

Exercise 6

Philipp Drebes

12.04.2023

Exercise 6.1

We visit the analysis of the yield of a chemical process and will have a look at the <http://stat.ethz.ch/Teaching/Datasets/WBL/yields.dat> time series and its autocorrelations.

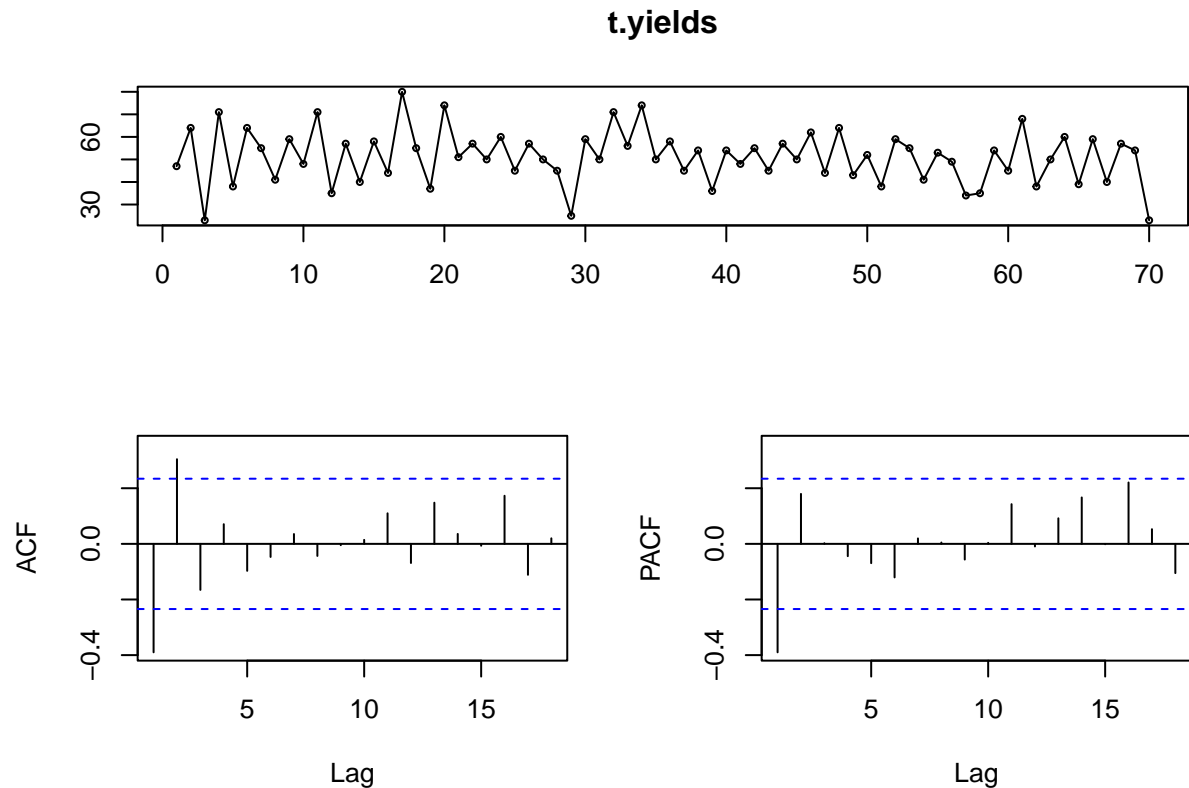
```
library(forecast)

## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo

yields <- read.table("http://stat.ethz.ch/Teaching/Datasets/WBL/yields.dat", header = FALSE)
t.yields <- ts(yields[, 1])
```

a) Could these data be generated by an AR-process? If yes, what is the order p ?

```
tsdisplay(t.yields)
```



It could be an AR(p) process as the ACF has an exponential decay. The PACF has no significant values after lag 1, so we assume $p = 1$.

- b) Using the autocorrelations, compute the Yule-Walker estimate of α by hand.
Recall the Yule-Walker equation for the estimated autocorrelation function at lag 1 reads:

$$\hat{p}(1) = \alpha \cdot \hat{p}(0)$$

Furthermore, find the estimated mean $\hat{\mu}_X$ as well as the innovation variance $\hat{\sigma}^2$.

