STATS 207 (Time Series Analysis) HW3 R Code

Chih-Hsuan ‘Carolyn’ Kao (chkao831 at stanford dot edu)

Nov 2, 2020 (Fall 2020)

All references (exercises, examples, equations, etc.) are to Shumway & Stoffer

Q5. Exercise 3.36

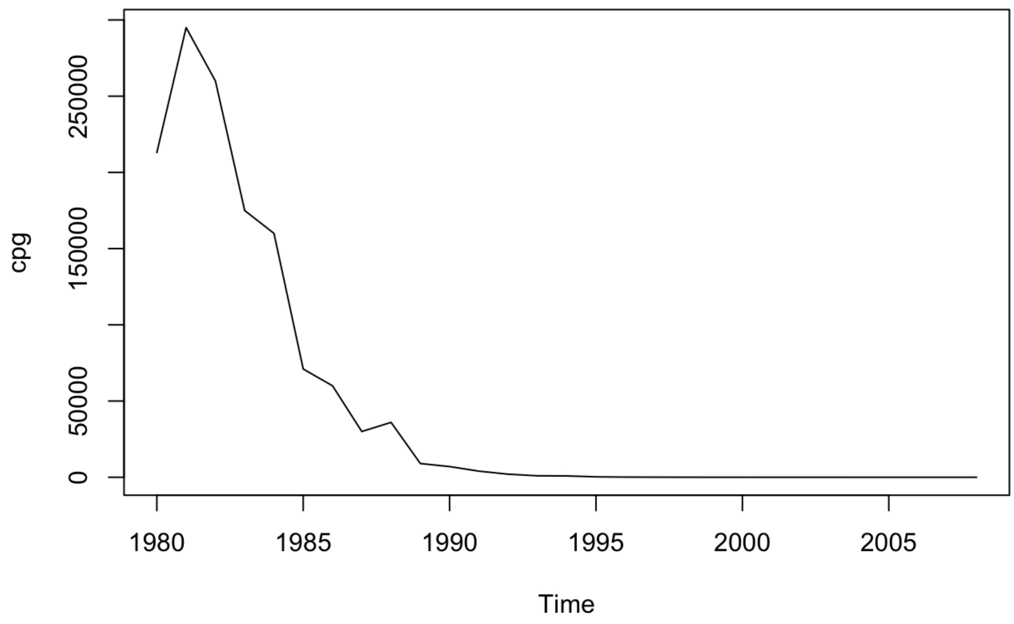
(a) Plot c_t and describe what you see

As we plot the c_t below, we see that the median annual retail price per GB of hard drives declines over time from year 1980 to 2008.

```
library(astsa)
```

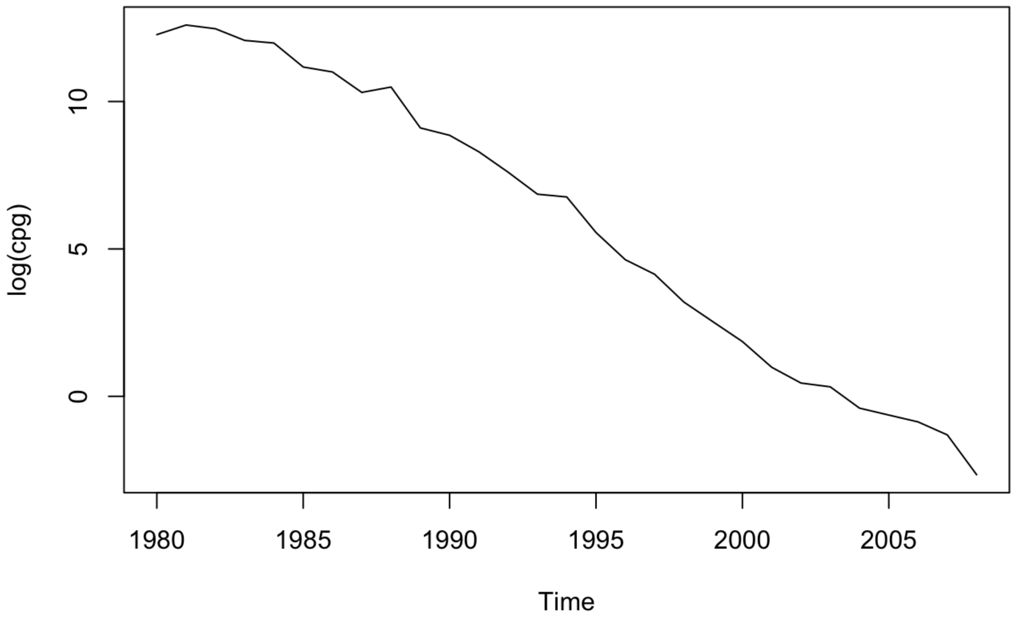
```
## Warning: package 'astsa' was built under R version 3.6.2
```

```
plot(cpg)
```



(b)

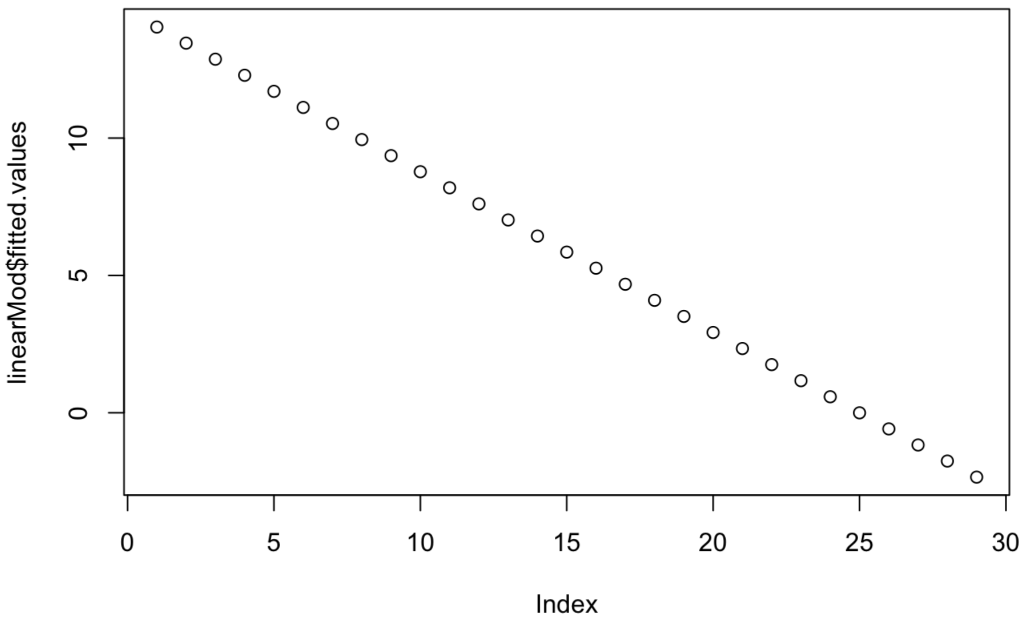
```
#Plotting logged data of c_t
plot(log(cpg))
```




```
#fitting a linear regression of log ct on t
linearMod <- lm(log(cpg) ~ c(1:length(cpg)))
linearMod
```

```
##
## Call:
## lm(formula = log(cpg) ~ c(1:length(cpg)))
##
## Coefficients:
##      (Intercept)  c(1:length(cpg))
##          14.6257           -0.5851
```

```
#and then plotting the fitted line
plot(linearMod$fitted.values)
```



