PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science

CSE DATA SCIENCE

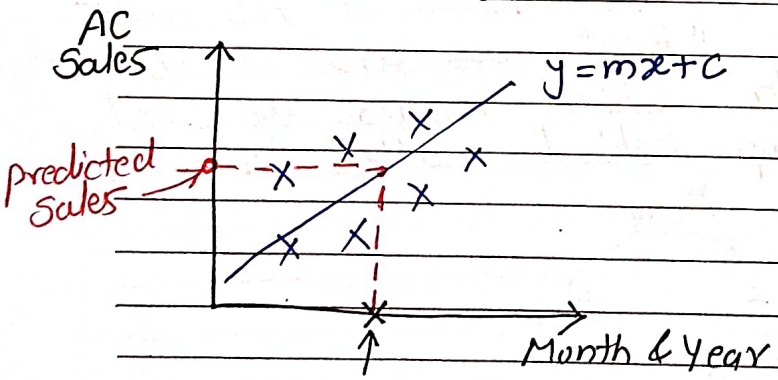
Semester : VI Subject :CSC601 Data Analytics and Visualization Academic Year: 2023- 2024

Time Series VS Non Time Series Problem

In this section we will discuss how time series problems are different than non time series problems. The concepts of interpolation problem statement and extrapolation problem statement.

Non time Series

Let's take the problem of regression from non time series. The use case of regression consists of two features "AC Sales" and "Month & Year". The datapoints we have for this case study is as shown below.



With the help of "moth & year" feature we are trying to predict the "AC sales".

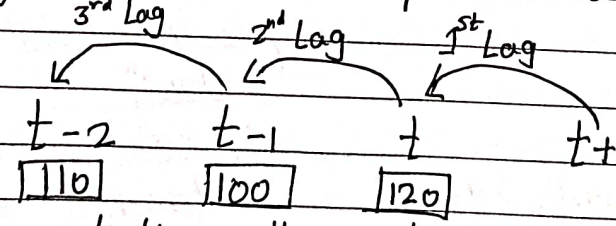
These kind of problem Statements are called as "Interpolation Problem Statement"

In interpolation problem statement we are specifically trying to predict the values in the given range. In our case for given range of "month & year" we are trying to predict the "AC sales" with the help of regression line.

Time Series Dat

Let's continue with the same example to understand the time series problem statement. and how the prediction is done.

The sales of AC on a particular day is "t". The To predict the value of AC sales on next day i.e. "t+1" we are dependent on previous data. which means we need to go one timestamp behind called "1 lag".



For predicting the sales of AC on next day we are dependent on previous time stamps.

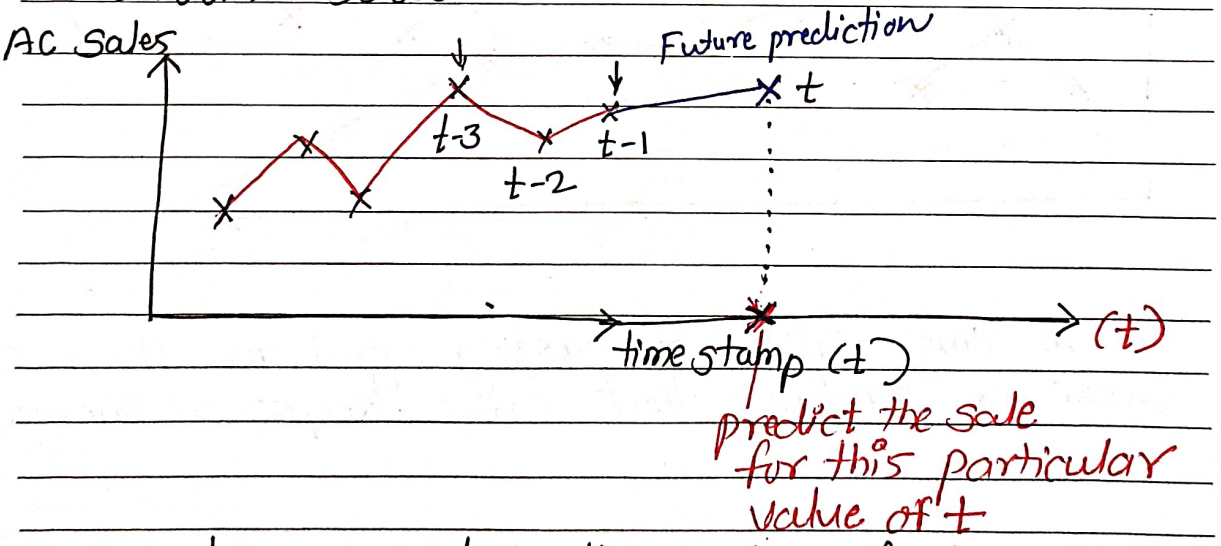
PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science
CSE DATA SCIENCE

Semester : VI Subject :CSC601 Data Analytics and Visualization Academic Year: 2023- 2024

the value of t+1 is determined/predicted from time stamps t, t-1, t-2.

How many lags should be considered that will be discussed in next section where we will be discussing auto regression moving average method and other approaches.