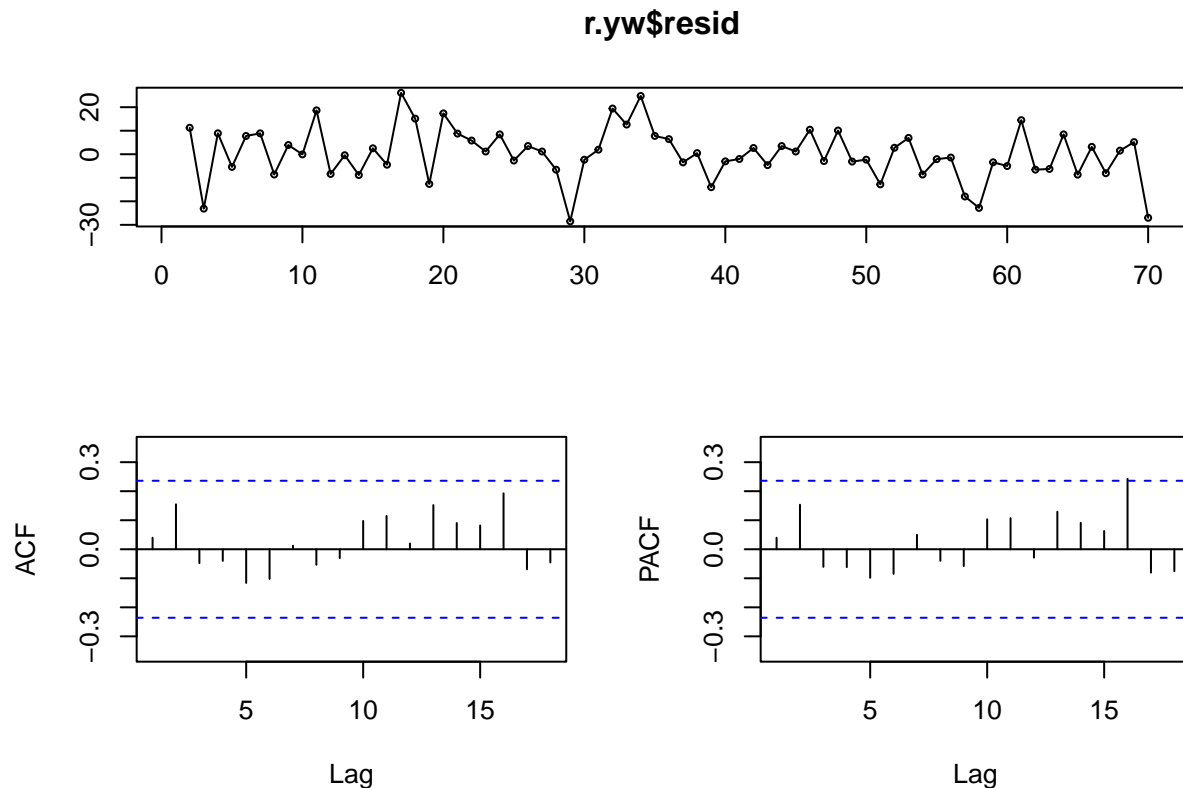
Check your results using R.

```
r.yw <- ar(t.yields, method = "yw", order.max = 1)
str(r.yw)
```

```
## List of 15
## $ order      : int 1
## $ ar         : num -0.39
## $ var.pred   : num 122
## $ x.mean     : num 51.1
## $ aic        : Named num [1:2] 9.54 0
## .. attr(*, "names")= chr [1:2] "0" "1"
## $ n.used     : int 70
## $ n.obs      : int 70
## $ order.max  : num 1
## $ partialacf : num [1, 1, 1] -0.39
## $ resid      : Time-Series [1:70] from 1 to 70: NA 11.26 -23.11 8.9 -5.38 ...
## $ method     : chr "Yule-Walker"
## $ series     : chr "t.yields"
```

```
## $ frequency : num 1
## $ call      : language ar(x = t.yields, order.max = 1, method = "yw")
## $ asy.var.coef: num [1, 1] 0.0125
## - attr(*, "class")= chr "ar"

tsdisplay(r.yw$resid)
```



α is estimated at $\alpha = -0.39$.

