ts19-sol8 code

TA Mazhaoxing

目录

```
\usepackage{ctex}
```

```
---
documentclass: ctexart
output: rticles::ctex
---
```

数据预处理

```r
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------

## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

## -- Conflicts ---------------------------------------------------------------------- t
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(readxl)
library(lmtest)
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(sandwich)
library(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
```

```
data_hw8 <- read_excel("/Users/mazhaoxing/Desktop/ts19-hw8-data.xlsx",
                               col_names = TRUE, range = "A3:G107" ) %>%
mutate(logY = log(Y),logK = log(K),logL = log(L)) %>%
mutate(dlogY = logY-lag(logY),dlogK = logK-lag(logK),dlogL = logL-lag(logL))
```

b. 对一阶差分回归模型进行 OLS 估计

i. 选取 dlogY,dlogK,dlogL 变量

```
df1 <- data_hw8 %>%
  select(dlogY, dlogK, dlogL) %>%
  filter(!is.na(dlogY))
```

ii. 求系数估计值的普通标准误和稳健标准误，$\mathbb{R}^2$

```r
lm1 <- lm(dlogY~., data = df1)
summary(lm1)
```

```
##
## Call:
## lm(formula = dlogY ~ ., data = df1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.012610 -0.004896 -0.000074  0.003123  0.017412
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.013205   0.003453   3.824 0.000229 ***
## dlogK        0.350976   0.124260   2.825 0.005716 **
## dlogL       -0.136650   0.115632  -1.182 0.240104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.006381 on 100 degrees of freedom
## Multiple R-squared:  0.08301,    Adjusted R-squared:  0.06467
## F-statistic: 4.526 on 2 and 100 DF,  p-value: 0.01313
```

```r
coeftest(lm1, vcov = vcovHC(lm1))
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  0.0132055  0.0031694  4.1666 6.576e-05 ***
## dlogK        0.3509761  0.1114971  3.1479  0.002168 **
## dlogL       -0.1366497  0.1463104 -0.9340  0.352568
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

综上，估计值 $\alpha^\top$ 与 $\beta^\top$ 的普通标准误分别为 0.124260 和 0.115632，稳健标准误分别为 0.1114971 和 0.1463104，R^2 为 0.08301，调整过的 R^2 为 0.06467。$[\hat{\alpha},\hat{\beta}]^\top$ 的协方差矩阵为

```
vcov(lm1)
```

```
##               (Intercept)         dlogK         dlogL
## (Intercept)  1.192645e-05 -0.0004210737 -5.010998e-06
## dlogK       -4.210737e-04  0.0154405055 -7.458727e-04
## dlogL       -5.010998e-06 -0.0007458727  1.337085e-02
```

```
vcovHAC(lm1)
```

```
##               (Intercept)        dlogK         dlogL
## (Intercept)  3.802972e-05 -0.001201717 -0.0003382021
## dlogK       -1.201717e-03  0.040178693  0.0064146112
## dlogL       -3.382021e-04  0.006414611  0.0380189792
```

```
se1 <- sqrt(diag(vcov(lm1)))
se1 #系数的普通标准误
```

```
## (Intercept)       dlogK       dlogL
## 0.003453469 0.124259831 0.115632390
```

```
se2 <- sqrt(diag(vcovHAC(lm1)))
se2 #系数的稳健标准误
```

```
## (Intercept)       dlogK       dlogL
## 0.006166824 0.200446235 0.194984561
```

iii. 对 $H_0:\alpha=0.5$ 及 $H_0:\beta=0.5$ 进行 t 检验

普通标准误下

对 α 进行 t 检验

```
a1 <- 0.350976
b1 <- -0.136650
mu1 <- 0.5
t_a1 <- (a1-mu1)/se1[2][[1]]
p_a1 <- 2*pt(t_a1,df=100)
p_a1
```

```
## [1] 0.2332479
```

p=0.2332479，接受原假设

对 β 进行 t 检验

```
t_b1 <- (b1-mu1)/se1[3][[1]]
p_b1 <- 2*pt(t_b1,df=100)
p_b1
```

```
## [1] 2.850408e-07
```

p=$2.850408e^{-7}$，拒绝原假设

稳健标准误下

对 α 进行 t 检验

```
a2 <- 0.3509761
b2 <- -0.1366497
mu2 <- 0.5
t_a2 <- (a2-mu2)/se2[2][[1]]
p_a2 <- 2*pt(t_a2,df=100)
p_a2
```

