TSA Assignment 5

Christian Hilscher - 1570550

4/2/2020

Question 9

a.) Constructing Dataset

I first read in the data and construct the growth rate of the seasonally adjusted GDP.

```r
df <- read_excel(data_path, skip=10)

# Generating time series object
gdp <- ts(df[2], frequency=4, start = c(1991, 1))

# Computing quarterly growth rate
r <- log(gdp) - log(lag(gdp))
```

b.) Setting up functions

```r
# Function for looping over all models
get_predictions <- function(data_train){

  # Initializing space
  prediction_list <- vector()

  # Looping over all models
  for (i in seq(from=0, to=3)){

    modl <- arima(data_train, order = c(i, 0, 0),
                  optim.control = list(maxit = 1000))
    prediction <- predict(modl, n.ahead=1, se.fit=FALSE)
    prediction_list <- c(prediction_list, prediction)
  }

  # Returning a list with all predictions givena dataset
  return(prediction_list)
}

# Function for generating the dataset. If I want to predict
# the value of observation 106, I use all obsrvations up to
# 105 for training the model.
get_dataset <- function(orig_data, go_back){
  data_out <- orig_data[0:(length(orig_data)- go_back)]
  return(data_out)
}
```

c.) Alternative 1 (The way I now think you meant it)

Now I start first predicting Q2 of 2016 up to Q3 2017, in total the last six observations. I always use all of the available data for making the next one-step ahead forecast. The predictions I then save in a matrix with the rows being the different models and the columns holding the predictions from the earliest to the last. Afterwards the RMSEs are calculated for each of the $AR(p)$ models.

```r
# Setting up space for the predictions
prediction_matrix <- matrix(NA, nrow = 4, ncol = 6)
rmse_vector <- vector()
# Saving the real values to compute the RMSEs at the end
real_values <- tail(r, 6)

# Getting the predictions.
# Need to reverse the counter since the highest number is
# the prediction for the earliest date
for (i in rev(seq(from=1, to=6))){
    df <- get_dataset(r, i)
    predictions <- get_predictions(df)

    # Need to fill the prediction matrix backwards to have the correct order
    prediction_matrix[,(7-i)] <- predictions
}

# Calculating the RMSE for each model
for(p in seq(1, 4)){
  error <- rmse(real_values, prediction_matrix[p,])
  rmse_vector <- c(rmse_vector, error)
}
```

The RMSEs I get for each model are then

Model	AR(0)	AR(1)	AR(2)	AR(3)
RMSE	0.0043977	0.0037531	0.0038073	0.0038563

Choosing based on the root mean squared errors, I'd recommend the $AR(1)$ model.

c.) Alternative 2 (The way I thought you meant it originally)

Here I'll do a six-step ahead forecast and choose the model then on the basis of the RMSE from those six predictions. I use the first 101 of the in total 107 observations to fit the $AR(p)$ models with $p = 0, 1, 2$ and 3. First I use all of the available data for each model to fit. In a second run I disregard the first $(3 - p)$ observations for each model sucht that all of them have the same amount of data avaiable to them. This way I schould be able to see if the additional data available for the smaller p models improves accuracy or not.

```r
obs_to_keep <- 6
data_validation <- tail(r, obs_to_keep)
data_train_full <- r[0:(length(r)-obs_to_keep)]

rmse_list_full <- vector()
rmse_list_restricted <- vector()

for (i in seq(from=0, to=3)){

  modl1 <- arima(data_train_full, order = c(i, 0, 0),
```

```
                    optim.control = list(maxit = 1000))

    # Restricting the information
    data_train_restricted <- r[(3-i):(length(r)-obs_to_keep)]
    modl2 <- arima(data_train_restricted, order = c(i, 0, 0),
                    optim.control = list(maxit = 1000))

    y_pred_full  <- predict(modl1, n.ahead=obs_to_keep, se.fit=FALSE)
    y_pred1 <- to_ts(y_pred_full)

    y_pred_restricted  <- predict(modl2, n.ahead=obs_to_keep, se.fit=FALSE)
    y_pred2 <- to_ts(y_pred_restricted)
    # Calculating RMSE
    err1 <-  rmse(data_validation, y_pred1)
    err2 <-  rmse(data_validation, y_pred2)

    #Appending RMSE
    rmse_list_full <- c(rmse_list_full, err1)
    rmse_list_restricted <- c(rmse_list_restricted, err2)
}
```