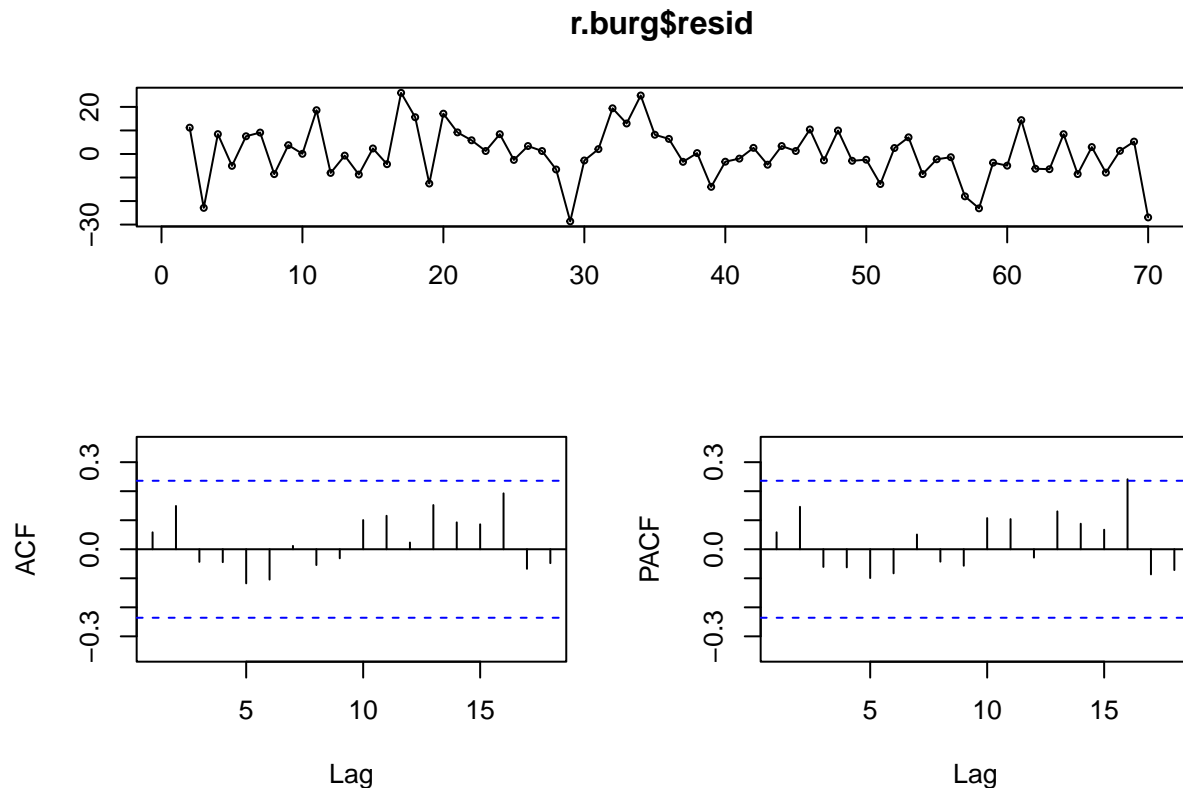
c) Use the Burg method to compute the parameters of the AR model. Check its residuals.

```
r.burg <- ar(t.yields, method = "burg", order.max = 1)
str(r.burg)
```

```
## List of 15
## $ order      : int 1
## $ ar         : num -0.407
## $ var.pred   : num 117
## $ x.mean     : num 51.1
## $ aic        : Named num [1:2] 10.7 0
## ..- attr(*, "names")= chr [1:2] "0" "1"
## $ n.used     : int 70
## $ n.obs      : int 70
## $ order.max  : num 1
## $ partialacf : num [1, 1, 1] -0.407
## $ resid      : Time-Series [1:70] from 1 to 70: NA 11.19 -22.89 8.43 -5.05 ...
## $ method     : chr "Burg"
## $ series     : chr "t.yields"
```

```
## $ frequency : num 1
## $ call      : language ar(x = t.yields, order.max = 1, method = "burg")
## $ asy.var.coef: num [1, 1] 0.0119
## - attr(*, "class")= chr "ar"

tsdisplay(r.burg$resid)
```



The autocorrelations of the residuals lie inside the confidence band. The residuals are stationary. α is estimated at $\alpha = -0.407$.

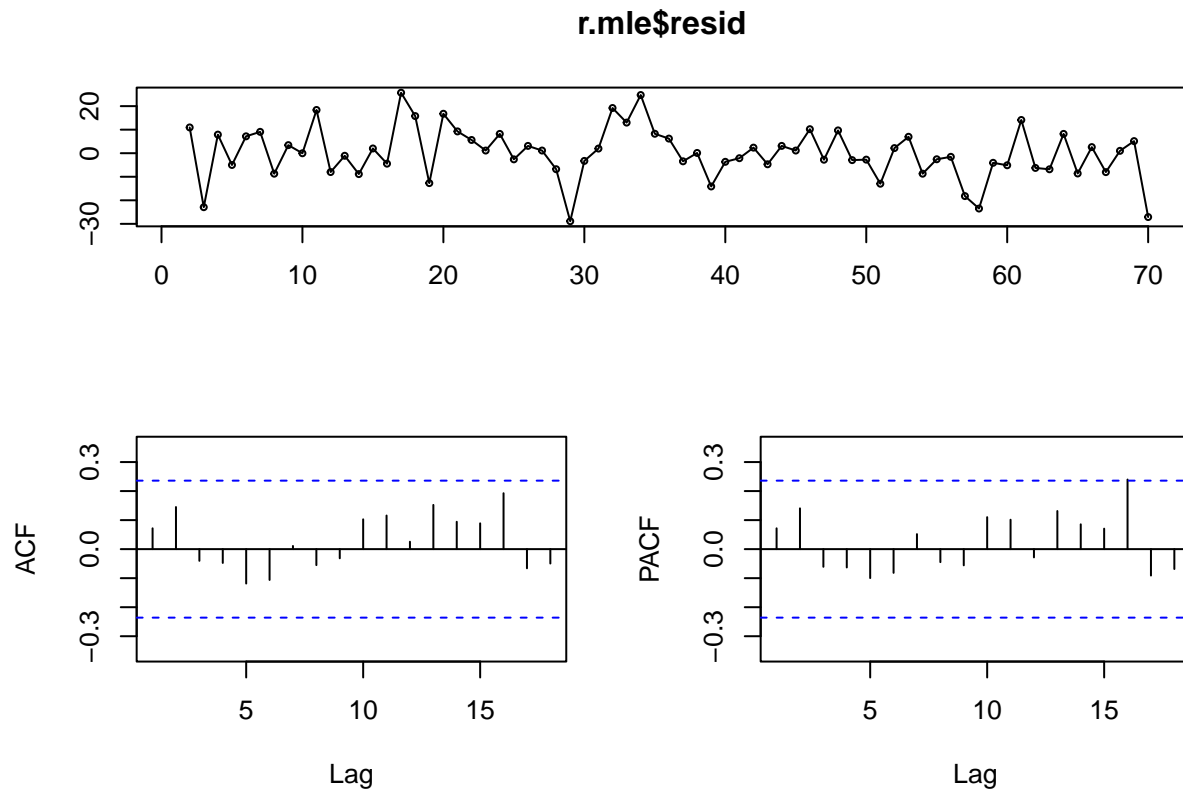
d) Use Maximum Likelihood to estimate these parameters.

```
r.mle <- ar(t.yields, method = "mle", order.max = 1)
str(r.mle)
```

```
## List of 15
## $ order      : int 1
## $ ar         : num -0.419
## $ var.pred   : num 117
## $ x.mean     : num 51.3
## $ aic        : Named num [1:2] 10.5 0
## ..- attr(*, "names")= chr [1:2] "0" "1"
## $ n.used     : int 70
## $ n.obs      : int 70
## $ order.max  : num 1
## $ partialacf : NULL
## $ resid      : Time-Series [1:70] from 1 to 70: NA 10.95 -22.93 7.89 -4.99 ...
## $ method     : chr "MLE"
```

```
## $ series      : chr "t.yields"
## $ frequency   : num 1
## $ call        : language ar(x = t.yields, order.max = 1, method = "mle")
## $ asy.var.coef: num [1, 1] 0.0119
## - attr(*, "class")= chr "ar"

tsdisplay(r.mle$resid)
```