```
## [1] 0.458946
```

p=0.458946，接受原假设

```
t_b2 <- (b2-mu2)/se2[3][[1]]
p_b2 <- 2*pt(t_b2,df=100)
p_b2
```

```
## [1] 0.001498798
```

p=0.001498798，拒绝原假设

c. 检验规模报酬不变

原假设为 $H_0 : \alpha + \beta = 1$

```
#在同方差假设下检验
a1 <- 0.350976
b1 <- -0.136650
c1 <- a1+b1
mu1 <- 1
t_c1 <- (c1-mu1)/sqrt(vcov(lm1)[2,2]+vcov(lm1)[3,3]+2*vcov(lm1)[2,3])
p_c1 <- 2*pt(t_c1,df=100)
p_c1
```

```
## [1] 6.718903e-06
```

p=6.718903e^{-8}，拒绝原假设

```
#在异方差假设下检验
a2 <- 0.3509761
b2 <- -0.1366497
c2 <- a2+b2
mu2 <- 1
t_c2 <- (c2-mu2)/sqrt(vcovHAC(lm1)[2,2]+vcovHAC(lm1)[3,3]+2*vcovHAC(lm1)[2,3])
p_c2 <- 2*pt(t_c2,df=100)
p_c2
```

```
## [1] 0.0106145
```

p=0.0106145，在 95% 的置信水平下拒绝原假设

d. 绘制 LogY,LogK,LogL 的时间序列图和其差分的时间序列图

```
par(mfrow=c(3,1))
plot(ts(data_hw8$dlogY,start=1992-03-31),main="dlogY~t",ylab="dlogY",xlim=c(1990,2020))
plot(ts(data_hw8$dlogK,start=1992-03-31),main="dlogK~t",ylab="dlogK",xlim=c(1990,2020))
plot(ts(data_hw8$dlogL,start=1992-03-31),main="dlogL~t",ylab="dlogL",xlim=c(1990,2020))
```

```
par(mfrow=c(3,1))
plot(ts(data_hw8$logY,start=1992-03-31),main="logY~t",ylab="logY",xlim=c(1990,2020))
plot(ts(data_hw8$logK,start=1992-03-31),main="logK~t",ylab="logK",xlim=c(1990,2020))
plot(ts(data_hw8$logL,start=1992-03-31),main="logL~t",ylab="logL",xlim=c(1990,2020))
```

可见其差分的时间序列图不存在明显的趋势性，而 $\log Y, \log K, \log L$ 的时间序列图存在明显的趋势性。

e.

```
#重复b.问步骤
df2 <- data_hw8 %>%
  select(logY, logK, logL) %>%
  filter(!is.na(logY))
lm2 <- lm(logY~., data = df2)
summary(lm2)
```

```
##
## Call:
## lm(formula = logY ~ ., data = df2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.144196 -0.037794 -0.004294  0.041603  0.112630
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.46900    3.45409  -0.715    0.476
## logK         0.79809    0.02527  31.576   <2e-16 ***
## logL         0.19342    0.28839   0.671    0.504
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.05967 on 101 degrees of freedom
## Multiple R-squared:  0.9928, Adjusted R-squared:  0.9927
## F-statistic:  6964 on 2 and 101 DF,  p-value: < 2.2e-16
```

```
coeftest(lm2, vcov = vcovHAC(lm2))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.469001   1.546909 -1.5961   0.1136
## logK         0.798089   0.024136 33.0665   <2e-16 ***
## logL         0.193423   0.139945  1.3821   0.1700
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#系数的协方差矩阵
vcov(lm2)
```

```
##             (Intercept)          logK         logL
## (Intercept)  11.9307386  0.0834634987 -0.995910618
## logK          0.0834635  0.0006388145 -0.007012089
## logL         -0.9959106 -0.0070120887  0.083170105
```

```
vcovHAC(lm2)
```

```
##             (Intercept)          logK        logL
## (Intercept)  2.39292713  0.0369427634 -0.21644119
## logK         0.03694276  0.0005825383 -0.00335106
## logL        -0.21644119 -0.0033510602  0.01958470
```

```
se1_new <- sqrt(diag(vcov(lm2)))
se1_new ##系数的普通标准误
```

```
## (Intercept)        logK        logL
##  3.45409013  0.02527478  0.28839228
```

```
se2_new <- sqrt(diag(vcovHAC(lm2)))
se2_new #系数的稳健标准误
```

```
## (Intercept)        logK        logL
##  1.54690890  0.02413583  0.13994534
```

α 的普通标准误为 0.02526467，β 的普通标准误为 0.28839228，α 的稳健标准误为 0.02413583，β 的稳健标准误为 0.13994534，R^2wei 0.9928，调整的 R^2 为 0.9927

