

1. 考虑在 2007 年 1 月 2 日到 2011 年 12 月 23 日期间，苹果公司股票每天的股价波动幅度（即当天的最高价减去最低价）。这个数据可以利用 quantmod 包从雅虎财经获得（股票代码 AAPL）。

（1）判断该序列是否平稳时间序列

（2）寻找适当的 ARIMA 模型拟合该序列，并做超前 1~12 步预测

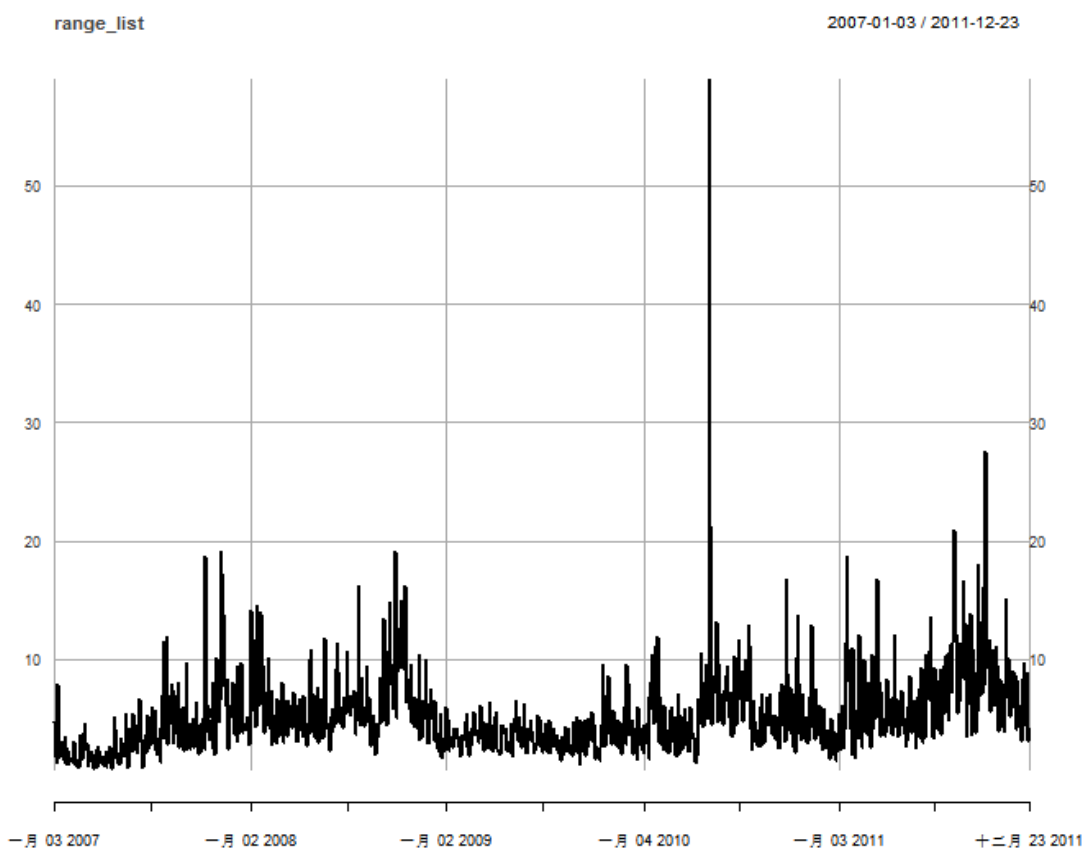
（3）对拟合的模型做残差检验

#获取整理数据

```
library(quantmod)
getSymbols("AAPL",src="yahoo",from="2007-01-02",to="2011-12-23")
chartSeries(AAPL,subset="2007-01-02::2011-12-23")
range_list=Hi(AAPL)-Lo(AAPL)
head(range_list)
tail(range_list)
plot(range_list)
```



```
> head(range_list)
      AAPL.High
2007-01-03    4.68
2007-01-04    2.13
2007-01-05    1.80
2007-01-08    1.25
2007-01-09    7.83
2007-01-10    4.35
> tail(range_list)
      AAPL.High
2011-12-16    4.58
2011-12-19    4.37
2011-12-20    8.84
2011-12-21    5.29
2011-12-22    3.03
2011-12-23    4.10
```



