第四作业：案例分析（二）

金科 201756010

要求：利用 R 软件分析 Intel 公司股票的月对数收益率数据，构建波动率模型并做向前 5 步预测，要求包含建模过程与代码。数据文件见附件，其中 `date` 表示日期，`intc` 变量即是要分析的对数收益率序列。

首先读取数据并进行初步的描述统计，

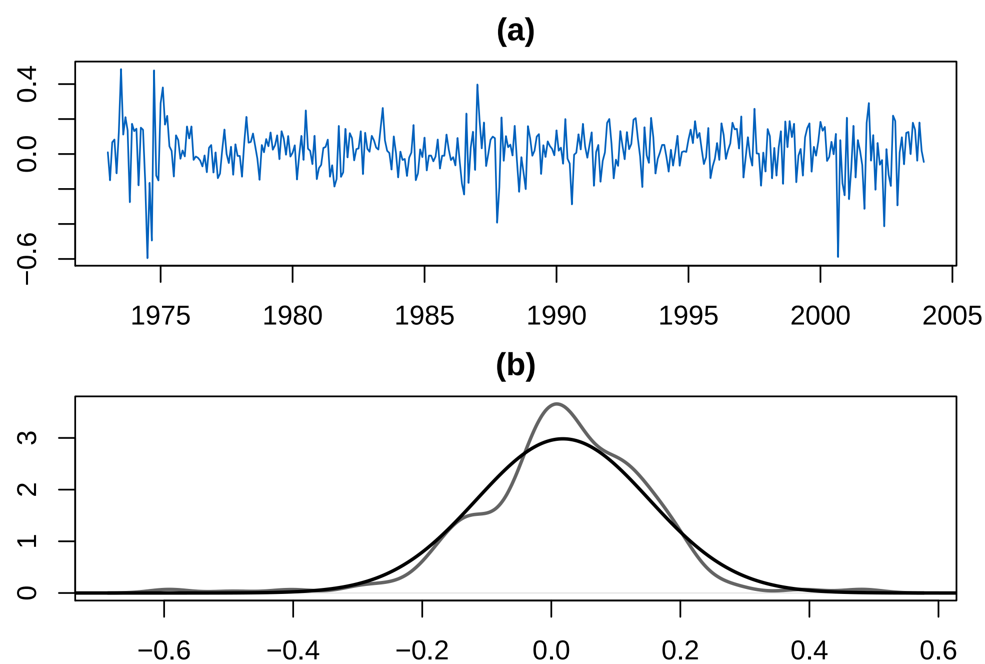
```
w <- read.table("m-intel.txt", sep = "", header = TRUE)
x <- ts(w$intc, start = c(1973, 01), frequency = 12)
```

```
stat <- psych::describe(x)
stat
```

```
##    vars   n mean   sd median trimmed  mad  min  max range skew kurtosis   se
## X1    1 372 0.02 0.13   0.02    0.02 0.12 -0.6 0.49  1.08 -0.6     2.89 0.01
```

