阅读参考书《时间序列分析及应用》的第 10 章并完成课后习题

P179——10.11　10.12

相应数据文件可以在 http://homepage.stat.uiowa.edu/~kchan/TSA.htm 下载

10.11　美国 Johnson & Johnson 公司于 1960～1980 年间每股收益的季度数据见于文件 JJ 中.

(a) 画出该序列及其取对数后的时间序列图．论证对序列进行对数变换的必要性.

(b) 序列明显是非平稳的．对其进行一次差分变换并画出序列图．现在序列平稳性有无合理性?

(c) 计算并画出经一次差分后序列的样本 ACF，并解释结果.

(d) 画出并解释经过一次差分和季节差分后的序列图．牢记季度数据一季的长度为 4.

(e) 画出并说明经过一次差分和季节差分后的序列的样本 ACF.

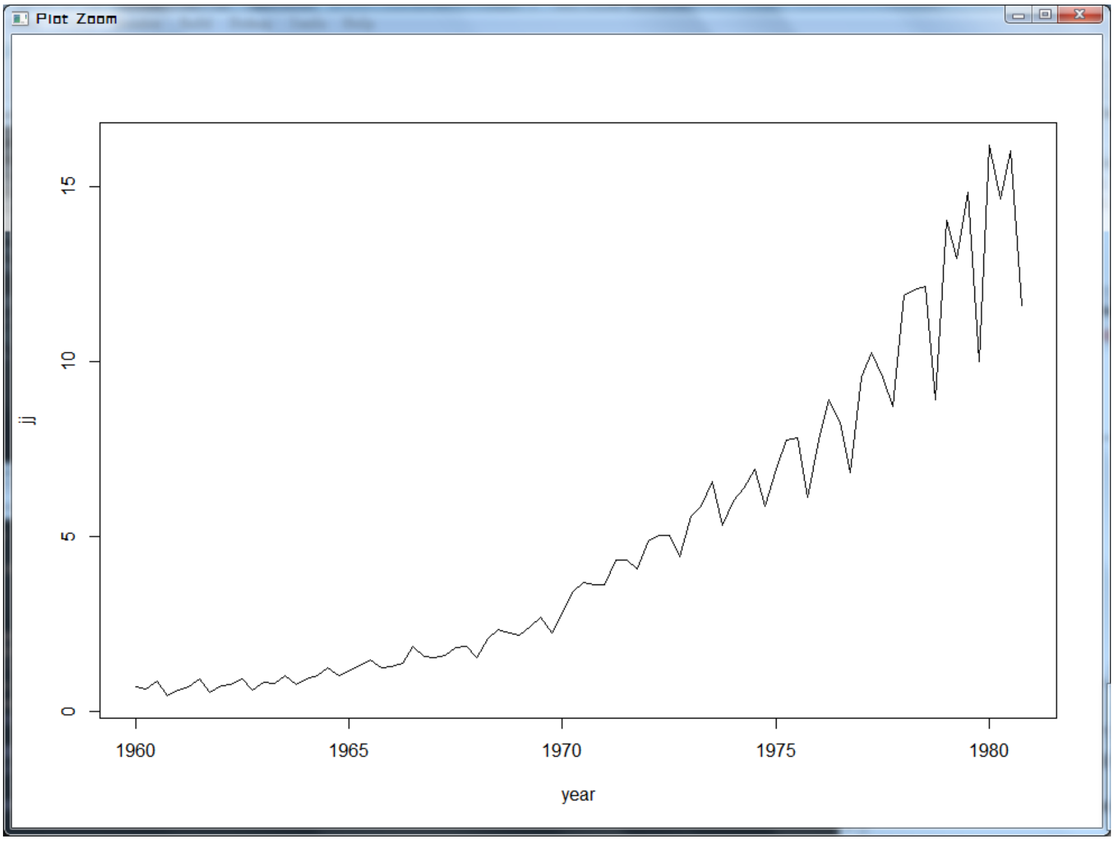
(f) 拟合 ARIMA(0，1，1)×(0，1，1)$_4$ 模型，并评估系数估计值的显著性.

(g) 对残差进行所有的诊断性检验.

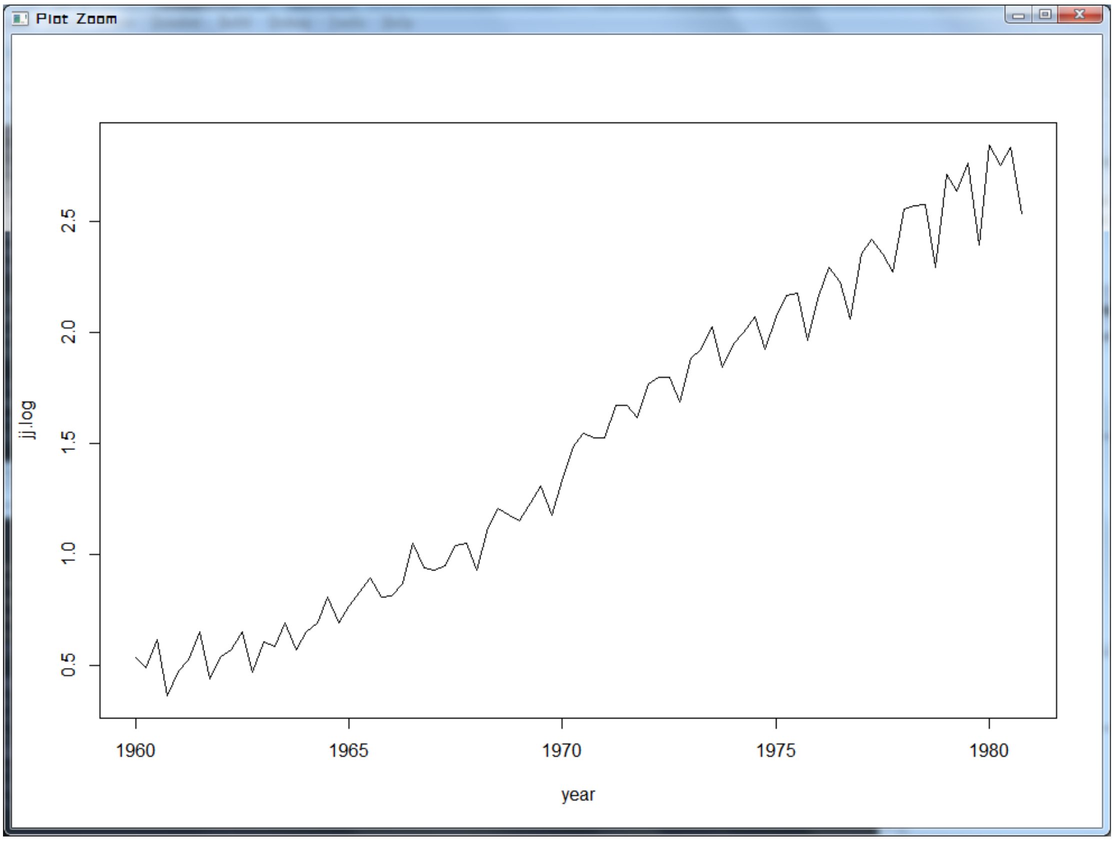
(h) 计算并画出序列未来两年的预测值，要求给出预测极限.