The autocorrelations of the residuals lie inside the confidence band. The residuals are stationary. α is estimated at $\alpha = -0.419$.

Exercise 6.2

In this exercise we examine measurements of the vertical force acting on a cylinder in a water tank. A total of 320 measurements were taken at intervals of 0.15 seconds. Load the data from <http://stat.ethz.ch/Teaching/Datasets/WBL/kraft.dat> and convert them to a time series.

```
d.force <- read.table("http://stat.ethz.ch/Teaching/Datasets/WBL/kraft.dat", header = FALSE)
ts.force <- ts(d.force[, 1])
```

It is already known that at the time of the experiment, the water in the tank formed waves with (randomly changing) periods of around 2 seconds.

- Create a subset of the data containing only the first 280 observations.

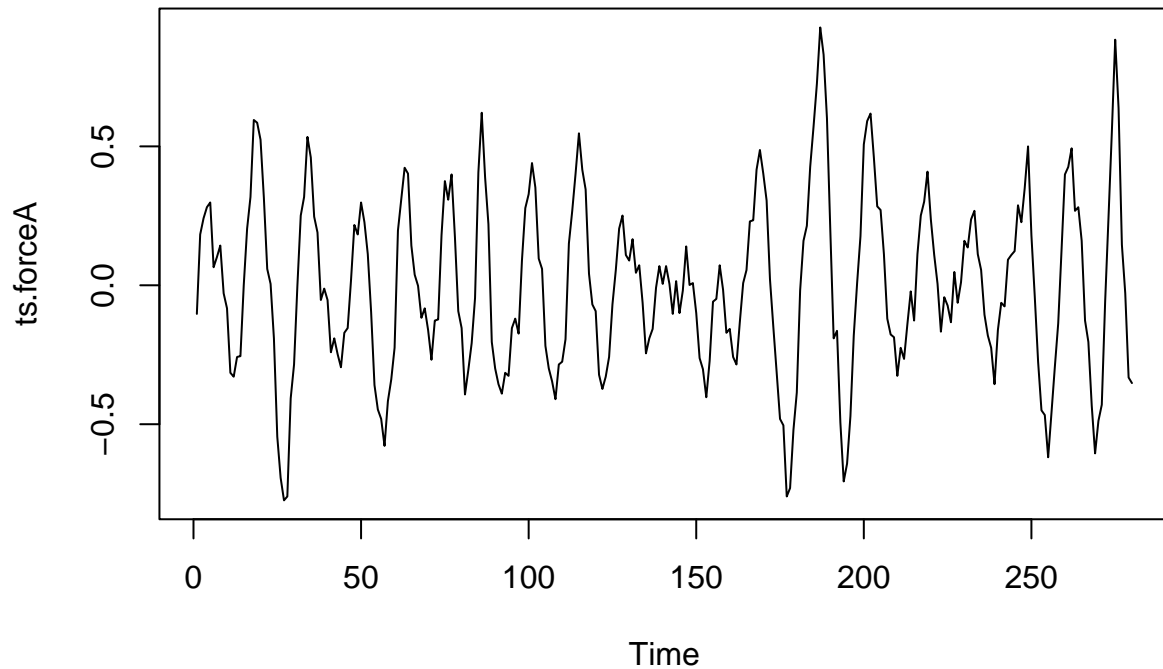
```
ts.forceA <- window(ts.force, end = 280)
```

Is a periodic behavior to be expected in these data? If so, what should the period be? Does the plot of the

times series agree with your expectations?

