STAT 443: Assignment 4

Winter 2023

Due: 11:59pm March 31 (Friday), 2023

Notes:

- Your assignment must be submitted by the due time via Crowdmark. You will receive an email with a personal submission link (if not, check your spam folder!).
- Make sure you include all R codes and outputs for questions related to data analysis. Use R.Markdown to prepare your R code reports. You will NOT receive marks for uncommented R code or output.
- It is your responsibility to ensure your solution to each question is uploaded in the correct section, that your work is legible, and that all pages are rotated correctly.

Problem 1. [20 points] Let $\{Y_t\}$ be a stationary process with mean 0 and let a and b be constants.

(a) [10 points] If $X_t = a + bt + s_t + Y_t$, where s_t is a seasonal component with period 12, show that $\nabla\nabla_{12}X_t = (1-B)(1-B^{12})X_t$ is stationary and express its autocovariance function in terms of that of $\{Y_t\}$.

(b) [10 points] If $X_t = (a + bt)s_t + Y_t$, where s_t is a seasonal component with period 12, show that $\nabla_{12}^2 X_t$ is stationary and express its autocovariance function in terms of that of $\{Y_t\}$.

Problem 2. [15 points] Let Z_t be a causal stationary solution of ARCH(p) process and $E[Z_t^4] \leq \infty$. Assuming that such a process exists, show that $Y_t = Z_t^2/\alpha_0$ satisfies the equations

$$Y_t = e_t^2\Big(1 + \sum_{i=1}^{p} \alpha_i Y_{t-i}\Big) \ , \ \{e_t\} \sim \mathrm{WN}(0,1) \ ,$$

and deduce that Y_t has the same autocorrelation function as the AR(p) process

$$W_t = \sum_{i=1}^{p} \alpha_i W_{t-i} + e_t \ , \ \{e_t\} \sim \mathrm{WN}(0,1) \ .$$

Problem 3. [15 points] For an MA(1) process $\{Y_t\}$, we have the PACF

$$\phi_{kk} = -\frac{\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}} \ .$$

Show that for $k \geq 1$, $-0.5 \leq \phi_{kk} \leq 0.5$.

Problem 4. [30 points] Consider the austa time series of total international visitors (in millions) to Australia over the period 1980 to 2015 in the R package fpp, which is the package associated to the Hyndman and Athanasopoulos textbook.

(a) [5 points] Plot the series and its empirical ACF and PACF. Plot the first difference series, and its ACF and PACF. Comment on any features of interest observed.

(b) [10 points] Forecast the series 10 steps ahead, and construct 95% prediction intervals for these forecasts, by fitting an appropriate ARIMA model to the series. Carefully describe how you settled on the model that you chose.

(c) [10 points] If x_t denotes this time series, then also forecast the series 10 steps ahead, and construct 95% prediction intervals for these forecasts, by fitting a trend+noise model of the form

$x_t = s_t + y_t$, where y_t follows an ARMA model. Carefully describe how you settled on your model for the trend and noise terms.

(d) [**5 points**] Comment on the differences between the ARIMA and trend+noise forecasts. Which do you think is better and why?

Problem 5. [**30 points**] We consider the quarterly United States Gross National Product (GNP) from 1947 to 2002, which is already seasonally adjusted. Load the data in R and answer the following questions.

(a) [**5 points**] Plot the raw data y_t and preprocess the data to fit ARIMA models via the 1st difference of log time series. Comment on both y_t and $\nabla \log(y_t)$, and explain the meaning of $\nabla \log(y_t)$ in real life.

(b) [**10 points**] Plot the sample ACF and PACF of $\nabla \log(y_t)$, and suggest the parameters p and q if we want to fit an ARIMA$(p, 1, q)$ model on the logrithmic data. Write the form of your ARIMA$(p, 1, q)$ model.

(c) [**5 points**] Estimate the parameters in your ARIMA model and do diagnostics.

(d) [**5 points**] Simulate the model with the estimated parameters, and comment on the ACF and PACF.

(e) [**5 points**] Forecast the GNP value, i.e. y_t, for the year 2003 using the estimated model. Include the 95% interval for the forecasted values.