1. Load libraries

```
library(tseries)
library(fBasics)
```

```
## Loading required package: timeDate
## Loading required package: timeSeries
##
## Rmetrics Package fBasics
## Analysing Markets and calculating Basic Statistics
## Copyright (C) 2005-2014 Rmetrics Association Zurich
## Educational Software for Financial Engineering and Computational Science
## Rmetrics is free software and comes with ABSOLUTELY NO WARRANTY.
## https://www.rmetrics.org --- Mail to: info@rmetrics.org
```

```
library(forecast)
library(lmtest)
```

```
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following object is masked from 'package:timeSeries':
##
##     time<-
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

2. Import Data

```
setwd("~/Desktop/CSC425/week6/casestudy3.1")
da=read.table("w-petroprice.txt",header=T)
da1=read.table("w-gasoline.txt", header=T)
head(da,10)
```

```
##    Mon Day Year World    US
## 1    1   3 1997 23.18 22.90
## 2    1  10 1997 23.84 23.56
## 3    1  17 1997 22.99 22.79
## 4    1  24 1997 22.05 21.83
## 5    1  31 1997 21.87 21.16
## 6    2   7 1997 21.56 20.97
## 7    2  14 1997 20.25 19.93
## 8    2  21 1997 19.78 19.38
## 9    2  28 1997 18.72 18.05
## 10   3   7 1997 18.54 17.74
```

```r
tail(da,10)
```

```
##     Mon Day Year World    US
## 708   7  23 2010 73.80 72.01
## 709   7  30 2010 74.69 73.15
## 710   8   6 2010 78.68 76.75
## 711   8  13 2010 77.28 75.90
## 712   8  20 2010 73.81 72.28
## 713   8  27 2010 71.28 69.39
## 714   9   3 2010 73.45 71.02
## 715   9  10 2010 74.42 71.72
## 716   9  17 2010 76.11 72.97
## 717   9  24 2010 75.14 71.11
```

```r
head(da1,10)
```

```
##      gas
## 1  1.225
## 2  1.241
## 3  1.241
## 4  1.238
## 5  1.236
## 6  1.234
## 7  1.227
## 8  1.223
## 9  1.215
## 10 1.206
```

```r
tail(da1,10)
```

```
##       gas
## 708 2.749
## 709 2.735
## 710 2.783
## 711 2.745
## 712 2.704
## 713 2.682
## 714 2.682
## 715 2.721
## 716 2.723
## 717 2.694
```

3. log transformation

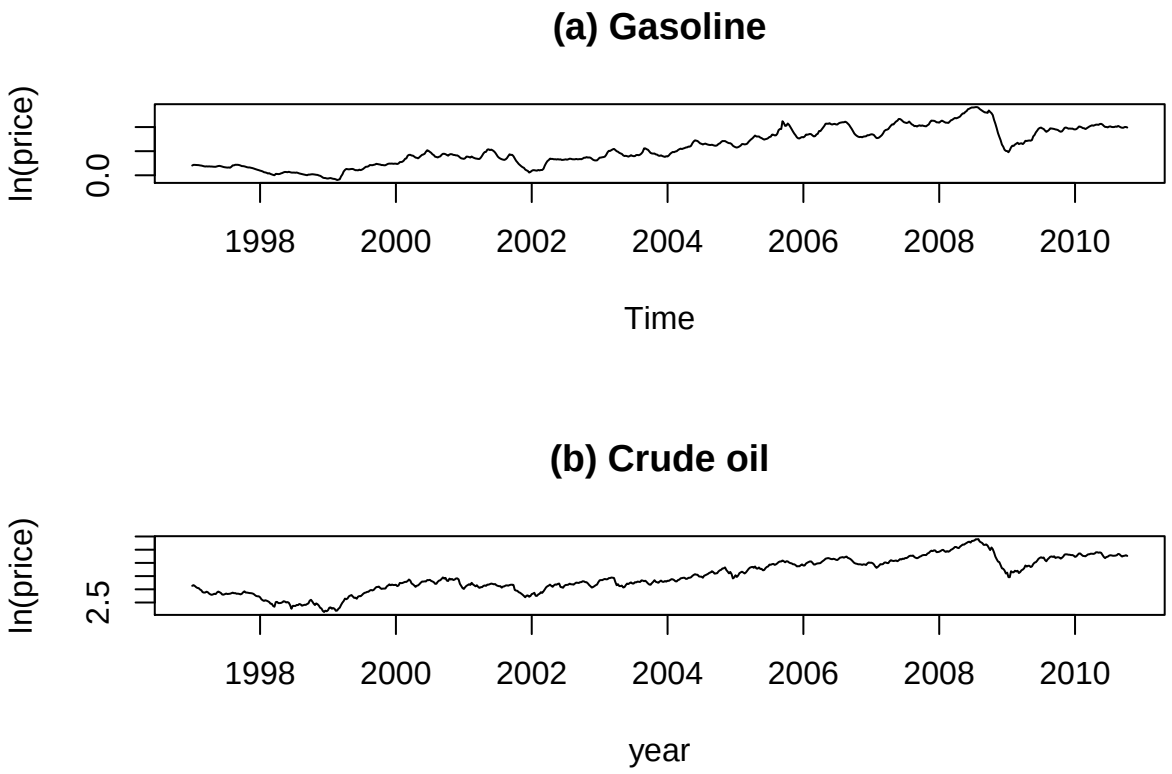
```r
#### log gasoline
pgs = log(da1[,1])
#### log crude oil
pus = log(da$US)
```

4. time series object

