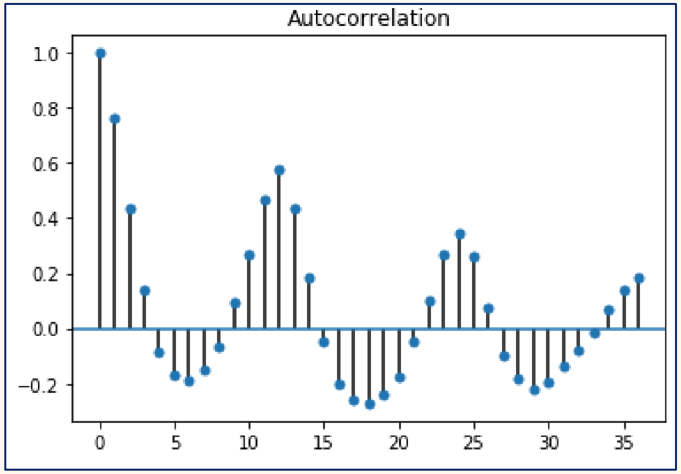
More or less random

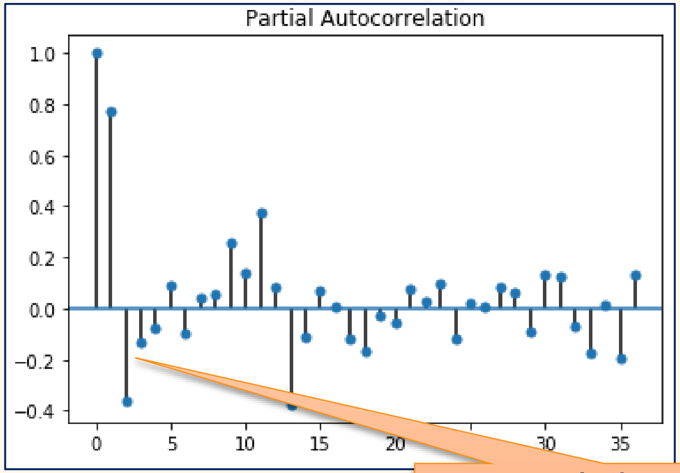
Autocorrelation over Time



Partial Autocorrelation

- Autocorrelation that is not explained by „carrying over“
  - x_t and x_{t+1} are correlated
  - x_{t+1} and x_{t+2} are correlated
  - How much of the correlation between x_t and x_{t+2} is not explained by the above correlations?
  - In other words, how much of the correlation between x_t and x_{t+2} is independent of the correlation between x_t / x_{t+1} and x_{t+1} / x_{t+2}?





Correlation at $t+2$ explained by auto correlation at $t+1$

ARMA Time Series Models

- Requires detrended and seasonally adjusted data
- Model for the autocorrelation part R_T of a time series
- $x_t = \underbrace{a_0 + \sum_{i=1}^{p} a_i x_{t-i}}_{\text{Autoregressive (AR)}} + \underbrace{\epsilon_t + \sum_{j=1}^{q} b_j \epsilon_{t-j}}_{\text{Moving Average (MA)}}$

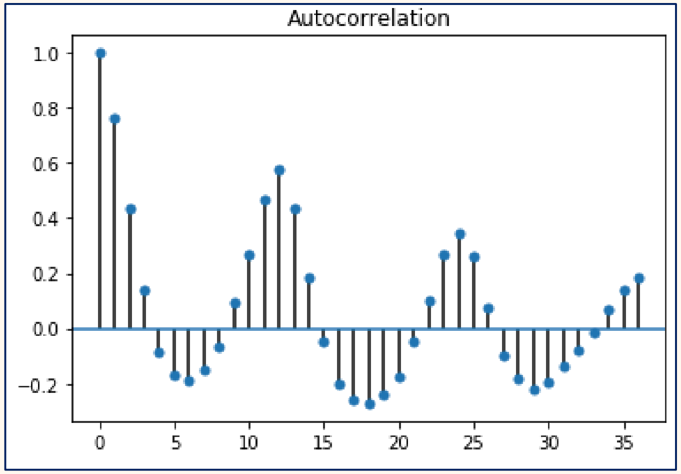
ϵ_i is a random variable with an expected value of 0
→ white noise

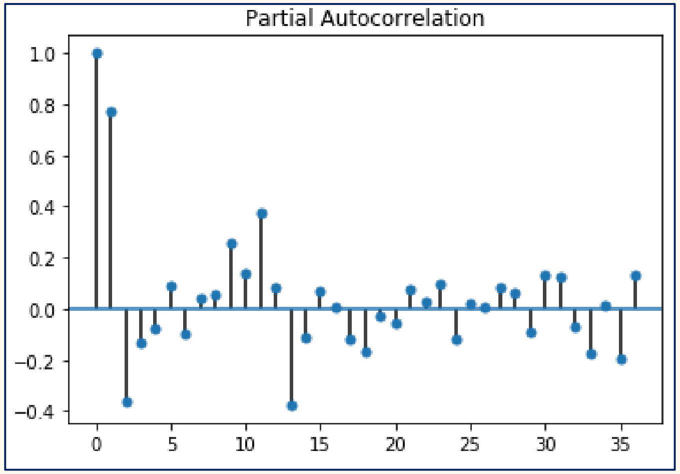
Constant plus linear combination of the p past values

Noise term + linear combination of last q noise values

Picking p and q

- Analyze (partial) autocorrelation function
  - $p = 1$ would model everything except the missing seasonal effect
  - $p = 12$ would capture missing seasonal effect at the cost of a more complex model
  - $q = 0$ or $q = 1$ to account for low random fluctuations





Outline

- Overview
- Methods for Time Series Analysis
- **Summary**

Summary

- Time series analysis considers data over time
  - Equal intervals
- More than just regression
  - Seasonal effects
  - Autocorrelation
- Complex topic with many options for modelling
  - Trend detection
  - Seasonal adjustment
  - Autocorrelation modelling
  - Completely different approaches, e.g., based on neural networks