（a）

```r
library(xts)
library(xtsExtra)
library(quantmod)
library(forecast)
library(urca)
library(FinTS)
library(rugarch)
library(TSA)
library(fUnitRoots)

# 原序列的时间序列图
jj<-read.table("E:/DATA/data mining/fts06/Datasets/JJ.dat",header=T)
rtn = ts(jj, frequency = 4, start = c(1960, 1))  #转化为时间序列
plot(rtn, type = "l", xlab = "year", ylab = "jj")  # 画时间序列
```




```r
# 取对数后的时间序列图
jj.log <- log(jj$JJ + 1)  #计算对数收益率
rtn = ts(jj.log, frequency = 4, start = c(1960, 1))  #转化为时间序列
plot(rtn, type = "l", xlab = "year", ylab = "jj.log")  # 画时间序列图
```

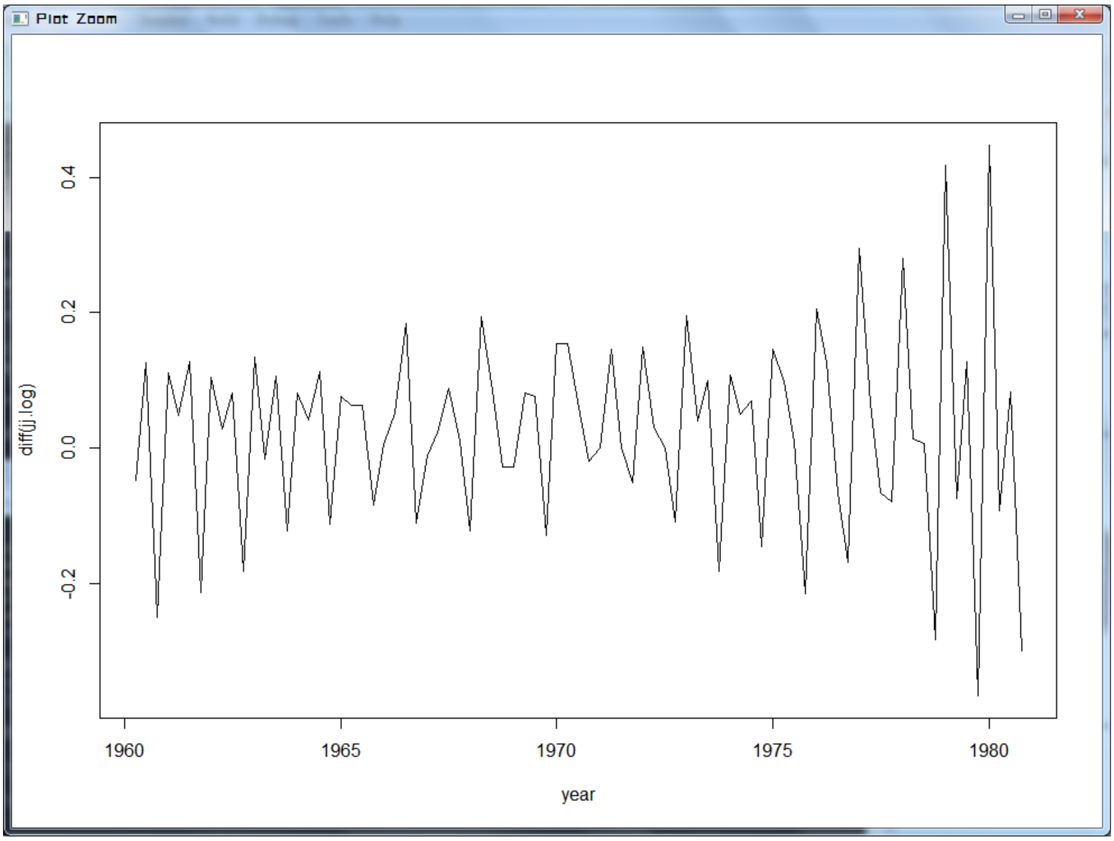


取对数后波动比较平稳

（b）

序列非平稳，经一阶差分后

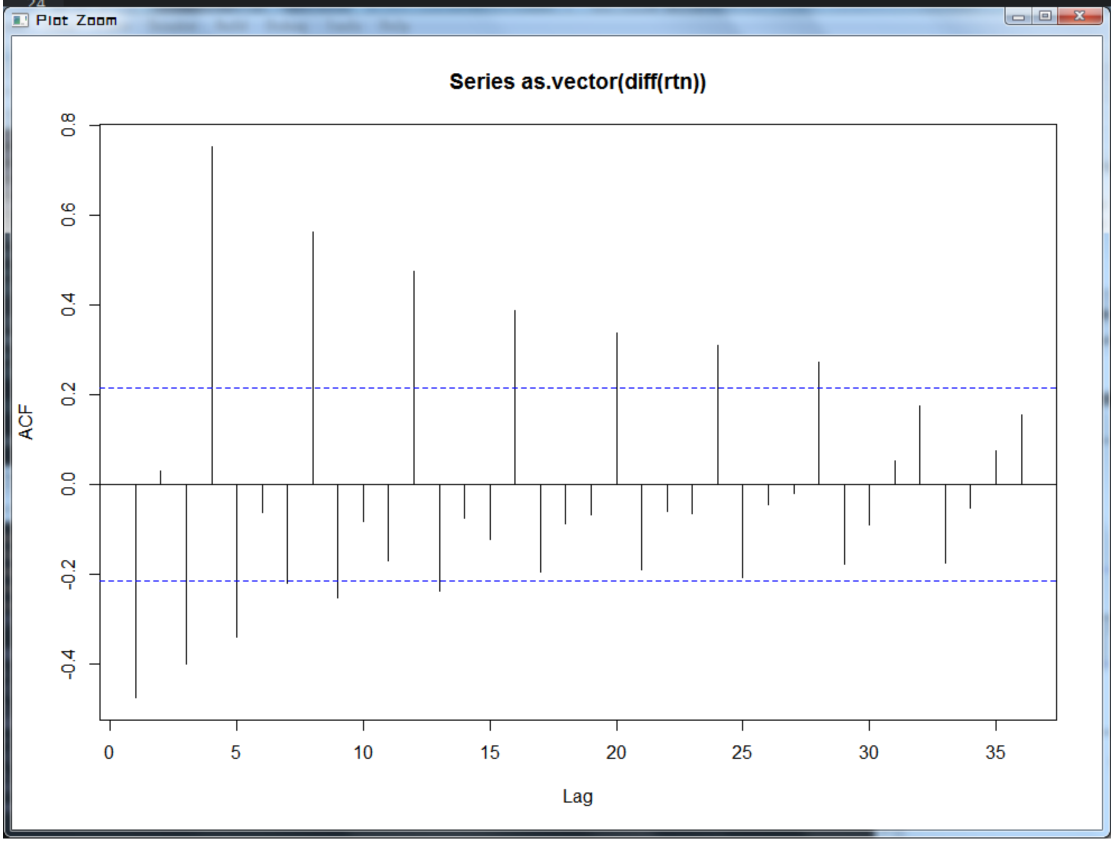
```
plot(diff(rtn), type = "l", xlab = "year", ylab = "diff(jj.log)")  # 画时间序列图
```



