时间序列分析Lab3 实验报告

2.10

读入数据后，输出统计量。

股票	均值	标准差	偏度	超额峰度	最大值	最小值
AAA	7.830109	2.418744	0.857092	0.578605	15.8500	4.1900
BAA	8.847122	2.717073	0.929779	0.760896	17.2900	4.7800

对AAA检验：

$$t = \frac{\hat{S}}{\sqrt{6/T}} = 17.37946, \qquad p = 0.000$$

p值接近0，于是在5%的显著性水平下拒绝了不偏斜的原假设。

$$t = \frac{\hat{K} - 3}{\sqrt{24/T}} = 5.8662626, \qquad p = 4.457 \times 10^{-9}$$

p值接近0，于是在5%的显著性水平下拒绝了没有超额峰度的原假设，且有很大的厚尾性。

对BAA检验：

$$t = \frac{\hat{S}}{\sqrt{6/T}} = 18.85335, \qquad p = 0.000$$

p值接近0，于是在5%的显著性水平下拒绝了不偏斜的原假设。

$$t = \frac{\hat{K} - 3}{\sqrt{24/T}} = 7.714438, \qquad p = 1.221 \times 10^{-14}$$

p值接近0，于是在5%的显著性水平下拒绝了没有超额峰度的原假设，且有很大的厚尾性。

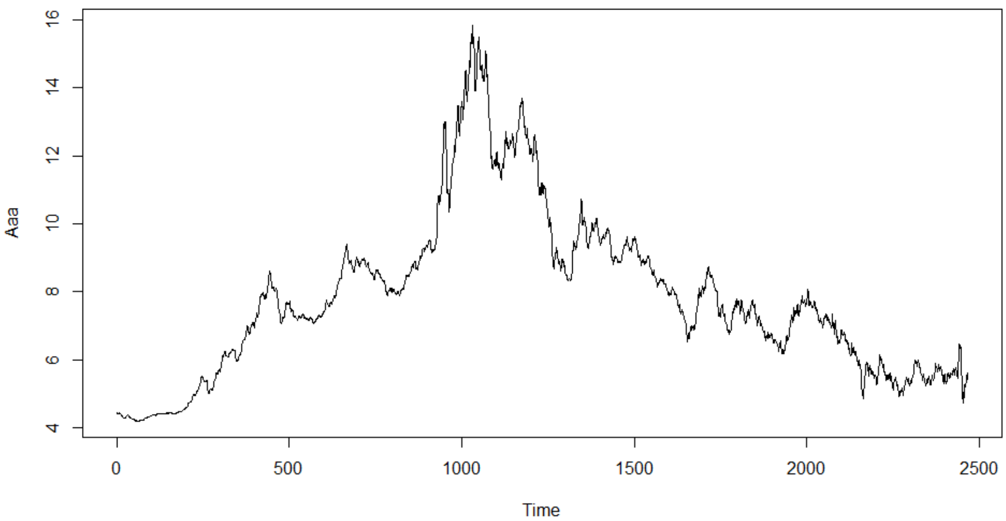
代码：

```
# 2.10
library(fBasics)

da = read.table("D:/USTC/时间序列分析/data/w-Aaa.txt", header = F)
Aaa = da[, 4]
```

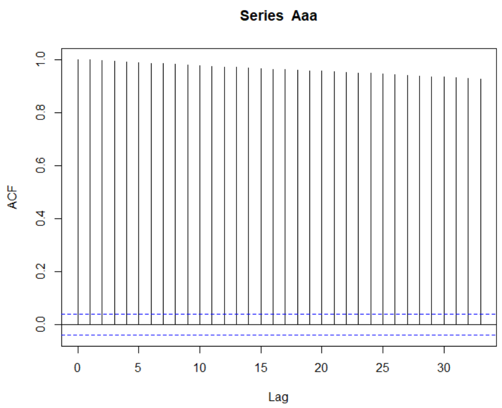
```
da = read.table("D:/USTC/时间序列分析/data/w-Baa.txt", header = F)
Baa = da[, 4]
basicStats(Aaa)
basicStats(Baa)

ts = skewness(Aaa) / sqrt(6 / length(Aaa))
ps = (1 - pnorm(ts)) * 2
tk = kurtosis(Aaa) / sqrt(24 / length(Aaa))
pk = (1 - pnorm(tk)) * 2

ts = skewness(Baa) / sqrt(6 / length(Baa))
ps = (1 - pnorm(ts)) * 2
tk = kurtosis(Baa) / sqrt(24 / length(Baa))
pk = (1 - pnorm(tk)) * 2
```