```
#在普通标准误下对系数做t检验
a1_new <- 0.79809
b1_new <- 0.193423
mu1 <- 0.5
t_a1_new <- (a1_new-mu1)/se1_new[2][[1]]
p_a1_new <- 2*(1-pt(t_a1_new,df=101))
p_a1_new
```

```
## [1] 0
```

```
t_b1_new <- (b1_new-mu1)/se1_new[3][[1]]
p_b1_new <- 2*pt(t_b1_new,df=101)
p_b1_new
```

```
## [1] 0.2902919
```

拒绝原假设 $H_0:\alpha=0.5$，接受原假设 $H_0:\beta=0.5$

```
#在稳健标准误下对系数做t检验
a2_new <- 0.798089
b2_new <- 0.193423
mu2 <- 0.5
t_a2_new <- (a2_new-mu2)/se2_new[2][[1]]
p_a2_new <- 2*(1-pt(t_a2_new, df=101))
p_a2_new
```

```
## [1] 0
```

```
t_b2_new <- (b2_new-mu2)/se2_new[3][[1]]
p_b2_new <- 2*pt(t_b2_new, df=101)
p_b2_new
```

```
## [1] 0.03077386
```

拒绝原假设 $_0 : \alpha = 0.5$，拒绝原假设 $H_0 : \beta = 0.5$

```
#重复c.的步骤，检验规模报酬不变
#在同方差下做t检验
c1_new <- a1_new+b1_new
mu1 <- 1
t_c1_new <- (c1_new-mu1)/sqrt(vcov(lm2)[2,2]+vcov(lm2)[3,3]+2*vcov(lm2)[2,3])
t_c1_new
```

```
## [1] -0.03212728
```

```
p_c1_new <- 2*pt(t_c1_new, df=101)
p_c1_new
```

```
## [1] 0.974434
```

p=0.974434，接受原假设

```
#在异方差下做t检验
c2_new <- a2_new+b2_new
mu2 <- 1
t_c2_new <- (c2_new-mu2)/sqrt(vcovHAC(lm2)[2,2]+vcovHAC(lm2)[3,3]+2*vcovHAC(lm2)[2,3])
t_c2_new
```

```
## [1] -0.07314764
```

```r
p_c2_new <- 2*pt(t_c2_new,df=101)
p_c2_new
```

```
## [1] 0.9418333
```

p=0.9418333，接受原假设。

使用水平值的回归模型的 $R^2=0.9928$，而在差分值的回归模型的 $R^2=0.08301$，使用水平值的回归模型更为合理。

f.

上述两组回归估计存在内生型偏误。(1) 可能存在遗漏变量，总产出不仅与资本存量和劳动投入有关。(2) 可能存在倒向因果，资本存量会影响总产出，同时总产出也会影响投资，影响资本存量。(3) 可能存在共同因素的问题，宏观经济变量会受到共同冲击的作用。

g.

由 $Y_t=A_tK_t^{\alpha}L_t^{1-\alpha}$ 得 $\log Y_t=\log A_t+\alpha\log K_t+(1-\alpha)\log L_t$，即 $\log Y_t-\log L_t=\log A_t+\alpha(\log K_t-\log L_t)$，令 $Y_1=\log Y_t-\log L_t, X_1=\log K_t-\log L_t$

```r
data_hw8 <- mutate(data_hw8,
                   Y1 = logY-logL,
                   X1 = logK-logL
                   )
df3 <- data_hw8 %>%
  select(Y1, X1)

#建立回归模型：
lm3 <- lm(Y1~., data = df3)

#回归系数的普通标准误，得到R^2和调整过的R^2，如下：
summary(lm3)
```

```
##
## Call:
## lm(formula = Y1 ~ ., data = df3)
```

```
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.144456 -0.037913 -0.004216  0.041521  0.112664
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.579985   0.018366  -140.5   <2e-16 ***
## X1           0.797313   0.007497   106.4   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05938 on 102 degrees of freedom
## Multiple R-squared:  0.9911, Adjusted R-squared:  0.991
## F-statistic: 1.131e+04 on 1 and 102 DF,  p-value: < 2.2e-16
```

```
#回归系数的稳健标准误，如下：
coeftest(lm3, vcov = vcovHAC(lm3))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value  Pr(>|t|)
## (Intercept) -2.579985   0.072964 -35.360 < 2.2e-16 ***
## X1           0.797313   0.028586  27.892 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#系数的协方差矩阵变化如下：
#普通标准误
vcov(lm3)
```