```
pgs.ts=ts(pgs,start=c(1997,1), frequency=52)
pus.ts=ts(pus,start=c(1997,1), frequency=52)
```

5.create time plots for gasoline and crude oil

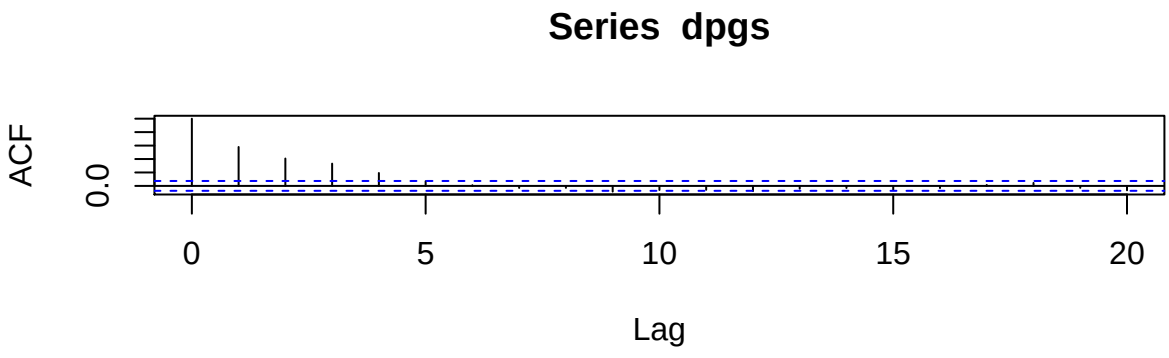
```
par(mfcol = c(2,1))
plot(pgs.ts, ylab = 'ln(price)',type = 'l')
title(main = '(a) Gasoline')

plot(pus.ts, xlab = 'year', ylab = 'ln(price)', type = 'l')
title(main = '(b) Crude oil')
```

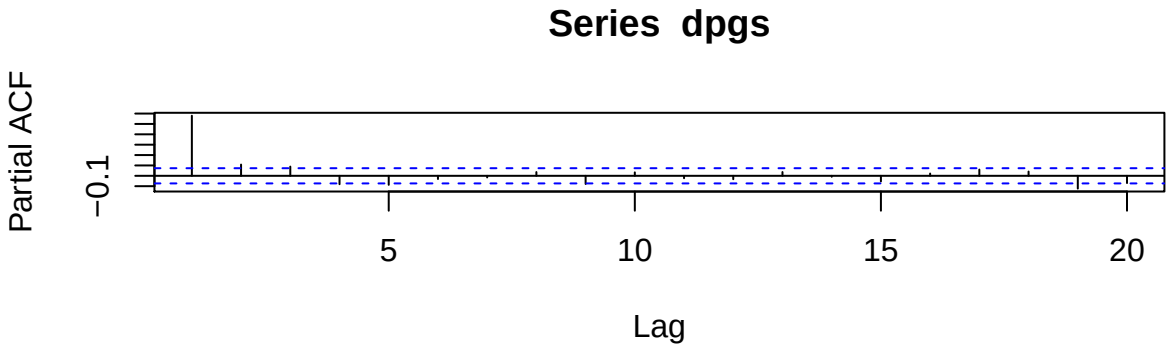



```
par(mfcol = c(1,1))
```

6. compute first difference

```
dpgs = diff(pgs)
par(mfcol = c(2,1))
acf(dpgs, lag = 20)
pacf(dpgs,lag = 20)
```





7. find best AR model for log(return) on gasoline

```
m1 = ar(diff(pgs), method = 'mle')
m1$order
```

```
## [1] 5
```

8. fit an AR(5) model

```
m1 = arima(dpgs, order = c(5,0,0), include.mean = F)
m1
```

```
##
## Call:
## arima(x = dpgs, order = c(5, 0, 0), include.mean = F)
##
## Coefficients:
##          ar1     ar2     ar3      ar4      ar5
##       0.5073  0.0788  0.1355  -0.0360  -0.0862
## s.e.  0.0372  0.0417  0.0415   0.0417   0.0372
##
## sigma^2 estimated as 0.0003262:  log likelihood = 1857.85,  aic = -3703.71
```

9. remove non-significant parameter at lag 4

