Date: Thursday, October 18, 2018.

It could be a class

TIME SERIES

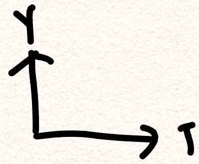
- Data with temporal measurement

Notation: $\{y_1, y_2, \dots, y_T\}$

y_t is the t^{th} observation of the variable

$\in t = 1, 2, \dots, T$.

Visualization: Scatter plot

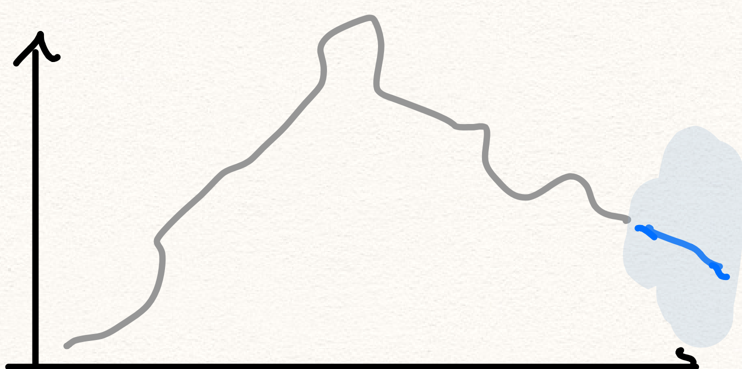


TIME SERIES ANALYSIS Model relationship between y and t.

FORECAST = PREDICTION (PURPOSE of SUBJECT)

T.S. Model characterises y_{t+1} and $\{y_1, \dots, y_t\}$

$$y_{t+1} = f(y_1, y_2, \dots, y_t)$$



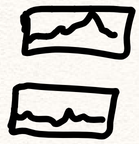
Prediction interval

Grows as its farther away because it predicts on predictions).

"Trying to understand the past to predict the future."

Multivariate: future values of y forecasted using past y's and $x_1, x_2, \ldots$

Correlation: mimic or inverse



Effectiveness:

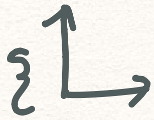
serial correlation

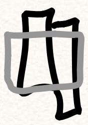
acf plots
autocorr fun

$corr(y_t, y_{t+1}) >> corr(y_t, y_{t+10})$
lag 1 lag 10

we need to account for this.

magnitude





Lag 0 = 1

Lag 1 = 0.9 (?)

Lag 20 = 0.1

Account for recent past

trend

consistent directional movement
("blur eyes")

↳ moving average

seasonality

regular and predictable fluctuations

$corr(y_t, y_{t+s})$ ↑↑↑
lag s

lag s

Random variation (errors) are unavoidable

trend + season : model → explains behavior on average

season - trend : errors

Classical decomposition model

$$\boxed{y_t = m_t + s_t + \varepsilon_t}$$

↑ observed time series ↑ trend component ↑ seasonal component ↑ error term

We can typically model trend with low-order polynomials, in which case

$$\boxed{m_t = \sum_{i=0}^{p} \beta_i t^i}$$

We can account for seasonality of period s by defining a categorical variable with s levels and model this with $s-1$ indicator variables, in which case

$$s_t = \sum_{j=1}^{s-1} \alpha_j x_j \qquad x_j = \begin{cases} 1 & \text{if } t \text{ is in season } j \\ 0 & \text{otherwise} \end{cases}$$

indicator variable,

This gives our model the following structure:

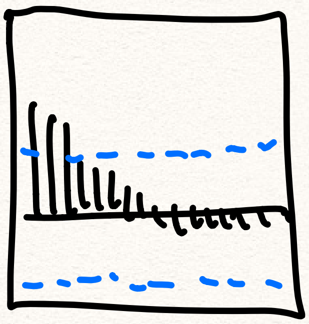
$$y_t = \sum_{i=0}^{p} \beta_i t^i + \sum_{j=1}^{s-1} q_j x_j + \varepsilon_t$$

R

Variance increasing → heteroscedasticity → log.

Fitting a line $\beta^{i=1}$

we need to undo transformation



Model not taking all correlation into account.

Assumptions about errors are not valid, inference not valid.
Estimates still work.

↳ $\hat{\beta} = (X^T X)^{-1} X^T Y$

It can be derived geometrically,