Solution for Problem Set 4

Mahrad Sharifvaghefi

The following lines call the required packages and source the required local functions for this problem set

```
source("r_functions/model_selection_function.R") # function for model selection
source("r_functions/ols_function.R") # function for OLS estimation
source("r_functions/t_test_function.R") # fucntion for t test
source("r_functions/expanding_window_forecast_function.R") # function for forecasting
#using expanding windows
source("r_functions/rolling_window_forecast_function.R") # function for forecasting
# using rolling windows
library("quantmod") # add quantmod to the list of Packages
library("fBasics") # add fBasics to the list of Packages
```

We can use the getSymbols command to fetch the data for GNP from Fred and then compute the GNP growth

```
getSymbols(Symbols ="GNP",src = "FRED",warnings = FALSE) # Download Quarterly data for GNP
```

```
## [1] "GNP"
```

```
Y <- as.matrix(GNP[,1])
DY <- diff(Y)/Y[1:(dim(Y)[1]-1),]

keep_data <- seq(from = as.Date("1960-04-01"), to = as.Date("2019-10-1"), by = "quarter")
DY_new = as.matrix(DY[as.Date(rownames(DY)) %in% keep_data,])
colnames(DY_new) = "GNP Growth"
n_obs = dim(DY_new)[1]
DY_new_date = as.Date(row.names(DY_new))
```

1.

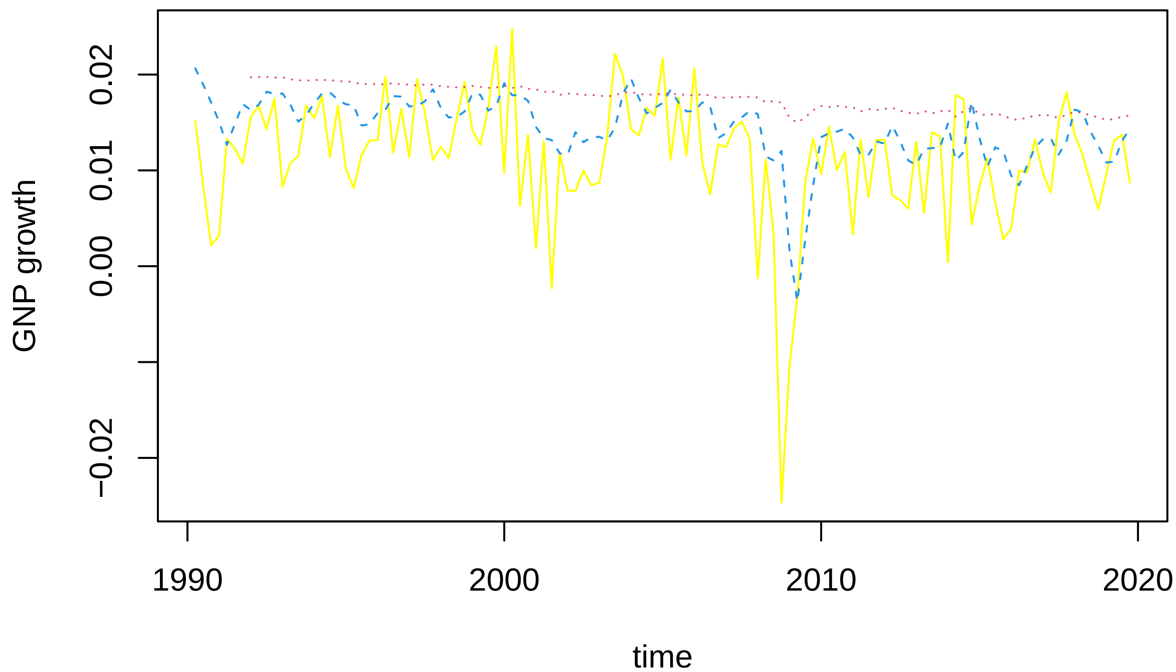
a)

```
lag_choice = NA
init_win_len = 120 # the first 30 years
num_step_ahead = 8 # 1 to 8 steps ahead forecastes
prediction_results = expanding_window(y = DY_new,
                                      init_win_len = init_win_len,
                                      pre_sel_num_lags = lag_choice,
                                      num_step_ahead = num_step_ahead,
                                      sel_method = 'aic'
                                      )
yhat_f_aic <- prediction_results$forecast
```

b)

```
y_f_aic <- prediction_results$actual_value
plot(x = DY_new_date[121:n_obs], y = y_f_aic,xlab='time',
     ylab='GNP growth',type='l',col="yellow")
```

```r
lines(x = DY_new_date[121:n_obs],y = yhat_f_aic[,1],lty=2, col = 4)
lines(x = DY_new_date[121:n_obs],y = yhat_f_aic[,8],lty=3, col = 2)
```