不是平稳时间序列

（c）

```
acf(as.vector(diff(rtn)), lag.max = 36)
```

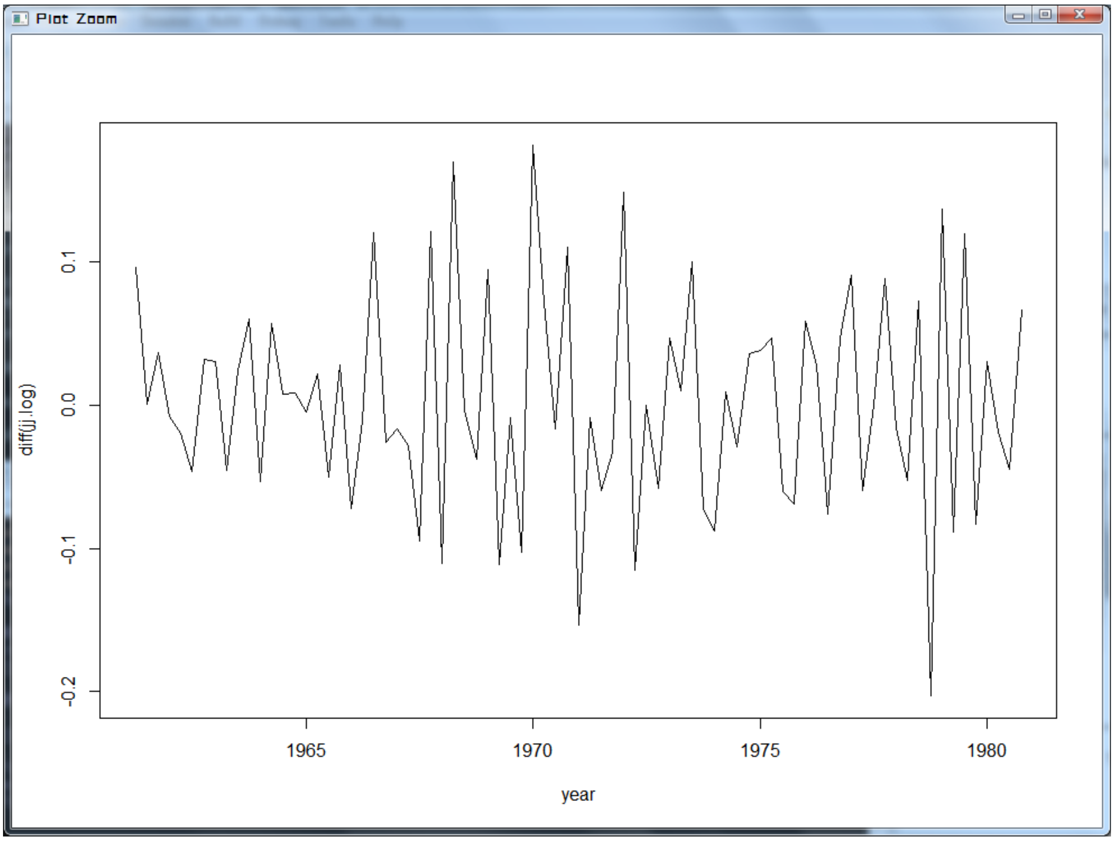


自相关性在 4，8，12，16，20，24，28 显著，呈现季节相关性

（d）

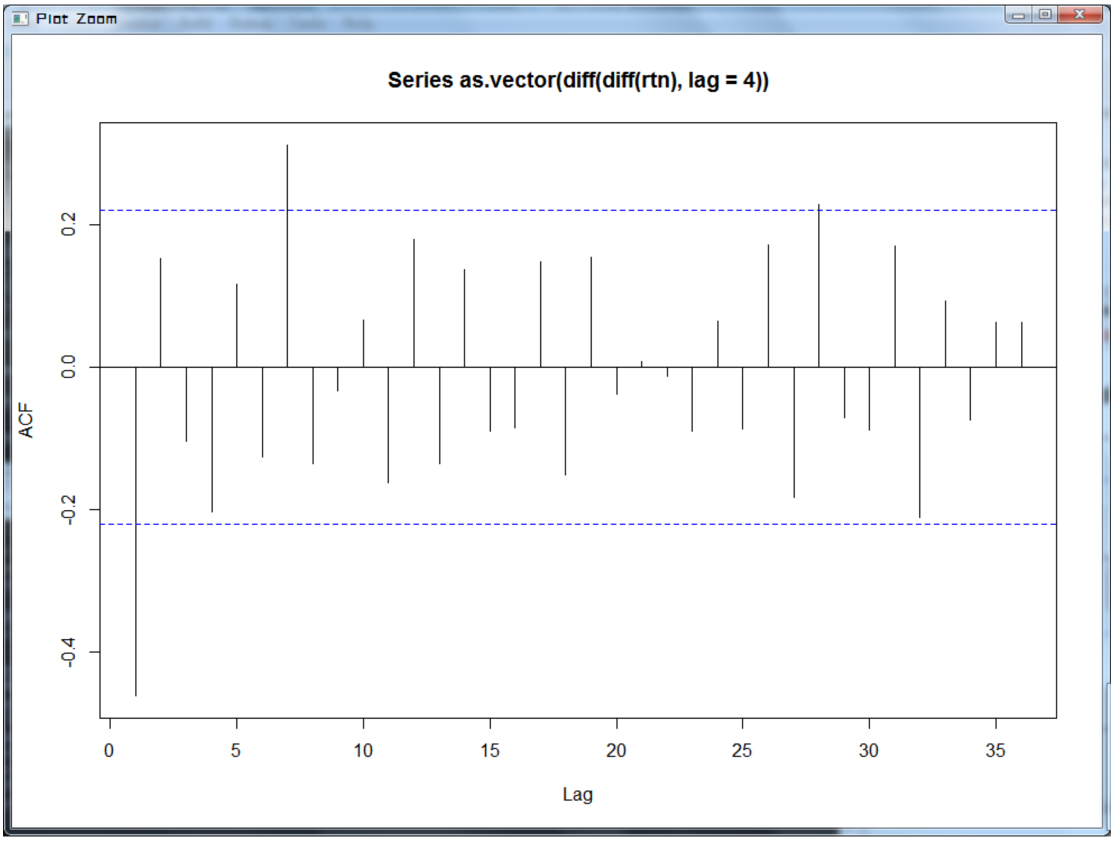
季节差分时间序列图

```r
plot(diff(diff(rtn), lag = 4), type = "l", xlab = "year", ylab = "diff(jj.log)")  # 画时间序列图
```



（e）

```r
acf(as.vector(diff(diff(rtn), lag = 4)), lag.max = 36)
```



对比一阶差分，季节差分的 ACF 图的季节性显著特征已经消除

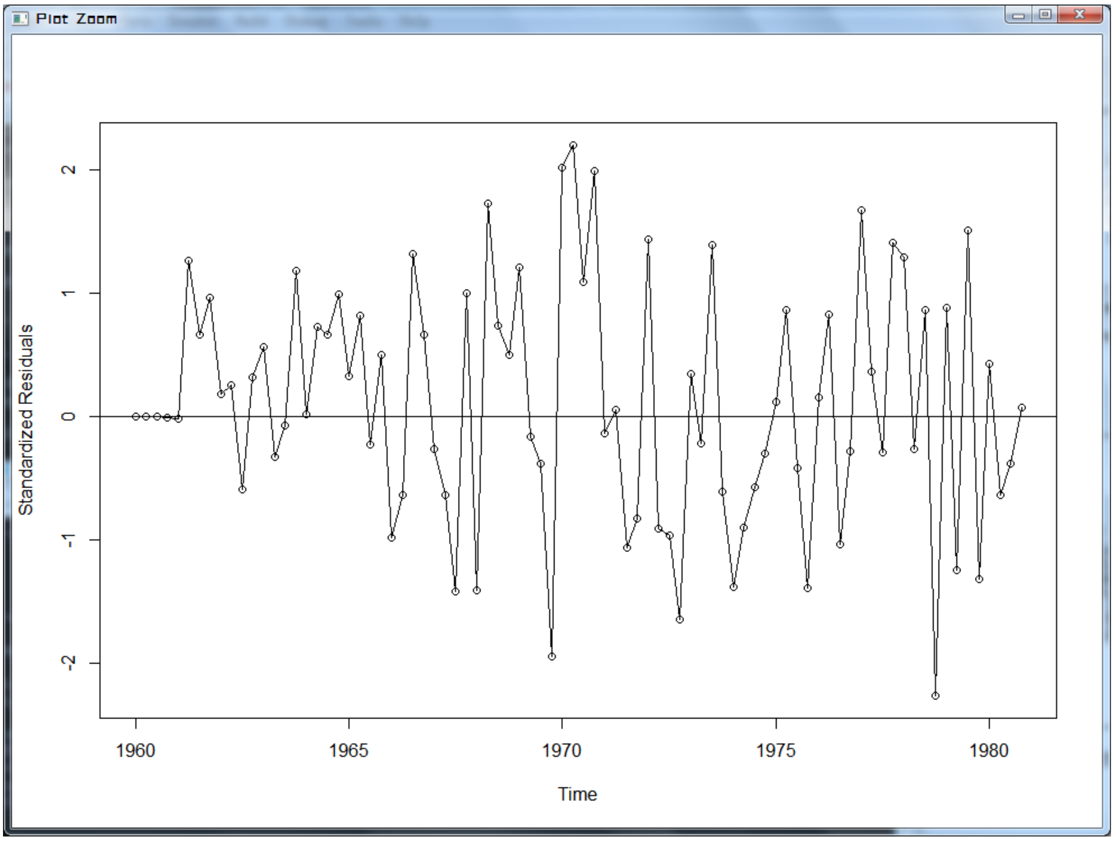
（f）

```
> m1.jj <- arima(rtn, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1),
+                                                      period = 4))
> m1.jj

Call:
arima(x = rtn, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1), period = 4))

Coefficients:
          ma1     sma1
      -0.6612  -0.3310
s.e.   0.0900   0.1093

sigma^2 estimated as 0.00365:  log likelihood = 109.03,  aic = -214.06
```