From this I get the following RMSEs:

Table 2: Comparison of errors

Model	RMSE using all data	RMSE using restricted data
$AR(0)$	0.0044262	0.0044693
$AR(1)$	0.0044001	0.0045036
$AR(2)$	0.004403	0.004403
$AR(3)$	0.0043273	0.0043273

For the first two models the additional information from using the whole dataset shows up in somewhat lower RMSEs but the difference is pretty small. This seems reasonable when keeping in mind that for the $AR(0)$ the three additional observations are an increase of something marginally above 3% and even less for the other models.

Either way, the RMSE in both specifications is the lowest for the $AR(3)$ model which based on this information is the one I'd take.

Question 10

a.) Show that $T^{-5/2} \sum_{t=1}^{T} t^2 \epsilon_t \xrightarrow{\mathbf{d}} \sigma W(1) - 2\sigma \int_0^1 rW(r)dr$**:**

First looking at $\sum_{t=1}^{T} t^2 \epsilon_t$ where rewriting yields

$$\sum_{t=1}^{T} t^2 \epsilon_t = \sum_{t=1}^{T} \left(\sum_{s=1}^{t} \right) t\epsilon_t$$

$$= \sum_{s=1}^{T} \left(\sum_{t=1}^{T} t\epsilon_t - \sum_{t=1}^{s-1} t\epsilon_t \right)$$

$$= \underbrace{T \left(\sum_{t=1}^{T} t\epsilon_t \right)}_{=A} - \underbrace{\sum_{s=1}^{T} \left(\sum_{t=1}^{s-1} t\epsilon_t \right)}_{=B}$$

Concentrating on B for a second one can see that

$$\sum_{s=1}^{T} \sum_{t=1}^{s-1} t\epsilon_t = \sum_{s=1}^{T} \sum_{t=1}^{s-1} \sum_{p=1}^{t} \epsilon_t$$

$$= \sum_{s=1}^{T} \sum_{p=1}^{s-1} \sum_{t=p}^{s-1} \epsilon_t$$

$$= \sum_{s=1}^{T} \sum_{p=1}^{s-1} \left(\sum_{t=1}^{s-1} \epsilon_t - \sum_{t=1}^{p-1} \epsilon_t \right)$$

$$= \sum_{s=1}^{T} \sum_{p=1}^{s-1} x_{s-1} - \sum_{s=1}^{T} \sum_{p=1}^{s-1} x_{p-1}$$

$$= \sum_{s=1}^{T} (s-1) x_{s-1} - \sum_{s=1}^{T} \left(\sum_{p=1}^{T} x_{p-1} - \sum_{p=s}^{T} x_{p-1} \right)$$

$$= \sum_{s=1}^{T} s x_{s-1} - \sum_{s=1}^{T} x_{s-1} - T \sum_{p=1}^{T} x_{p-1} + \sum_{s=1}^{T} \sum_{p=s}^{T} x_{p-1}$$

$$= \sum_{s=1}^{T} s x_{s-1} - \sum_{s=1}^{T} x_{s-1} - T \sum_{p=1}^{T} x_{p-1} + \sum_{s=1}^{T} s x_{s-1}$$

Now I rewrite all indices as t to make everything coherent and change x_{t-1} to x_t:

$$\sum_{t=1}^{T} x_{t-1} = \sum_{t=1}^{T} x_t - T(x_0 + x_T) = \sum_{t=1}^{T} x_t + o_p(1)$$

and hence

$$B = 2 \sum_{t=1}^{T} t x_t - \sum_{t=1}^{T} x_t - T \sum_{t=1}^{T} x_t$$

where the $o_p(1)$ comes from the multiplication of $T^{-5/2}$ which I will later add but I don't want to carry around those terms with me so I omit them now.