Since, current information are not very informative of what is going to happen 2 years later and hence our forecasts get closer the unconditional mean of GNP growth.

c)

```r
forecast_error =  kronecker(matrix(1,ncol = num_step_ahead),y_f_aic) - yhat_f_aic
rmsfe_ar_aic = sqrt(colMeans(forecast_error^2, na.rm = TRUE, dims = 1))
rmsfe_ar_aic
```

```
## [1] 0.006730889 0.007249419 0.007820273 0.008042676 0.008252251 0.008475482
## [7] 0.008639188 0.008738704
```

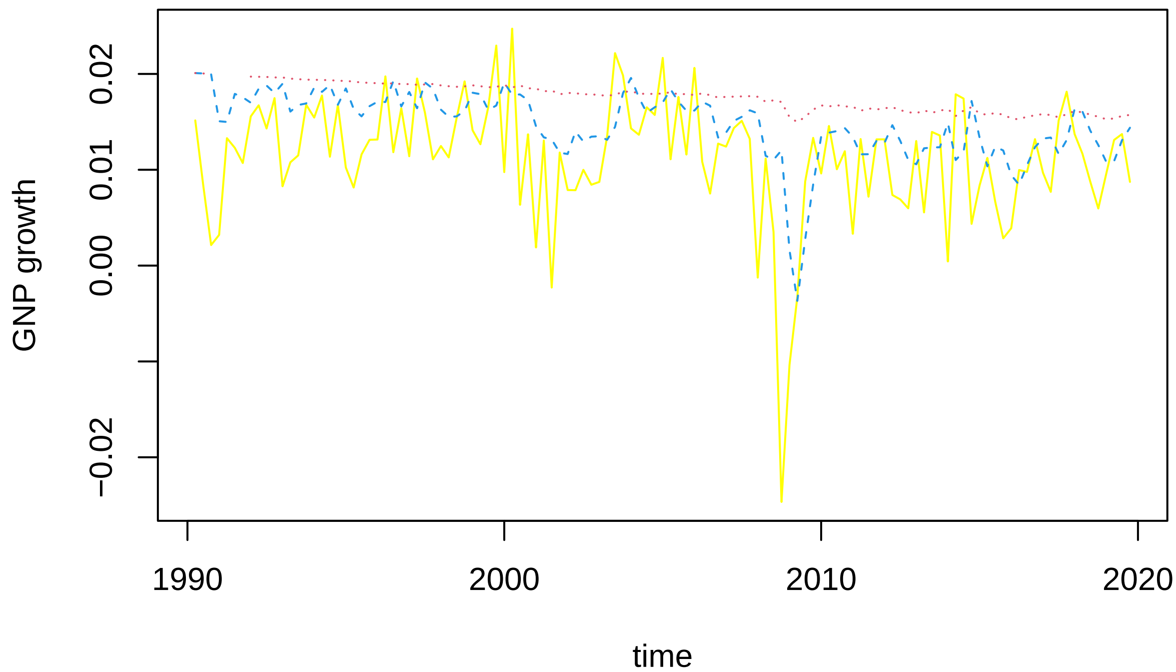
2

a)

```r
lag_choice = NA
init_win_len = 120 # the first 30 years
num_step_ahead = 8 # 1 to 8 steps ahead forecastes
prediction_results = expanding_window(y = DY_new,
                                      init_win_len = init_win_len,
                                      pre_sel_num_lags = lag_choice,
                                      num_step_ahead = num_step_ahead,
                                      sel_method = 'bic')
yhat_f_bic <- prediction_results$forecast
```

b)

```r
y_f_bic <- prediction_results$actual_value
plot(x = DY_new_date[121:n_obs], y = y_f_bic,xlab='time',ylab='GNP growth',type='l',col="yellow")
lines(x = DY_new_date[121:n_obs],y = yhat_f_bic[,1],lty=2, col = 4)
lines(x = DY_new_date[121:n_obs],y = yhat_f_bic[,8],lty=3, col = 2)
```