```
##              (Intercept)           X1
## (Intercept) 0.0003373146 1.305855e-04
## X1          0.0001305855 5.620265e-05
```

```r
#稳健标准误
vcovHAC(lm3)
```

```
##             (Intercept)           X1
## (Intercept) 0.005323708 0.0019625155
## X1          0.001962515 0.0008171597
```

```r
#系数的普通标准误可表示为:
se31 <- sqrt(diag(vcov(lm3)))
se31
```

```
## (Intercept)          X1
## 0.018366127 0.007496843
```

```r
#系数的稳健标准误可表示为:
se32 <- sqrt(diag(vcovHAC(lm3)))
se32
```

```
## (Intercept)          X1
##  0.07296374  0.02858601
```

```r
#在普通标准误下对系数做t检验
a31 <- 0.797313
mu3 <- 0.5
t_a31 <- (a31-mu3)/se31[2][[1]]
p_a31 <- 2*(1-pt(t_a31, df=102))
p_a31
```

```
## [1] 0
```

```r
#在稳健标准误下对系数做t检验
a32 <- 0.797313
mu3 <- 0.5
t_a32 <- (a32-mu3)/se32[2][[1]]
p_a32 <- 2*(1-pt(t_a32, df=102))
p_a32
```

```
## [1] 0
```

h.

```
#加入logA变量
data_hw8 <- mutate(data_hw8,
                   logA = logY-a31*logK-(1-a31)*logL
                   )
#logA的时间序列图
ggplot(data_hw8,mapping=aes(x=time,y=logA))+
  geom_line()
```

```
#计算logA的样本均值和方差
mean(data_hw8$logA)
```

```
## [1] -2.579986
```

```
var(data_hw8$logA)
```

```
## [1] 0.003491629
```

```
#自协方差函数图
acf(data_hw8$logA,lag.max = 100)
```

i.

```
data_hw8 <- data_hw8 %>%
  mutate(laglogA = lag(logA))

#选取需要的数据:
df4 <- data_hw8 %>%
  select(logA, laglogA) %>%
```

```r
  filter(!is.na(laglogA))

#建立回归模型：
lm4 <- lm(logA~., data = df4)

#回归系数的普通标准误
summary(lm4)
```

```
##
## Call:
## lm(formula = logA ~ ., data = df4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0112131 -0.0056734 -0.0000595  0.0043109  0.0173880
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.05209    0.03032  -1.718   0.0888 .
## laglogA      0.97967    0.01175  83.361   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00693 on 101 degrees of freedom
## Multiple R-squared:  0.9857, Adjusted R-squared:  0.9855
## F-statistic:  6949 on 1 and 101 DF,  p-value: < 2.2e-16
```

```r
#回归系数的稳健标准误
coeftest(lm4, vcov = vcovHAC(lm4))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -0.052090   0.086118 -0.6049   0.5466
## laglogA      0.979667   0.033608 29.1498   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#系数的协方差矩阵变化如下：
#普通标准误
vcov(lm4)
```

```
##              (Intercept)      laglogA
## (Intercept) 0.0009190435 0.0003561853
## laglogA     0.0003561853 0.0001381135
```

```r
#稳健标准误
vcovHAC(lm4)
```

```
##             (Intercept)     laglogA
## (Intercept) 0.007416297 0.002893781
## laglogA     0.002893781 0.001129499
```

```r
#系数的普通标准误为:
se41 <- sqrt(diag(vcov(lm4)))
se41
```

```
## (Intercept)     laglogA
##  0.03031573  0.01175217
```

```r
#系数的稳健标准误为:
se42 <- sqrt(diag(vcovHAC(lm4)))
se42
```

```
## (Intercept)     laglogA
##  0.08611792  0.03360802
```

在平稳条件下，有 $\mathbb{E}\log A_t = \mu + \rho\mathbb{E}\log A_{t-1}$，则 $\log A_t$ 的理论期望值 $\mathbb{E}\log A_t = \frac{\mu}{1-\rho}$

```
mu<-lm4$coefficients[[1]]
rho<-lm4$coefficients[[2]]
ElogA <- mu/(1-rho)
ElogA #理论期望值
```

```
## [1] -2.561844
```

```
mean(data_hw8$logA) #样本均值
```

```
## [1] -2.579986
```

可以看到 $\log A_t$ 的样本均值和理论期望值接近，可以认为上述模型设定和估计比较合理。

j.

