Time Series 413, Assignment 2
Introduction to Gaussian-Based Time Series Models (TS2)

Introduction

For this assignment, you will use the datasets and R script (TS2.R) included in the zip file (TS2.zip), posted to the Module 2 Overview page of Canvas. You will read the data files into R and conduct the requested analyses.

The instructions for submitting your assignment follow the Procedure section below.

The following list defines the data sets and their respective variables.

- Consumer Sentiment data of the University of Michigan. This survey series is widely used to indicate the consumer confidence about the U.S. economy. The data are available from Surveys of Consumers from the University of Michigan (urlhttp://www.sca.isr.umich.edu/tables.html) and also in the file m-umcsent.txt. The sample period may be determined from the earliest and latest dates of the data set.

    - Month: The month given as the full name.
    - YYYY: The four-digit year.
    - ICS_ALL: U.S. consumer sentiment index.

Procedure

The following steps are necessary to complete this assignment. Address each and every part and ensure that you cover all the details specified in the questions.

1. **EDA** (16 points) Consider the monthly series of Consumer Sentiment from the University of Michigan. This survey series is widely used to indicate the consumer confidence about the U.S. economy. The data are available from Surveys of Consumers from the University of Michigan. The sample period may be determined from the earliest and latest dates of the data set.

    1.1. Import the data and identify the years and months for which the sentiment data are missing. Remove the missing data records.

    1.2. As per Assignment 1, perform a complete EDA on the subset data with no missing records. This includes interpretations and an assessment of whether a log transformation, differencing, or both are necessary.

2. **Accounting for a Linear Trend in the Time Series** (16 points) Continuing with the Consumer Sentiment of the University of Michigan, denote the sentiment index as $x_t = $ Index.

Let $d_t = x_{t+1} - x_t$ or let $d_t = \log(x_{t+1}) - \log(x_t)$ depending on what you determined from your EDA.

2.1. For the transformed series of the sentiment data, test the hypothesis $H_0 : \mu = 0$ vs. $H_a : \mu \neq 0$; i.e., the expected change of sentiment is zero versus the alternative that the expected change is non-zero. Interpret the test outcome.

2.2. Is the transformed time series a Gaussian time series? Justify your answer.

3. **Identifying Autocorrelation, Stationarity, and White Noise** (16 points) With the Consumer Sentiment data:

3.1. Construct a mean model $d_t = \mu$ and give the parameter estimates. Interpret the estimates. Give the model degrees of freedom (df) which includes counting the intercept, the order of p, d, and q.

3.2. Perform model checking on the fitted model. Interpret the diagnostics for autocorrelation, stationarity, independence of lags, and normality of the residuals..

3.3. What are the business cycles in consumer sentiment? What do they mean?

3.4. Is the first difference mean model a white noise process? Justify your answer.

4. **Autoregressive (AR) Models** (16 points) Still with the Consumer Sentiment data:

4.1. Test the null hypothesis $H_0 : \rho_1 = \rho_2 = \cdots = \rho_{12} = 0$ versus the alternative $H_a : \rho_i \neq 0$ for some $i \in \{1, 12\}$ autocorrelation coefficients for the transformed time series data from part 2. Compare this with the ACF and PACF of the d_t series.

4.2. Construct an AR(12) model, $d_t = \phi_1 d_{t-1} + \cdots + \phi_{12} d_{t-12}$, and assess the fit with model diagnostics. Give the model degrees of freedom (df).

4.3. Compare the AR(12) model with the mean model from part 3 by giving your preference with a justification.

4.4. Simplify the fitted AR by removing parameter estimates with p-values greater than 0.05 using the **fixed** option in the **arima** function. Is the model adequate? Why?

4.5. What are the simplified model's business cycles? Are they different from the AR(12) model?

4.6. Compare the simplified model with the mean model. In terms of model diagnostics, which model is preferred? Justify your choice.

4.7. Compare the simplified model with the AR(12) model. In terms of model diagnostics and in-sample fitting, which model is preferred? Justify your choice.

4.8. Use **backtest** to compare the two AR models. You may start the forecast origin at $t = 400$. Which model is preferred? Why?

5. **Report** (16 points) Describe your choice of consumer sentiment model as if to a lay stakeholder; i.e, what are the forecasts and what do they mean? The report requires information from which the stakeholder can make decisions or take action, and not encumbered with statistics jargon.

.

Deliverables

Your instructor may modify these and all the following directions. See Section Submission Directions below. The assignment deliverables, each in pdf format, are as follows:

- *Only if requested by instructor*

    - The program or script
    - Logs
    - Outputs

- Mandatory
  Data analysis write-up: no programs, logs, or just code outputs; **complete EDA and model diagnostics are expected unless otherwise instructed; I will be looking for innovative interpretations in the assignments over and above the rote adherence to assignment requirements. Only partial credit will be awarded for rote adherence to assignment requirements.**.

The data analysis must follow and use the item numbering of each assignment, i.e., use the numbers, say, 1 - 5, with the sub-lettering if used. These deliverables are provided according to the instructions in the Submission Directions section below.

Submission Directions

Title Page

Include a title page with your name and the assignment designation. Leave room for instructor comments.

File Names

The assignment write-up file shall be submitted to Canvas according to the schedule in the syllabus using the item (1) naming convention below. The naming convention is case sensitive. Use letters and numbers as given. **The file name parts have no spaces or other separator characters**. TS2Lastname.pdf (submit via Canvas)

The parts are the assignment code, TS2; your last name with only the first letter capitalized; a period, and lastly, the extension "pdf". Generically,

TS2Lastname.pdf

For example: Suppose your name is Student McStats. Your filename then is:

TS2Mcstats.pdf

The analysis write-up file must be submitted for grading. Each write-up requires a title page for instructor comments. The analysis may use either R or any other statistics package you wish, or if you use more than one package, you must use the germane tables, plots, etc., in a single report. If you use more than one package, differences and similarities should be indicated.

email

ONLY IF REQUESTED email your instructor the program (script or code), log and output as separate pdf files. The R log and output may be combined. The file names shall be as follows:

- The program or script file names

    - TS2LastnameRprog.pdf

- The log file names

    - TS2LastnameRlog.pdf