(1) 判断该序列是否平稳时间序列

```
> library(urca)
> summary(ur.df(range_list,lags=6,type='drift'))
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression drift

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-10.064  -1.502  -0.549   1.007   51.637

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.20857     0.19112   6.323 3.56e-10 ***
z.lag.1       -0.22863     0.03279  -6.972 5.05e-12 ***
z.diff.lag1   -0.43597     0.03863 -11.286 < 2e-16 ***
z.diff.lag2   -0.34883     0.03953  -8.824 < 2e-16 ***
z.diff.lag3   -0.22917     0.03890  -5.890 4.95e-09 ***
z.diff.lag4   -0.19745     0.03724  -5.302 1.36e-07 ***
z.diff.lag5   -0.09407     0.03391  -2.774 0.00562 **
z.diff.lag6   -0.05124     0.02832  -1.810 0.07060 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.882 on 1241 degrees of freedom
Multiple R-squared:  0.3192, Adjusted R-squared:  0.3154
F-statistic: 83.14 on 7 and 1241 DF, p-value: < 2.2e-16
```

Value of test-statistic is: **-6.9724** 24.3077

Critical values for test statistics:

| | 1pct | 5pct | 10pct |
|------|-------|-------|-------|
| tau2 | -3.43 | -2.86 | -2.57 |
| phi1 | 6.43 | 4.59 | 3.78 |

```
10 library(tseries)
11 adf.test(range_list)
> adf.test(range_list)

Augmented Dickey-Fuller Test

data: range_list
Dickey-Fuller = -5.4855, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary
```

Warning message:

In adf.test(range_list) : p-value smaller than printed p-value

#-6.9724 小于-3.43、-2.86、-2.57，且 p 值足够小，所以可以拒绝 H_0 假设，是平稳的时间序列

(2) 寻找适当的 ARIMA 模型拟合该序列，并做超前 1~12 步预测

```
12 library(forecast)
13 range_list.arma<-auto.arima(range_list,stationary=TRUE,seasonal=FALSE,ic="aic")
14 summary(range_list.arma)
15 confint(range_list.arma)
```

```

> summary(range_list.arma)
Series: range_list
ARIMA(2,0,3) with non-zero mean

Coefficients:
      ar1      ar2      ma1      ma2      ma3  intercept
    0.0233  0.9108  0.2956 -0.7672 -0.1208   5.2201
s.e.  0.0418  0.0414  0.0509  0.0389  0.0335   0.4890

sigma^2 estimated as 8.136:  log likelihood=-3099.04
AIC=6212.09   AICC=6212.18   BIC=6248.04

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE
Training set 0.01000129 2.852327 1.761872 -20.55835 40.00176
      MASE      ACF1
Training set 0.8098619 0.003593337
> confint(range_list.arma)
      2.5 %      97.5 %
ar1    -0.05850686  0.10517226
ar2    0.82966632  0.99190966
ma1    0.19588139  0.39524392
ma2   -0.84338509 -0.69091776
ma3   -0.18647357 -0.05512852
intercept 4.26176420 6.17849484

> pred.last12<-t(predict(range_list.arma,n.ahead=12)$pred[1:12])
> pred.last12
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
[1,] 5.10333 5.247323 5.229492 5.245115 5.22924 5.243099 5.228963
      [,8]      [,9]     [,10]     [,11]     [,12]
[1,] 5.241256 5.228668 5.23957 5.22836 5.238028
>

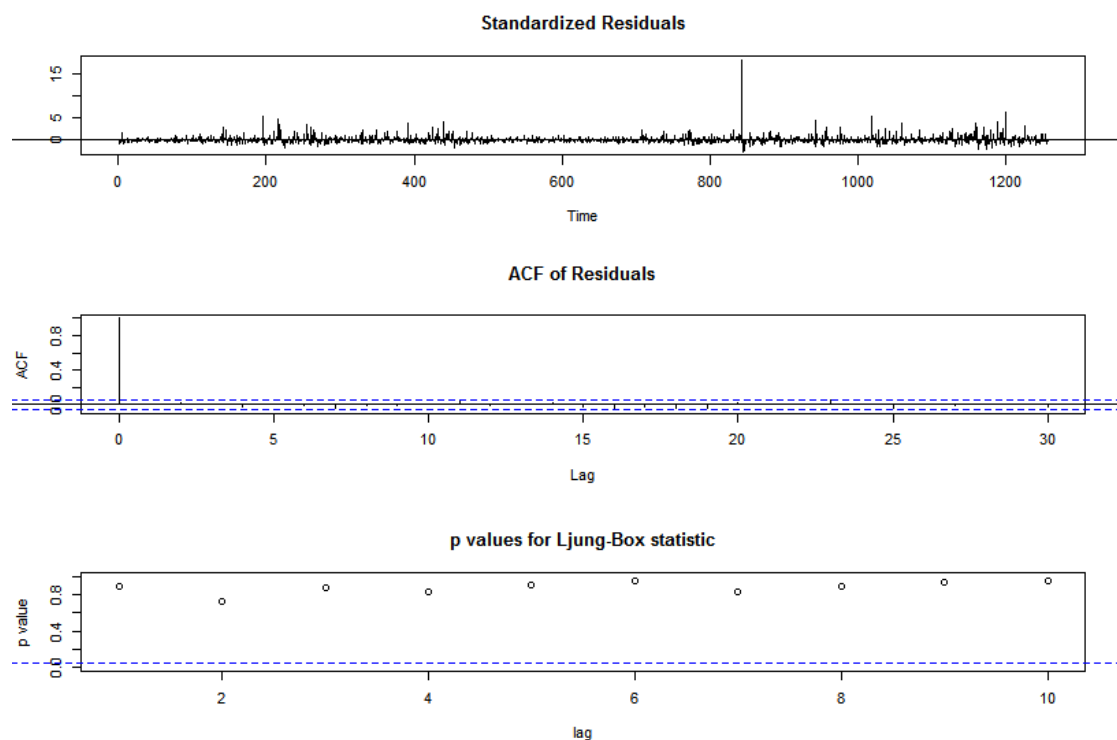
```

(3) 对拟合的模型做残差检验

```

> accuracy(range_list.arma)
      ME      RMSE      MAE      MPE      MAPE
Training set 0.01000129 2.852327 1.761872 -20.55835 40.00176
      MASE      ACF1
Training set 0.8098619 0.003593337
> tsdiag(range_list.arma)