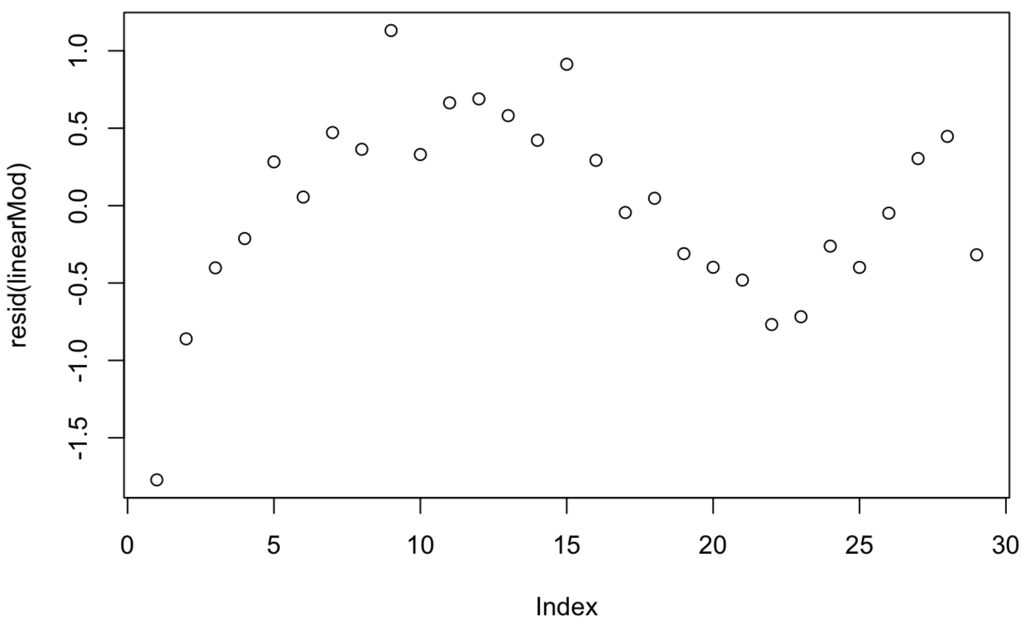
From the two figures above, we are able to see a pretty good alignment between the fitted linear line and the logged data. Hence, it is possible to argue that the curve c_t versus t behaves like $c_t = \alpha e^{\beta t}$.

(c) Inspect the residuals of the linear regression fit and comment.

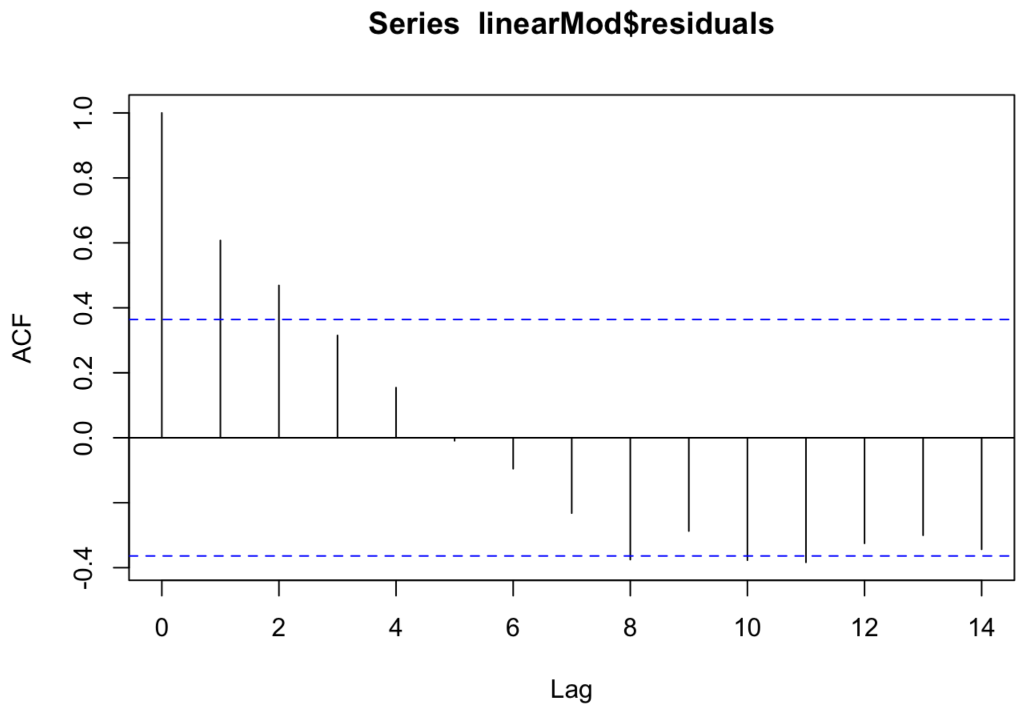
```
resid(linearMod)
```

```
##           1           2           3           4           5           6 
## -1.77156094 -0.86080012 -0.40201622 -0.21283425  0.28263122  0.05521491 
##           7           8           9          10          11          12 
##  0.47195722  0.36388767  1.13128685  0.33007012  0.66383332  0.68929516 
##          13          14          15          16          17          18 
##  0.58122560  0.42186275  0.91320790  0.29238409 -0.04463736  0.04725744 
##          19          20          21          22          23          24 
## -0.31053798 -0.39840482 -0.48119657 -0.76816142 -0.71782497 -0.26173946 
##          25          26          27          28          29 
## -0.39922290 -0.04854598  0.30390936  0.44715423 -0.31769486
```

```
plot(resid(linearMod))
```




```
acf(linearMod$residuals)
```



From the plot above along with the acf, it is clearly seen that the residuals so far do not exhibit iid behavior in the linear fit.

