

Time Series

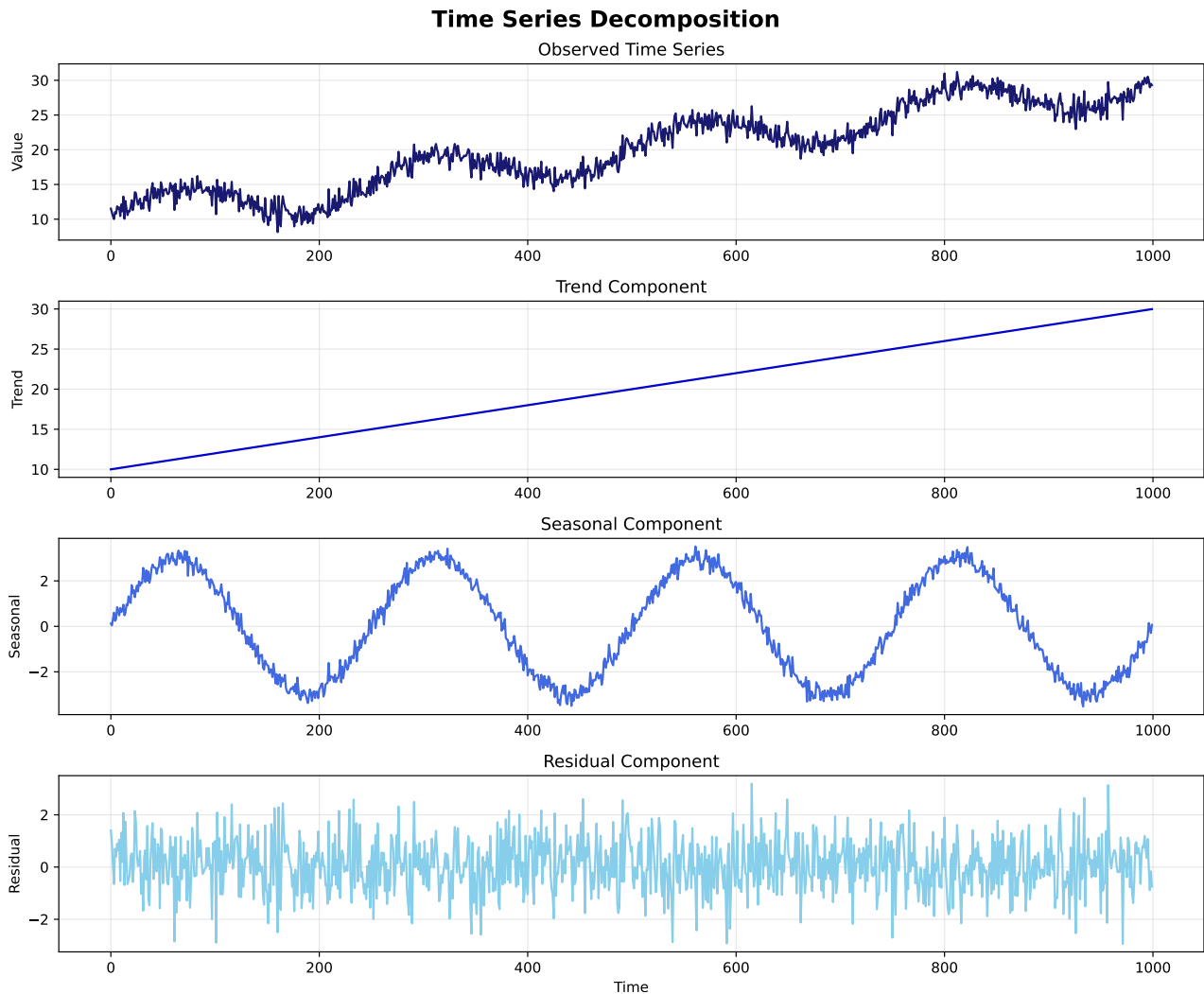
Oliver Gregory

Time series decomposition

Time series decomposition involves breaking down a time series in three main components: trend, seasonality and residual. The trend (T_t) is the long term progression of the series. The seasonality (S_t) is the recurring patterns within the data. The residual (R_t) is the random, unexplained variation in the data.

$$y_t = T_t + S_t + R_t$$

For example, in daily electricity usage data, the trend might reflect rising consumption over the years, the seasonality might be higher usage during winter, and the residuals would be the random, unexplained changes.



Stationarity

A time series is stationary when its behaviour remains constant over time which means that its statistical properties, like the mean and variance, don't change over time. Stationarity is crucial because many time series models, like ARMA, assume a stationary input. Non stationary data can be things like cumulative revenue over

time which will have a clear upward trend. Techniques like differencing are able to transform trending data to stationary data.

Differencing

Differencing removes trends to make a time series stationary by calculating the difference between consecutive observations.

$$\Delta y_t = y_t - y_{t-1}$$

For higher order differencing, the process is repeated, for example:

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} + y_{t-2}$$

We can also perform seasonal differencing to remove repeating patterns, such as weekly or yearly seasonality, by subtracting the value from the same period in the previous cycle.

$$\Delta y_t = y_t - y_{t-p} \text{ where } p \text{ is the period of the seasonal behaviour.}$$

Autoregression and Moving Averages

Autoregressive (AR) models

An AR model predicts the current value using past values. For example, an AR(2) model forecasts using the previous two lags.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t$$

Moving Average (MA) models

An MA model predicts the current value using previous forecast errors.

$$y_t = \mu + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \epsilon_t$$

Autoregressive Moving Average (ARMA) models

ARMA models combine AR and MA components to model stationary data. An ARMA(1,1) model contains one lag and one past error.

$$y_t = \phi_1 y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

This allows the model to capture both lagged relationships and the impact of past errors. This could be useful in demand forecasting for example where the demand on a given day may depend on the previous day's demand and unexpected supply chain issues.

AIC and BIC for optimising model complexity

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used to compare models by balancing goodness of fit with model complexity. They are defined as,

$$AIC = 2k - 2\ln(L)$$

$$BIC = k \ln(n) - 2\ln(L)$$

where k is the number of parameters in the model, n is the number of observations in the model and L is the likelihood of the model (i.e. the probability of observing the data, assuming the model is true).

! How to calculate likelihood !

ADF and KPSS too!