```
m1 = arima(dpgs, order = c(5,0,0), include.mean = F, fixed = c(NA,NA,NA,0,NA))
```

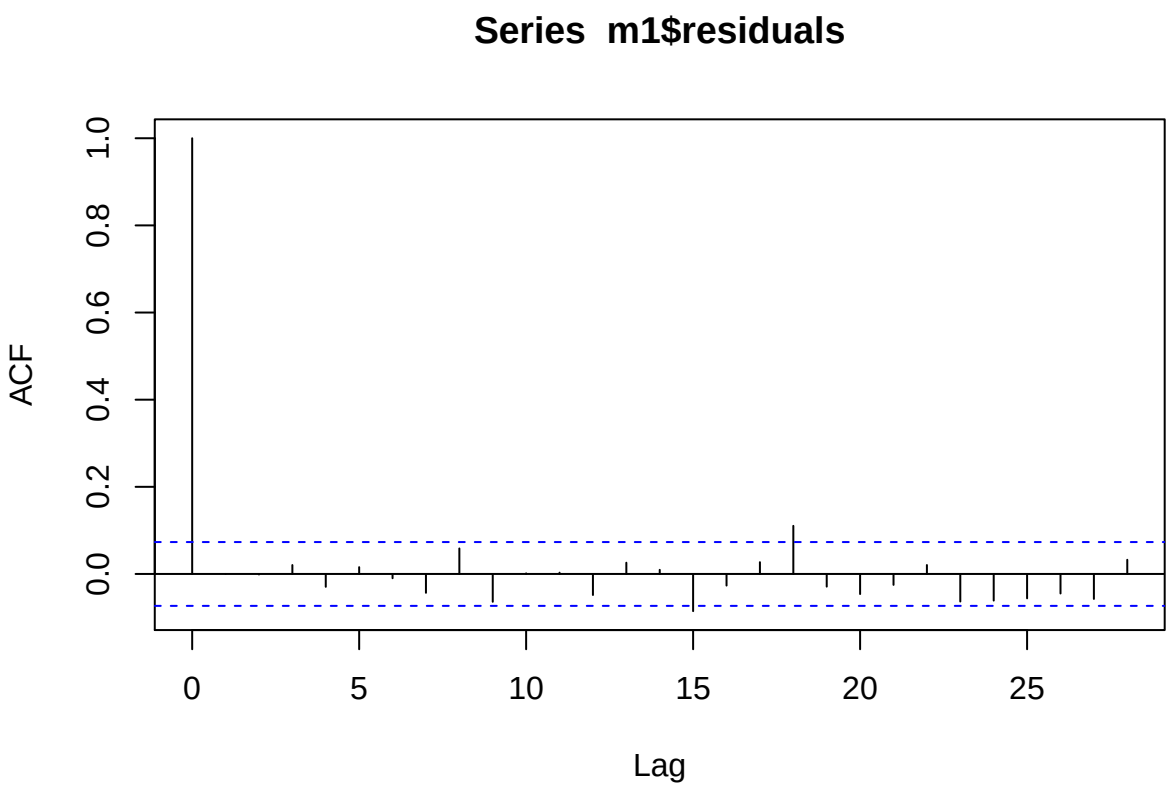
```
## Warning in arima(dpgs, order = c(5, 0, 0), include.mean = F, fixed =
## c(NA, : some AR parameters were fixed: setting transform.pars = FALSE
```

```
m1
```

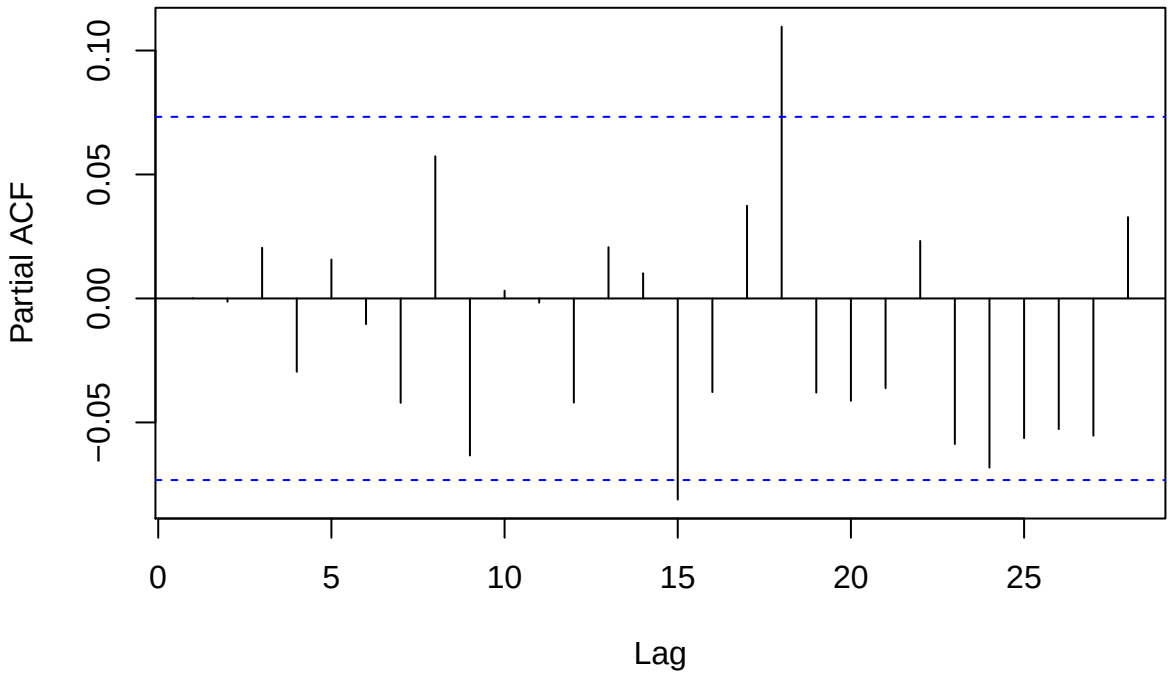
```
##
## Call:
## arima(x = dpgs, order = c(5, 0, 0), include.mean = F, fixed = c(NA, NA, NA,
##     0, NA))
##
## Coefficients:
##          ar1     ar2     ar3  ar4      ar5
##       0.5036  0.0789  0.1220    0  -0.1009
## s.e.  0.0370  0.0418  0.0385    0   0.0330
##
## sigma^2 estimated as 0.0003265:  log likelihood = 1857.48,  aic = -3704.96
```

10. analysis of residuals

