Time Series

Topics covered so far

- Time Series
  - Introduction
  - Decomposition
  - Stationarity
  - Autoregressive models
  - Moving Average models
  - ARMA
  - ARIMA

Discussion questions

1. What is a time series and what are the features of a time series?
2. What is decomposition of time series and why do we do decomposition?
3. How do you define a stationary time series and how do we check stationarity?
4. What is the difference between Autoregressive and Moving Average models?
5. What do you understand by ARMA & ARIMA models?

Time Series

- A time series is a sequence of measurements on the same variable collected over time.
- The measurements are made at regular time intervals.
- Examples include - predicting the number of churning customers, explaining seasonal patterns in sales, etc.

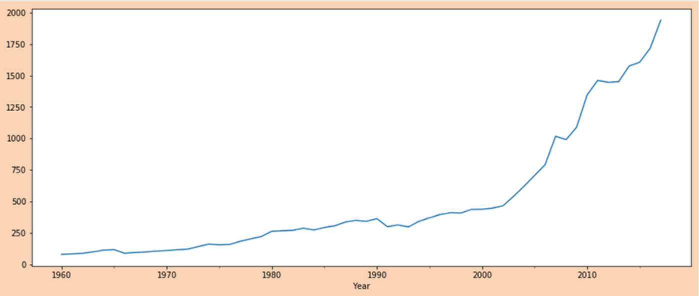
What is not a time series?

- Data collected on multiple items at the same point of time. For eg, dow jones average on a single day
- When the time periods are not the same. for e.g. in a single time series both quarterly and yearly data

Features of a time series: The following are the features of a time series-

- Data is not independent
- Ordering is very important because there is dependency and changing the order will change the data structure

Yearly time series of per capita GDP of India



Decomposition of a Time Series

Decomposition refers to the task of deconstructing the time series into several components like trend, seasonality etc.

There are two types of models used to decompose time series-

1. **Additive model:** It is useful when the seasonal variation is relatively constant over time.

Observation = Trend + Seasonality + Error

1. **Multiplicative model:** It is more realistic in nature and can be used when there are some variations over time

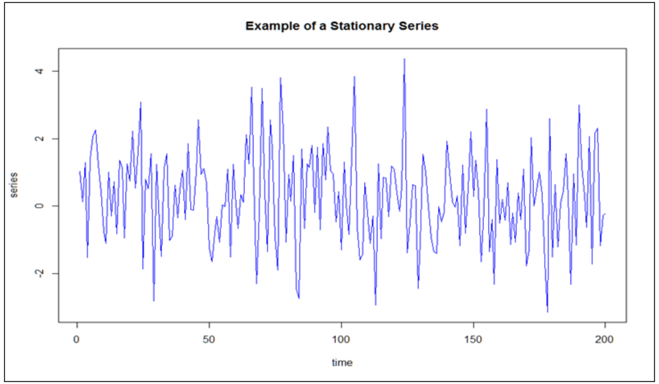
Observation = Trend * Seasonality * Error

Why do we need to decompose a time series:

- To compare the long-term movement of the series (Trend) vis-a-vis short-term movement (seasonality) to understand which has the higher influence
- To compare it amongst multiple sectors to avoid non-uniformity

Stationary Time Series

- In time series analysis, stationarity is a characteristic property of having **constant statistical measures** such as mean, variance, co-variance etc over a period of time.
- In other words, a time series is said to be stationary if the marginal distribution of y at a time p(yt) is the same at any other point in time. For time series analysis to be performed on a dataset, it should be stationary.
- It is also known as **white noise.**
- We can check the stationarity
  - By visualizing the rolling mean and standard deviation of the series
  - By using the Augmented Dickey-Fuller test



Autoregressive model vs Moving Average Model

Autoregressive models: Autoregressive models are based on the idea that the current value of the series, y_t, can be explained as a linear combination of p past values, $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ of the same series.

