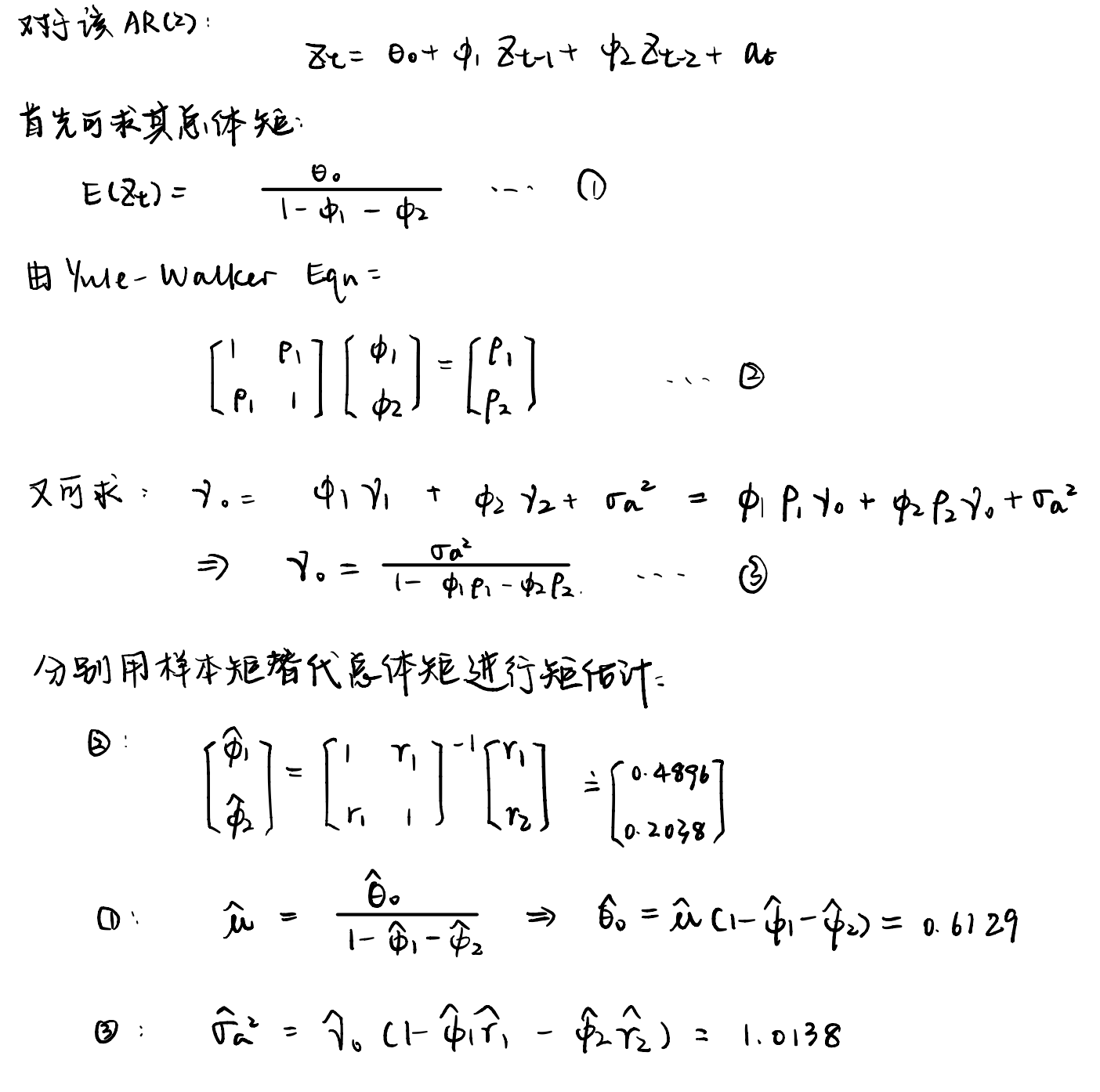
第三次作业

辛柏嬴 2020111753 20金统2班

2023-03-31

## 第一题

**长度为n = 100 的时间序列{Zt}，已知r1 = 0.615、r2 = 0.505、r3 = 0.364，样本均值 为1.999，样本方差为1.701。为该序列识别一个AR(2) 模型，请计算模型参数的矩估计。**



## 第二题

**使用R 软件模拟样本量为n = 36 的零均值MA(1) 序列，其中θ = −0.6，白噪声 at ~ N(0, 1)，且模拟设定随机种子数为1。**

# Z\_t = a\_t + 0.6a\_t-1 library(TSA)  
set.seed(1)  
offset <- 300  
n <- 36  
theta <- -0.6  
at <- rnorm(n + offset)  
Zt <- NULL  
Zt[1] <- at[1] - theta \* 0  
for (i in 2:(n + offset)) {  
 Zt[i] <- at[i] - theta \* at[i - 1]  
}  
Zt <- Zt[(offset + 1):(offset + n)]  
plot(Zt, type = "o", main = "", xlab = "t", ylab = expression(Z[t]))