```
acf(m1$residuals)
```




```
pacf(m1$residuals)
```




```
Box.test(m1$residuals, lag = 6, type = "Ljung-Box", fitdf = 4)
```

```
##
##  Box-Ljung test
##
## data:  m1$residuals
## X-squared = 1.1787, df = 2, p-value = 0.5547
```

```
Box.test(m1$residuals, lag = 12, type = "Ljung-Box", fitdf = 4)
```

```
##
##  Box-Ljung test
##
## data:  m1$residuals
## X-squared = 9.7116, df = 8, p-value = 0.2859
```

11. fit an Arima(1,3) model.

```
m2 = arima(dpgs, order = c(1,0,3), include.mean = F)
m2
```

```
##
## Call:
## arima(x = dpgs, order = c(1, 0, 3), include.mean = F)
##
## Coefficients:
##          ar1      ma1     ma2     ma3
##       0.5917  -0.0919  0.0413  0.1547
## s.e.  0.0750   0.0774  0.0489  0.0435
```

```
##
## sigma^2 estimated as 0.0003273:  log likelihood = 1856.65,  aic = -3703.29
```

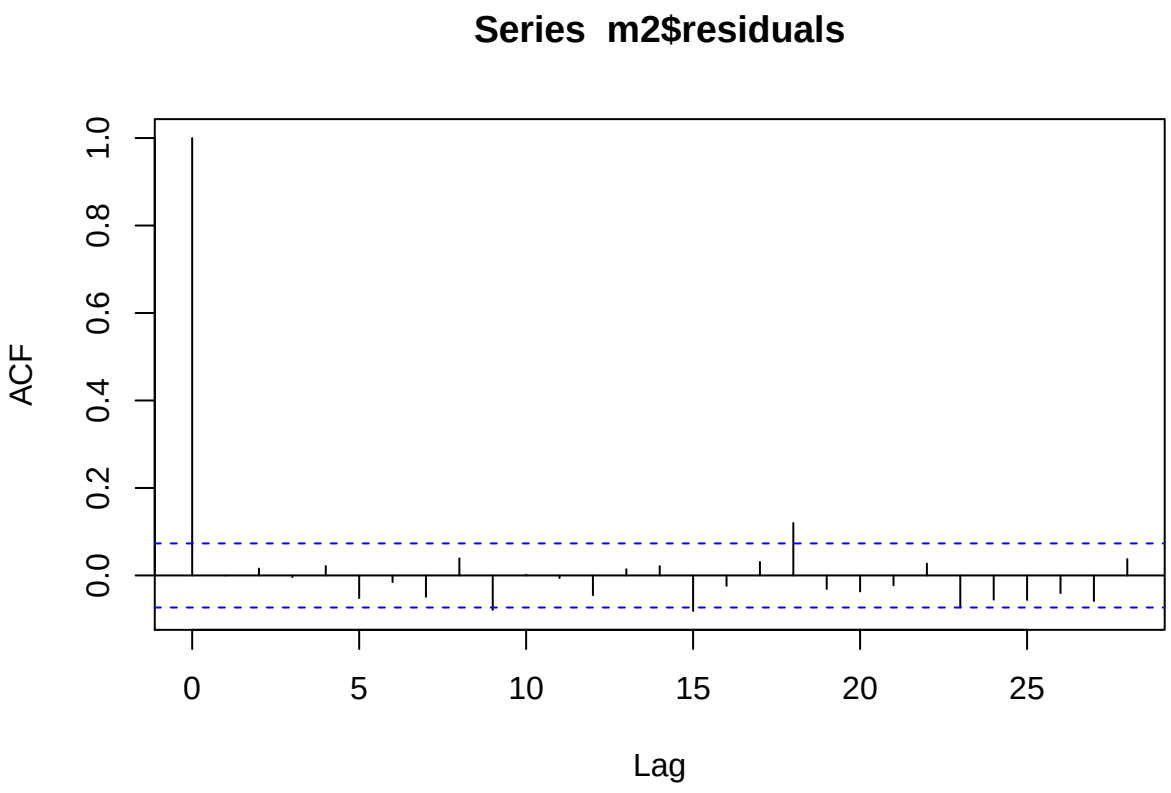
12. remove non-significant MA parameter at lag 2