An autoregressive model of order p, abbreviated AR(p), is of the form

$$AR(p): \; y_t = f(y_{t-1}, y_{t-2}, \ldots y_{t-p})$$

The simplest AR process is AR(0), which has no dependence between the terms. In fact, AR(0) is essentially white noise. If y depends on more than one of its previous values then it is denoted by p parameters.

Moving Average models: While the autoregressive model considered the past values of the target variable for prediction, Moving average makes use of the white noise error terms. It can be represented as follows:

$$Y_t = \beta_0 + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \phi_3 \varepsilon_{t-3} + \ldots + \phi_q \varepsilon_{t-q}$$

Where, εt, εt−1,..., εt−q are the white noise error terms.

ARMA Model

The Autoregressive Moving Average model is a combination of the autoregressive model and the moving average model which uses both the past values as well as the error terms to predict for the future time series.
If the process has terms from both an AR(p) and MA(q) process, then the process ARMA(p, q) can be expressed as

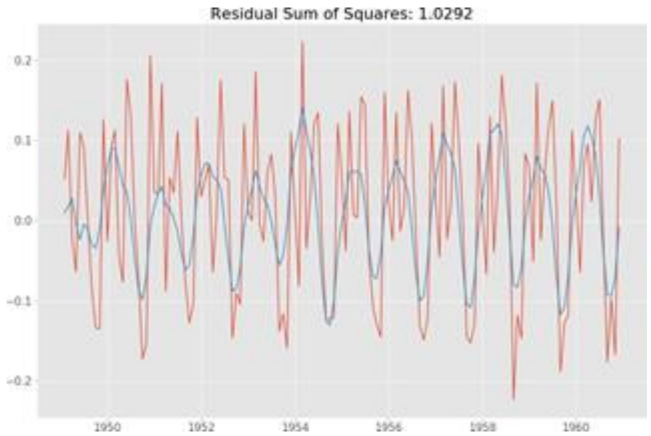
$$y_i = \phi_0 + \phi_1 y_{i-1} + \phi_2 y_{i-2} + \cdots + \phi_p y_{i-p} + \varepsilon_i + \theta_1 \varepsilon_{i-1} + \cdots + \theta_q \varepsilon_{i-q}$$

The ARMA model makes use of 2 parameters as given below:

p: Lag order or the number of past orders to be included in the model

q: The order of moving average

This plot shows the ARMA model fitted on the air passenger dataset. It gives the Residual Sum of Squares as 1.0292.



Source

ARIMA Model

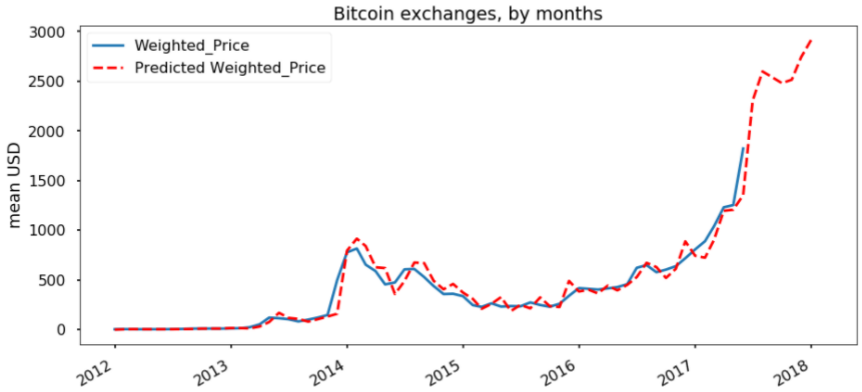
ARIMA stands for AutoRegressive Integrated Moving Average. It is a generalization of ARMA model and adds the notion of integration. The ARIMA model makes use of 3 parameters, which are given below:

p: Lag order or the number of past orders to be included in the model

d: The number of times differencing was applied in the original series in order to make the series stationary.

q: The order of moving average

This plot shows the ARIMA model fitted on the bitcoin price prediction dataset.



Case Study

Great Learning

Happy Learning !