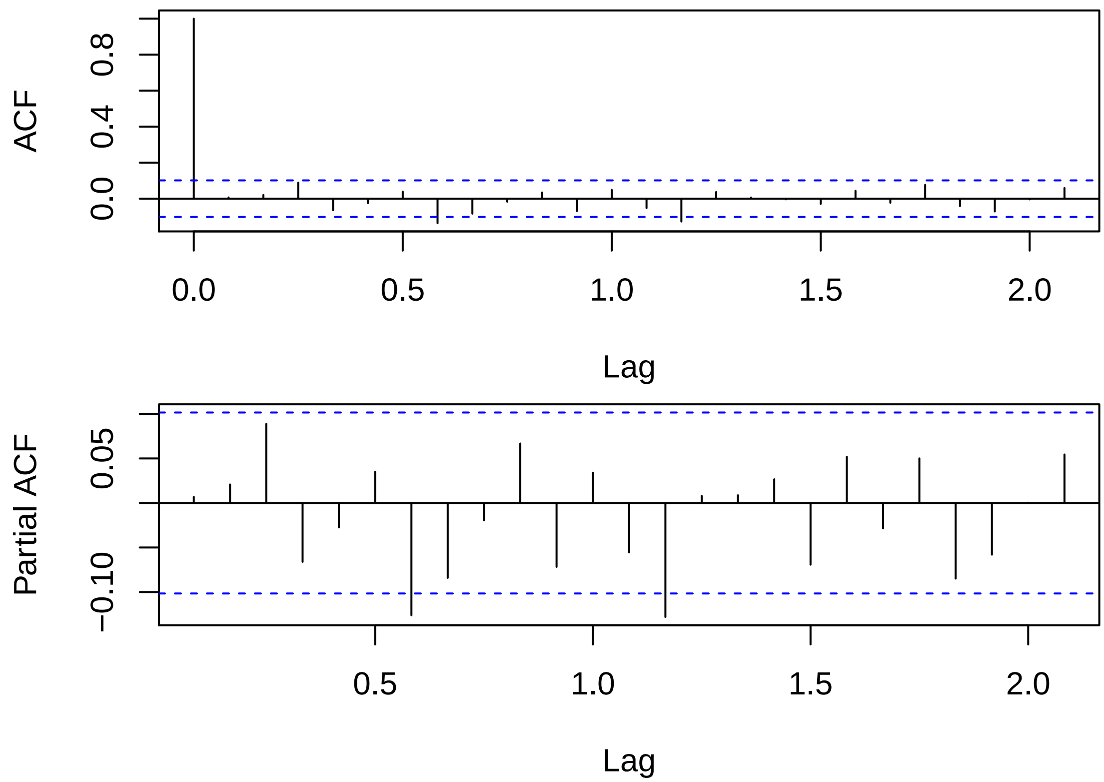
绘制时序图，如 (a) 所示，可以看出序列围绕着稳定的均值水平，并无明显的趋势，但是其波动性却明显呈现出聚集性的特点；图 (b) 与同均值方差的正态分布密度曲线相比较可以看出，月对数收益率具有尖峰的特征。

```
par(mfrow = c(2, 1), mar = c(2,2,2,1), oma = c(1,0.5,1,2))
plot(x, col = "#0061bd", main = "(a)")
d <- density(x)
plot(d, col = "#646464", main = "(b)", lwd = 2)
s2 <- seq(from = -4, to = 4, length.out = 10000)
lines(s2, dnorm(s2, mean = stat$mean, sd = stat$sd), col = "black", lwd = 2)
```



接下来拟合均值方程，作 ACF、PACF 图如下所示，从两幅图中都可以看出呈现拖尾的情形，因此可以选择 **ARMA(1,1)** 模型。

```r
par(mfrow = c(2, 1), mar = c(4,4,0.5,1), oma = c(1,1,1,2))
acf(x, main = "")
pacf(x, main = "")
```



拟合 **ARMA(1,1)** 模型，并对残差序列进行白噪声检验，原假设 $H_0:,$ 序列为白噪声，输出对应的 P 值

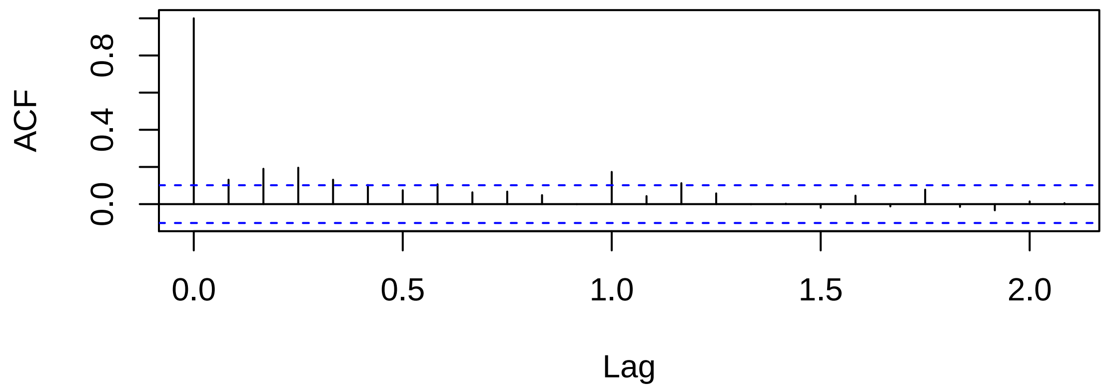
```r
x.fit <- arima(x, order = c(1, 0, 1))

sapply(1:6, function(i) {
  Box.test(x.fit$residual, type = "Ljung-Box", lag = i) $ p.value
})
```