```
m2 = arima(dpgs, order = c(1,0,3), include.mean = F, fixed = c(NA,NA,0,NA))
m2
```

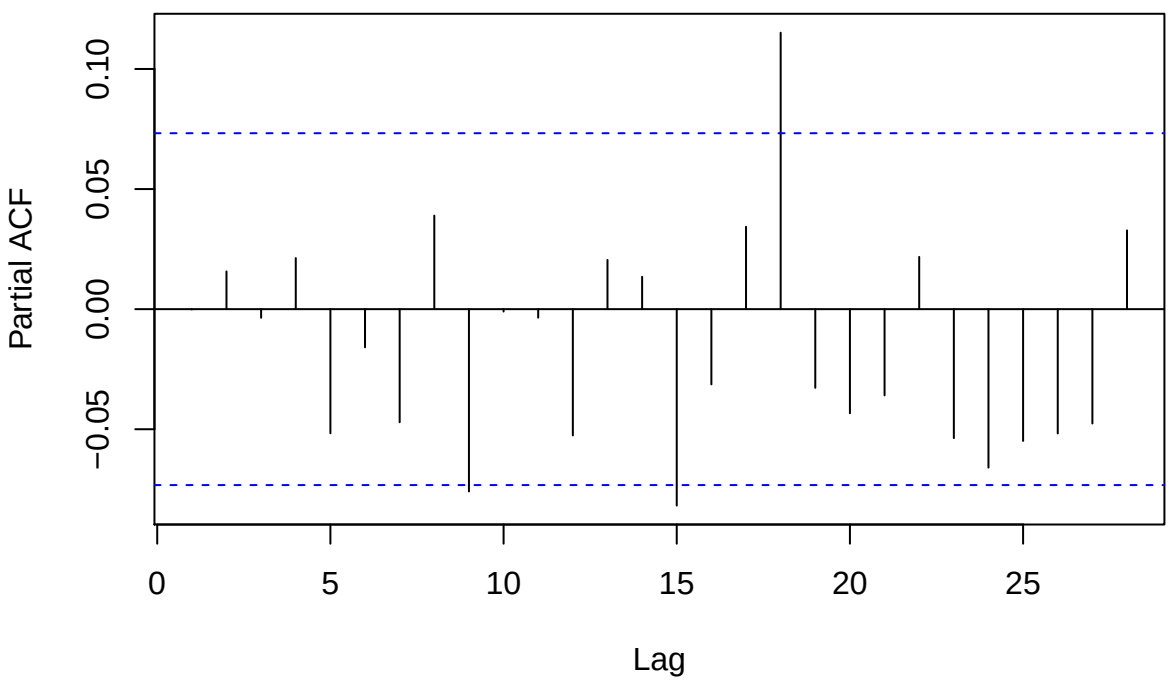
```
##
## Call:
## arima(x = dpgs, order = c(1, 0, 3), include.mean = F, fixed = c(NA, NA, 0, NA))
##
## Coefficients:
##          ar1      ma1  ma2     ma3
##       0.6332  -0.1266    0  0.1411
## s.e.  0.0507   0.0595    0  0.0408
##
## sigma^2 estimated as 0.0003276:  log likelihood = 1856.3,  aic = -3704.6
```

13. Analysis of residuals

```
acf(m2$residuals)
```




```
pacf(m2$residuals)
```




```
Box.test(m2$residuals, lag = 6, type = "Ljung-Box", fitdf = 4)
```

```
##
##  Box-Ljung test
##
## data:  m2$residuals
## X-squared = 2.6266, df = 2, p-value = 0.2689
```

```
Box.test(m2$residuals, lag = 12, type = "Ljung-Box", fitdf = 4)
```

```
##
##  Box-Ljung test
##
## data:  m2$residuals
## X-squared = 11.477, df = 8, p-value = 0.1761
```

14. Fit simple regression

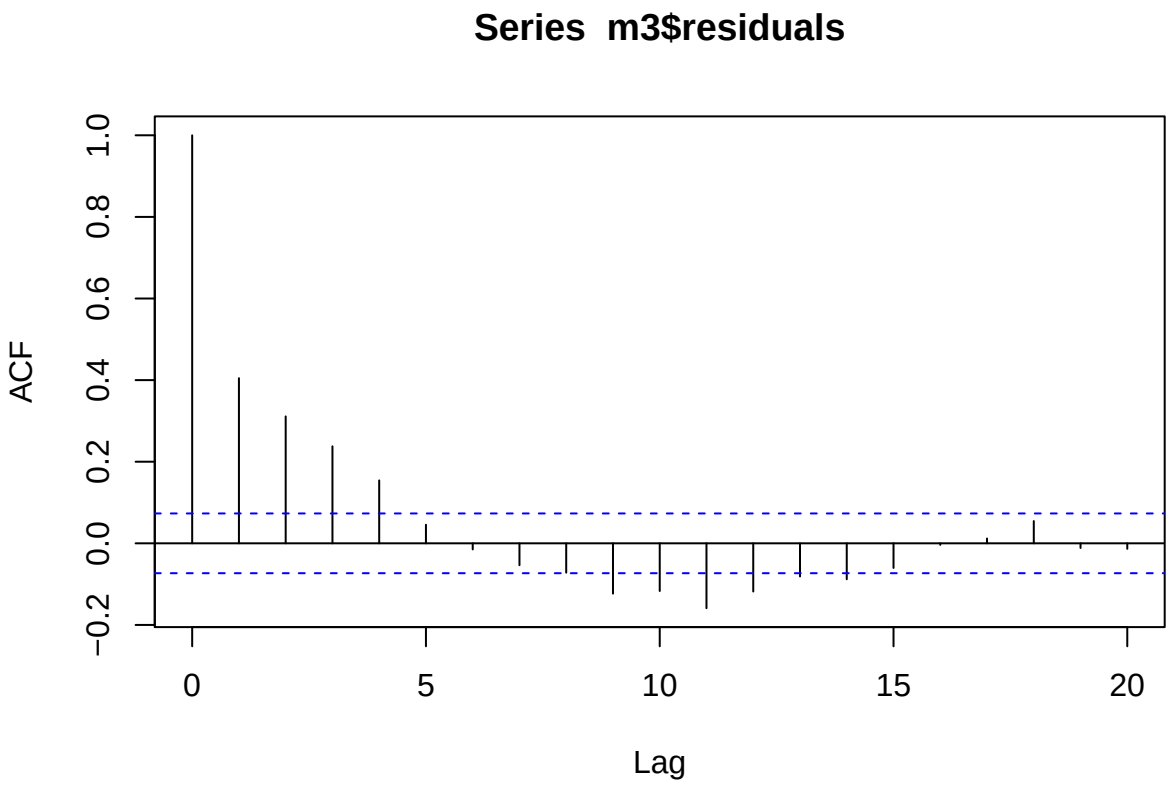
```
dpus = diff(pus)
m3 = lm(dpgs~1+dpus)
summary(m3)
```