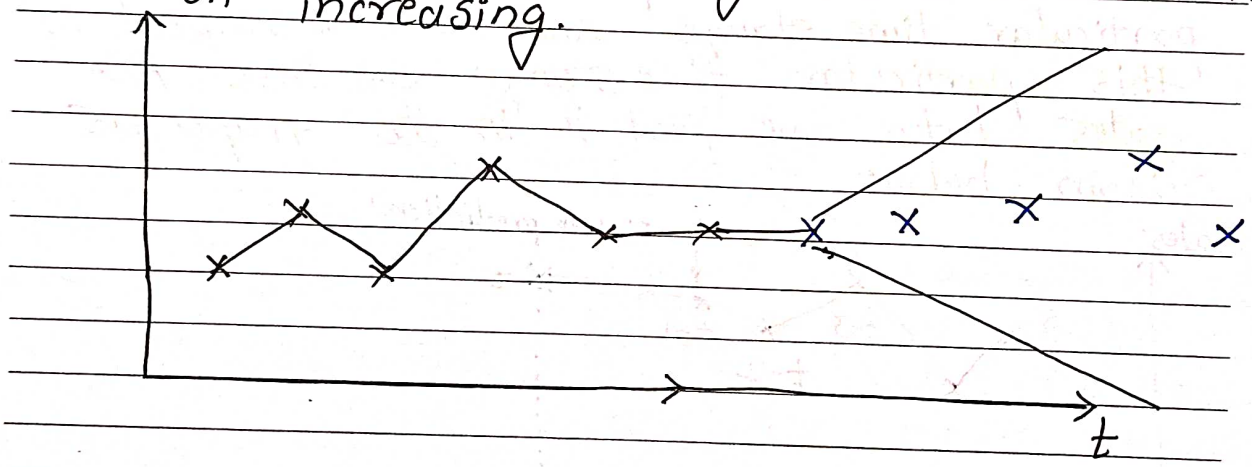
Thus the time series problem where we have particular time stamps and with respect to this particular time stamp we have AC sales data we plot it in the graph as shown below.



In order to predict the value of t we are dependent on 3 lags t-1, t-2 & t-3

This is called extrapolation. In extrapolation we only dependent on ~~timestamps to predict~~ previous timestamps for future forecasting.

from the data available of previous time stamps we are predicting value of t. While calculating the value of t using future forecasting there will be some error as we are predicting the value. & it keeps on increasing.



So in time series forecasting problem there is ~~error~~ a tendency that error keeps on increasing.

PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science

CSE DATA SCIENCE

Semester : VI Subject :CSC601 Data Analytics and Visualization Academic Year: 2023- 2024

Classical Time Series Models

The classical time series models consists of
1) AR Model (Autoregressive Model)
2) MA Model
3) ARMA Model
4) ARIMA Model

AR Model

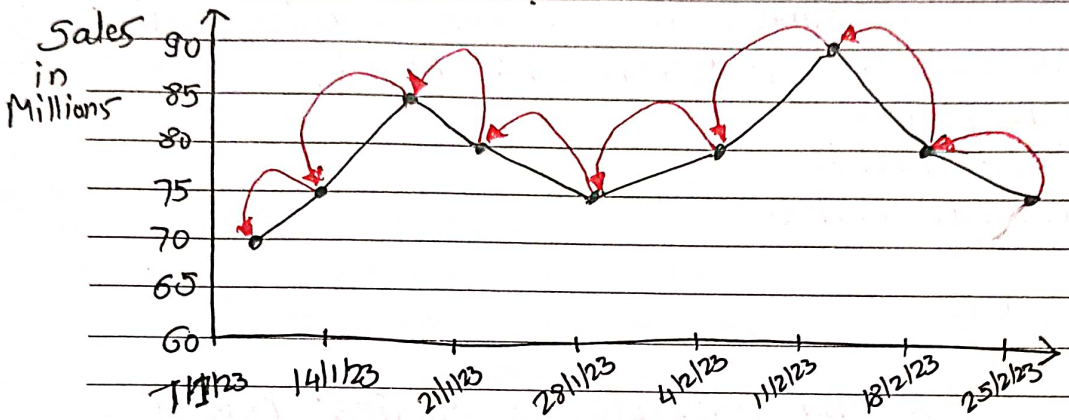
- This model ~~lag~~ forecast a series based solely on the past values in the series called lags.

- The model that depends only on one lag in the past is called AR(1) model.

$$Y_t = w + \Theta Y_{t-1} + e_t$$

Target — Y_t
Intercept — w
coefficient — Θ
Lagged Target — Y_{t-1}
Error — e_t

What is Y_{t-1} ?



This recursion in time goes back until the beginning of the series so these are called long memory models

$$Y_t = w + \phi Y_{t-1} + e_t$$

$$Y_{t-1} = w + \phi Y_{t-2} + e_{t-1}$$

Substitute Y_{t-1} in eqn of Y_t

$$Y_t = w + \phi(w + \phi Y_{t-2} + e_{t-1}) + e_t$$

$$Y_t = w + w\phi + \phi^2 Y_{t-2} + \phi e_{t-1} + e_t$$

If we continue with substitution method till we reach to very first point in time series the equation we get is

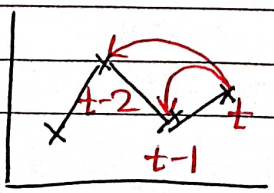
$$Y_t = \frac{w}{1-\phi} + \phi^t Y_1 + \phi^{t-1} e_2 + \phi^{t-2} e_3 + \dots + e_t$$

We call these models as long memory models as the effect of shocks that happen long ago have a little effect on the present if $|\phi| < 1$

AR(2) Model

A time series that is a linear function of 2 past values plus error is called an autoregressive process of order 2 - AR(2)

$$Y_t = w + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$



A time series that is a linear function of p past values plus error is called an autoregressive process of order p - AR(p)

$$Y_t = w + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots\ldots + \phi_p Y_{t-p} + e_t$$