Time Series HW2

Ze YANG

September 15, 2018

Problem 1

(a)

```
# Load data
library(forecast)
library(Ecdat)
```

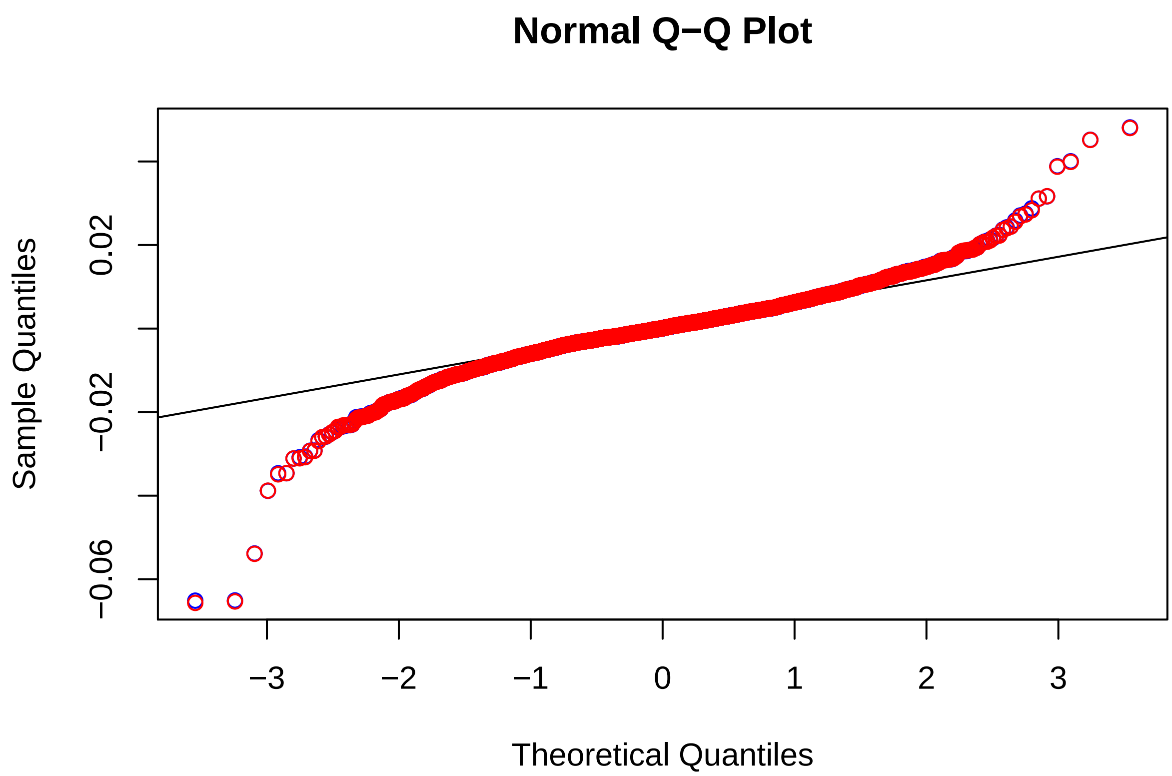
```
## Loading required package: Ecfun

##
## Attaching package: 'Ecfun'

## The following object is masked from 'package:forecast':
##
##     BoxCox

## The following object is masked from 'package:base':
##
##     sign

##
## Attaching package: 'Ecdat'

## The following object is masked from 'package:datasets':
##
##     Orange
```

```
data(CRSPday)
crsp = CRSPday[,7]
```

```
# Model fitting
ar1 = arima(crsp, order=c(1,0,0))
ar2 = arima(crsp, order=c(2,0,0))
```

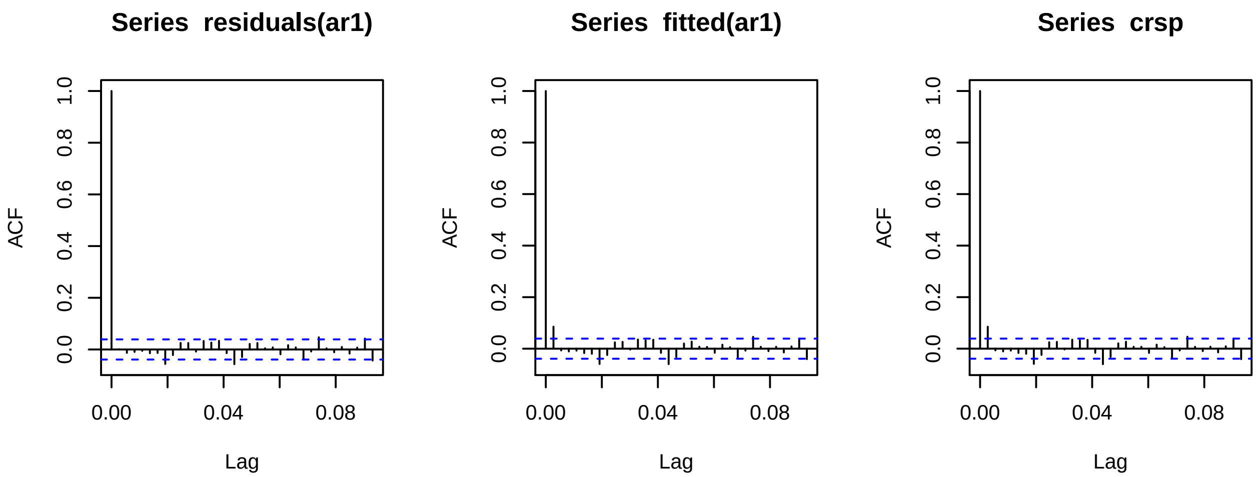
The normal probability plot of residuals

```
# model assessments
qqnorm(residuals(ar1), col='blue')
secondpts = qqnorm(residuals(ar2), plot.it=FALSE)
qqline(residuals(ar1))
points(secondpts, col="red")
```