```
##
## Call:
## lm(formula = dpgs ~ 1 + dpus)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.076840 -0.009456 -0.000279  0.008804  0.150721
```

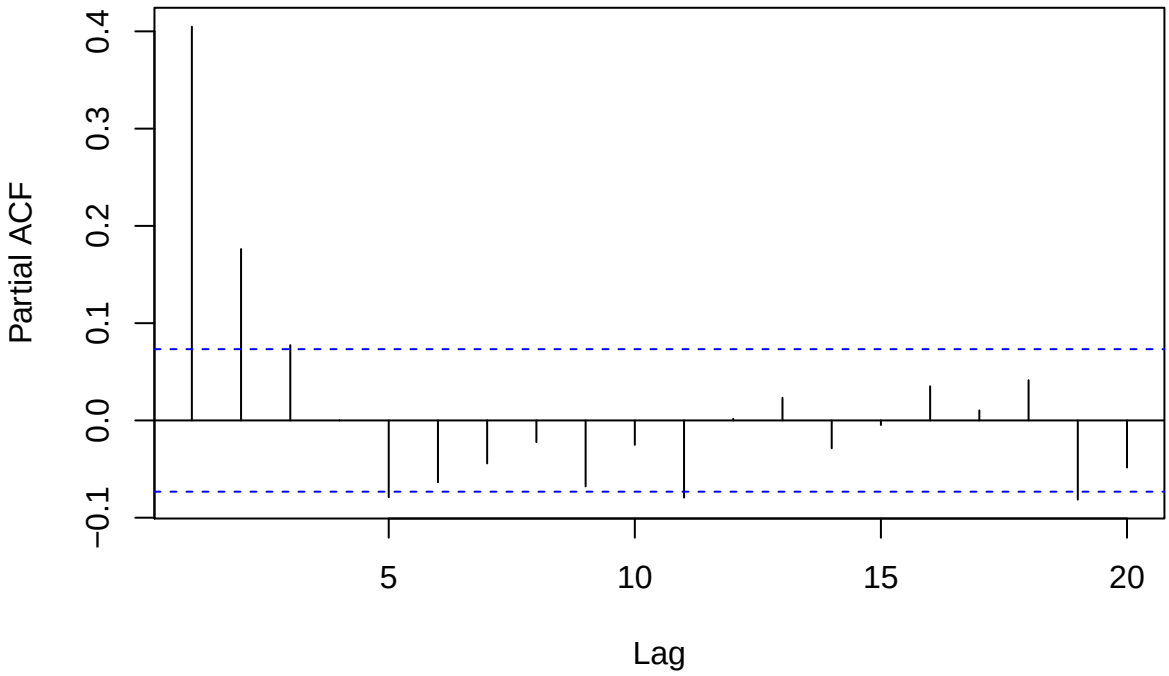
```
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0006472  0.0006876   0.941    0.347
## dpus        0.2865347  0.0150791  19.002   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01839 on 714 degrees of freedom
## Multiple R-squared:  0.3359, Adjusted R-squared:  0.3349
## F-statistic: 361.1 on 1 and 714 DF,  p-value: < 2.2e-16
```

15. residual analysis shows residuals are serially correlated

```
acf(m3$residuals, lag = 20)
```




```
pacf(m3$residuals, lag = 20)
```



16. Find AR model for residuals to identify order of AR(p) model

```
m4 = ar(m3$residuals, method = 'mle')
m4$order
```

```
## [1] 6
```

17. fit fegression model with time series errors.