Using the result from the lecture, I know that

$$A = T \left(\sum_{t=1}^{T} t\epsilon_t \right)$$

$$= T \left(T \sum_{t=1}^{T} \epsilon_t - \sum_{t=1}^{T} x_t \right)$$

4

Combining A and B then yields

$$A - B = T^2 \sum_{t=1}^{T} \epsilon_t - T \sum_{t=1}^{T} x_t - 2 \sum_{t=1}^{T} t x_t + \sum_{t=1}^{T} x_t + T \sum_{t=1}^{T} x_t$$

$$= T^2 \sum_{t=1}^{T} \epsilon_t - 2 \sum_{t=1}^{T} t x_t + \sum_{t=1}^{T} x_t$$

Multiplying with $T^{-5/2}$ results in

$$T^{-5/2}(A - B) = T^{-1/2} \sum_{t=1}^{T} \epsilon_t - 2(T^{-3/2}) \sum_{t=1}^{T} \frac{t}{T} x_t + T^{-5/2} \sum_{t=1}^{T} x_t$$

$$= \underbrace{T^{-1/2} \sum_{t=1}^{T} \epsilon_t}_{\Lambda} - \underbrace{2(T^{-3/2}) \sum_{t=1}^{T} \frac{t}{T} x_t}_{\Phi} + o_p(1)$$

The for the convergence of the first term I just use the result from the lecture. Regarding the second term, the interval for r gets smaller and small as $T \to \infty$ and finally converges to $\frac{t}{T}$. Then,

$$\Lambda \xrightarrow{d} \sigma W(1) \quad \text{and} \quad \Phi \xrightarrow{d} 2 \int_0^1 \sigma r W(r) dr$$

Putting everything together:

$$T^{-5/2} \sum_{t=1}^{T} t^2 \epsilon_t \quad = \quad T^{-5/2}(A - B) \quad \xrightarrow{d} \quad \sigma W(1) - 2 \int_0^1 \sigma r W(r) dr$$

b.) Find the variance of the limit

I can rewrite the variance as

$$var \left(\sigma W(1) - 2 \int_0^1 \sigma r W(r) dr \right) = var \left(\sigma W(1) \right)$$

$$+ var \left(2 \int_0^1 \sigma r W(r) dr \right)$$

$$- 2cov \left(\sigma W(1), 2 \int_0^1 \sigma r W(r) dr \right)$$

where I use the result from the lecture regarding the first term such that

$$var \left(\sigma W(1) \right) = \frac{1}{3} \sigma^2$$

For the second term I rewrite it a little bit

$$var \left(2 \int_0^1 \sigma r W(r) dr \right) = 4\sigma^2 var \left(\int_0^1 r W(r) dr \right)$$

$$= 4\sigma^2 var \left(\int_0^1 (1 - r^2) dW(r) \right)$$

$$= 4\sigma^2 \int_0^1 (1 - r^2)^2 dr$$

$$= 4\sigma^2 \left[r - \frac{2}{3} r^3 + \frac{1}{5} r^5 \right]_0^1$$

$$= 4\sigma^2 \frac{8}{15}$$

Regarding the covariance, rearranging results in

$$cov \left(\sigma W(1), 2 \int_0^1 \sigma r W(r) dr \right) = E \left[2\sigma^2 W(1) \int_0^1 r W(r) dr \right] - \sigma \underbrace{E \left[W(1) \right]}_{=0} E \left[2\sigma \int_0^1 r W(r) dr \right]$$

$$= 2\sigma^2 E \left[W(1) \int_0^1 (1 - r^2) dW(r) \right]$$

$$= 2\sigma^2 E \left[W(1) \left[(r - \frac{1}{3} r^3) \right]_0^1 \right]$$

$$= \frac{1}{3} \sigma^2 E \left[W(1) \right] = 0$$

because the expectation of a Wiener process is zero. Putting the variance together then yields

$$var \left(\sigma W(1) - 2 \int_0^1 \sigma r W(r) dr \right) = \frac{1}{3} \sigma^2 + \frac{32}{15} \sigma^2 = \frac{37}{15} \sigma^2$$