Since, current information are not very informative of what is going to happen 2 years later and hence our forecasts get closer the unconditional mean of GNP growth.

c)

```r
forecast_error =  kronecker(matrix(1,ncol = num_step_ahead),y_f_bic) - yhat_f_bic
rmsfe_ar_bic = sqrt(colMeans(forecast_error^2, na.rm = TRUE, dims = 1))
rmsfe_ar_bic
```

```
## [1] 0.006902092 0.007404226 0.007954118 0.008264777 0.008432854 0.008624285
## [7] 0.008769194 0.008859158
```

3

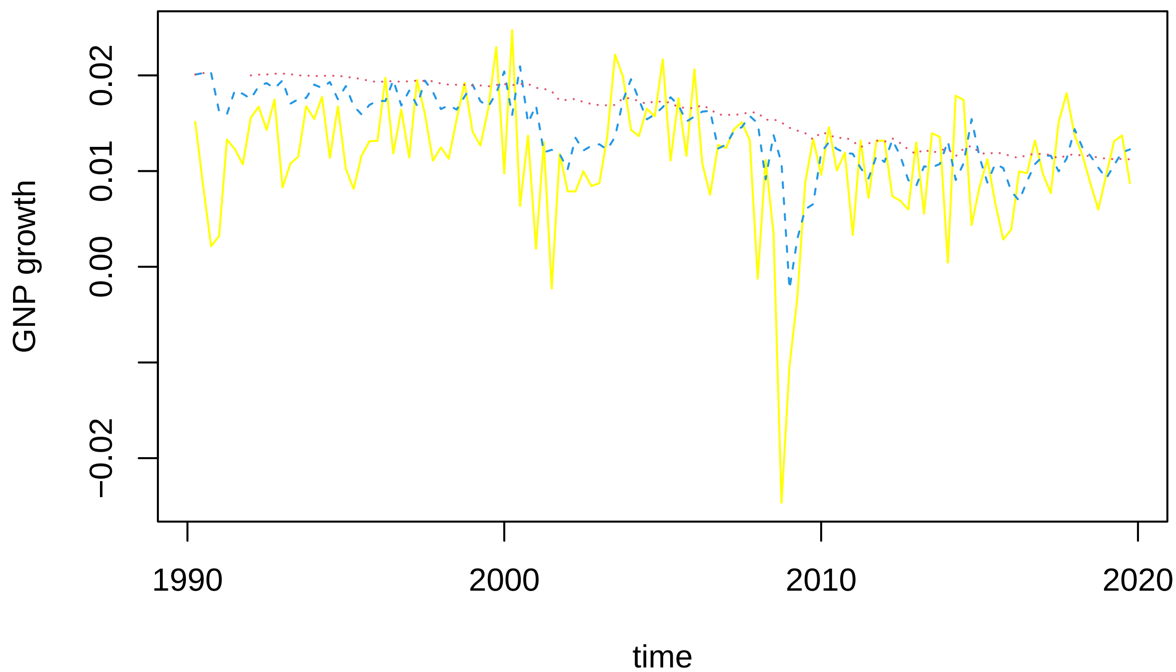
a)

```r
lag_choice = NA
init_win_len = 120 # the first 30 years
num_step_ahead = 8 # 1 to 8 steps ahead forecasts

prediction_results_bic = rolling_window(y = DY_new,
                                        init_win_len = init_win_len,
                                        pre_sel_num_lags = lag_choice,
```