```
# xreg defines the list of regressor to be included in the arima model
m4 = arima(dpgs, order = c(6,0,0), include.mean = F, xreg = c(dpus))
m4
```

```
##
## Call:
## arima(x = dpgs, order = c(6, 0, 0), xreg = c(dpus), include.mean = F)
##
## Coefficients:
##          ar1     ar2     ar3     ar4      ar5      ar6  c(dpus)
##       0.3953  0.1634  0.0946  0.0297  -0.0873  -0.0525   0.1927
## s.e.  0.0389  0.0400  0.0404  0.0405   0.0400   0.0373   0.0136
##
## sigma^2 estimated as 0.0002524:  log likelihood = 1949.61,  aic = -3883.21
```

```
m4 = arima(dpgs, order = c(5,0,0), include.mean = F, xreg = c(dpus))
m4
```

```
##
```

```
## Call:
## arima(x = dpgs, order = c(5, 0, 0), xreg = c(dpus), include.mean = F)
##
## Coefficients:
##          ar1     ar2     ar3     ar4      ar5  c(dpus)
##       0.4022  0.1621  0.0899  0.0209  -0.1086   0.1914
## s.e.  0.0387  0.0401  0.0403  0.0400   0.0371   0.0136
##
## sigma^2 estimated as 0.0002531:  log likelihood = 1948.62,  aic = -3883.23
```

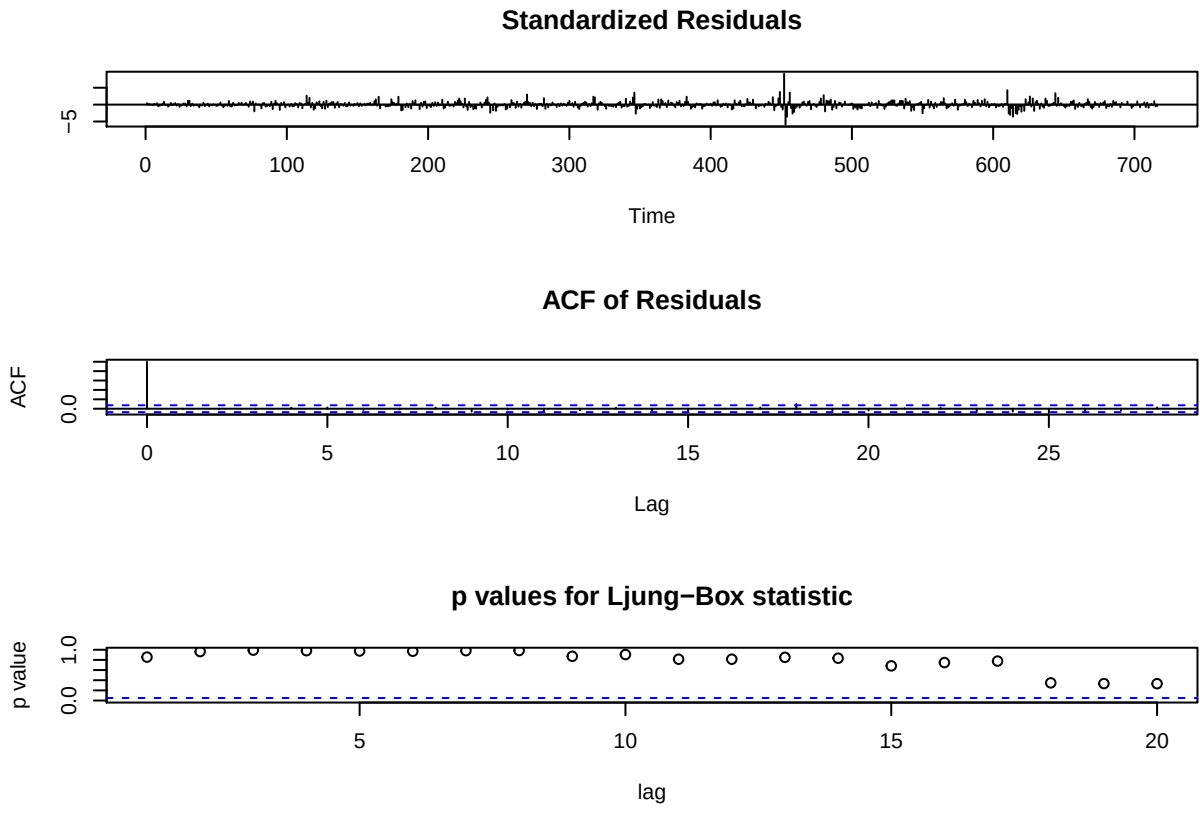
```
m4 = arima(dpgs, order = c(5,0,0), include.mean = F, xreg = c(dpus), fixed = c(NA,NA,NA,0,NA,NA))
```

```
## Warning in arima(dpgs, order = c(5, 0, 0), include.mean = F, xreg =
## c(dpus), : some AR parameters were fixed: setting transform.pars = FALSE
```