(d) Fit again, now with autocorrelated errors

```r
#library(nlme)
#new_fit1 <- gls(log(cpg) ~ c(1:length(cpg)))
#new_fit1
```

```r
#plot(new_fit1)
```

```r
require(forecast)
```

```
## Loading required package: forecast
```

```
## Registered S3 method overwritten by 'xts':
##   method     from
##   as.zoo.xts zoo
```

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

```
## Registered S3 methods overwritten by 'forecast':
##   method             from
##   fitted.fracdiff    fracdiff
##   residuals.fracdiff fracdiff
```

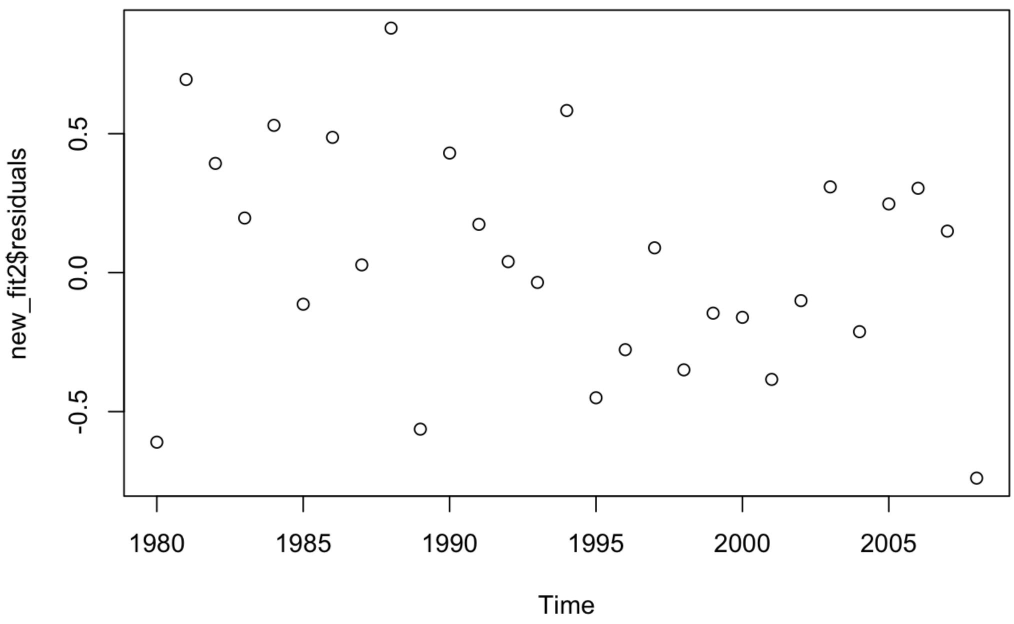
```
##
## Attaching package: 'forecast'
```

```
## The following object is masked from 'package:astsa':
##
##     gas
```

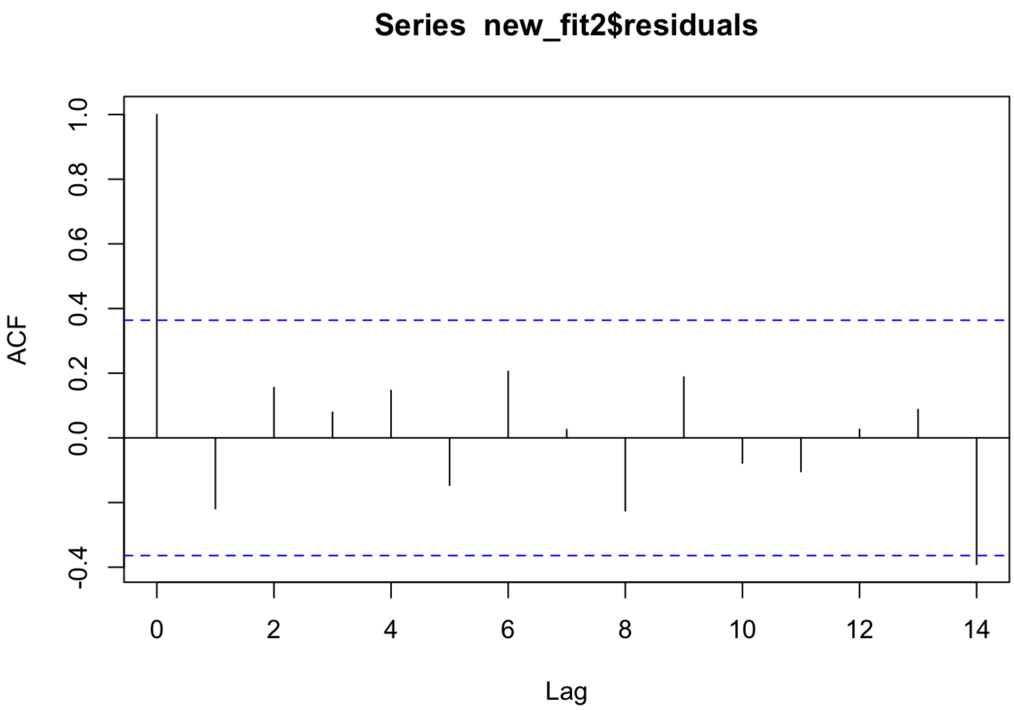
```r
new_fit2 <- auto.arima(log(cpg), xreg=c(1:length(cpg)))
new_fit2
```

```
## Series: log(cpg)
## Regression with ARIMA(1,0,0) errors
##
## Coefficients:
##          ar1  intercept     xreg
##       0.8297    13.9174  -0.5554
## s.e.  0.1177     0.7357   0.0368
##
## sigma^2 estimated as 0.181:  log likelihood=-15.37
## AIC=38.73   AICc=40.4   BIC=44.2
```

```
plot(new_fit2$residuals, type='p')
```




```
acf(new_fit2$residuals)
```

