Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

Department of Computer Science and Engineering (Data Science)

Subject: Time Series Analysis

Experiment 7

NAME:DIVYESH KHUNT SAPID:60009210116 BATCH:D12

(Stochastic Model - Vector Autoregression Model)

Aim: Implement a model on stochastic situation.

Theory:

Vector autoregression (VAR) is a stochastic process model utilized to seize the linear relation among the multiple variables of time-series data. In other words, it is a multivariate forecasting method utilized when two or more time-series variables have a strong internal relationship with each other. VAR is a bidirectional model, while others are unidirectional models. In a unidirectional model, a predictor influences the target, but not vice versa. In a bidirectional model, variables influence each other.

The normal AR(p) model equation looks like this:

$$\hat{Y}t = \mu + \phi 1 Yt\text{-}1 + \phi 2 Yt\text{-}2 \ldots + \phi p Yt\text{-}p + \varepsilon t$$

where μ is intercepting, and ϕ1, ϕ2…. ϕn are the coefficient of the lags of Y. In the VAR model, every single variable is modeled as a linear grouping of its past values and the past values of other variables in the time series. If you have multiple time series, which is determined to each other. So, one variable per equation will be designed. For instance, imagine that we have two variables of a time series, Y1, Y2. We want to forecast the value of these at time (t).

Here is the VAR (1) model with two time series (Y1 and Y2):

$$\hat{Y}1,t = \mu 1 + \phi 11 Y1,t\text{-}1 + \phi 12 Y1,t\text{-}2 + \varepsilon 1,t$$

$$\hat{Y}2,t = \mu 2 + \phi 21 Y2,t\text{-}1 + \phi 22 Y2,t\text{-}2 + \varepsilon 2,t$$

where y1,t-1, y2,t-1 are the first lag of the time series Y1 and Y2.

Lab Assignments to complete:

Perform the following tasks using the datasets mentioned. Download the datasets from the link given:

Link:

https://drive.google.com/drive/folders/1dbqJuZJULas76_Zzkqs-yRd2DbJReJup?usp=sharing

Colab Links:

https://colab.research.google.com/drive/1SqthK0TJ9xeYYdlS9uflc_bjAz89orCh?usp=sharing

Dataset 1: ice_cream_vs_heater.csv

1. Implement the stochastic model on the given dataset.
    a. Normalize the dataset\
    b. Take First Difference to Remove Trend
    c. Remove Increasing Volatility
    d. Remove Seasonality

e. Plot the PACF for Heater attribute
f. Select an appropriate model based on the correlation plot.
g. Calculate the Correlation between "heater" and lagged "ice cream".
h. Fit a VAR Model.
i. Analyze the Correlation matrix of residuals
j. State the final model used based on the coefficients and lag values.

COLAB LINK:

https://colab.research.google.com/drive/13_CrhFTt_iuCou8BmqJwk9LbzgrxSwlm?usp=sharing