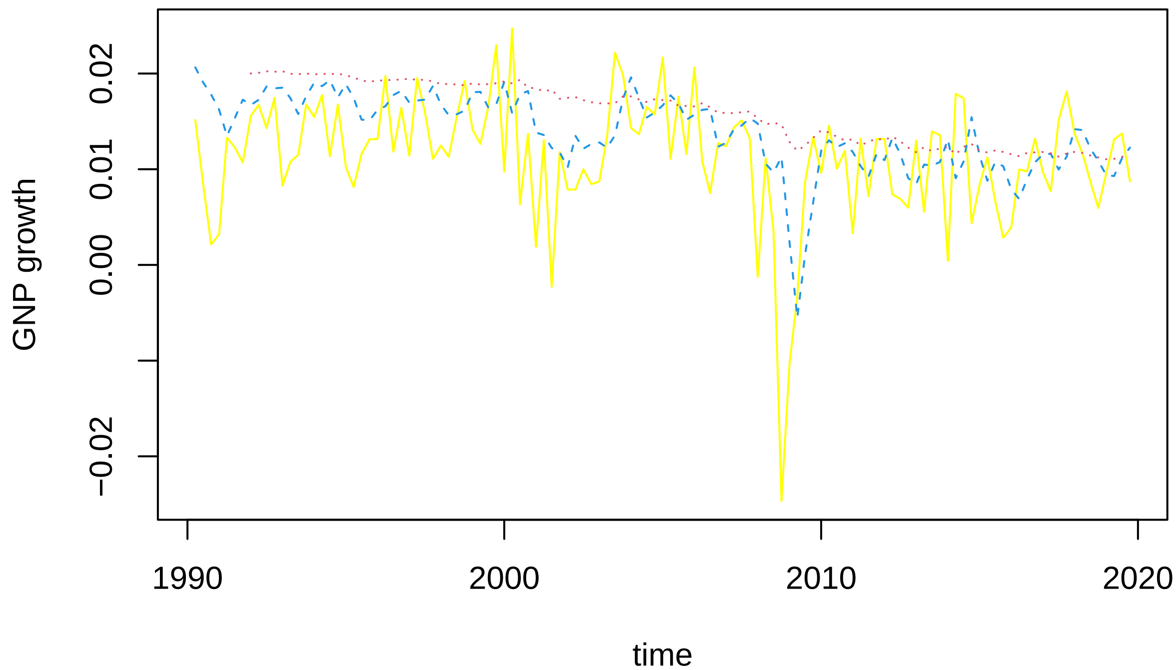
```
                                            num_step_ahead = num_step_ahead,
                                            sel_method = 'bic')
yhat_f_roll_bic <- prediction_results_bic$forecast

prediction_results_aic = rolling_window(y = DY_new,
                                        init_win_len = init_win_len,
                                        pre_sel_num_lags = lag_choice,
                                        num_step_ahead = num_step_ahead,
                                        sel_method = 'aic')
yhat_f_roll_aic <- prediction_results_aic$forecast
```

b)

```
y_f_roll_bic <- prediction_results_bic$actual_value
plot(x = DY_new_date[121:n_obs], y = y_f_roll_bic,xlab='time',ylab='GNP growth',type='l',col="yellow")
lines(x = DY_new_date[121:n_obs],y = yhat_f_roll_bic[,1],lty=2, col = 4)
lines(x = DY_new_date[121:n_obs],y = yhat_f_roll_bic[,8],lty=3, col = 2)
```




```
y_f_roll_aic <- prediction_results_aic$actual_value
plot(x = DY_new_date[121:n_obs], y = y_f_roll_aic,xlab='time',ylab='GNP growth',type='l',col="yellow")
lines(x = DY_new_date[121:n_obs],y = yhat_f_roll_aic[,1],lty=2, col = 4)
lines(x = DY_new_date[121:n_obs],y = yhat_f_roll_aic[,8],lty=3, col = 2)
```



Since, current information are not very informative of what is going to happen 2 years later and hence our forecasts get closer the unconditional mean of GNP growth.

c)

```
forecast_error =  kronecker(matrix(1,ncol = num_step_ahead),y_f_roll_bic) - yhat_f_roll_bic
rmsfe_ar_roll_bic = sqrt(colMeans(forecast_error^2, na.rm = TRUE, dims = 1))
rmsfe_ar_roll_bic
```

```
## [1] 0.006820480 0.007497703 0.007810063 0.007985793 0.007950125 0.008013522
## [7] 0.008047356 0.008057507
```

```
forecast_error =  kronecker(matrix(1,ncol = num_step_ahead),y_f_roll_aic) - yhat_f_roll_aic
rmsfe_ar_roll_aic = sqrt(colMeans(forecast_error^2, na.rm = TRUE, dims = 1))
rmsfe_ar_roll_aic
```

```
## [1] 0.006622253 0.006989234 0.007309363 0.007375393 0.007437334 0.007578259
## [7] 0.007677265 0.007729170
```