```



2. 自己随意找一组时间序列，按照时间序列分析的步骤对该序列做分析并检验模型

```
> library(quantmod)
> getSymbols("^SSEC")
As of 0.4-0, 'getSymbols' uses env=parent.frame() and
auto.assign=TRUE by default.

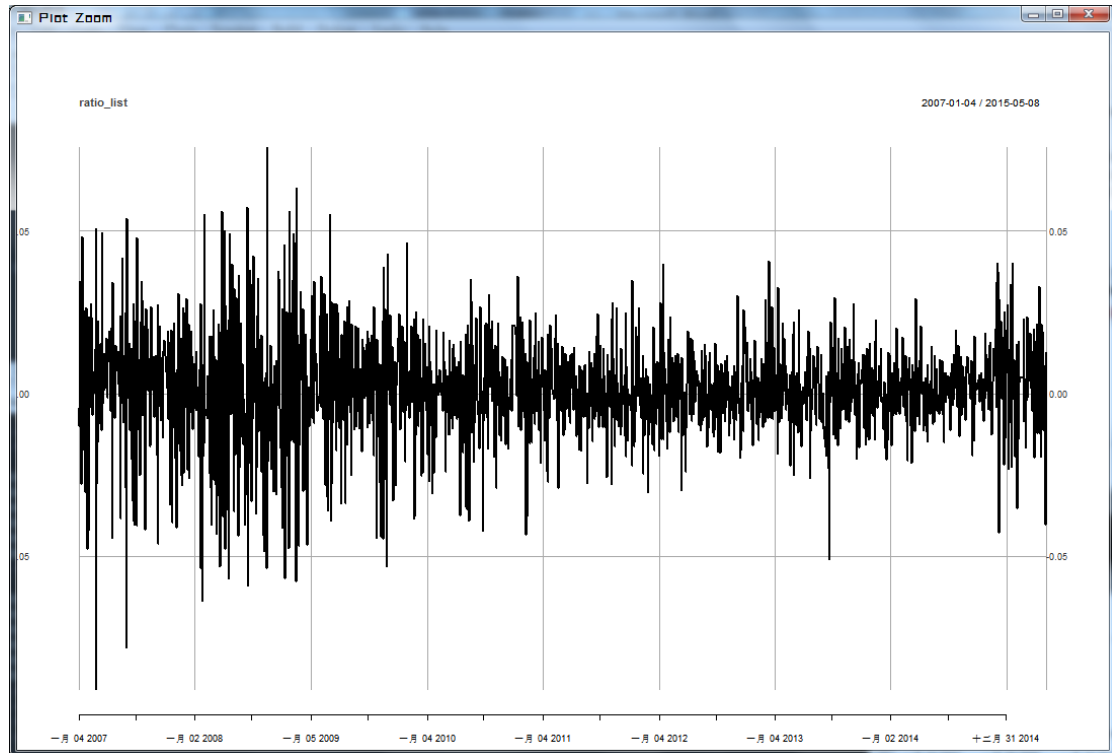
This behavior will be phased out in 0.5-0 when the call will
default to use auto.assign=FALSE. getOption("getSymbols.env") and
getOption("getSymbols.auto.assign") are now checked for alternate defaults

This message is shown once per session and may be disabled by setting
options("getSymbols.warning4.0"=FALSE). See ?getSymbols for more details.
[1] "SSEC"
warning message:
in download.file(paste(yahoo.URL, "s=", Symbols.name, "&a=", from.m, " :
downloaded length 145479 != reported length 200
> ratio_list = (Cl(SSEC) - Op(SSEC))/Hi(SSEC)
```

```
> head(ratio_list)
      SSEC.close
2007-01-04 -0.004378752
2007-01-05 -0.010143302
2007-01-08  0.031800860
2007-01-09  0.034439772