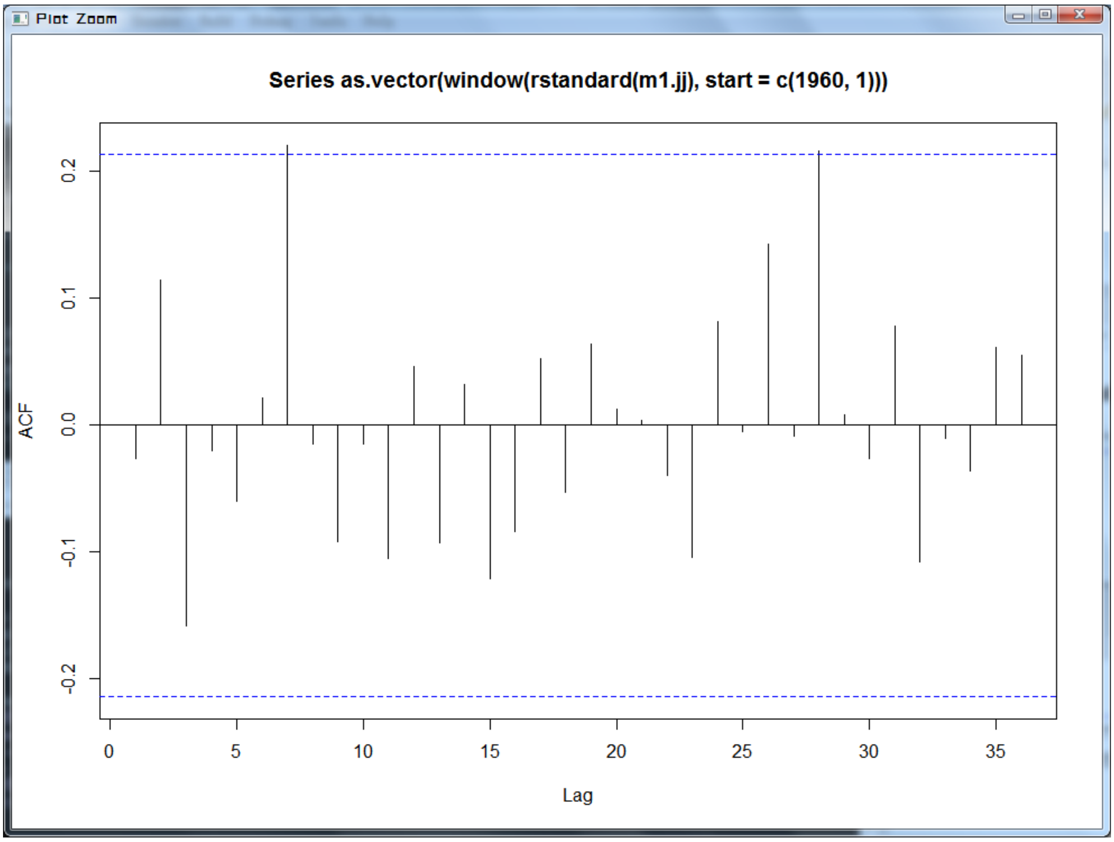
系数估计值高度显著

（g）

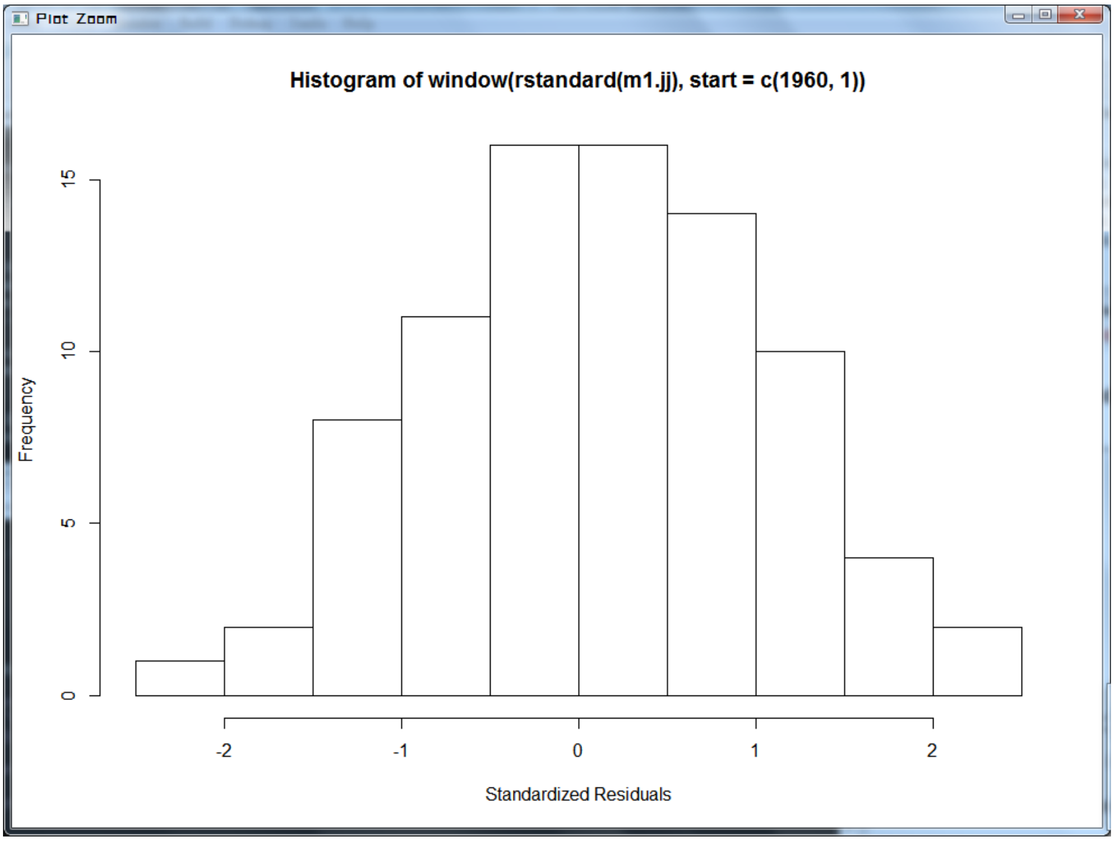
```
# 模型的残差图
plot(window(rstandard(m1.jj), start = c(1960, 1)), ylab = "Standardized Residuals",
     type = "o")
abline(h = 0)
```