4

a)

```
yhat_f_ave = (yhat_f_aic + yhat_f_roll_aic)/2
forecast_error =  kronecker(matrix(1,ncol = num_step_ahead),y_f_aic) - yhat_f_ave
rmsfe_ave = sqrt(colMeans(forecast_error^2, na.rm = TRUE, dims = 1))
rmsfe_ave
```

```
## [1] 0.006649454 0.007082377 0.007512828 0.007649116 0.007777164 0.007954953
## [7] 0.008082559 0.008155535
```

b)

```r
rmsfe_all = rbind(rmsfe_ar_aic,rmsfe_ar_bic,rmsfe_ar_roll_aic,rmsfe_ar_roll_bic, rmsfe_ave)
rmsfe_all
```

```
##                          [,1]        [,2]        [,3]        [,4]        [,5]
## rmsfe_ar_aic      0.006730889 0.007249419 0.007820273 0.008042676 0.008252251
## rmsfe_ar_bic      0.006902092 0.007404226 0.007954118 0.008264777 0.008432854
## rmsfe_ar_roll_aic 0.006622253 0.006989234 0.007309363 0.007375393 0.007437334
## rmsfe_ar_roll_bic 0.006820480 0.007497703 0.007810063 0.007985793 0.007950125
## rmsfe_ave         0.006649454 0.007082377 0.007512828 0.007649116 0.007777164
##                          [,6]        [,7]        [,8]
## rmsfe_ar_aic      0.008475482 0.008639188 0.008738704
## rmsfe_ar_bic      0.008624285 0.008769194 0.008859158
## rmsfe_ar_roll_aic 0.007578259 0.007677265 0.007729170
## rmsfe_ar_roll_bic 0.008013522 0.008047356 0.008057507
## rmsfe_ave         0.007954953 0.008082559 0.008155535
```

i) In this case, the rolling forecast with lags chosen by AIC has the lowest root mean square forecast error and performs best for each horizon.

ii) The average model never has the lowest root mean square forecast error and hence never performs the best. There also exists no forecasting horizon over which the average model has the highest root mean square forecast error and hence performs the worst.

\

5 (a) $y^f_{t+\tau|t} = \mathbb{E}(y_{t+\tau}|I_t) = 1 + 0.5\mathbb{E}(y_{t+\tau-2}|I_t) + \mathbb{E}(u_{t+\tau}|I_t) = 1 + 0.5y^f_{t+\tau-2|t}$.

(b) $e^f_{t+\tau|t} = y_{t+\tau} - y^f_{t+\tau|t} = 1 + 0.5y_{t+\tau-2} + u_{t+\tau} - (1 + 0.5y^f_{t+\tau-2|t}) = 0.5e^f_{t+\tau-2|t} + u_{t+\tau}$

(c) We know that $\lim_{\tau\to\infty} y^f_{t+\tau|t} = \mathbb{E}(y_t)$. And $\mathbb{E}(y_t) = 1 + 0.5\mathbb{E}(y_{t-2}) = 1 + 0.5\mathbb{E}(y_t)$. Therefore, $\mathbb{E}(y_t) = 2$. So, $\lim_{\tau\to\infty} y^f_{t+\tau|t} = 2$

(d) We know that $\lim_{\tau\to\infty} \mathrm{var}(e^f_{t+\tau|t}) = \mathrm{var}(y_t)$. And $\mathrm{var}(y_t) = 0.25\mathrm{var}(y_{t-2}) + \mathrm{var}(u_t) = 0.25\mathrm{var}(y_t) + 1$. Therefore, $\mathrm{var}(y_t) = \frac{4}{3}$. So, $\lim_{\tau\to\infty} \mathrm{var}(e^f_{t+\tau|t}) = \frac{4}{3}$