2007-01-10 -0.004411771
2007-01-11 -0.017336751
> tail(ratio_list)
      SSEC.close
2015-05-01 -0.009173839
2015-05-04  0.008717658
2015-05-05 -0.040353649
2015-05-06 -0.018822750
2015-05-07 -0.020333869
2015-05-08  0.012582568
> #蜡烛图
> chartSeries(SSEC)
```



```
> #指数变化率
> plot(ratio_list)
```



```
> ratio_list.numeric = as.numeric(ratio_list$SSEC.close)
> ar(ratio_list.numeric,method='mle')$order #base on AIC准则
[1] 4
```

```
> library(funitroots)
> adfTest(ratio_list.numeric,lags=4,type="c")

Title:
Augmented Dickey-Fuller Test

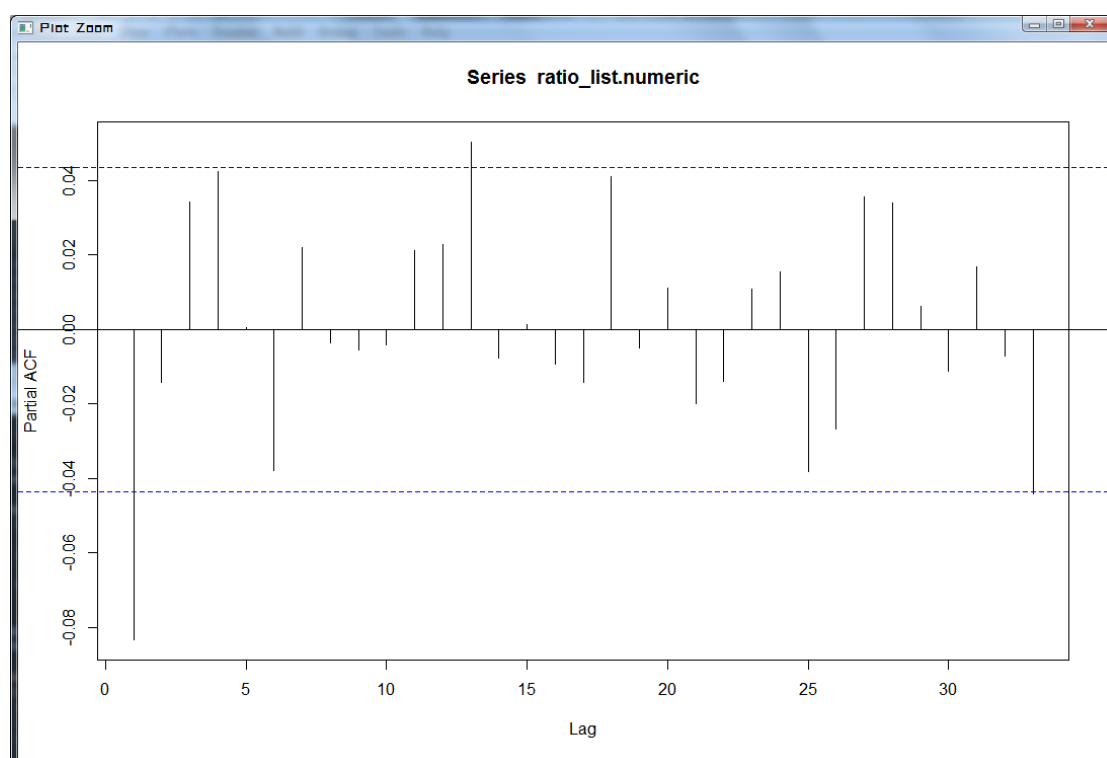
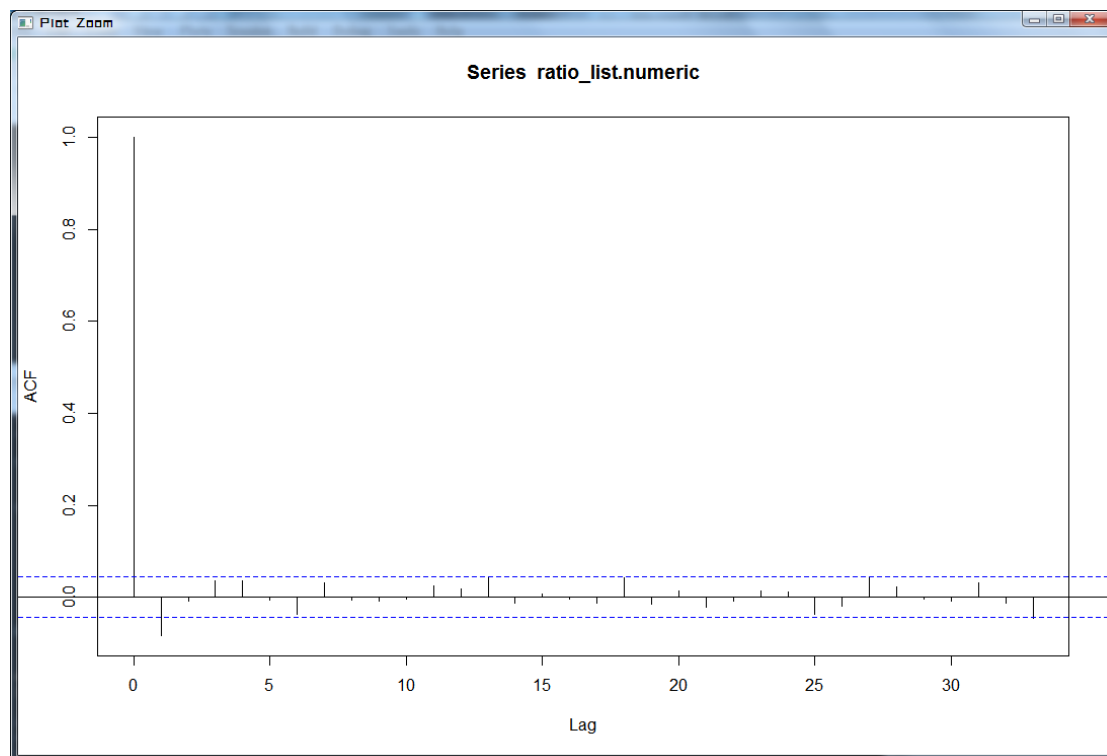
Test Results:
PARAMETER:
Lag order: 4
STATISTIC:
Dickey-Fuller: -19.4049
P VALUE:
0.01

Description:
Mon May 11 20:35:35 2015 by user: Administrator

warning message:
In adfTest(ratio_list.numeric, lags = 4, type = "c") :
p-value smaller than printed p-value
```

