MA 641 Time Series Analysis
Homework 4

Nicolas Jorquera

Problem 1

Identify the following as specific ARIMA models. That is, what are p, d, and q and what are the values of the parameters (the ϕ's and θ's)?

ARIMA models are denoted by $ARIMA(p, d, q)$: p is the order of the autoregressive term (AR term). d is the number of differences required to make the series stationary. q is the order of the moving average term (MA term).

Part A

a. $Y_t = Y_{t-1} - 0.25Y_{t-2} + e_t - 0.1e_{t-1}$

The AR terms are Y_{t-1} and Y_{t-2}, and the MA terms are e_t and e_{t-1}.

- $p = 2$ (since there are two AR terms: Y_{t-1} and Y_{t-2})
- $q = 1$ (since there is one MA term: e_{t-1})
- $d = 0$ (since there is no differencing)

So, the model is $ARIMA(2, 0, 1)$ with: $\phi_1 = 1, \phi_2 = -0.25, \theta_1 = -0.1$

Part B

b. $Y_t = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$

The AR terms are Y_{t-1} and Y_{t-2}, and the MA terms are e_t, e_{t-1}, and e_{t-2}.

- $p = 2$ (since there are two AR terms: Y_{t-1} and Y_{t-2})
- $q = 2$ (since there are two MA terms: e_{t-1} and e_{t-2})
- $d = 0$ (since there is no differencing)

So, the model is $ARIMA(2, 0, 2)$ with: $\phi_1 = 0.5, \phi_2 = -0.5, \theta_1 = -0.5, \theta_2 = 0.25$

Problem 3

Suppose that $\{Y_t\}$ is generated according to $Y_t = e_t + c(e_{t-1} + e_{t-2} + e_{t-3} + ... + e_0)$ for $t > 0$, and $c > 0$.

Part A

a. Find the mean and covariance functions for $\{Y_t\}$. Is $\{Y_t\}$ stationary?

Mean of Y_t:

Given that e_t is a white noise process with mean 0:

$$E(Y_t) = E(e_t + c\sum_{i=0}^{t-1} e_i) = E(e_t) + cE(\sum_{i=0}^{t-1} e_i) = 0 + c \times 0 = 0$$

Covariance function:

$$\gamma_k = Cov(Y_t, Y_{t-k})$$

Using the given formula for Y_t, for $k = 0$:

$$\gamma_0 = Var(Y_t) = Var(e_t + c\sum_{i=0}^{t-1} e_i)$$

This is a sum of variances since e_t is independent across time. However, this variance increases with t, and thus is not constant. For $k > 0$: The covariance would involve terms from the error process at different times, so the structure would still depend on t. Since both the mean and the covariance are not independent of time (especially the covariance), $\{Y_t\}$ is not stationary.

Part B

b. Find the mean and covariance functions for $\{\nabla Y_t\}$. Is $\{\nabla Y_t\}$ stationary?

The differenced series $\nabla Y_t = Y_t - Y_{t-1}$:

$$\nabla Y_t = e_t + c(e_{t-1} + e_{t-2} + \ldots + e_0) - (e_{t-1} + c(e_{t-2} + e_{t-3} + \ldots + e_0))$$

$$\nabla Y_t = e_t - e_{t-1} + ce_{t-1} = e_t$$

Mean of ∇Y_t:

$$E(\nabla Y_t) = E(e_t) = 0$$

Covariance function of ∇Y_t:

$$\gamma_k^{\nabla} = Cov(\nabla Y_t, \nabla Y_{t-k})$$

Given that $\nabla Y_t = e_t$, the covariance is 0 for all $k > 0$ (since e_t is a white noise process) and the variance of e_t for $k = 0$. Since the mean is constant and the covariance is independent of time, $\{\nabla Y_t\}$ is stationary.

Part C

c. Identify $\{Y_t\}$ as a specific ARIMA process.

The series Y_t can be seen as an infinite order MA process (MA(∞)), with the weights decreasing geometrically. This is not a classic ARIMA process in the usual p, d, q format but can be approximated by a high-order ARIMA process.

Problem 4

Consider the stationary model $X_t = 0.5X_{t-2} + Z_t$ in which X_i and Z_j are uncorrelated for all $i < j$ and $\{Z_t\} \sim WN(0, \sigma^2)$. This is a special case of the general AR(2) process. Calculate the ACF of $\{X_t\}$.

