Time Series Analysis

Basics of Time Series Analysis

Nicoleta Serban, Ph.D.

Professor

Stewart School of Industrial and Systems Engineering

Decomposition: Seasonality Estimation

About This Lesson

Time Series: Basics

Data: Y_t, where t indexes time, e.g. minute, hour, day, month

Model: $Y_t = m_t + s_t + X_t$

- m_t is a trend component;
- s_t is a seasonality component with known periodicity *d* ($s_t = s_{t+d}$) such that $\sum_{j=1}^{d} s_j = 0$
- X_t is a stationary component, i.e. its probability distribution does not change when shifted in time

Approach: m_t and s_t are first estimated and subtracted from Y_t to have left the stationary process X_t to be model using time series modeling approaches.

Time Series: Seasonality

Elimination of Seasonality when there is no Trend

1. Estimate seasonality and remove it, or
2. Difference the data to remove the seasonality directly.

Seasonality Estimation Methods

1. Seasonal Average
2. Parametric Regression
   - Fit a mean for each seasonality group (e.g. month) using linear regression
   - Use a cosine-sin curve to fit the seasonal component

Seasonality: Averaging

For $\mathrm{k} = 1, 2, \ldots, d$, compute the average w_k of

$$\{Y_{k+jd}, with\ k + jd\ in\ the\ time\ domain\}$$

Then

$$\hat{s}_k = w_k - \frac{1}{d}\sum_{j=1}^{d} w_j$$

Seasonality: Seasonal Means Model

Model: $Y_t = \mu + s_t + X_t$ with $\sum_{j=1}^{d} s_j = 0$

Approach: Fit a mean for each seasonality group (e.g. month) using linear regression

- ANOVA model: Group k: Y_t for $\mathrm{t} = k + jd$
- Dummy Variables: C_k= 1 if $\mathrm{t} = k + jd$ and 0 otherwise
- Fit a linear regression model with d-1 dummy variables if a model with intercept or with d dummy variables if a model without intercept

Seasonality: Cosine-Sine Model

Model: Y_t = μ +s_t+X_t with $\sum_{j=1}^{d} s_j = 0$

Approach: Assume $s_t = \beta\cos(2\pi f t + \varphi)$ where β is the amplitude, f is the frequency (1/ f is the period) and φ is the phase (sets the set the arbitrary origin on the time axis).

- $s_t = \beta\cos(2\pi f t + \varphi) = \beta_1\cos(2\pi f t) - \beta_2\sin(2\pi f t)$ with $\beta_1 = \beta\cos(\varphi)$ and $\beta_2 = \beta\sin(\varphi)$
- Fit a linear regression: Y_t~ $\beta_1\cos(2\pi f t) - \beta_2\sin(2\pi f t)$ where β_1 and β_2 regression coefficients
- If seasonality has multiple frequencies (e.g. month, week), we can use different values of f (two predicting variables for each f)

Time Series: Trend & Seasonality

Step 1. Estimate the trend $\widehat{m}_t$ for $t = q + 1 \ldots n - q$

Step 2. Estimate seasonal components:

For $\mathrm{k} = 1, 2, \ldots, d$, compute the average w_k of

$$\{x_{k+jd} - \widehat{m}_{k+jd}, q < k + jd \leq n - q\}$$

Then $\hat{s}_k = w_k - \frac{1}{d}\sum_{j=1}^{d} w_j$

Step 3. Re-estimate the trend from the “deseasonalized data"

$$d_t = x_t - \hat{s}_t$$

A new set of estimates $\widehat{m}_t$ of the trend based on the deseasonalized data.

Time Series: Trend & Seasonality (cont'd)

Seasonality: Set the predicting variables

- Dummy variables for the seasonal effects (ANOVA)
- Cosine and sine variables

Trend: Set the approach

- Parametric Regression: Polynomial predicting variables
- Nonparametric Regression

Trend and Seasonality: Joint modeling

- Linear regression: Seasonality predicting variables and polynomial predicting variables in t
- Semiparametric Regression: Nonparametric model for the trend with linear predicting variables for seasonality

Differencing to Remove Trend & Seasonality

Define $\nabla_d Y_t = Y_t - Y_{t-d} = \left(1 - B^d\right) Y_t$

Then apply the differencing operator

$$\begin{aligned}\nabla_d Y_t &= \nabla_d m_t + \nabla_d s_t + \nabla_d X_t \\ &= m_t - m_{t-d} + s_t - s_{t-d} + X_t - X_{t-d} \\ &= m_t - m_{t-d} + X_t - X_{t-d}\end{aligned}$$

deseasonalized data.

This method is recommended when the time series is observed over a long period of time to allow for differencing over long periodicities/seasonality

Summary