p 值=0.01 很小，拒绝 H0 假设，证明是一个平稳序列

```
> acf(ratio_list.numeric)
> pacf(ratio_list.numeric)
```



因此不需要差分，原时间序列已经是一个平稳时间序列

```
> library("TSA")
> eacf(ratio_list.numeric)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x o o o o o o o o o o o o o
1 x o o o o o o o o o o o x o
2 x x o o o o o o o o o x o
3 x x x o o o o o o o o x o
4 o x o x o o o o o o o o o
5 o x x x x o o o o o o o o
6 x x o x x o o o o o o o o
7 x x o x x o o o o o o o o
```

根据结果我们最终选择 $p=0, q=1$, 即确定 ARIMA(0,0,1)

```
> m2=arima(ratio_list.numeric,order=c(0,0,1))
> m2

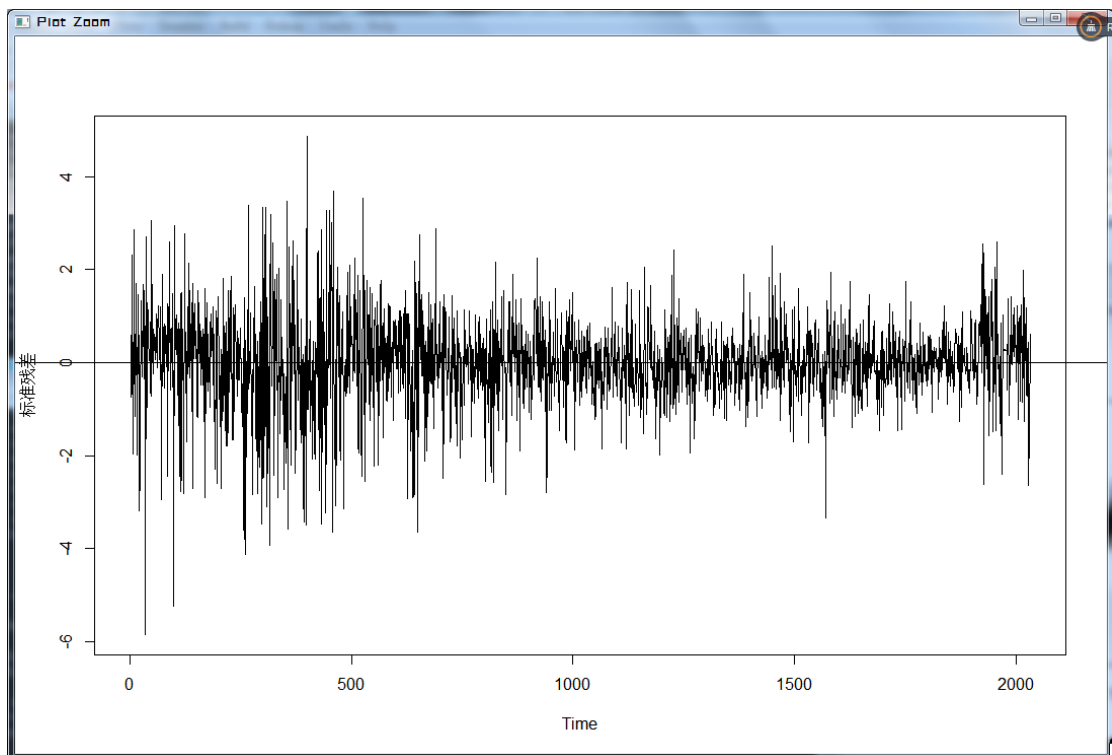
Call:
arima(x = ratio_list.numeric, order = c(0, 0, 1))