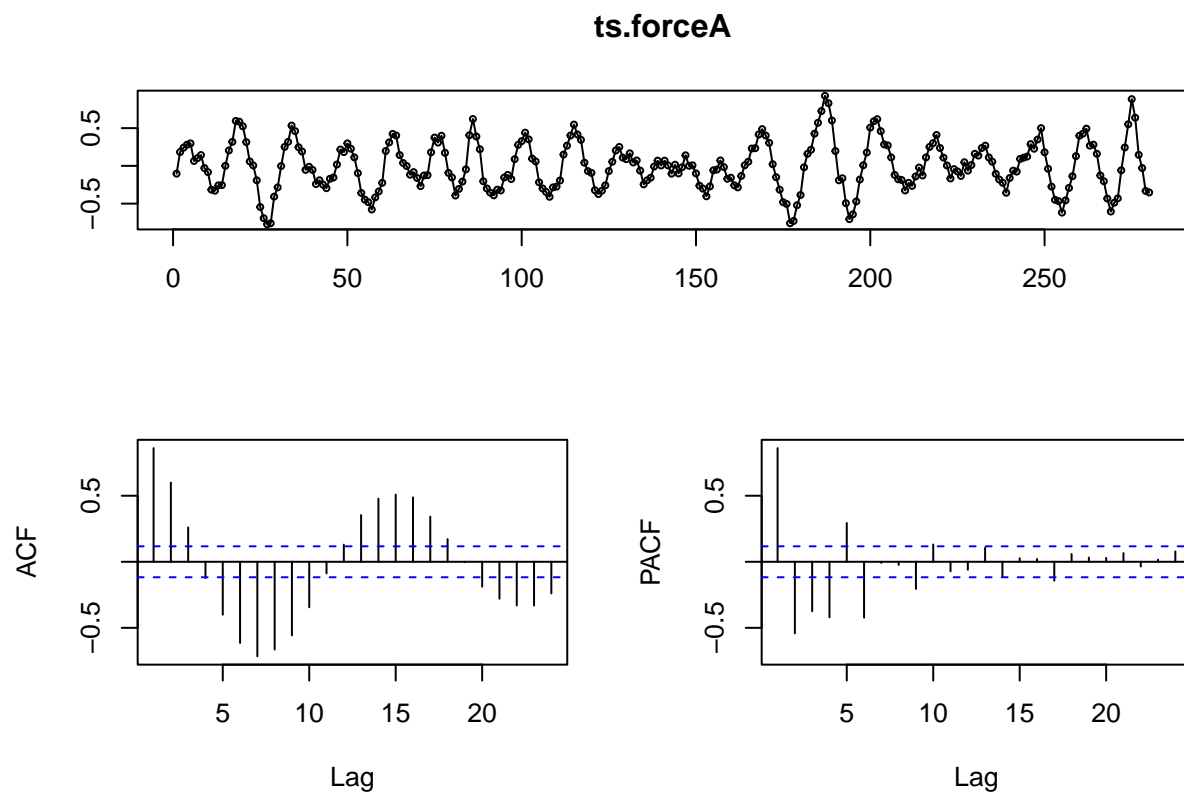
Yes, a periodic behavior is expected, as the water in the tank forms waves. It is expected that the pressure on the cylinder increases and decreases with the pattern of the waves.

```
plot(ts.forceA)
```

