Feature Engineering for Time Series

- By describing a time series not with a series of numbers detailing the step-by-setp outputs of a process but rather by describing it with a set of features, we can access ML methods designed for cross sectional data
- The features can be computed over the entire time series of as rolling or expanding window functions

Considerations for extracting features from Time series

- Stationarity
- Length of the time series - features may become unstable as the length of time series increases
- Domain knowledge
- External considerations

Catalog of common features

- Mean and variance

- Maximum and minimum
- Difference between last and first values
- Number of local maxima and minima
- Smoothness of the time series
- Periodicity and autocorrelation of the time series

Packages for feature generation

- tsfresh
- The following categories of features are computed
    - Descriptive statistics
    - Physics inspired category of indicators - nonlinearity (C3), complexity(cid_ce), friedrich_coefficients(returns coefficients of a model fitted to describe complex nonlinear motion) etc
    - History-compressing counts

*Cesium. This library also has a web-based GUI for feature generation

Feature selection

- Automatic feature selection based on automatic feature generation