TIME SERIES ANALYSIS AND FORECASTING (MScA 31006)

Arnab Bose, Ph.D.

Description

The objective of this course is two-folds – first, to understand time series concepts and methodologies. Second, to develop the ability to apply those concepts and methodologies to diverse practical applications and evaluate the results.

This course covers time series data characteristics (stationarity, unit-root), different modeling methodologies (regression, Box-Jenkins models, ARCH/GARCH, co-integration), and statistical tools to quantitatively select and evaluate the best modeling approach. The course focuses on industry applications via case studies and a class project. The software implementations are done in R.

Having completed this course, each student is able to apply time series modeling and forecasting tools to practical applications using the statistical software R. Furthermore, students develop an appreciation for the role of statistical modeling and machine learning in time series forecasting.

Course Materials

Class textbooks are:

[FPP] Hyndman, R. and Athanasopoulos, G. *Forecasting: Principles and Practice*. OTexts. 2014.

[MJK] Montgomery, D.C., Jennings, C.L. and Kulahci, M. *Introduction to Time Series Analysis and Forecasting*. 2nd Edition, Wiley, 2015.

Recommended reference books are:

[TSA] Cryer, Jonathan D. and Kung-Sik Chan. *Time Series Analysis with Applications in R*. 2nd Edition. Springer, 2008.

Box, George EP, Gwilym M. Jenkins, and Gregory C. Reinsel. *Time Series Analysis: Forecasting and Control*. 4th Edition, Wiley, 2008.

Additionally, students are given supplemental papers and case studies. These cover topics such as applied research and applications not covered in the class. The purpose is to illustrate practical applications of different time series methodologies and concepts.

Prerequisites

Students need to have taken Linear Algebra and Basic Statistics courses. Students need to be familiar with R software and basic R commands for linear algebra, statistics, and to read/write from/to files, and manipulate data frames.

Syllabus

If covered material is available in class textbooks, then relevant chapters are mentioned.

Important Note: *Changes may occur to the syllabus at the instructor's discretion. Students will be notified of changes by in-class announcement and/or email.*

Week 1 (FPP: 1 – 3, MJK: 1 – 2, TSA: 1 – 2)

Introduction to Time Series

White noise and random walk

Correlation, independence, orthogonality and autocorrelation

Stationarity and non-stationarity

Data modeling workflow

Exploratory data analysis

Getting started with R software

Week 2 (FPP: 4 – 7, MJK: 3 – 4, TSA: 3)

Smoothing Time Series (Moving Average / Exponential)

Holt Winters

Regression analysis

Univariate and multivariate regression modeling – model assumptions & multicollinearity

Week 3 (FPP: 8, MKJ: 5, TSA: 4 – 6)

Data transformations

Box-Jenkins ARMA models

Box-Jenkins ARIMA models

Stationarity and invertibility

Model specification

Week 4 (FPP: 8, MJK: 5, TSA: 7, 8, 10)

Akaike and Bayesian Information Criterion

Model estimation

Yule-Walker equations

Ljung Box test

Trend and seasonal decompositions

Seasonal ARIMA models

Final Project: Choose a time-series data application and develop a model solution

Week 5 (MJK: 2, TSA: 7 – 9)

Model diagnostics – residual analysis

Bootstrapping

Cross-Validation methodologies

Bagging, Random Forest and Boosting

Forecasting updates

Forecasting non-zero mean processes

Forecasting evaluation & accuracy

Prediction intervals

Week 6 (FPP: 9, MKJ: 4, TSA: 9)

Regression with ARMA errors

Cochrane Orcutt procedure

ARMAX models

Long memory ARIMA models

Model selection uncertainty

Week 7 (FPP: 9, MJK: 6, TSA: 11)

State space model representations

Multivariate linear models

Transfer function models

Week 8 (MJK: 7, TSA: 11 – 12)

Intervention & outlier analysis

ARCH/GARCH models

McLeod-Li and Jarque-Bera tests

Week 9 (FPP: 9, MJK: 7, TSA: 13)

CoIntegration & Error-Correction models

Granger causality

Spectral analysis

Polynomial and nonlinear regressions

Modeling multiple seasonality

Neural Network Autoregression (NNAR)

Deep Learning modeling using Recurrent Neural Network (RNN)

Real-world examples

Week 10

Final Project Presentation

Assignments, Quizzes and Class Project

Students are evaluated as follows. There are 8 assignments, each posted on canvas with a due date. The assignment problems are aimed at testing the students' understanding of the materials presented in the class. A student is allowed one assignment due date extension (no more than 3 days from due date) and two delayed submissions (no more than 2 weeks from due date) with a 20% penalty. Beyond these 3 allowed extensions/delays, no credit is given for other late assignments.

There are 4 quizzes held during class hours. The best 3 out of 4 will count towards grade. Students are asked questions based on the materials covered in class till that point.

There is a class project that is done by groups of students. Depending on the class size, each group consists of 1 - 3 students. The purpose of the project is to test the students' ability to apply the time series methodologies and tools presented in the class to diverse practical problems and interpret the results. The class project covers:

- Choose a real-world time series data problem and a forecasting objective
- Procure and analyze the time series data
- Propose 3 forecasting models
- Select final model based on selection criteria
- Evaluate forecasting metrics and accuracy
- Determine how the model forecasting can be improved
- Effectively communicate analytical work

For the class project, students are graded on 8 criteria outlined below. Each criterion is 5 points for a total of 40 points. The criteria are:

Presentation, Problem Statement, Assumptions, Data Properties & Transformations, Proposed Approaches (Models), Proposed Solution (Model Selection), Results (Accuracy), Future Work

I am available to work with each group on their project problem formulation and objectives. Each group member is asked to explicitly state his or her contribution to the project.

Software Requirements

The statistical software package R is used for this course. It is freely available from http://cran.rstudio.com/. Documentation is available at http://www.r-project.org/.

Grading

To determine that a student has met the objectives of this course, the grading includes homework assignments and class project weighted as follows:

Assignments	45%
Quizzes	15%
Class Project	40%

The grading scale is:

A = 93%-100%
A- = 90%-92%
B+ = 87%-89%
B = 83%-86%
B- = 80%-82%
C+ = 77%-79%
C = 73%-76%
C- = 70%-72%
F = 0%-69%

Additional Reading Materials

1. James Surowiecki, *The Wisdom of CrowdsThe Wisdom of Crowds*, Anchor Books, 2004.
2. Nassim N. Taleb, *Fooled by Randomness*, Penguin Books, 2004.
3. Sam L. Savage, *The Flaw of Averages*, Wiley, 2009.
4. Nassim N. Taleb, *Antifragile*, Random House, 2012.
5. Nate Silver, *The Signal and The Noise*, Penguin Press, 2012.
6. Philip Tetlock and Dan Gardner, *Superforecasting*, Random House, 2015.
7. Garry Kasparov, *Deep Thinking*, John Murray, 2017.