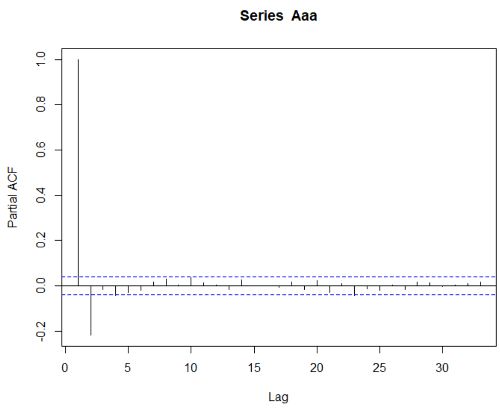
2.11

时序图如下:

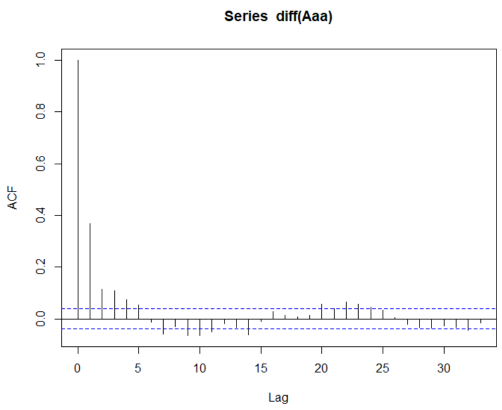


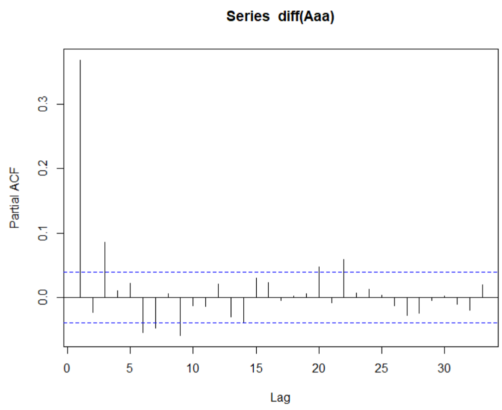
可以看到序列不平稳，绘制ACF和PACF:





ACF显示出序列间具有很强的相关性，进行一阶差分：

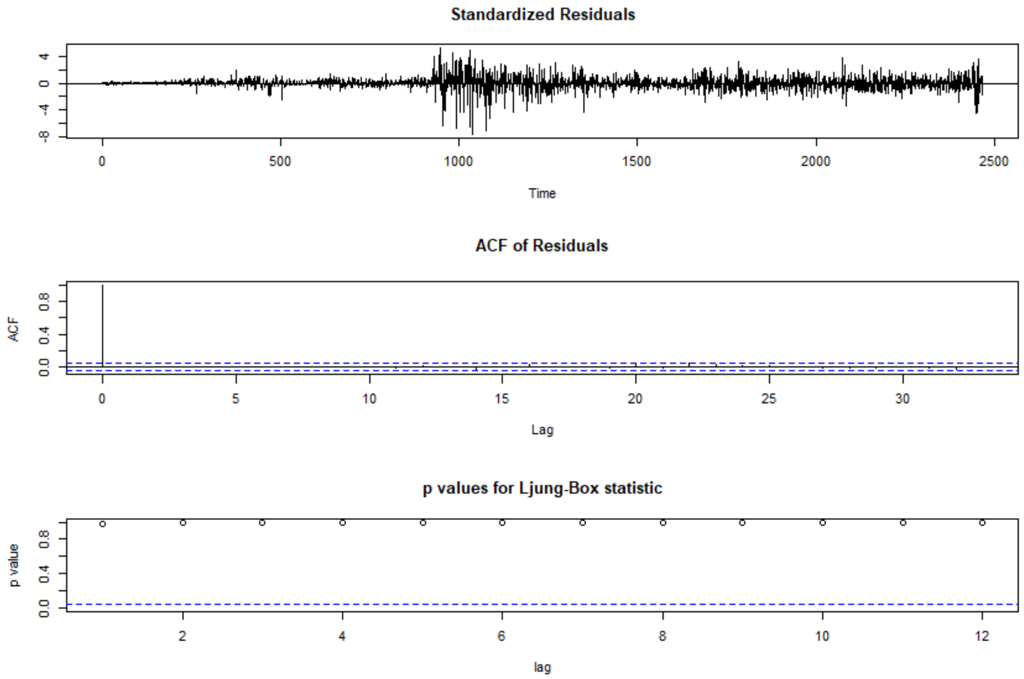




使用ar函数的极大似然法确定阶数为9，建立AR(9)模型：

```
arima(x = Aaa, order = c(9, 1, 0))

Coefficients:
         ar1      ar2     ar3     ar4     ar5      ar6      ar7     ar8      ar9
      0.3772  -0.0579  0.0843  0.0054  0.0392  -0.0314  -0.0531  0.0274  -0.0592
s.e.  0.0201   0.0215  0.0215  0.0216  0.0215   0.0215   0.0215  0.0215   0.0201

sigma^2 estimated as 0.00793:  log likelihood = 2464.86,  aic = -4909.72
```

对模型进行检验：



残差不存在序列相关性。

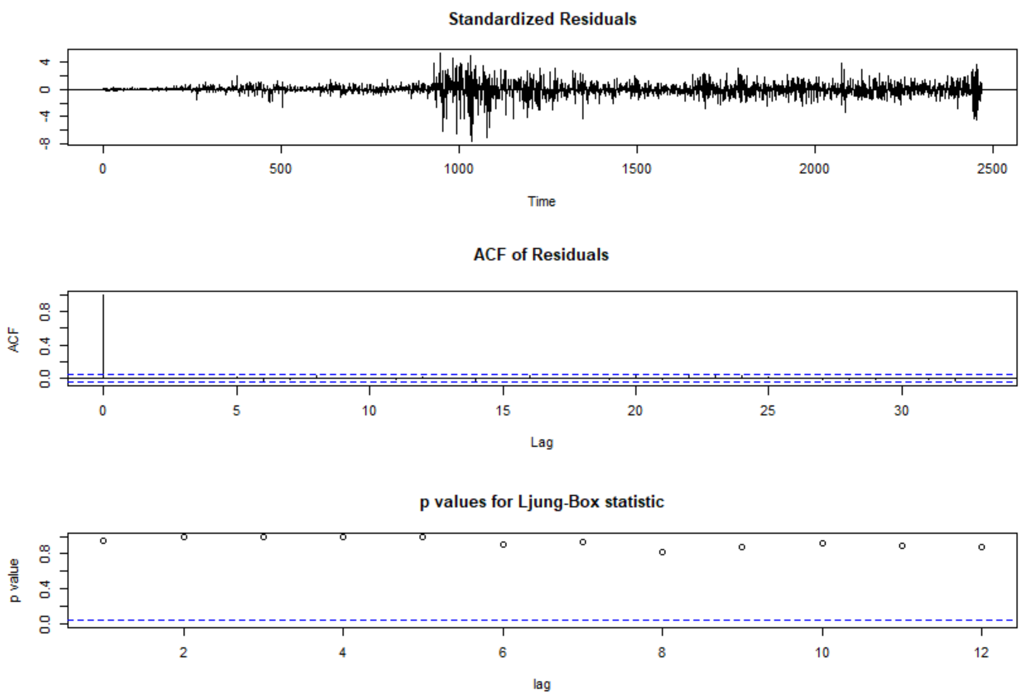
模型复杂度过高，对参数进行显著性检验：

阶数	1	2	3	4	5	6	7	8	9
p-value	0.000	0.01	0.001	0.403	0.046	0.085	0.014	0.113	0.006

重新建模:

```
arima(x = Aaa, order = c(9, 1, 0), fixed = c(NA, NA, NA, 0, NA, 0, NA, 0, NA))

Coefficients:
         ar1      ar2     ar3  ar4     ar5  ar6      ar7  ar8      ar9
      0.3753  -0.0587  0.0859    0  0.0322    0  -0.0541    0  -0.0516
s.e.  0.0200   0.0214  0.0201    0  0.0189    0   0.0189    0   0.0188

sigma^2 estimated as 0.007944:  log likelihood = 2462.77,  aic = -4911.53
```

对模型进行检验:



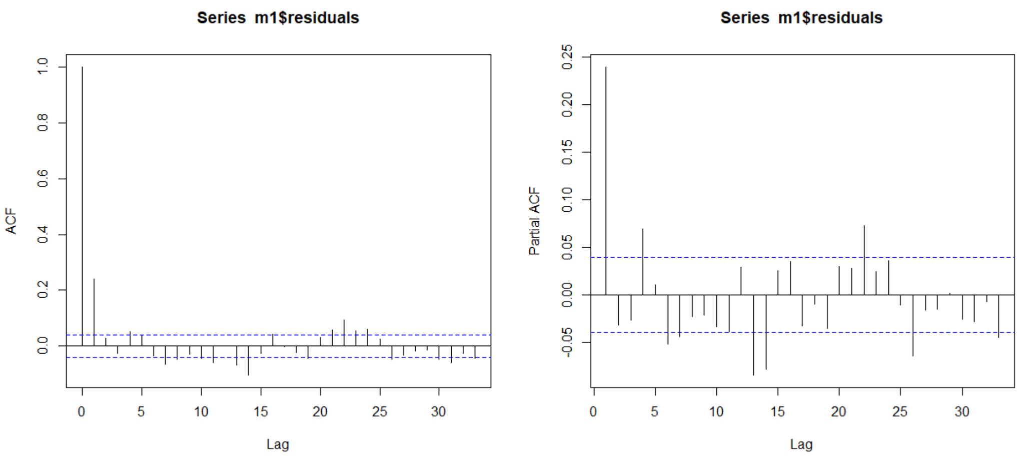
代码:

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

da = read.table("D:/USTC/时间序列分析/data/w-Aaa.txt", header = F)
Aaa = da[, 4]
plot(ts(Aaa), xlab = 'Time', ylab = 'Aaa')

par(mfrow = c(1, 2))
acf(Aaa)
pacf(Aaa)
```

```r
acf(diff(Aaa))
pacf(diff(Aaa))

m=ar(diff(Aaa), method = 'mle')
m$order
m1 = arima(Aaa, order = c(9, 1, 0))
m1
tsdiag(m1, gof = 12)

t1=-0.0579/0.0215
pt(t1,df=12,lower.tail=T)

m2 = arima(Aaa, order = c(9, 1, 0), fixed = c(NA, NA, NA, 0, NA, 0, NA, 0, NA))
m2
tsdiag(m2, gof = 12)
```

2.12

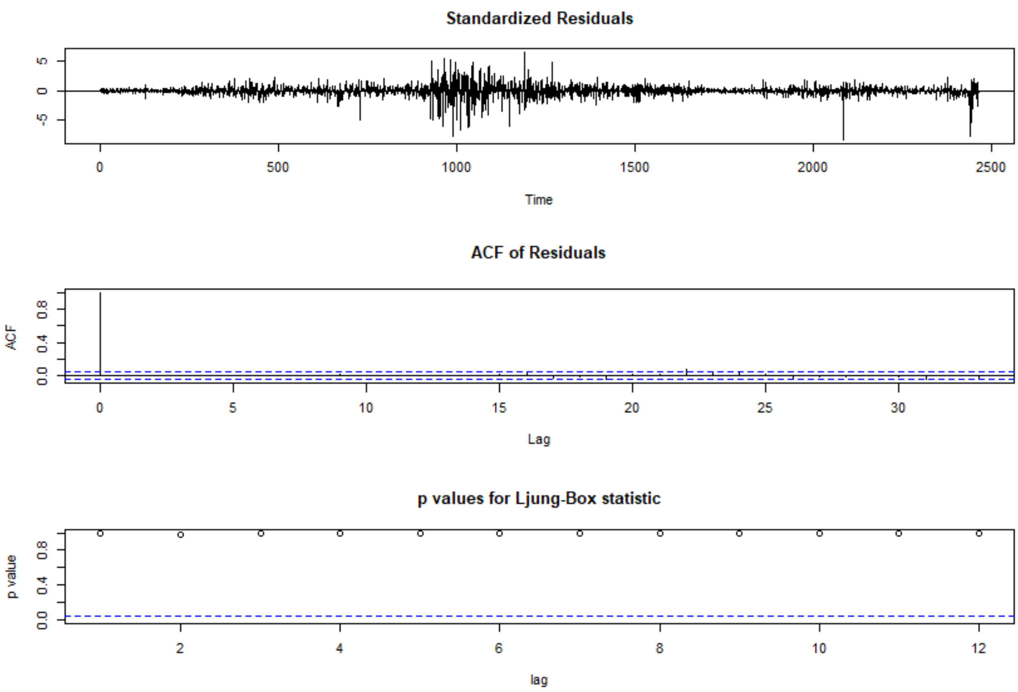
定义 $y_t=(1-B)Y_t$和$x_t=(1-B)X_t$，拟合模型为$y_t=0.946x_t+e_t$，但是残差具有很强的相关性，绘制残差的ACF和PACF:



选取模型ARMA(14,1):

