Time Series Forecasting

Armine Hayrapetyan

Outline

- Time series
- Time series analysis
- Time series forecasting

Time Series

A time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time. Thus it is a sequence of discrete-time data.

Describing vs Predicting

We have different goals depending on whether we are interested in understanding a dataset or making predictions.

*In descriptive modeling, or **time series analysis**, a time series is modeled to determine its components in terms of seasonal patterns, trends, relation to external factors, and the like. … In contrast, **time series forecasting** uses the information in a time series (perhaps with additional information) to forecast future values of that series.*

Time Series Analysis

When using classical statistics, the primary concern is the analysis of time series.

Time series analysis involves developing models that best capture or describe an observed time series in order to understand the underlying causes. This field of study seeks the "why" behind a time series dataset.

This often involves making assumptions about the form of the data and decomposing the time series into constitution components.

The primary objective of time series analysis is to develop mathematical models that provide plausible descriptions from sample data

Time Series Forecasting

Making predictions about the future is called extrapolation in the classical statistical handling of time series data.

Forecasting involves taking models fit on historical data and using them to predict future observations.

The purpose of time series analysis is generally twofold: to understand or model the stochastic mechanisms that gives rise to an observed series and to predict or forecast the future values of a series based on the history of that series

Components of Time Series

- **level** - the baseline value for the series if it were a straight line
- **trend** - the optional and often linear increasing or decreasing behavior of the series over time
- **seasonality** - the optional repeating patterns or cycles of behavior over time
- **Noise** - the optional variability in the observations that cannot be explained by the model

All time series have a level, most have noise, and the trend and seasonality are optional.

Time Series

The main features of many time series are trends and seasonal variations … another important feature of most time series is that observations close together in time tend to be correlated (serially dependent)

y = level + trend + seasonality + noise

Concerns of Forecasting

- **How much data do you have available and are you able to gather it all together?** More data is often more helpful, offering greater opportunity for exploratory data analysis, model testing and tuning, and model fidelity
- **What is the time horizon of predictions that is required? Short, medium or long term?** Shorter time horizons are often easier to predict with higher confidence
- **Can forecasts be updated frequently over time or must they be made once and remain static?** Updating forecasts as new information becomes available often results in more accurate predictions
- **At what temporal frequency are forecasts required?** Often forecasts can be made at a lower or higher frequencies, allowing you to harness down-sampling, and up-sampling of data, which in turn can offer benefits while modeling

Caution

Time series data often requires cleaning, scaling, and even transformation.

- **Frequency** - perhaps data is provided at a frequency that is too high to model or is unevenly spaced through time requiring resampling for use in some models
- **Outliers** - perhaps there are corrupt or extreme outlier values that need to be identified and handled
- **Missing** - perhaps there are gaps or missing data that need to be interpolated or imputed

Summary

In these slides we've learned:

- About time series data and the difference between time series analysis and time series forecasting
- The constituent components that a time series may be decomposed into when performing an analysis

Reference: time-series-forecasting