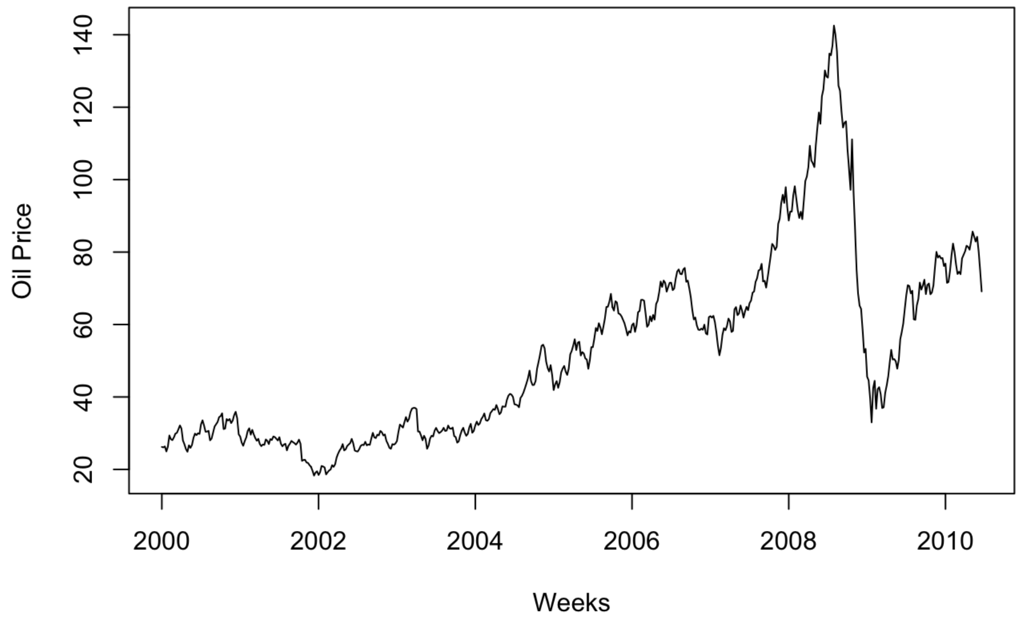


Now, as we take the fact that the errors are autocorrelated into consideration, and then fit the regression again with ARIMA(1,0,0) errors, the residual now behaves relatively more with iid assumption.

Q6. Exercise 3.32

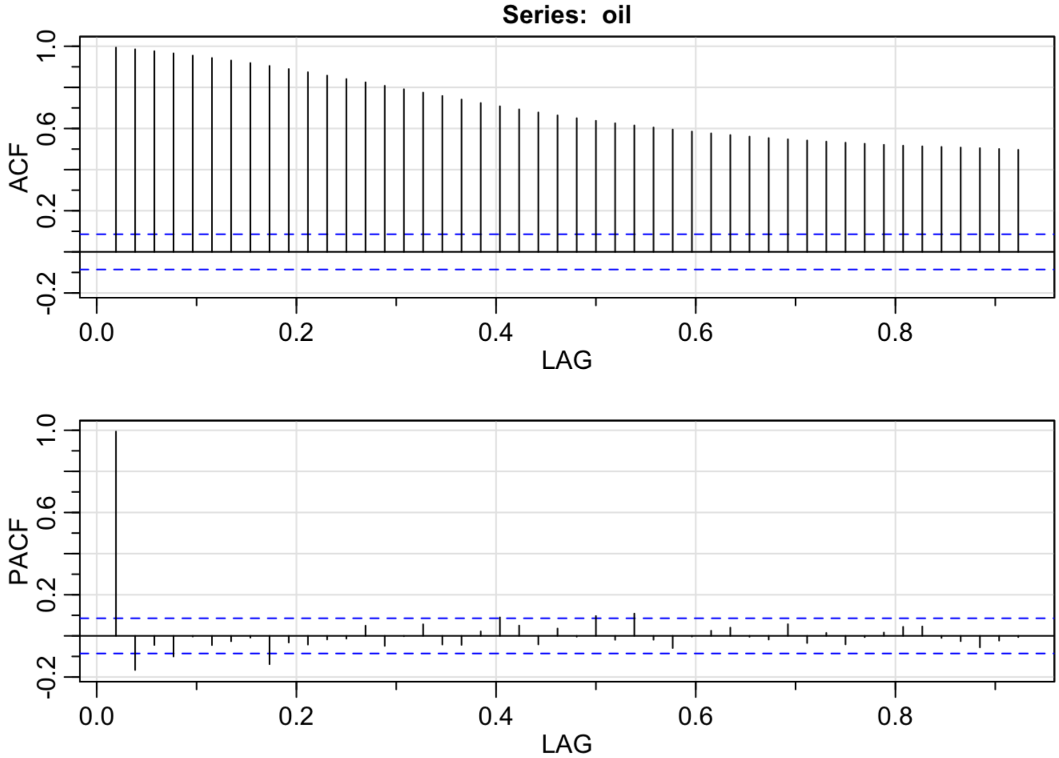
Crude oil prices in dollars per barrel are in oil.

```
plot.ts(oil, xlab = "Weeks", ylab = "Oil Price")
```



###(a) Fit an ARIMA(p,d,q) model to the growth rate performing all necessary diagnostics. Comment. From the upward trend in the plot above, it is clear that some differencing method is required. Firstly, I would inspect acf and pacf.

```
acf2(oil, 48)
```