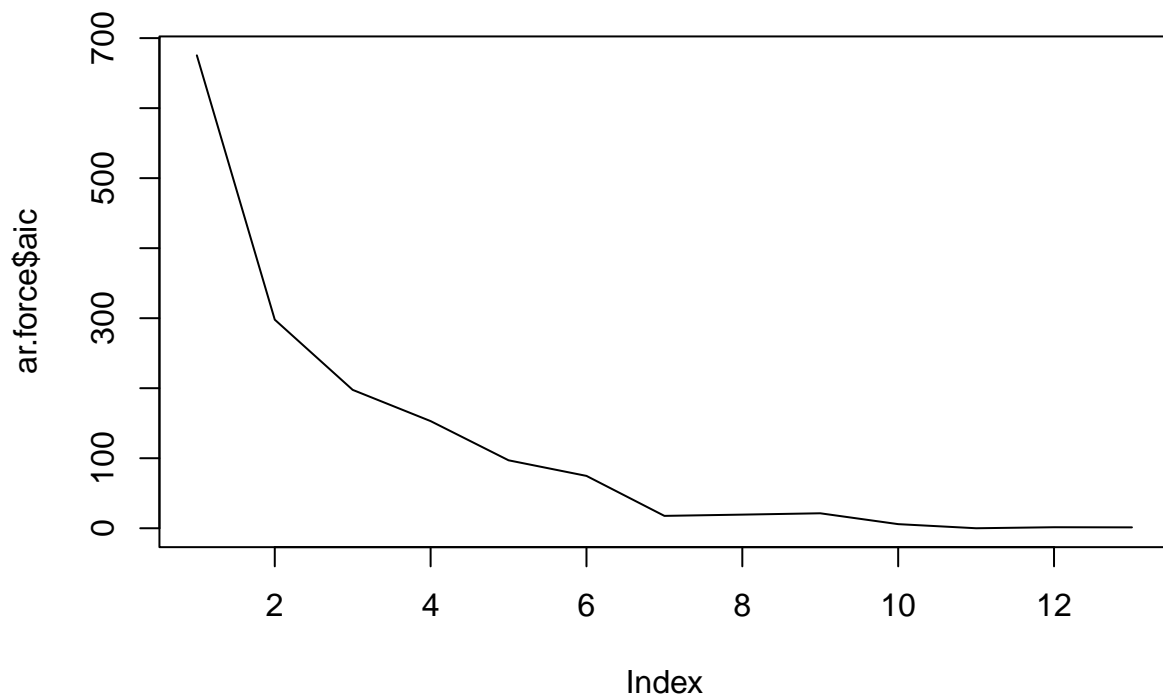


- b) Suppose you would like to fit the time series `ts.forceA` by an $AR(p)$ model. Which order p should this model have? Choose a suitable order once by looking at the partial autocorrelations, and once by using the Akaike information criterion (AIC).

```
tsdisplay(ts.forceA)
```



```
ar.force <- ar(ts.forceA, method = 'mle')  
plot(ar.force$aic, type = 'l')
```

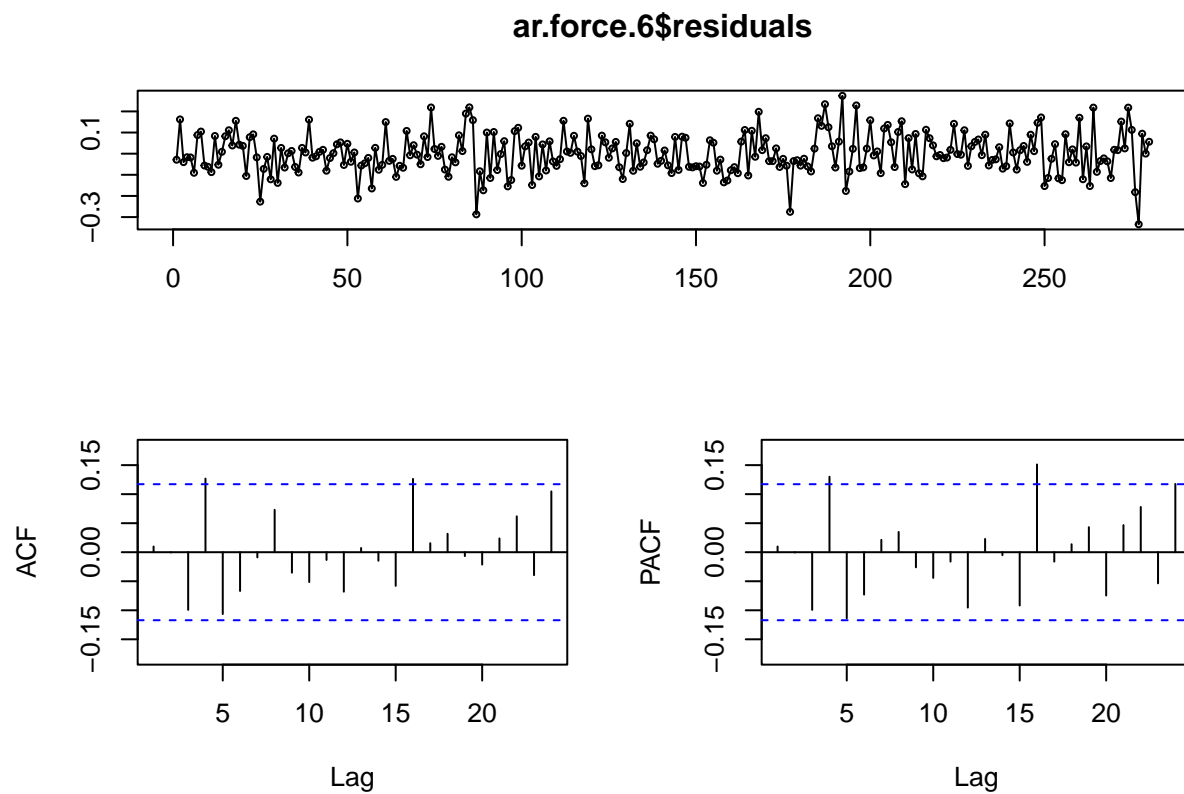


The PACF has no significant values after a lag of 6. We could assume that this is the order of the AR(p) process. In the AIC with the method MLE we observe the lowest value at lag 7. When fitting a model in step c) we will therefore compare the models with $p = 6$ and $p = 7$ and then decide which one has the better fit.