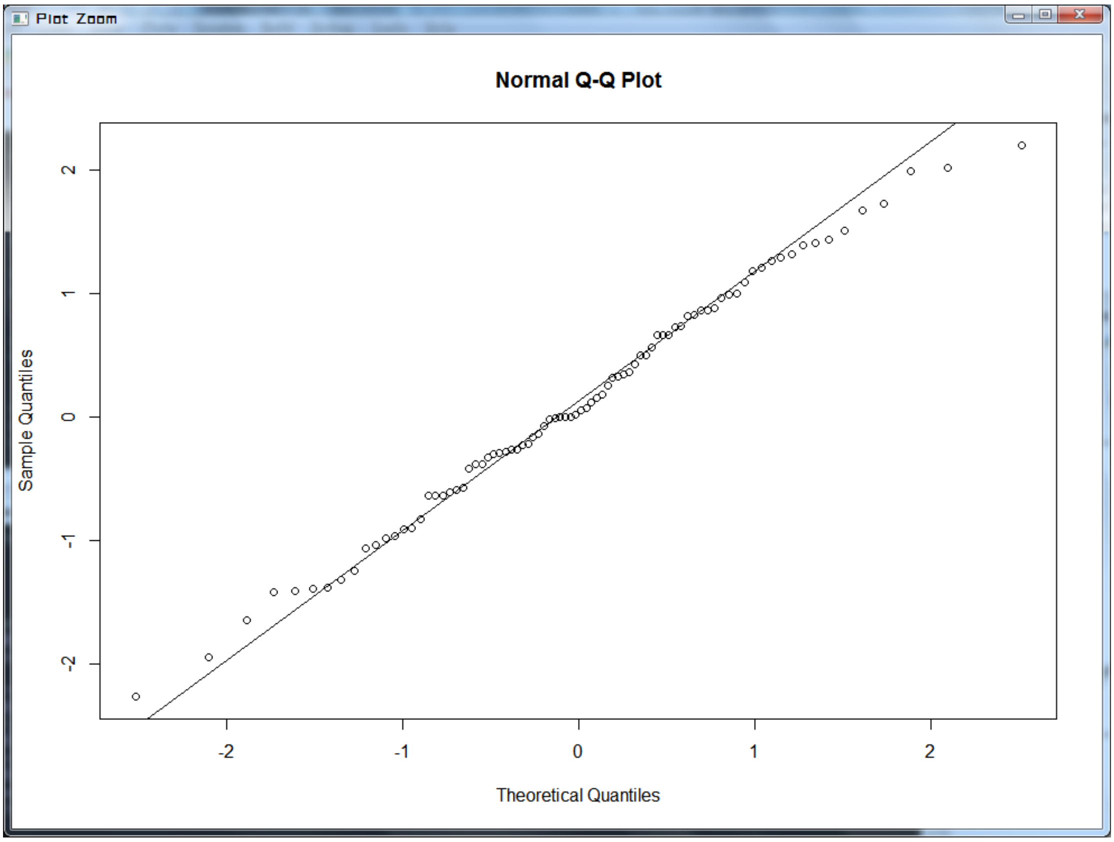

```
# 残差的ACF图
acf(as.vector(window(rstandard(m1.jj), start = c(1960, 1))), lag.max = 36)
```




```
# 残差的直方图
hist(window(rstandard(m1.jj), start = c(1960, 1)), xlab = "Standardized Residuals")
```