**(a) 求θ 的矩估计；**

ma1.mme <- function(x) {  
 r = TSA::acf(x, plot = F)$acf[1]  
 if (abs(r) < 0.5)  
 return((-1 + sqrt(1 - 4 \* r^2))/(2 \* r)) else return(NA)  
}  
mme <- ma1.mme(Zt)  
cat(paste("MME result:", mme, ""))

## MME result: -0.417855739742642

**(b) 求θ 的条件最小二乘估计，并且与(a) 中的结果进行比较；**

arima(Zt, order = c(0, 0, 1), method = "CSS")$coef[1]

## ma1   
## 0.5892244

由代码可知，条件最小二乘的计算结果是-0.5892，要明显优于矩估计。

**(c) 求θ 的极大似然估计，并且与(a) 和(b) 中的结果进行比较；**

arima(Zt, order = c(0, 0, 1), method = "ML")$coef[1]

## ma1   
## 0.7224018

由代码可知，MLE的估计结果为-0.7224.因此矩估计的效果最差，误差最大；条件最小二乘的估计效果最好，误差最小。

**(d) 设定随机种子数为5，选取相同参数和样本量产生新的模拟序列，使用新的模拟序 列重复(a)、(b) 和(c)，并将两次模拟结果进行比较。**

set.seed(5)  
offset <- 300  
n <- 36  
theta <- -0.6  
et <- rnorm(n + offset)  
Yt <- NULL  
Yt[1] <- et[1] - theta \* 0  
for (i in 2:(n + offset)) {  
 Yt[i] <- et[i] - theta \* et[i - 1]  
}  
Yt <- Yt[(offset + 1):(offset + n)]  
mme <- ma1.mme(Yt)  
print("MME result:")  
mme  
print("CSS result:")  
arima(Yt, order = c(0, 0, 1), method = "CSS")  
print("MLE result:")  
arima(Yt, order = c(0, 0, 1), method = "ML")

## [1] "MME result:"  
## [1] -0.4922402  
## [1] "CSS result:"  
##   
## Call:  
## arima(x = Yt, order = c(0, 0, 1), method = "CSS")  
##   
## Coefficients:  
## ma1 intercept  
## 0.7102 0.0604  
## s.e. 0.1221 0.2626  
##   
## sigma^2 estimated as 0.8707: part log likelihood = -48.59  
## [1] "MLE result:"  
##   
## Call:  
## arima(x = Yt, order = c(0, 0, 1), method = "ML")  
##   
## Coefficients:  
## ma1 intercept  
## 0.7661 0.0031  
## s.e. 0.1099 0.2633  
##   
## sigma^2 estimated as 0.8196: log likelihood = -47.94, aic = 101.89

由本次模拟可知（具体输出结果见上），矩估计的估计效果依然是最差的，MLE和CSS的估计效果大致相似，其中CSS在本题中效果略好。此外，MLE和CSS的估计结果由其s.e.可知，真实值均落在其95%置信区间内。

## 第三题

**使用R 软件模拟样本量为n = 72 的零均值ARMA(1,1) 序列，其中，ϕ = 0.7，θ = 0.4， 白噪声at ~ N(0, 1)，且模拟设定随机种子数为1。**

# Z\_t=0.7Z\_t-1 + a\_t - 0.4a\_t-1 library(TSA)  
set.seed(1)  
offset <- 300  
n <- 72  
theta <- 0.4  
phi <- 0.7  
at <- rnorm(n + offset)  
Zt <- NULL  
Zt[1] <- at[1] - theta \* 0  
for (i in 2:(n + offset)) {  
 Zt[i] <- at[i] - theta \* at[i - 1] + phi \* Zt[i - 1]  
}  
Zt <- Zt[(offset + 1):(offset + n)]  
plot(Zt, type = "o", main = "", xlab = "t", ylab = expression(Z[t]))

**(a) 求ϕ 和θ 的矩估计；**

rho1 <- acf(Zt, plot = FALSE)$acf[2]  
rho2 <- acf(Zt, plot = FALSE)$acf[3]  
# phi MME:  
phi\_hat <- rho2/rho1  
cat("phi's MME: ", phi\_hat, "\n")  
  
# solve the eqn to get theta\_hat MME, note that only keep the reversable  
# solution.  
func <- function(theta) rho1 - (1 - theta \* phi\_hat) \* (phi\_hat - theta)/(1 - 2 \*  
 theta \* phi\_hat + theta \* theta)  
theta\_hat <- uniroot(func, c(0, 1), tol = 1e-05)$root  
cat("theta's MME:", theta\_hat)

## phi's MME: 0.7487895   
## theta's MME: 0.4320485

MME估计结果如上所示

**(b) 求ϕ 和θ 的条件最小二乘估计，并且与(a) 中的结果进行比较；**

arima(Zt, order = c(1, 0, 1), method = "CSS")

##   
## Call:  
## arima(x = Zt, order = c(1, 0, 1), method = "CSS")  
##   
## Coefficients:  
## ar1 ma1 intercept  
## 0.6145 -0.2544 0.0875  
## s.e. 0.1911 0.2265 0.2287  
##   
## sigma^2 estimated as 0.9894: part log likelihood = -101.78

本次CSS的估计结果，ϕ\_hat = 0.6145，θ\_hat = 0.2544；与真实值ϕ = 0.7，θ = 0.4相比，MME的估计效果此处要优于CSS.