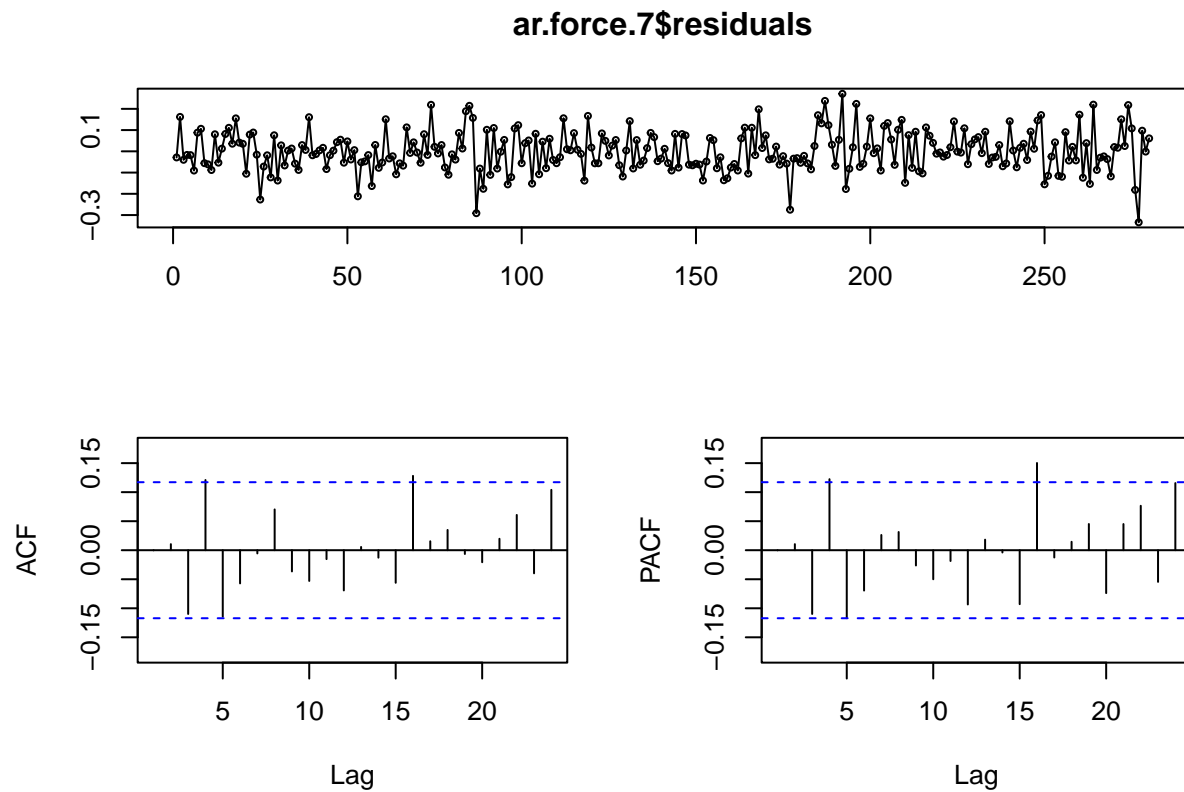
- c) Fit an AR(p) model using maximum likelihood for the time series ts.forceA, where p is the order specified in Part b). Analyze the residuals. Is the model appropriate for this time series?

```
ar.force.6 <- arima(ts.forceA, order = c(6, 0, 0), method = 'ML')
ar.force.7 <- arima(ts.forceA, order = c(7, 0, 0), method = 'ML')

tsdisplay(ar.force.6$residuals)
```

```
tsdisplay(ar.force.7$residuals)
```

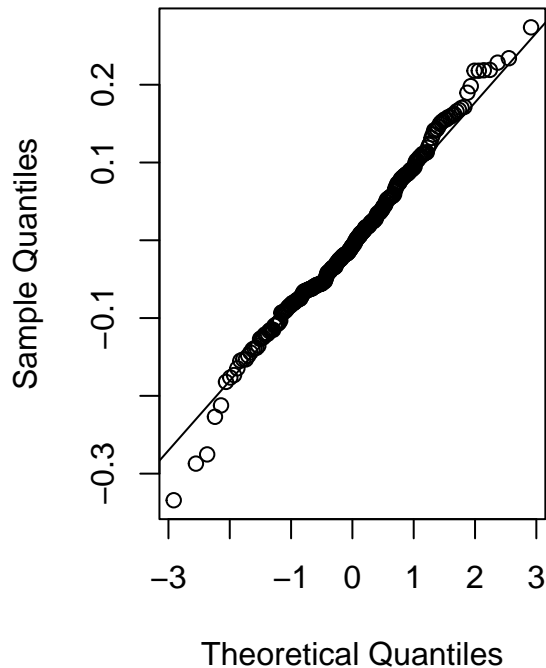


Just by comparing the residuals of the models fit with order 6 vs order 7, we cannot clearly conclude which model has the more optimal fit.