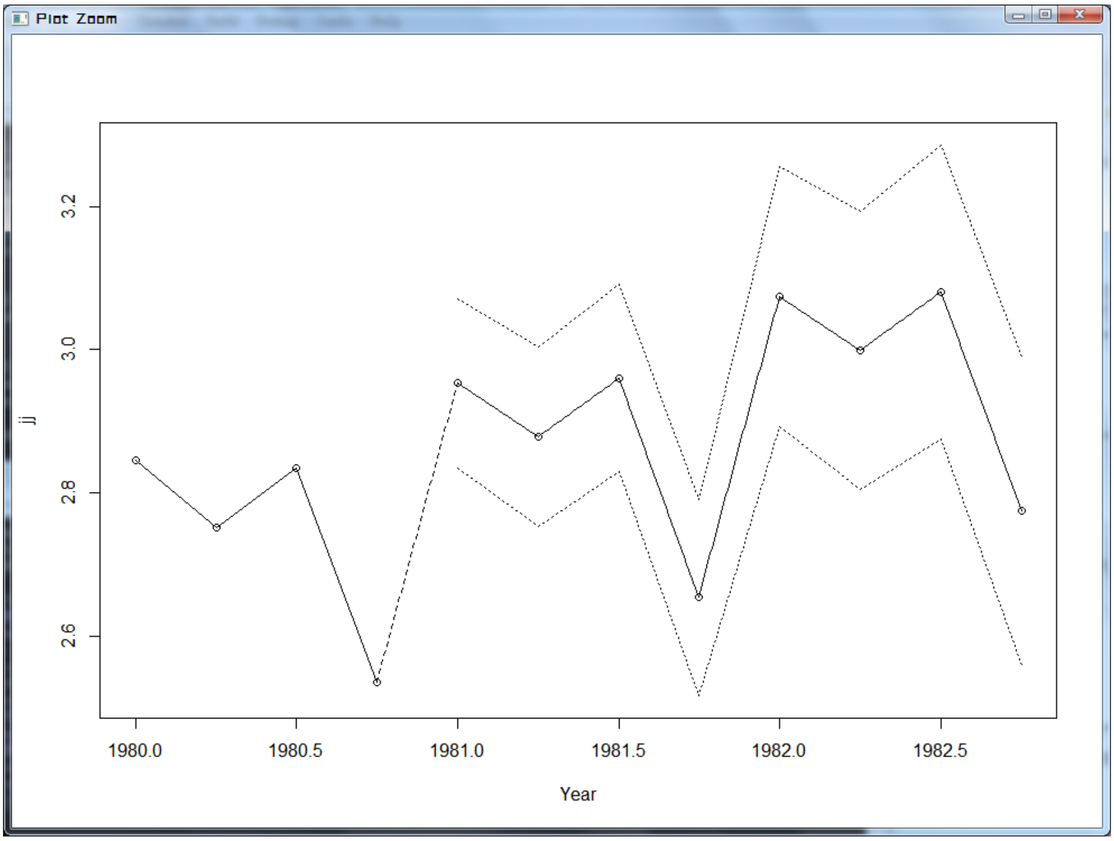



```
# 残差的QQ图
qqnorm(window(rstandard(m1.jj), start = c(1960, 1)))
qqline(window(rstandard(m1.jj), start = c(1960, 1)))
```



（h）

```
plot(m1.jj, n1 = c(1980, 1), n.ahead = 8, xlab = "Year", type = "o", ylab = "jj")
```



10.12 文件 boardings 中保存的是 2000 年 8 月至 2005 年 12 月间在美国科罗拉多州的丹佛搭乘交通工具（多乘轻轨火车和城市巴士）人数的月度数据.

(a) 画出时间序列图. 要求使用有助于评估季节性的绘图符号. 应用平稳模型合理吗?

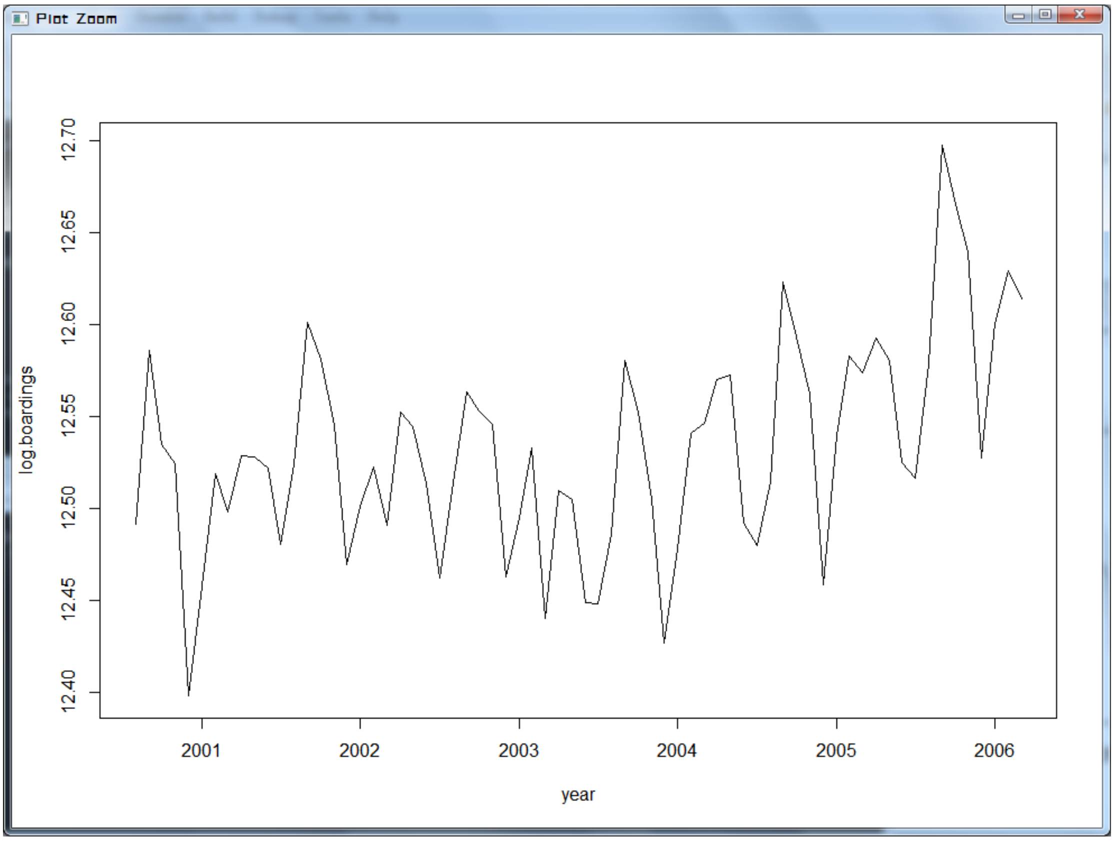
(b) 计算并画出序列的样本 ACF. 当滞后为多少时，存在显著的自相关性?