```
arima(x = diff(Aaa), order = c(14, 0, 1), xreg = diff(Baa), include.mean = F)

Coefficients:
         ar1     ar2      ar3     ar4     ar5      ar6      ar7      ar8      ar9     ar10     ar11    ar12
      0.0920  0.0149  -0.0520  0.0531  0.0302  -0.0391  -0.0516  -0.0289  -0.0092  -0.0237  -0.0548  0.0404
s.e.  0.1644  0.0448   0.0205  0.0214  0.0225   0.0205   0.0213   0.0213   0.0204   0.0202   0.0206  0.0215
         ar13     ar14     ma1  diff(Baa)
      -0.0572  -0.0921  0.1479     0.9369
s.e.   0.0215   0.0233  0.1645     0.0131

sigma^2 estimated as 0.002611:  log likelihood = 3834.77,  aic = -7635.53
```

模型检验:



残差不存在相关性。

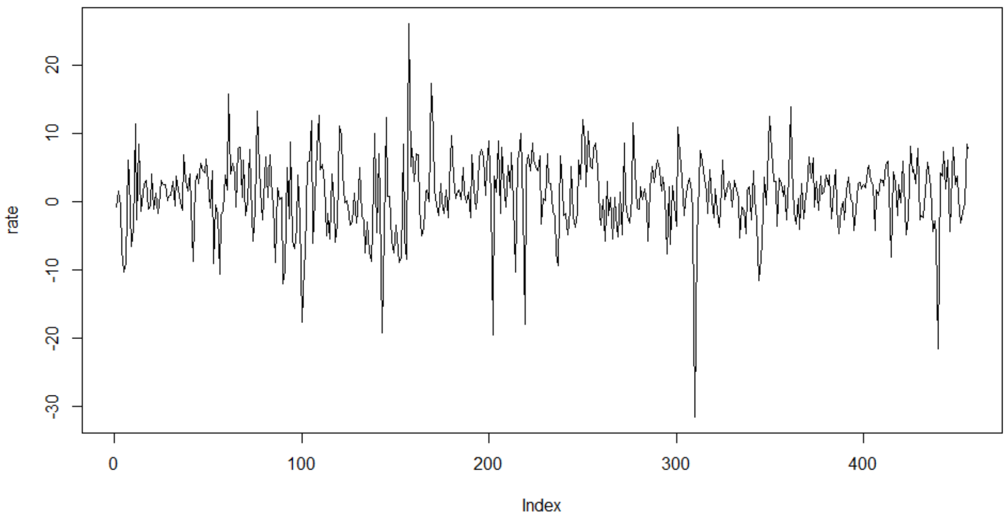
代码:

```
library(fBasics)

da = read.table("D:/USTC/时间序列分析/data/w-Aaa.txt", header = F)
Aaa = da[, 4]
da = read.table("D:/USTC/时间序列分析/data/w-Baa.txt", header = F)
Baa = da[, 4]
m1 = lm(diff(Aaa) ~ diff(Baa))
summary(m1)
```

```
par(mfrow = c(1, 2))
acf(m1$residuals)
pacf(m1$residuals)

