Spring 2022 semester

DSA 5203 Time Series Analysis

S. Lakshmivarahan

varahan@ou.edu

I. Goal: Time series from different walks of life – monthly unemployment in various industrial sectors, seven-day average of COVID-19 infection rate and the resulting death rate in each of the countries in the world, daily foreign exchange rate between US dollars vs. all the major currencies in the world, annual rate of rise of carbon in the atmosphere, price of your favorite stock, to name a few, is one of the most ubiquitous form of data organization. This elective course epitomizes the basic principles of data analysis in that it rests on five pillars big data analytics. In the first data mining step, we extract the inherent temporal correlation structure by decomposing it into trend, seasonal and the stochastic components. In the second step, we catalogue the correlation structure of wide class of discrete time stochastic dynamic models from the autoregressive, integrated and moving average family, known as ARIMA models. By comparing the observed temporal correlations in each time series with those of the different types of ARIMA, in this third pattern recognition step, we then narrow down a small class of "useful" models that can capture the observed correlation. The fourth step deals with the critical task of estimating the unknown parameters of the chosen models using one of many known statistical methods and deals with the various criteria for selecting a "good" model. In the fifth and the last step, we then generate both short and long-range prediction and quantify the prediction error.

II. Intended Audience: This three-hour on-line course would be of interest to all those graduate and senior undergraduate students from all branches of Engineering, Computer Science, Data Science, Economics and Finance, Mathematics, Geophysics and Geology, Petroleum Engineering, Meteorology and Geography.

III. Modus operandi: This course will be based on prerecorded on-line lectures that you can watch and learn from. The assessment will be based on series of programming projects and some classical pencil paper calculations to enrich the understanding of the basic concepts. The preferred programming platform is MATLAB or any equivalent platform with graphical interface.

IV. Detailed Course Contents:

Part 1 Introduction

- Module 1.1 Data Mining and Time Series Analysis
- Module 1.2 Examples of Time Series
- Module 1.3 Basic Concepts in Time Series

Part 2 Statistical Estimation

- Module 2.1 Some Standard distributions
- Module 2.2 Statistical Estimation - basic concepts

Module 8.2 Estimation in AR (p) models: Maximum likelihood and least squares approach

Module 8.3 Estimation in MA (q) models

Module 8.4 Estimation in ARMA (p, q) models

Part 9 Partial Autocorrelation function (PACF)

Module 9.1 Computation of PACF

Part 10 Wold's decomposition

10.1 Deterministic component

10.2 Nondeterministic component

Part 11 Model Selection – Box and Jenkin's approach

Module 11.1 Model Selection Criteria

Module 11.2 Examples

Part 12 Volatility models

Module 12.1 ARCH models

Module 12.2 GARCH models

V. References:

P.J. Brockwell and R.A. Davis (2013) ***Time Series: theory and methods,*** Springer, Second Edition

J.D. Hamilton (1995) ***Time Series***, Princeton University Press

W. A. Fuller (2009) **Introduction to Statistical Time Series**, Wiley & sons, Second Edition

VI. Data Sources:

1) Department of Labor Statistics

2) http://www.qlik.com/us/products/data-market

provides examples from a number of application domains.

IV Programming environ:

1) MATLAB Time Series Toolbox

2) Open-source programs in R

****