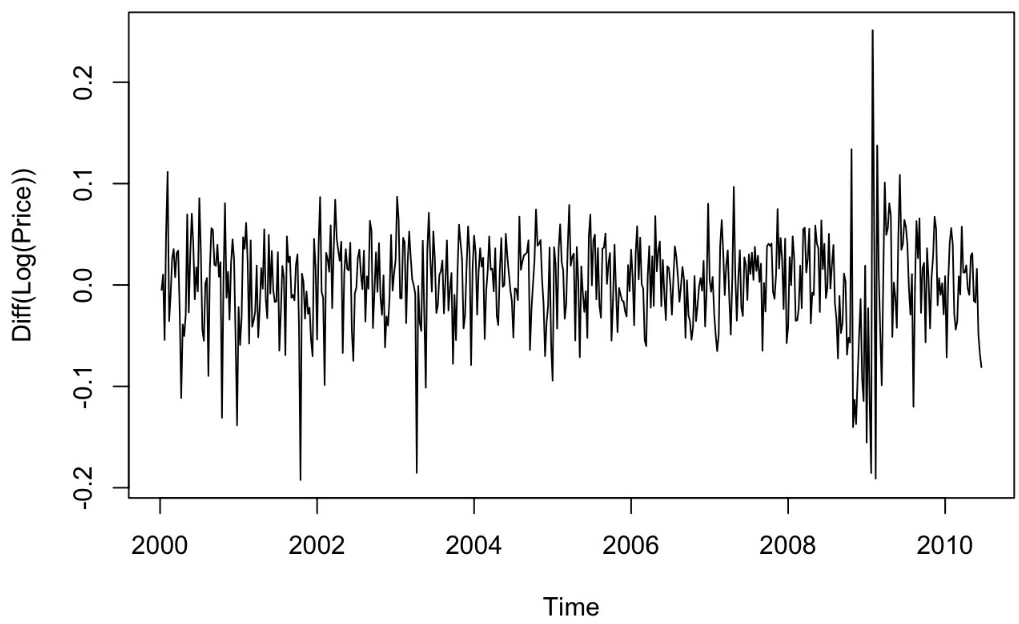


```
##      [,1]  [,2]  [,3]  [,4] [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11] [,12]
## ACF  0.99  0.99  0.98  0.97 0.95  0.94  0.93  0.92  0.90  0.89  0.87  0.86
## PACF 0.99 -0.17 -0.04 -0.10 0.00 -0.04 -0.03 -0.01 -0.14 -0.03 -0.04 -0.02
##      [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23]
## ACF   0.84  0.82  0.81  0.79  0.77  0.76  0.74  0.72  0.71  0.69  0.68
## PACF -0.01  0.05 -0.05  0.00  0.06 -0.04 -0.04  0.02  0.09  0.05 -0.04
##      [,24] [,25] [,26] [,27] [,28] [,29] [,30] [,31] [,32] [,33] [,34]
## ACF   0.66  0.65  0.64  0.63  0.61  0.61  0.59  0.59  0.58  0.57  0.56
## PACF  0.04  0.00  0.10 -0.02  0.11 -0.02 -0.06  0.00  0.03  0.04  0.00
##      [,35] [,36] [,37] [,38] [,39] [,40] [,41] [,42] [,43] [,44] [,45]
## ACF   0.55  0.55  0.54  0.54  0.53  0.53  0.52  0.52  0.51  0.51  0.51
## PACF -0.02  0.06 -0.03  0.01 -0.04 -0.01  0.02  0.04  0.05 -0.01 -0.03
##      [,46] [,47] [,48]
## ACF   0.50  0.50  0.50
## PACF -0.06 -0.02 -0.01
```

The ACF is constantly falling down over time; however, PACF is only significant on lag 1. Hence, from this, we see that significance prevails only until lag 1, and that conclude AR = 1.

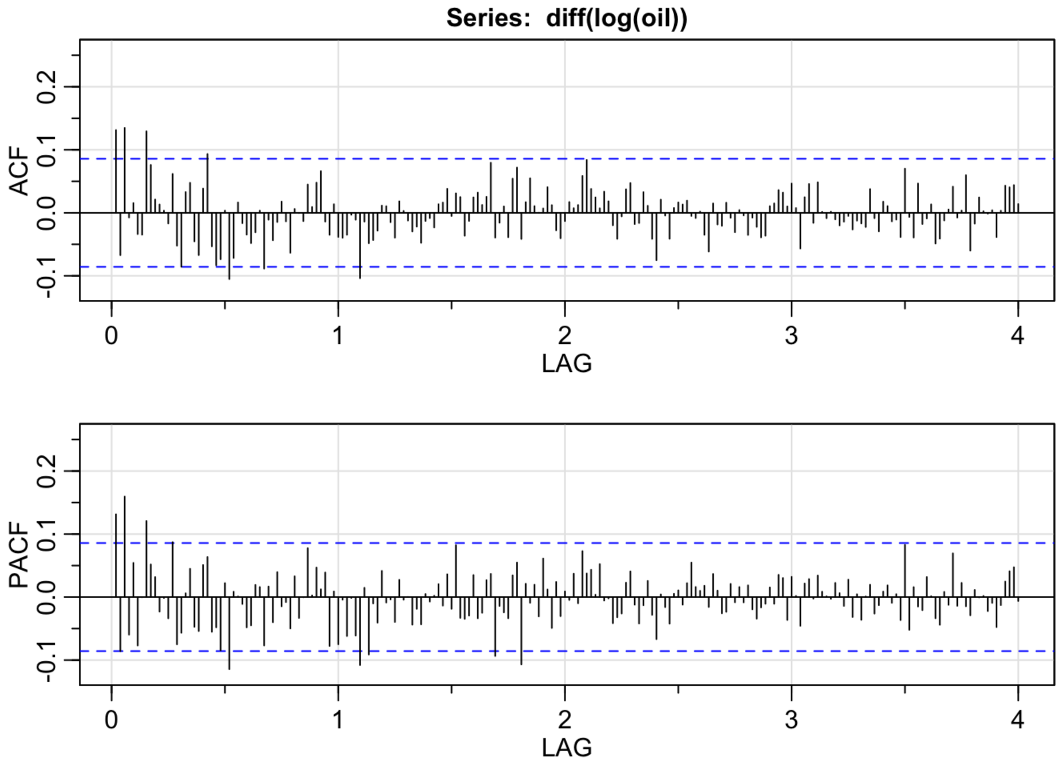
Here, I firstly take log to perform logistic transformation to remove variance in time. Then, I create a differenced data to achieve data stationarity.

```
plot(diff(log(oil)), ylab = 'Diff(Log(Price))')
```



From this, we could further argue that I = 1 as the differenced price oscillates across mean of zero.

```
acf2(diff(log(oil)))
```