m2 = arima(diff(Aaa), order = c(14, 0, 1), xreg = diff(Baa),
include.mean = F)
m2
tsdiag(m2, gof = 12)
```

2.13

时序图如下：



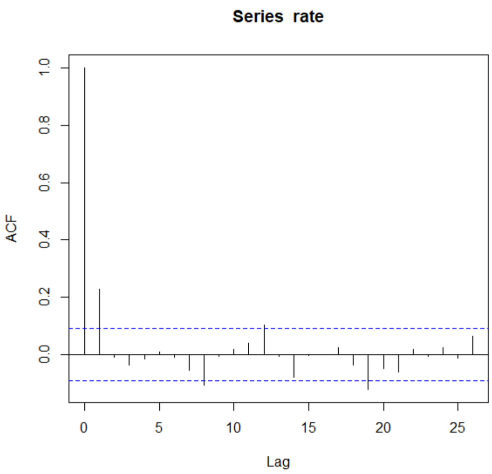
平稳性检验：

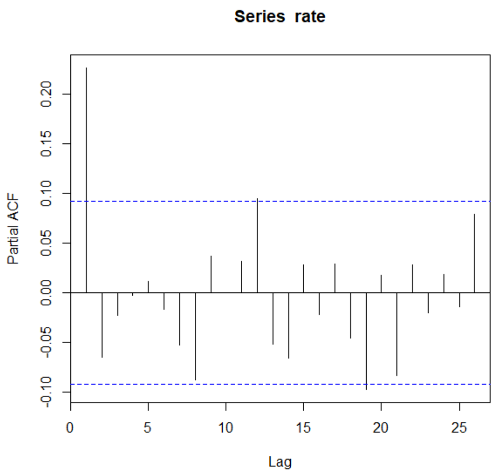
```
        Augmented Dickey-Fuller Test

data:  rate
Dickey-Fuller = -8.498, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

序列平稳，进行白噪声检验：

```
         Box-Ljung test

data:  rate
X-squared = 24.424, df = 6, p-value = 0.0004362


         Box-Ljung test

data:  rate
X-squared = 37.302, df = 12, p-value = 0.0001995


         Box-Ljung test

data:  rate
X-squared = 41.103, df = 18, p-value = 0.001473
```

序列不是白噪声，进行建模，绘制ACF、PACF：

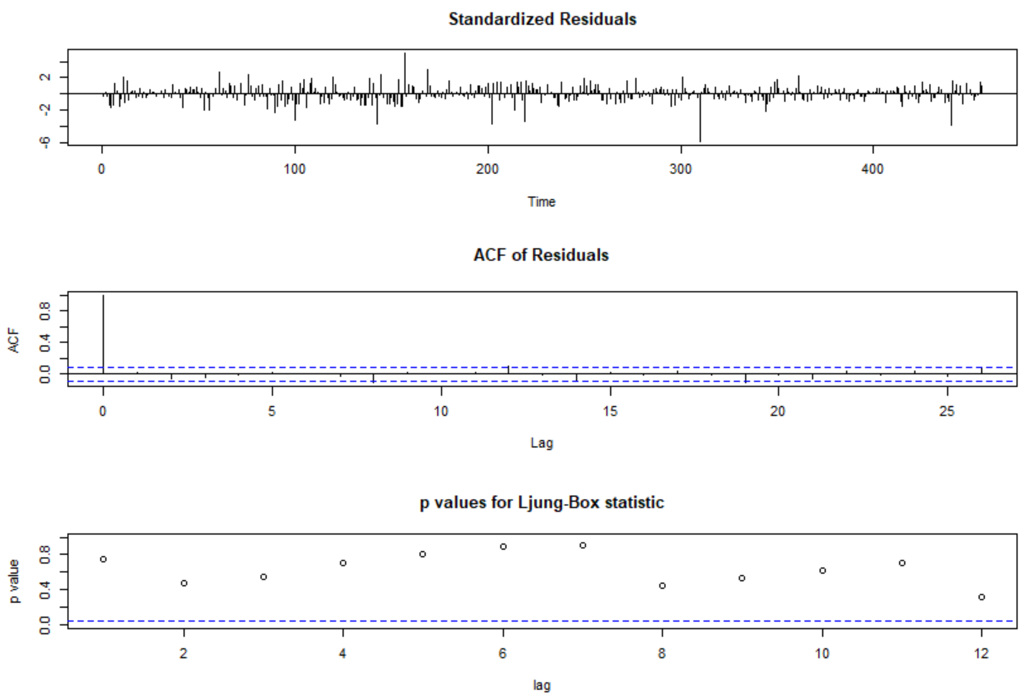




建立AR(1)模型：

```
arima(x = rate, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.2267     1.0626
s.e.  0.0456     0.3297

sigma^2 estimated as 29.68:  log likelihood = -1420.11,  aic = 2846.22
```