(c) 为数据拟合一个 ARMA(0, 3)×(1, 0)$_{12}$模型. 评估系数估计值的显著性.

(d) ARMA(0, 4)×(1, 0)$_{12}$模型是过度拟合的，解释这一结果.

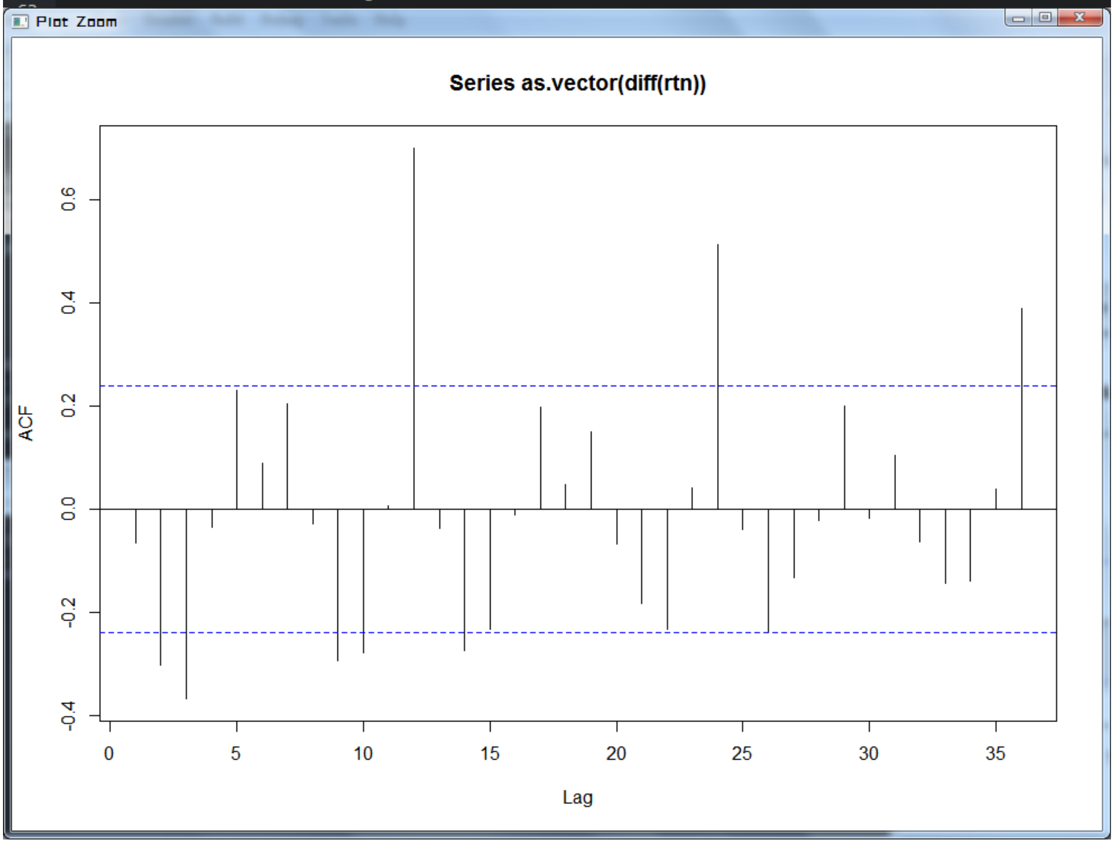
（a）

```
boardings <- read.table("E:/DATA/data mining/fts06/Datasets/boardings.dat", header = T)
rtn = ts(boardings$log.boardings, frequency = 12, start = c(2000, 8))  #转化为时间序列
plot(rtn, type = "l", xlab = "year", ylab = "log.boardings")  # 画时间序列
```



（b）

```
61  acf(as.vector(diff(rtn)), lag.max = 36)
```



12 阶具有明显自相关性

（c）

```
m1.boardings <- arima(rtn, order = c(0, 0, 3), seasonal = list(order = c(1,
                                                                    0, 0), period = 12))
m1.boardings
```

```
> m1.boardings <- arima(rtn, order = c(0, 0, 3), seasonal = list(order = c(1,
+                                                                     0, 0), period = 12))
> m1.boardings

Call:
arima(x = rtn, order = c(0, 0, 3), seasonal = list(order = c(1, 0, 0), period = 12))

Coefficients:
         ma1     ma2     ma3    sar1  intercept
      0.7290  0.6116  0.2950  0.8776    12.5455
s.e.  0.1186  0.1172  0.1118  0.0507     0.0354

sigma^2 estimated as 0.0006543:  log likelihood = 143.54,  aic = -277.09
```

（d）

```
m2.boardings <- arima(rtn, order = c(0, 0, 4), seasonal = list(order = c(1,
                                                                    0, 0), period = 12))
m2.boardings
```

```
> m2.boardings <- arima(rtn, order = c(0, 0, 4), seasonal = list(order = c(1,
+                                                                     0, 0), period = 12))
> m2.boardings

Call:
arima(x = rtn, order = c(0, 0, 4), seasonal = list(order = c(1, 0, 0), period = 12))

Coefficients:
         ma1     ma2     ma3     ma4    sar1  intercept
      0.7276  0.6685  0.4244  0.1414  0.8918    12.5459
s.e.  0.1212  0.1327  0.1681  0.1228  0.0445     0.0419

sigma^2 estimated as 0.0006279:  log likelihood = 144.22,  aic = -276.45
```

两次参数变化不大，可以认为 ARMA(0,4)*(1,0)12 模型是过度拟合