Series: diff(log(oil))
ACF
-0.1
0.0
0.1
0.2
0
1
2
3
4
LAG
PACF
-0.1
0.0
0.1
0.2
0
1
2
3
4
LAG

```
##      [,1]  [,2] [,3]  [,4] [,5]  [,6]  [,7] [,8] [,9] [,10] [,11] [,12]
## ACF  0.13 -0.07 0.13 -0.01 0.02 -0.03 -0.03 0.13 0.08  0.02  0.01     0
## PACF 0.13 -0.09 0.16 -0.06 0.05 -0.08  0.00 0.12 0.05  0.03 -0.02     0
##      [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23]
## ACF  -0.02  0.06 -0.05 -0.09  0.03  0.05 -0.05 -0.07  0.04  0.09 -0.05
## PACF -0.03  0.09 -0.07 -0.06  0.01  0.04 -0.05 -0.05  0.05  0.06 -0.06
##      [,24] [,25] [,26] [,27] [,28] [,29] [,30] [,31] [,32] [,33] [,34]
## ACF  -0.08 -0.07  0.00 -0.11 -0.07  0.02 -0.02 -0.03 -0.05 -0.03  0.00
## PACF -0.05 -0.08  0.02 -0.11  0.01  0.00 -0.01 -0.05 -0.04  0.02  0.02
##      [,35] [,36] [,37] [,38] [,39] [,40] [,41] [,42] [,43] [,44] [,45]
## ACF  -0.09 -0.01 -0.04 -0.01  0.02 -0.01 -0.06  0.01  0.00 -0.01  0.04
## PACF -0.08  0.02 -0.04  0.04 -0.01 -0.01 -0.05  0.03 -0.03  0.00  0.08
##      [,46] [,47] [,48] [,49] [,50] [,51] [,52] [,53] [,54] [,55] [,56]
## ACF   0.01  0.05  0.07 -0.01 -0.03  0.01 -0.04 -0.04 -0.03     0 -0.01
## PACF  0.00  0.05  0.01  0.04 -0.08  0.01 -0.07  0.00 -0.06     0 -0.06
##      [,57] [,58] [,59] [,60] [,61] [,62] [,63] [,64] [,65] [,66] [,67]
## ACF  -0.10 -0.01 -0.05 -0.04 -0.03  0.01  0.01 -0.01 -0.04  0.02     0
## PACF -0.11  0.01 -0.09 -0.01 -0.04  0.04 -0.01  0.00 -0.04  0.03     0
##      [,68] [,69] [,70] [,71] [,72] [,73] [,74] [,75] [,76] [,77] [,78]
## ACF  -0.01 -0.03 -0.02 -0.05 -0.01 -0.01 -0.02  0.01  0.02  0.04 -0.01
## PACF  0.00 -0.04 -0.02 -0.04  0.00 -0.01  0.00  0.02 -0.01  0.04 -0.02
##      [,79] [,80] [,81] [,82] [,83] [,84] [,85] [,86] [,87] [,88] [,89]
## ACF   0.03  0.02 -0.04 -0.01  0.02  0.03  0.01  0.03  0.08 -0.04 -0.02
## PACF  0.08 -0.03 -0.03 -0.03  0.03 -0.03 -0.02  0.03  0.04 -0.09 -0.01
##      [,90] [,91] [,92] [,93] [,94] [,95] [,96] [,97] [,98] [,99] [,100]
## ACF   0.01 -0.04  0.05  0.07 -0.04  0.02  0.05  0.01  0.00  0.01   0.04
## PACF -0.02 -0.03  0.03  0.05 -0.11  0.02 -0.01  0.02 -0.03  0.06   0.01
##      [,101] [,102] [,103] [,104] [,105] [,106] [,107] [,108] [,109] [,110]
## ACF    0.01  -0.03  -0.04  -0.01   0.02   0.01   0.01   0.06   0.08   0.04
## PACF  -0.05   0.02  -0.03   0.01   0.00   0.04  -0.01   0.07   0.04   0.04
##      [,111] [,112] [,113] [,114] [,115] [,116] [,117] [,118] [,119] [,120]
## ACF    0.02   0.01   0.03   0.02  -0.02  -0.04  -0.01   0.04   0.05  -0.02
## PACF   0.00   0.05  -0.01   0.00  -0.04  -0.03  -0.03   0.02   0.04  -0.01
##      [,121] [,122] [,123] [,124] [,125] [,126] [,127] [,128] [,129] [,130]
## ACF   -0.02   0.03   0.01  -0.04  -0.08   0.02   0.00  -0.04   0.01   0.02
## PACF  -0.04  -0.01   0.03  -0.03  -0.07   0.00  -0.02  -0.04   0.01   0.01
##      [,131] [,132] [,133] [,134] [,135] [,136] [,137] [,138] [,139] [,140]
## ACF    0.01   0.02   0.00  -0.01   0.00  -0.03  -0.06   0.01  -0.02  -0.02
## PACF  -0.01   0.02   0.05   0.02   0.01   0.02  -0.02   0.04   0.01  -0.03
##      [,141] [,142] [,143] [,144] [,145] [,146] [,147] [,148] [,149] [,150]
## ACF    0.02  -0.01  -0.03   0.00   0.00  -0.04  -0.01  -0.02  -0.04  -0.04
## PACF  -0.02   0.02  -0.01   0.02  -0.01   0.02  -0.02  -0.03  -0.02  -0.01
##      [,151] [,152] [,153] [,154] [,155] [,156] [,157] [,158] [,159] [,160]
## ACF    0.01   0.01   0.04   0.03   0.01   0.05   0.01  -0.06   0.02   0.05
## PACF   0.02  -0.01   0.04   0.03  -0.04   0.03   0.00  -0.05   0.02   0.03
##      [,161] [,162] [,163] [,164] [,165] [,166] [,167] [,168] [,169] [,170]
## ACF   -0.02   0.05   0.00  -0.01      0  -0.01  -0.02  -0.01   0.00  -0.03
## PACF   0.00   0.03   0.01   0.00      0   0.02   0.01  -0.01   0.03  -0.03
##      [,171] [,172] [,173] [,174] [,175] [,176] [,177] [,178] [,179] [,180]
## ACF   -0.01  -0.02  -0.02   0.04  -0.01  -0.03   0.02   0.01  -0.01  -0.01
## PACF   0.00  -0.04   0.00   0.02  -0.03  -0.01   0.01   0.02  -0.01   0.00
##      [,181] [,182] [,183] [,184] [,185] [,186] [,187] [,188] [,189] [,190]
## ACF   -0.04   0.07  -0.01  -0.04   0.05  -0.02  -0.01   0.01  -0.05  -0.04
## PACF  -0.04   0.08  -0.05   0.02  -0.01  -0.02   0.03   0.00  -0.03  -0.04
##      [,191] [,192] [,193] [,194] [,195] [,196] [,197] [,198] [,199] [,200]
## ACF   -0.01   0.01   0.04  -0.01   0.00   0.06  -0.06  -0.02   0.02      0
## PACF   0.01  -0.01   0.07  -0.01   0.02  -0.01  -0.03   0.01   0.00      0
##      [,201] [,202] [,203] [,204] [,205] [,206] [,207] [,208]
## ACF    0.00   0.00  -0.04   0.00   0.04   0.04   0.04   0.01
## PACF  -0.02  -0.01  -0.05  -0.01   0.02   0.04   0.05  -0.01
```

