📖

Select number of harmonics

Let y_t denote time series data with multiple frequencies. For example, if y_t represents hourly data, then it will have two frequencies: $m_1 = 24$ (daily) and $m_2 = 24 * 7$ (weekly). Let $i = 1, 2$.

Step 1: Decompose the series for each frequency

Apply a $2 \times m_i$ moving average to estimate the trend f_i, t and then compute the detrended series $z_{i,t} = y_t - f_{i,t}$.

Step 2: Model each season component

For each series, we'll model the corresponding detrented series using a Fourier series. For $i = 1, 2$,

$$z_{i,t} \approx \sum_{j=1}^{k_i} a_{i,j} cos\left(\frac{2\pi jt}{m_i}\right) + b_{i,j} sin\left(\frac{2\pi jt}{m_i}\right) \tag{1}$$

Here k_i denotes the number of terms in the Fourier series, also known as harmonics.

Step 3: Estimate coefficients of Fourier series

For each frequency and a given harmonic k_i, perform multiple linear regression to estimate the coefficients. The regression model for each frequency $i = 1, 2$ is given in equation (2)

$$z_{i,t} = \beta_0 + \sum_{j=1}^{k_i} \left(\beta_{1,j} cos\left(\frac{2\pi jt}{m_i}\right) + \beta_{2,j} sin\left(\frac{2\pi jt}{m_i}\right)\right) + \epsilon_{i,t} \tag{2}$$

Here

- β_0 is the intercept.
- $\beta_{1,j}$ and $\beta_{2,j}$ are the coefficients for the cosine and the sine terms of the jth harmonic.
- $\epsilon_{i,t}$ is the error term.

Performing separate regressions for each frequency is a common approach when dealing with multiple frequencies since this allows us to better isolate the effects of each frequency. It can also simplify the interpretation of results.

Step 4: Combine the models

Combine the models for each frequency to obtain a model for the seasonal component z_t.

$$z_t = \sum_{i=1}^{2}\sum_{j=1}^{k_i} a_{i,j} cos\left(\frac{2\pi jt}{m_i}\right) + b_{i,j} sin\left(\frac{2\pi jt}{m_i}\right) \tag{3}$$

How to select the number of harmonics k_i

To select the number of harmonics, we can use an information criteria like the AIC. We can test different values for k_i, starting with 1, and then select the model with the lowest AIC.

When dealing with periodic functions with period M, the exact representation is achieved when

$$k = \begin{cases} \frac{M}{2} & \text{if M is even} \\ \frac{M-1}{2} & \text{if M is odd} \end{cases} \tag{4}$$

This is due to the Nyquist-Shannon samling theorem.

In practice, however, a smaller number of harmonics is enough to model the seasonality.

References

chatGPT: https://chat.openai.com/share/1c20b6ce-0bde-4c6a-b9a6-b993342bb51c

- Fourier Analysis of Time Series: An Introduction, by Peter Bloomfield.

Search words

- Fourier decomposition of a time series.