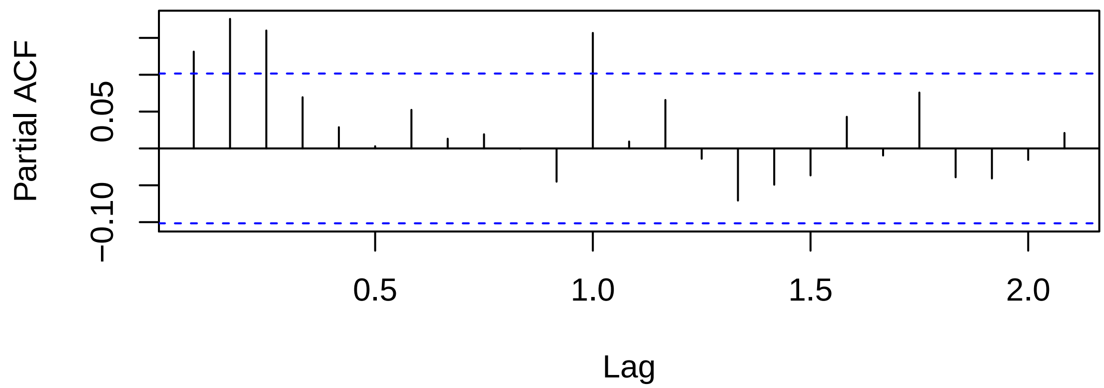
```
## [1] 0.7790473 0.9449843 0.4138305 0.3314833 0.4381149 0.4789648
```

可以认为 **ARMA(1,1)** 基本提取了序列中的自相关结构。

接下来检验对数收益率序列是否包含 ARCH 效应，绘制残差平方序列的 ACF 图和 PACF 图

```r
par(mfrow = c(2, 1), mar = c(4,4,0.5,1), oma = c(1,1,1,2))
acf(x.fit$residuals^2, main = "")
pacf(x.fit$residuals^2, main = "")
```





对残差平方序列进行 Portmantea Q 检验，原假设 $H_0:,$ 残差平方序列纯随机，即方差齐性，P 值结果如下

```
sapply(1:6, function(i) {
  Box.test(x.fit$residual^2, type = "Ljung-Box", lag = i) $ p.value
})
```

```
## [1] 1.095670e-02 4.419376e-05 1.517263e-07 2.598256e-08 1.386019e-08
## [6] 1.699952e-08
```

前六阶检验的 P 值都特别小，说明存在低阶的显著的自相关性，存在 ARCH 效应。根据 ACF 和 PACF 图的情况，将波动方程定阶为 **GARCH(1,1)**，结合均值方程，模型为 **ARMA(1,1)+GARCH(1,1)**

```
library(fGarch)
```

```
m1 <- garchFit(intc ~ arma(1, 1) + garch(1, 1), data = w, trace = FALSE)
summary(m1)
```

```
## Coefficient(s):
##         mu         ar1         ma1       omega      alpha1       beta1
##  0.0096245   0.4098051  -0.3832953   0.0010924   0.0808624   0.8547395
##
## Error Analysis:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.009624    0.012183    0.790  0.42952
## ar1     0.409805    0.711036    0.576  0.56438
## ma1    -0.383295    0.720425   -0.532  0.59470
## omega   0.001092    0.000530    2.061  0.03928 *
## alpha1  0.080862    0.028380    2.849  0.00438 **
```

```
## beta1   0.854739    0.046413   18.416  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  239.6704    normalized:  0.6442751
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  159.0297  0
##  Shapiro-Wilk Test  R    W      0.967951  2.727838e-07
##  Ljung-Box Test     R    Q(10)  9.558936  0.4800025
##  Ljung-Box Test     R    Q(15)  16.37201  0.3577674
##  Ljung-Box Test     R    Q(20)  17.70885  0.606581
##  Ljung-Box Test     R^2  Q(10)  0.5568293 0.9999889
##  Ljung-Box Test     R^2  Q(15)  9.771177  0.8338812
##  Ljung-Box Test     R^2  Q(20)  11.34174  0.9368743
##  LM Arch Test       R    TR^2   8.768857  0.7225361
##
## Information Criterion Statistics:
##       AIC       BIC       SIC      HQIC
## -1.256292 -1.193084 -1.256802 -1.231191
```