From the sample ACF above, after logistic transformation, we could also see that the ACF with a significant autocorrelation only at lag 1 is an indicator of a possible MA(1) model. Hence, AR = 1, I = 1, MA = 1. Now, we could then fit the model after all necessary diagnostics.

```
fit <- arima(log(oil), order = c(1,1,1))
fit
```

```
##
## Call:
## arima(x = log(oil), order = c(1, 1, 1))
##
## Coefficients:
##           ar1     ma1
##       -0.5253  0.7142
## s.e.   0.0872  0.0683
##
## sigma^2 estimated as 0.002104:  log likelihood = 904.58,  aic = -1803.15
```

```
require(sarima)
```

```
## Loading required package: sarima
```

```
## Loading required package: FitAR
```

```
## Loading required package: lattice
```

```
## Loading required package: leaps
```

```
## Loading required package: ltsa
```

```
## Loading required package: bestglm
```

```
##
## Attaching package: 'FitAR'
```

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

```
## Loading required package: stats4
```

```
##
## Attaching package: 'sarima'
```

```
## The following object is masked from 'package:astsa':
##
##     sarima
```

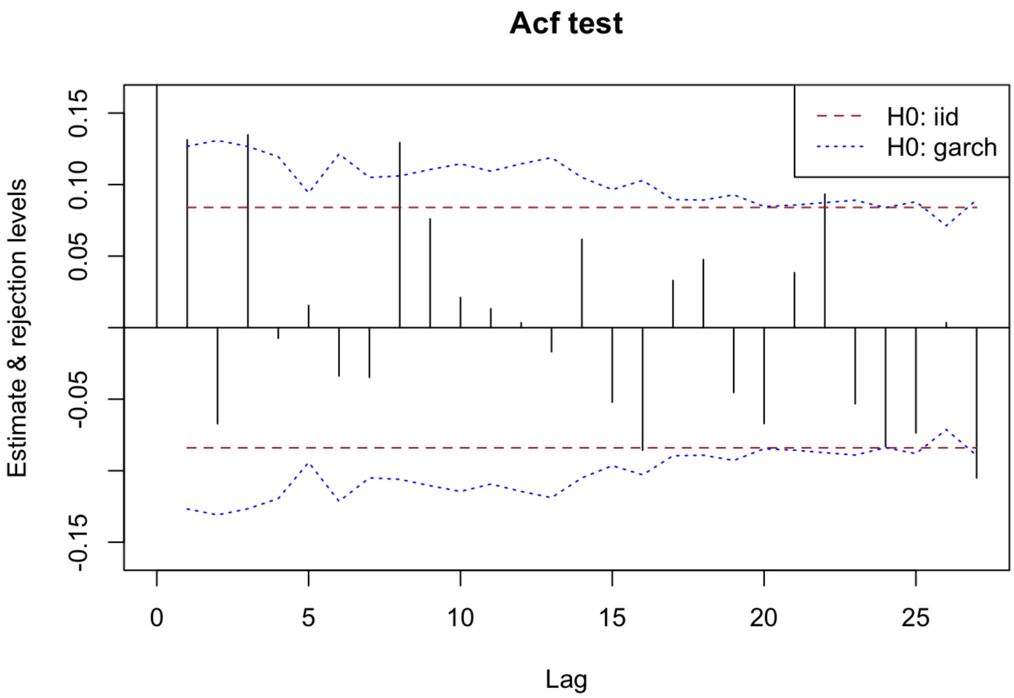
```
arima.fit = arima.sim(list(order=c(1,1,1), ar = -0.5253, ma = 0.7142), n = 100)
```

```
#sarima(log(oil), 1,1,1)
```

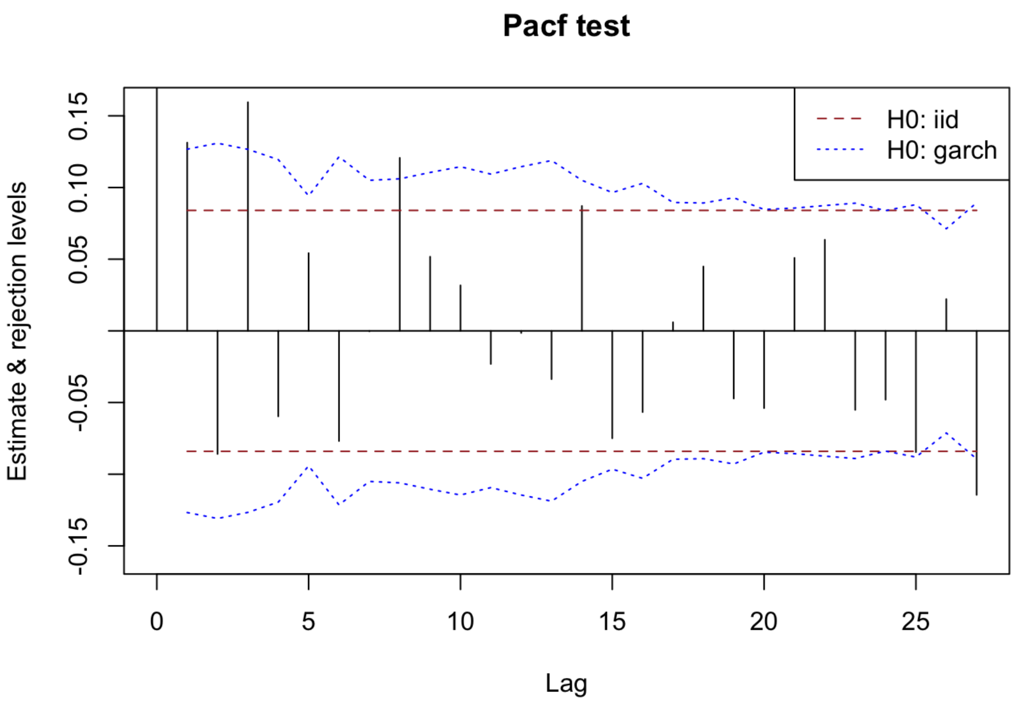
The plot above behaves like white noise with some outliers especially around year 2009. Using ARIMA(1,1,1), the residuals seem pretty much like iid as there's no significant trend. The ACF lies in the band for most lags, too. Also, the flying tails for QQ Plots indicate the presence of some outliers. Lastly, looking at the P-Values for Ljung-Box Statistic, we would see that we REJECT the null hypothesis that suggests this model as most p-values are below the 0.0 dotted line. Hence, some other model needs to be considered.