To calculate the autocorrelation function (ACF) for the AR(2) process, we use the Yule-Walker equations. However, X_t only depends on X_{t-2} and not X_{t-1} First, we calculate the ACF:

For lag 0:

$$\rho(0) = 1$$

For lag 1:

$$\rho(1) = 0$$

(since X_t does not depend on X_{t-1})

For lag 2: Using the model equation:

$$\gamma(2) = Cov(X_t, X_{t-2})$$

$$= Cov(0.5X_{t-2} + Z_t, X_{t-4} + Z_{t-2})$$

Given that X_i and Z_j are uncorrelated for all $i < j$ and $\{Z_t\}$ is white noise:

$$\gamma(2) = 0.5\gamma(0)$$

$$\rho(2) = \frac{\gamma(2)}{\gamma(0)} = 0.5$$

For lags $k > 2$, the autocorrelations are 0 since X_t only depends on X_{t-2}.

So, the ACF for $\{X_t\}$ is:

$$\rho(0) = 1, \rho(1) = 0, \rho(2) = 0.5, \rho(k) = 0 \text{ for } k > 2$$

Problem 6

The data file winnebago contains monthly unit sales of recreational vehicles (RVs) from Winnebago, Inc., from November 1966 through February 1972.

(a) **Display and interpret the time series plot for these data.**

There's a clear upward trend in the sales data, indicating an increase in the popularity or demand for these RVs over the observed period. The sales data exhibits periodic peaks and troughs, suggesting a seasonal pattern. This could be attributed to factors such as holidays, vacation seasons, or other cyclical events that influence RV purchases. The sales figures show a degree of volatility, with some months experiencing sharp increases or drops.

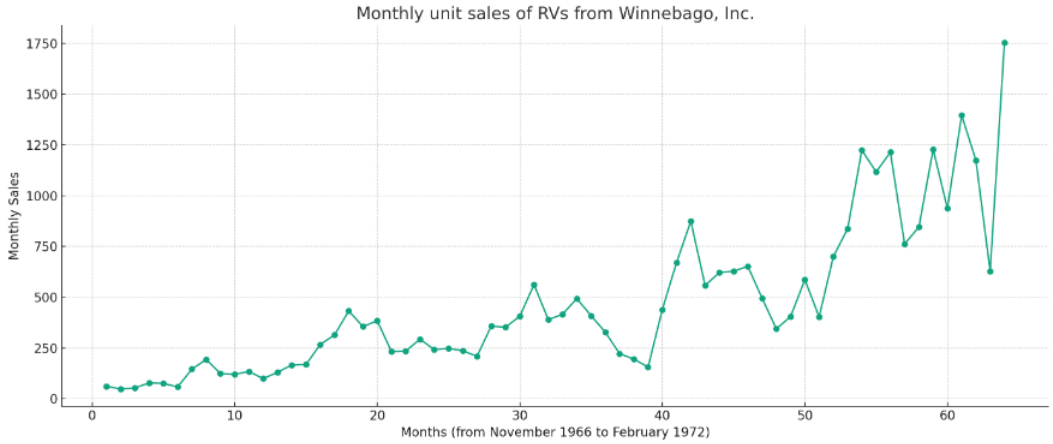

Figure 1: Monthly Unit Sells

(b) **Now take natural logarithms of the monthly sales figures and display the time series plot of the transformed values. Describe the effect of the logarithms on the behavior of the series.**

One of the primary effects of the logarithm transformation is the stabilization of variance. In the original time series, the amplitude of the fluctuations seemed to grow over time (heteroscedasticity). The log transformation helps in reducing this effect, making the variance more constant across time (homoscedasticity). The upward trend and the seasonal pattern evident in the original series are still visible in the transformed data. However, the magnitude of the peaks and troughs appears more consistent after the transformation.