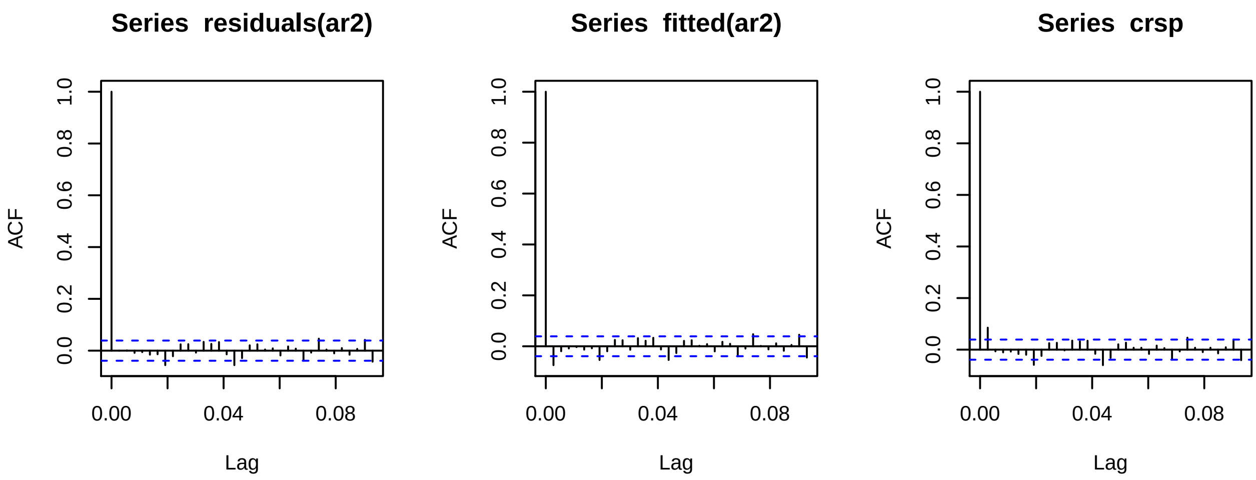


ACF of residuals, fitted values, and original sample

```
par(mfrow=c(1,3))
acf(residuals(ar1))
acf(fitted(ar1))
acf(crsp)
```




```
par(mfrow=c(1,3))
acf(residuals(ar2))
acf(fitted(ar2))
acf(crsp)
```



Ljung-Box test on residuals

```
Box.test(ar1$residuals, fitdf=1, lag=10, type="Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  ar1$residuals
## X-squared = 14.163, df = 9, p-value = 0.1166
```

```
Box.test(ar2$residuals, fitdf=2, lag=10, type="Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  ar2$residuals
## X-squared = 13.607, df = 8, p-value = 0.09259
```

AICs

```
c(ar1$aic, ar2$aic)
```

```
## [1] -17406.37 -17404.87
```

Conclusions

- The normal probability plots of residuals of both models overlay closely. For both of them, the residuals have heavier tails than normal on two sides.
- The ACF of residuals displays no strong evidence of autocorrelation at any lag for both $AR(1)$ and $AR(2)$.
- The sample ACF display strong correlation only at lag $= 1$, no strong evidence of correlation beyond. The ACF of fitted values of $AR(1)$ displays a similar pattern. However, the ACF of fitted values of $AR(2)$ has a correlation of opposite (negative) direction at Lag=1.
- The p-value of Ljung-Box tests for the residuals of both models are above 0.05, indicating that we fail to reject $\{\rho_\epsilon(h) = 0$ for some $h \geq 1\}$ at a confidence level of 0.05: There is no strong evidence of serial

correlation in residuals up to lag 10. But if we increase the number of lags to 20, the null hypothesis can be rejected. Therefore, there might still be long lagged correlations in the residuals.
- The AIC of AR(1) model is -17406.37, and the AIC of AR(2) model is -17404.87. The level of validity of model assumptions is acceptable, so we choose the AR(1) model based on the minimum AIC.

Problem 2

```
auto.arima(crsp, max.p=20, max.q=0, ic='aic', approximation=T)
```

```
## Series: crsp
## ARIMA(2,0,0) with non-zero mean
##
## Coefficients:
##          ar1      ar2   mean
##       0.0865  -0.0141  7e-04
## s.e.  0.0199   0.0199  2e-04
##
## sigma^2 estimated as 5.979e-05:  log likelihood=8706.43
## AIC=-17404.87   AICc=-17404.85   BIC=-17381.53
```

```
auto.arima(crsp, max.p=20, max.q=0, ic='aic', approximation=F)
```

```
## Series: crsp
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##          ar1   mean
##       0.0853  7e-04
## s.e.  0.0198  2e-04
##
## sigma^2 estimated as 5.978e-05:  log likelihood=8706.18
## AIC=-17406.37   AICc=-17406.36   BIC=-17388.86
```

Comments

- With `approximation = F`, the procedure chooses AR(1), while with `approximation = T` the procedure will choose AR(2): the results are different.