```
#加入时间趋势
df5 <- df4 %>%
  mutate(t = 1:length(logA))

#建立回归模型：
lm5 <- lm(logA~., data = df5)

#回归系数的普通标准误
summary(lm5)
```

```
##
## Call:
## lm(formula = logA ~ ., data = df5)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.014628 -0.003272 -0.000237  0.003260  0.015659
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -2.832e-02  2.425e-02  -1.168    0.246
## laglogA      9.860e-01  9.361e-03 105.330   <2e-16 ***
## t           -1.422e-04  1.829e-05  -7.772    7e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.005499 on 100 degrees of freedom
## Multiple R-squared:  0.9911, Adjusted R-squared:  0.9909
## F-statistic:  5548 on 2 and 100 DF,  p-value: < 2.2e-16
```

```r
#回归系数的稳健标准误:
coeftest(lm5, vcov = vcovHAC(lm5))
```

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value  Pr(>|t|)
## (Intercept) -2.8317e-02  4.5293e-02 -0.6252    0.5333
## laglogA      9.8602e-01  1.7634e-02 55.9170 < 2.2e-16 ***
## t           -1.4218e-04  2.6928e-05 -5.2798 7.549e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

可知 κ, ρ 显著异于 0，而 μ 不能拒绝 $\mu = 0$ 的原假设

k.

```r
#包含一阶差分的数据集
df6 <- df5 %>%
  mutate(dlogA = logA-lag(logA),
         dlaglogA = laglogA-lag(laglogA)) %>%
  select(dlogA, dlaglogA) %>%
  filter(!is.na(dlogA))

#建立回归模型:
```

```r
lm6 <- lm(dlogA~., data = df6)

#回归系数的普通标准误
summary(lm6)
```

```
##
## Call:
## lm(formula = dlogA ~ ., data = df6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0159449 -0.0031137 -0.0005345  0.0031879  0.0146649
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.914e-05  5.175e-04  -0.037    0.971
## dlaglogA     6.595e-01  7.405e-02   8.906 2.46e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.005218 on 100 degrees of freedom
## Multiple R-squared:  0.4423, Adjusted R-squared:  0.4368
## F-statistic: 79.32 on 1 and 100 DF,  p-value: 2.463e-14
```

```r
#回归系数的稳健标准误
coeftest(lm6, vcov = vcovHAC(lm6))
```

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value  Pr(>|t|)
## (Intercept) -1.9135e-05  4.6471e-04 -0.0412    0.9672
## dlaglogA     6.5947e-01  1.0174e-01  6.4822 3.454e-09 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

j. 中模型的估计系数都有较大的 t 值，且 j. 问模型的 R^2 和调整过的 R^2 都达到了 0.99，而 k. 问模型 $R^2 = 0.4423$，调整过的 $R^2 = 0.4368$，故 j. 中模型的估计结果更可靠。

l.

设 v_t 为全要素生产率的增速，$\frac{A_t}{A_{t-s}} = 1 + v_t$

则 $\log A_t - \log A_{t-1} = \log(1 + v_t) \approx v_t$

$\mathbb{E}\Delta \log A_t \approx \mathbb{E}v_t$，即 $\mathbb{E}\Delta \log A_t$ 近似等于全要素生产率增速的平均值

```r
mean(df6$dlogA)
```

```
## [1] 0.0002442609
```

即全要素生产率增速的平均值约等于 0.024%

m.

```r
df7 <- data.frame(
  dlogA=diff(data_hw8$logA))
auto.arima(df7$dlogA,max.p = 5,max.q = 0)
```

```
## Series: df7$dlogA
## ARIMA(1,1,0)
##
## Coefficients:
##           ar1
##       -0.3825
## s.e.   0.0909
##
## sigma^2 estimated as 2.782e-05:  log likelihood=390.67
## AIC=-777.35   AICc=-777.23   BIC=-772.1
```

由 AIC 及 BIC 可知，当滞后阶数为 1-阶时，有最小的 AIC 和 BIC。

```r
#使用ar()函数进行自回归
ar(df7$dlogA,ic="aic")
```

```
##
## Call:
## ar(x = df7$dlogA, ic = "aic")
##
## Coefficients:
##       1        2        3        4
##  0.5125   0.1904  -0.1269   0.2112
##
## Order selected 4  sigma^2 estimated as  2.626e-05
```

```r
ar(df7$dlogA,ic="bic")
```

```
##
## Call:
## ar(x = df7$dlogA, ic = "bic")
##
## Coefficients:
##       1        2        3        4
##  0.5125   0.1904  -0.1269   0.2112
##
## Order selected 4  sigma^2 estimated as  2.626e-05
```

可知滞后阶数 4-阶最好。

综上，可以认为 k 中的 AR(1) 足以捕捉 $\Delta \log A_t$ 的动态特征