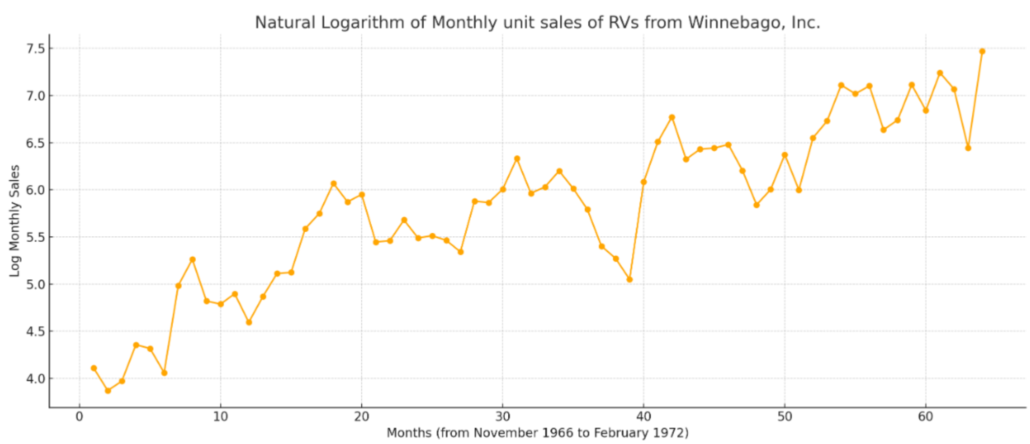

Figure 2: Monthly Unit Sells - Log

(c) Calculate the fractional relative changes, $\frac{Y_t - Y_{t-1}}{Y_{t-1}}$, and compare them with the differences of (natural) logarithms, $\Delta \log(Y_t) = \log(Y_t) - \log(Y_{t-1})$.

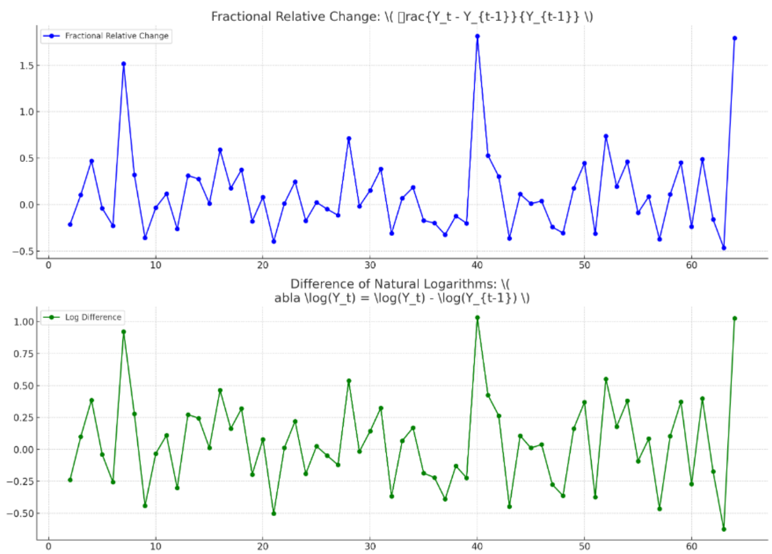

Figure 3: Fractional Relative Change - Log Difference

Both the fractional relative changes and the differences of natural logarithms show a similar pattern in their movements. This is expected since the difference of logarithms approximates percentage changes, especially for small changes. While the patterns are similar, the magnitudes differ. The difference in logarithms usually has a smaller range compared to the fractional relative changes. For smaller changes in the sales figures, the fractional relative change and the difference in logarithms are quite similar in magnitude. For larger changes, the difference in logarithms tends to be smaller in magnitude compared to the fractional relative change.

```
# Load necessary libraries
library(ggplot2)

# Read data
winnebago_data <- read.table("path_to_winnebago.dat", header=FALSE, skip=1)
colnames(winnebago_data) <- c("Monthly Sales")

# Plot
ggplot(data=winnebago_data, aes(x=seq_along(`Monthly Sales`), y=`Monthly Sales`)) +
  geom_line() +
  geom_point() +
  labs(title="Monthly unit sales of RVs from Winnebago, Inc.", x="Months (from November 
  theme_minimal()

# Calculate log sales
winnebago_data$Log_Sales <- log(winnebago_data$`Monthly Sales`)

# Plot
ggplot(data=winnebago_data, aes(x=seq_along(Log_Sales), y=Log_Sales)) +
  geom_line(color="orange") +
```

```
  geom_point(color="orange") +
  labs(title="Natural Logarithm of Monthly unit sales of RVs from Winnebago, Inc.", x="M
  theme_minimal()

# Calculate fractional relative change and log difference
winnebago_data$Fractional_Relative_Change <- c(NA, diff(winnebago_data$`Monthly Sales`)/
winnebago_data$Log_Difference <- c(NA, diff(winnebago_data$Log_Sales))

# Plot
par(mfrow=c(2,1))  # To plot 2 graphs in a single column

# Fractional relative change
plot(winnebago_data$Fractional_Relative_Change, type="o", col="blue", ylab="Fractional R

# Log difference
plot(winnebago_data$Log_Difference, type="o", col="green", ylab="Log Difference", xlab="
```