```
m4
```

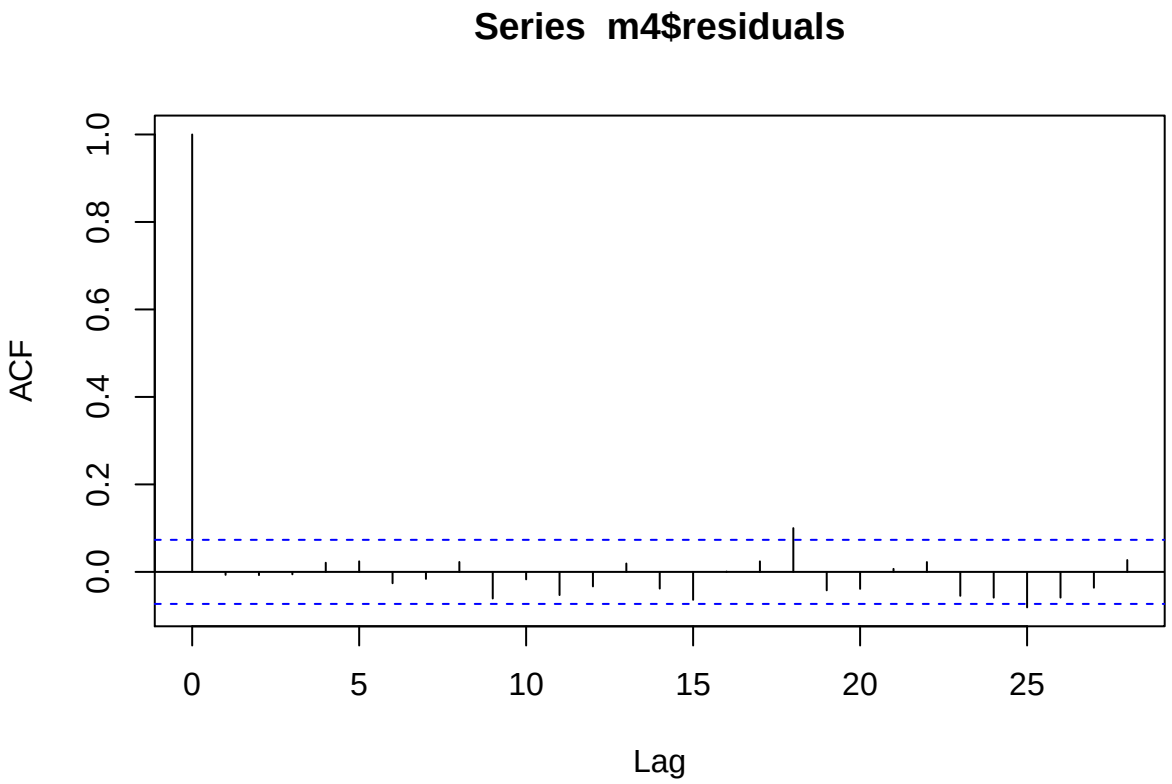
```
##
## Call:
## arima(x = dpgs, order = c(5, 0, 0), xreg = c(dpus), include.mean = F, fixed = c(NA,
##     NA, NA, 0, NA, NA))
##
## Coefficients:
##          ar1     ar2     ar3  ar4      ar5  c(dpus)
##       0.4037  0.1642  0.0961    0  -0.1014   0.1911
## s.e.  0.0386  0.0399  0.0386    0   0.0345   0.0136
##
## sigma^2 estimated as 0.0002532:  log likelihood = 1948.48,  aic = -3884.95
```

```
tsdiag(m4, gof = 20)
```



18. analysis of residuals

```
acf(m4$residuals)
```




```
Box.test(m4$residuals, lag = 6, type = "Ljung-Box", fitdf = 4)
```

```
##
##  Box-Ljung test
##
## data:  m4$residuals
## X-squared = 1.3176, df = 2, p-value = 0.5175
```

```
Box.test(m4$residuals, lag = 12, type = "Ljung-Box", fitdf = 4)
```

```
##
##  Box-Ljung test
##
## data:  m4$residuals
## X-squared = 7.6239, df = 8, p-value = 0.471
```

19. model regressing gasoline prce vs lagged crude oil

```
m5 = arima(dpgs[2:716], order = c(9,0,0), include.mean = F, xreg = dpus[1:715])
m5
```

```
##
## Call:
## arima(x = dpgs[2:716], order = c(9, 0, 0), xreg = dpus[1:715], include.mean = F)
##
```

```
## Coefficients:
##          ar1     ar2     ar3      ar4      ar5      ar6      ar7     ar8
##       0.4559  0.0888  0.1679  -0.0468  -0.0653  -0.0195  -0.0362  0.0797
## s.e.  0.0425  0.0410  0.0423   0.0415   0.0416   0.0414   0.0410  0.0408
##           ar9  dpus[1:715]
##       -0.0882       0.0454
## s.e.   0.0373       0.0174
##
## sigma^2 estimated as 0.0003204:  log likelihood = 1861.55,  aic = -3701.1
```

```
m5 = arima(dpgs[2:716], order = c(9,0,0), include.mean = F, xreg = dpus[1:715], fixed = c(NA,NA,NA,0,NA
```

```
## Warning in arima(dpgs[2:716], order = c(9, 0, 0), include.mean = F, xreg
## = dpus[1:715], : some AR parameters were fixed: setting transform.pars =
## FALSE
```

```
m5
```

```
##
## Call:
## arima(x = dpgs[2:716], order = c(9, 0, 0), xreg = dpus[1:715], include.mean = F,
##     fixed = c(NA, NA, NA, 0, NA, 0, 0, 0, NA, NA))
##
## Coefficients:
##          ar1     ar2     ar3  ar4      ar5  ar6  ar7  ar8      ar9
##       0.4544  0.0877  0.1415    0  -0.0830    0    0    0  -0.0640
## s.e.  0.0427  0.0413  0.0393    0   0.0345    0    0    0   0.0318
##       dpus[1:715]
##            0.0406
## s.e.       0.0176
##
## sigma^2 estimated as 0.000323:  log likelihood = 1858.7,  aic = -3703.4
```