(b) Investigate whether the growth rate of the weekly oil price exhibits GARCH behavior. If so, fit an appropriate model to the growth rate. Comment.

```
log_oil.acf <- autocorrelations(diff(log(oil)))
plot(log_oil.acf, data = diff(log(oil)))
```




```
log_oil.pacf <- partialAutocorrelations(diff(log(oil)))
plot(log_oil.pacf, data = diff(log(oil)))
```



From the acf and pacf plots that test against garch hypothesis (in blue dotted lines), we see that although several (partial) autocorrelations seem significant under the garch hypothesis for certain lags, there is no evidence against the GARCH hypothesis (i.e. we FAIL TO REJECT the null hypothesis that the autocorrelations are zeroes in population), as most values lie within the bands.

This further suggests the use of GARCH model may be considered.

```
require(tseries)
```

```
## Loading required package: tseries
```

```
oil.garch <- garch(diff(log(oil)))  # Fit a GARCH(1,1)
```

```
##
##  ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
##
##
##      I     INITIAL X(I)        D(I)
##
##      1     1.988229e-03     1.000e+00
##      2     5.000000e-02     1.000e+00
##      3     5.000000e-02     1.000e+00
##
##     IT   NF      F         RELDF    PRELDF    RELDX   STPPAR   D*STEP   NPRELDF
##      0    1 -1.399e+03
##      1    6 -1.400e+03  3.43e-04  5.48e-04  1.0e-03  7.7e+07  1.0e-04  2.10e+04
##      2    7 -1.400e+03  2.54e-05  2.80e-05  9.5e-04  2.0e+00  1.0e-04  7.22e+00
##      3   13 -1.406e+03  4.31e-03  7.24e-03  4.2e-01  2.0e+00  7.2e-02  7.14e+00
##      4   15 -1.413e+03  5.06e-03  7.26e-03  6.9e-01  2.0e+00  2.7e-01  5.37e-01
##      5   21 -1.414e+03  3.38e-04  7.46e-04  1.3e-04  8.8e+00  8.2e-05  1.45e-02
##      6   22 -1.414e+03  5.50e-06  5.26e-06  1.2e-04  2.0e+00  8.2e-05  2.99e-03
##      7   29 -1.416e+03  1.34e-03  2.36e-03  2.7e-01  6.0e-01  2.5e-01  3.05e-03
##      8   30 -1.416e+03  4.33e-04  4.98e-04  1.7e-02  0.0e+00  1.9e-02  4.98e-04
##      9   32 -1.417e+03  9.37e-04  5.04e-04  3.9e-02  0.0e+00  5.8e-02  5.04e-04
##     10   34 -1.418e+03  6.65e-04  7.39e-04  3.9e-02  1.9e+00  5.5e-02  1.41e-02
##     11   36 -1.419e+03  6.18e-04  6.23e-04  3.6e-02  9.6e-01  5.5e-02  5.16e-03
##     12   38 -1.421e+03  1.08e-03  1.21e-03  6.7e-02  8.6e-01  1.1e-01  6.23e-03
##     13   40 -1.421e+03  5.68e-05  3.96e-04  1.6e-02  1.6e+00  2.8e-02  2.11e-03
##     14   41 -1.422e+03  4.97e-04  6.00e-04  1.6e-02  1.6e+00  2.8e-02  1.32e-03
##     15   42 -1.422e+03  4.84e-04  6.59e-04  1.2e-02  2.7e-01  2.8e-02  6.90e-04
##     16   43 -1.422e+03  1.26e-04  2.78e-04  1.6e-02  8.6e-01  2.8e-02  4.32e-04
##     17   44 -1.423e+03  2.70e-05  9.44e-05  7.8e-03  0.0e+00  1.4e-02  9.44e-05
##     18   45 -1.423e+03  1.98e-05  1.77e-05  3.1e-04  0.0e+00  6.1e-04  1.77e-05
##     19   48 -1.423e+03  3.17e-06  1.93e-06  1.4e-03  0.0e+00  2.5e-03  1.93e-06
##     20   63 -1.423e+03 -3.68e-15  2.71e-15  1.1e-14  2.4e+06  1.8e-14  7.65e-08
##
##  ***** FALSE CONVERGENCE *****
##
##  FUNCTION    -1.422570e+03   RELDX        1.052e-14
##  FUNC. EVALS      63         GRAD. EVALS      20
##  PRELDF       2.709e-15      NPRELDF      7.645e-08
##
##      I      FINAL X(I)        D(I)          G(I)
##
##      1    1.221742e-04     1.000e+00    -2.101e+02
##      2    6.713111e-02     1.000e+00     1.415e-01
##      3    8.721110e-01     1.000e+00     1.375e-01
```

```
summary(oil.garch)
```

```
##
## Call:
## garch(x = diff(log(oil)))
##
## Model:
## GARCH(1,1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.0008 -0.5693  0.1627  0.7685  3.0145
##
## Coefficient(s):
##     Estimate  Std. Error  t value Pr(>|t|)
## a0 1.222e-04   5.909e-05    2.068 0.038667 *
## a1 6.713e-02   1.913e-02    3.509 0.000449 ***
## b1 8.721e-01   4.278e-02   20.384  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 113.29, df = 2, p-value < 2.2e-16
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 0.017169, df = 1, p-value = 0.8958
```

```
#plot(oil.garch)
```

From the Box-Ljung test, p-value for Box-Ljung Test is very big, as big as 0.8958. This indicates weak evidence against the null hypothesis that GARCH is suggested, so we fail to reject the null hypothesis. Hence, the fitted GARCH model here is an appropriate choice to the [illegible]s, the growth rate of the weekly oil price exhibits GARCH behavior.