模型检验：

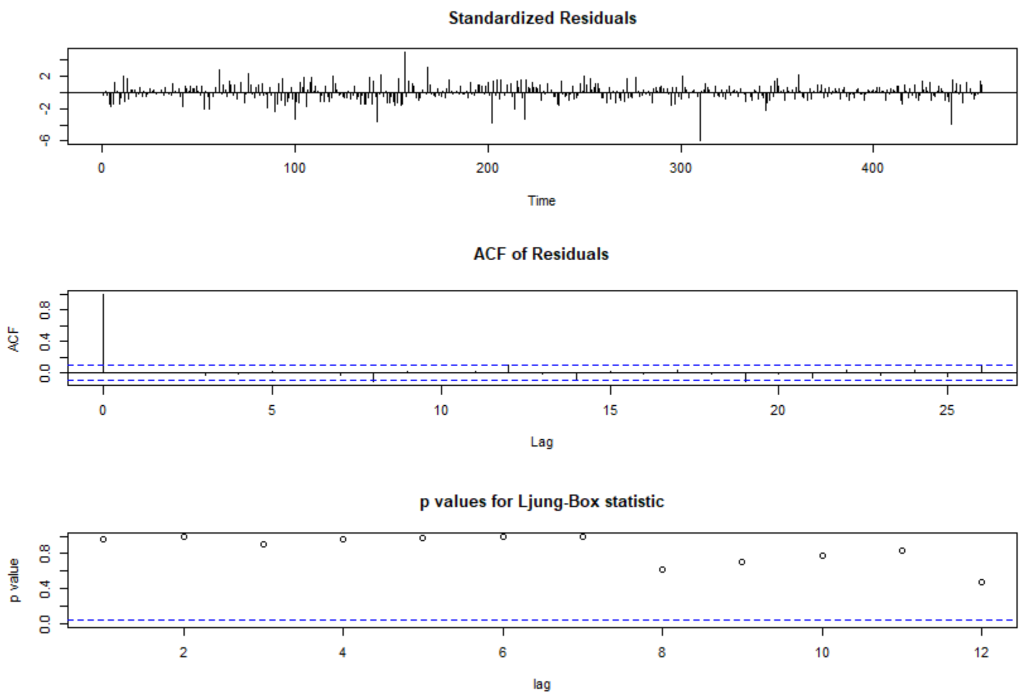


残差没有相关性。

建立MA(1)模型：

```
arima(x = rate, order = c(0, 0, 1))

Coefficients:
         ma1  intercept
      0.2385     1.0605
s.e.  0.0449     0.3153

sigma^2 estimated as 29.59:  log likelihood = -1419.37,  aic = 2844.73
```

模型检验：



残差没有相关性。

预测：

AR(1)：

```
$pred
Time Series:
Start = 457
End = 458
Frequency = 1
[1] 2.601682 1.411453

$se
Time Series:
Start = 457
End = 458
Frequency = 1
[1] 5.448175 5.586364
```

MA(1)：

```
$pred
Time Series:
Start = 457
End = 458
Frequency = 1
[1] 2.250303 1.060512

$se
Time Series:
Start = 457
End = 458
Frequency = 1
[1] 5.439245 5.591797
```

将AR(1)写为：

$$(1 - 0.227B)r_t = 0.826 + a_t$$

将MA(1)写为：

$$r_t = 1.061 + a_t + 0.239a_{t-1}$$

对AR(1)利用长除法：

$$r_t = 1.069 + a_t + 0.227a_{t-1} + 0.052a_{t-2} + 0.012a_{t-3} + 0.003a_{t-4} + \cdots$$

与MA(1)模型非常接近。因此，这两个模型在本质上是等价的。

代码：

```r
library(fBasics)

da = read.table("D:/USTC/时间序列分析/data/m-ew6299.txt", header = F)
rate = da[, 1]
plot(rate, type = 'l', ylab = 'rate')

adf.test(rate)
```

```r
for(i in 1:3) print(Box.test(rate,type = "Ljung-Box",lag=6*i))
par(mfrow = c(1, 2))
acf(rate)
pacf(rate)

m1 = arima(rate, order = c(1, 0, 0))
m1
tsdiag(m1, gof = 12)

m2 = arima(rate, order = c(0, 0, 1))
m2
tsdiag(m2, gof = 12)
predict(m1, 2)
predict(m2, 2)
```