Coefficients:
      ma1      intercept
    -0.0842      0.0011
s.e.    0.0220      0.0003

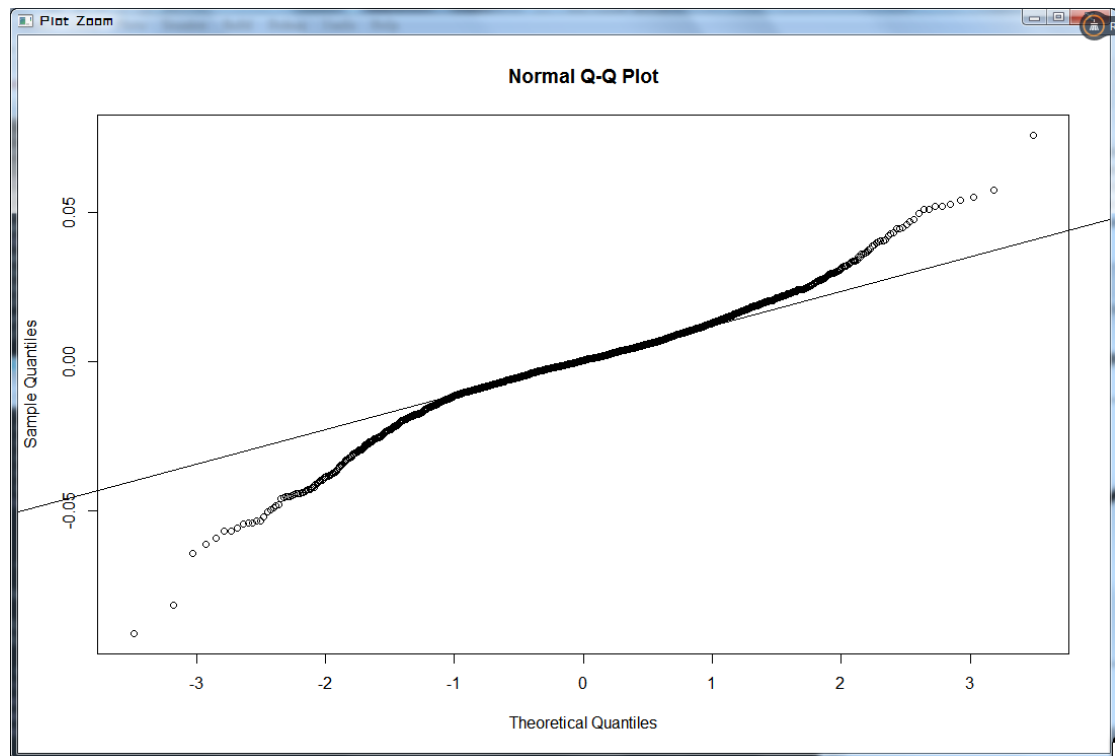
sigma^2 estimated as 0.0002418:  log likelihood = 5577.3,  aic = -11150.6
```

检查残差是否为白噪声序列

```
> plot(rstandard(m2),ylab='标准残差',type='l')
> abline(h=0)
```



```
> qqnorm(residuals(m2))
> qqline(residuals(m2))
```

```
> predict(m2,12)$pred
Time Series:
Start = 2033
End = 2044
Frequency = 1
[1] 0.0002930005 0.0010943606 0.0010943606 0.0010943606 0.0010943606 0.0010943606 0.0010943606
[8] 0.0010943606 0.0010943606 0.0010943606 0.0010943606 0.0010943606
```

通过预测当然不能得到准确的未来大盘指数变化率。但是可以看到市场已经进入一个不稳定状态并且收益率待定状态。