其中 ar1、ma1 项的系数并不显著，说明均值方程相对整个模型波动的影响并不明显；从序列原始的 ACF 图也可以认为其自相关系数 0 阶截尾，均值方程为常数。因此将模型调整为 **GARCH(1,1)** 查看结果

```r
m1 <- garchFit(intc ~ garch(1, 1), data = w, trace = FALSE)
summary(m1)
```

```
## Coefficient(s):
##        mu      omega     alpha1      beta1
## 0.0163276  0.0010918  0.0802716  0.8553014
##
## Error Analysis:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.0163276   0.0062624    2.607  0.00913 **
## omega   0.0010918   0.0005291    2.063  0.03907 *
## alpha1  0.0802716   0.0281162    2.855  0.00430 **
## beta1   0.8553014   0.0461374   18.538  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  156.5138  0
##  Shapiro-Wilk Test  R    W      0.9676933 2.471139e-07
##  Ljung-Box Test     R    Q(10)  9.805485  0.4577215
##  Ljung-Box Test     R    Q(15)  16.54435  0.346824
##  Ljung-Box Test     R    Q(20)  17.8005   0.6005484
##  Ljung-Box Test     R^2  Q(10)  0.5130171 0.9999925
##  Ljung-Box Test     R^2  Q(15)  10.24557  0.8040151
##  Ljung-Box Test     R^2  Q(20)  11.77988  0.9234441
##  LM Arch Test       R    TR^2   9.334459  0.6741288
##
## Information Criterion Statistics:
```

```
##        AIC       BIC       SIC      HQIC
## -1.266231 -1.224092 -1.266459 -1.249496
```

此时所有参数都表现为显著，并且 AIC 和 BIC 的数值也有所下降。

对残差项进行 Kolmogorov-Smirnov 正态性检验，原假设 $H_0:,$ 残差序列服从正态分布。

```
ks.test(x.fit$residuals,
        dnorm(s2, mean = stat$mean, sd = stat$sd),
        alternative = c("two.sided", "less", "greater"),
        exact = NULL)
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  x.fit$residuals and dnorm(s2, mean = stat$mean, sd = stat$sd)
## D = 0.49194, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

P 值几乎接近 0，因此误差项分布显著异于正态分布，设定为 t 分布重新拟合模型

```
m1 <- garchFit(intc ~ garch(1, 1), data = w,
               cond.dist = "std", trace = FALSE)
summary(m1)
```

```
## Coefficient(s):
##        mu      omega     alpha1      beta1      shape
## 0.0219121  0.0014771  0.1036656  0.8085661  6.6892311
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##         Estimate  Std. Error  t value Pr(>|t|)
## mu      0.0219121   0.0059313    3.694 0.000220 ***
## omega   0.0014771   0.0007917    1.866 0.062064 .
## alpha1  0.1036656   0.0408173    2.540 0.011093 *
## beta1   0.8085661   0.0685649   11.793  < 2e-16 ***
## shape   6.6892311   1.9602897    3.412 0.000644 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
##  251.7896    normalized:  0.6768539
##
## Standardised Residuals Tests:
##                                 Statistic p-Value
##  Jarque-Bera Test   R    Chi^2  184.8699  0
##  Shapiro-Wilk Test  R    W      0.9647736 8.306688e-08
##  Ljung-Box Test     R    Q(10)  9.545124  0.4812646
##  Ljung-Box Test     R    Q(15)  16.54094  0.3470391
##  Ljung-Box Test     R    Q(20)  17.98089  0.5886671
##  Ljung-Box Test     R^2  Q(10)  0.6650225 0.9999743
##  Ljung-Box Test     R^2  Q(15)  10.71018  0.7728558
##  Ljung-Box Test     R^2  Q(20)  12.09555  0.912748
##  LM Arch Test       R    TR^2   9.870077  0.6273576
##
```

```
## Information Criterion Statistics:
##       AIC       BIC       SIC      HQIC
## -1.326826 -1.274153 -1.327181 -1.305908
```

至此模型 AIC 和 BIC 的结果最小，模型效果达到相对最优。最后的模型为

$$x_t = 0.02191 + \varepsilon_t, \quad \sigma_t^2 = 0.00148 + 0.1037\varepsilon_{t-1} + 0.8086\sigma_{t-1}^2$$

进行向前五步预测

```
pre <- predict(m1, 5)
pre
```

```
##   meanForecast meanError standardDeviation
## 1   0.02191214 0.1194190         0.1194190
## 2   0.02191214 0.1203592         0.1203592
## 3   0.02191214 0.1212105         0.1212105
## 4   0.02191214 0.1219819         0.1219819
## 5   0.02191214 0.1226813         0.1226813
```