Department of Statistical Sciences

Time Series Analysis

STA457H1/STA2202H1 S-LEC0201

STA457H1/STA2202H S-LEC0101 & STA457H1S LEC2001

Winter 2021

Assignment 4

Due: Apr. 8th at 10.00 pm EST

Submit an RMarkdown file and the corresponding pdf report. Please confirm the code's readability by leaving an appropriate comment on each line and add title and axis labels for the plots to get the full credit.

Q1. Crude oil prices in dollars per barrel are in oil. $library(astsa)$; $data(oil)$

a. Identifying the dependence orders of an ARIMA(p, d, q) model

b. Fit an ARIMA(p, d, q) model, estimate the parameters, and test the significance of the parameter estimates.

c. Performing all necessary diagnostics for the proposed model

d. After deciding on an appropriate model, forecast (with limits) the next 10 weeks.

e. Discuss the limitations your model may suffer from.

Q2. Repeat the simulations and analyses in Example 4.1 and Example 4.2 with the following changes:

a. Change the sample size to $n = 128$ and generate and plot the same series as in Example 4.1:

$$x_{t1} = 2\cos(2\pi\ 0.06t) + 3\sin(2\pi\ 0.06t),$$

$$x_{t2} = 4\cos(2\pi\ 0.10t) + 5\sin(2\pi\ 0.10t),$$

$$x_{t3} = 6\cos(2\pi\ 0.40t) + 7\sin(2\pi\ 0.40t),$$

$$x_t = x_{t1} + x_{t2} + x_{t3}.$$

What is the major difference between these series and the series generated in Example 4.1?

b. As in Example 4.2, compute and plot the periodogram of the series, x_t , generated in (a) and comment.

c. Repeat the analyses of (a) and (b) but with $n = 100$ (as in Example 4.1), and adding noise to x_t ; that is

$$x_t = x_{t1} + x_{t2} + x_{t3} + w_t$$

where $w_t \sim iid\ N(0, 25)$. That is, you should simulate and plot the data, and then plot the periodogram of x_t and comment.

Q3. The levels of salt concentration known to have occurred over rows, corresponding to the average temperature levels for the soil science data considered in Figs. 1.18 and 1.19, are in salt and saltemp. $library(astsa)$; $data(salt)$; $data(saltemp)$.

a. Identify the first three dominant frequencies by performing separate spectral analyses on the two series.

b. Include 95% confidence intervals for the dominant frequencies in part(a) and interpret your findings.

Late assignments will be accepted but subject to a 20% penalty per day late. Late submissions will not be allowed beyond 48 hours of the due date.