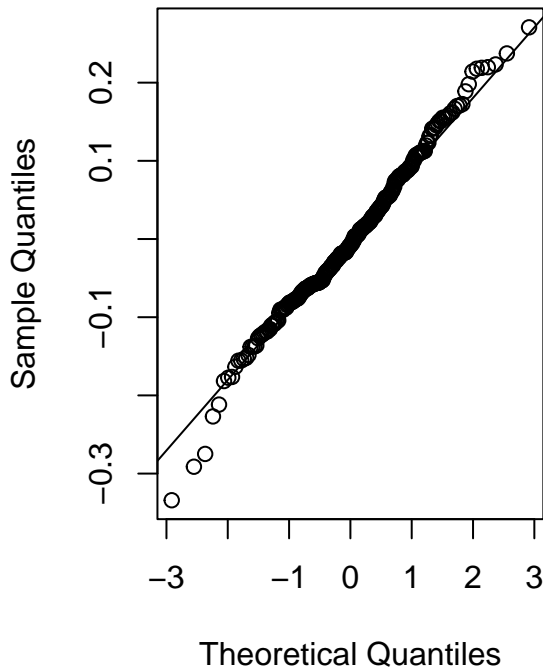
```
par(mfrow=c(1,2))
qqnorm(ar.force.6$residuals, main='Normal Q-Q Plot: AR(6)')
qqline(ar.force.6$residuals)

qqnorm(ar.force.7$residuals, main='Normal Q-Q Plot: AR(7)')
qqline(ar.force.7$residuals)
```

Normal Q–Q Plot: AR(6)



Normal Q–Q Plot: AR(7)

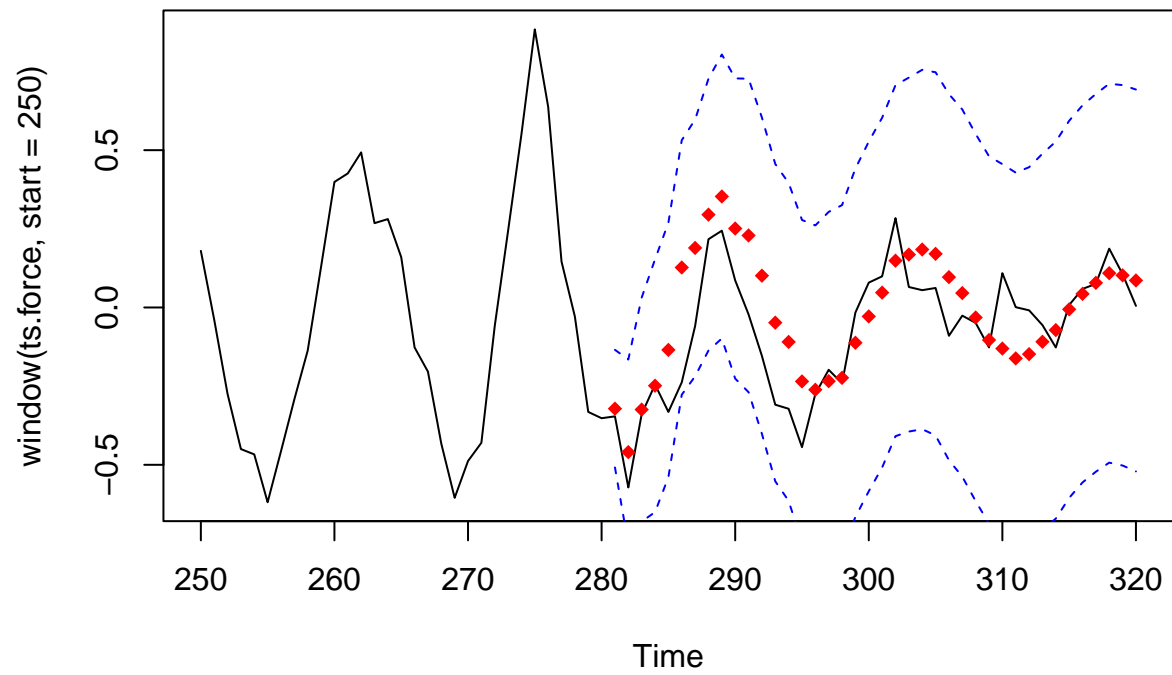


The model of order 7 does appear to fit a little better as the dots in the upper right corner of the Q-Q plot align a little better with the line. However, the difference between order 6 and 7 is miniscule. Going forward we will now use the model with order 7.

- d) Use the model fitted in Part c) to compute point predictions and prediction intervals for the next 40 measurements. Compare these graphically to the actual measurements.

Then, plot the point predictions and the confidence intervals into the plot.

```
force.pred <- predict(ar.force.7, n.ahead = 40)
plot(window(ts.force, start = 250))
lines(force.pred$pred, col = 'red', type = 'p', pch = 18)
lines(force.pred$pred + 1.96 * force.pred$se, col = 'blue', lty = 2)
lines(force.pred$pred - 1.96 * force.pred$se, col = 'blue', lty = 2)
```



The predicted values are relatively close to the actual measurements. Depending on the actual use-case, this could be considered a usable fit. It could be however argued, that the confidence interval is quite wide and predictions are to be interpreted with caution.