**(c) 求ϕ 和θ 的极大似然估计，并且与(a) 和(b) 中的结果进行比较；**

arima(Zt, order = c(1, 0, 1), method = "ML")

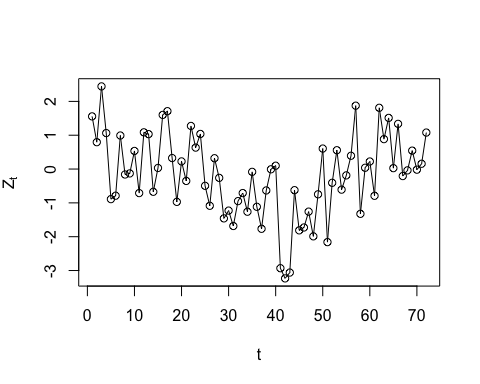
##   
## Call:  
## arima(x = Zt, order = c(1, 0, 1), method = "ML")  
##   
## Coefficients:  
## ar1 ma1 intercept  
## 0.6194 -0.2628 0.1279  
## s.e. 0.1773 0.2074 0.2223  
##   
## sigma^2 estimated as 0.9822: log likelihood = -101.62, aic = 211.23

本次MLE的估计结果，ϕ\_hat = 0.6194，θ = 0.2628；与真实值ϕ = 0.7，θ = 0.4相比，MME的估计效果此处要优于MLE, MLE与CSS的估计效果近似相同.

**(d) 设定随机种子数为5，选取相同参数和样本量产生新的模拟序列，使用新的模拟序 列重复(a)、(b) 和(c)，并将两次模拟结果进行比较。**

# Z\_t=0.7Z\_t-1 + a\_t - 0.4a\_t-1 library(TSA)  
set.seed(5)  
offset <- 300  
n <- 72  
theta <- 0.4  
phi <- 0.7  
at <- rnorm(n + offset)  
Zt <- NULL  
Zt[1] <- at[1] - theta \* 0  
for (i in 2:(n + offset)) {  
 Zt[i] <- at[i] - theta \* at[i - 1] + phi \* Zt[i - 1]  
}  
Zt <- Zt[(offset + 1):(offset + n)]  
plot(Zt, type = "o", main = "", xlab = "t", ylab = expression(Z[t]))  
rho1 <- acf(Zt, plot = FALSE)$acf[2]  
rho2 <- acf(Zt, plot = FALSE)$acf[3]  
# phi MME:  
phi\_hat <- rho2/rho1  
cat("phi's MME: ", phi\_hat, "\n")  
  
# solve the eqn to get theta\_hat MME, note that only keep the reversable  
# solution.  
func <- function(theta) rho1 - (1 - theta \* phi\_hat) \* (phi\_hat - theta)/(1 - 2 \*  
 theta \* phi\_hat + theta \* theta)  
theta\_hat <- uniroot(func, c(0, 1), tol = 1e-05)$root  
cat("theta's MME:", theta\_hat)  
  
# CSS  
arima(Zt, order = c(1, 0, 1), method = "CSS")  
# MLE  
arima(Zt, order = c(1, 0, 1), method = "ML")

## phi's MME: 0.661981   
## theta's MME: 0.2837916  
## Call:  
## arima(x = Zt, order = c(1, 0, 1), method = "CSS")  
##   
## Coefficients:  
## ar1 ma1 intercept  
## 0.8854 -0.6561 -0.3286  
## s.e. 0.0728 0.1206 0.3734  
##   
## sigma^2 estimated as 1.025: part log likelihood = -103.07  
##   
## Call:  
## arima(x = Zt, order = c(1, 0, 1), method = "ML")  
##   
## Coefficients:  
## ar1 ma1 intercept  
## 0.9238 -0.6833 0.0319  
## s.e. 0.0657 0.1209 0.4688  
##   
## sigma^2 estimated as 1.045: log likelihood = -104.03, aic = 216.06



输出结果如上所示，从上至下依次为参数的MME，CSS与MLE。在本次模拟中，MME的估计效果依然优于MLE或CSS，但整体效果普遍差于第一次估计。

## 第四题

**名为robot 的数据文件中包含了一个来自工业机器人的时间序列。机器人需要完成一系列动作，并以英寸为单位记录与理想终点的距离，重复此过程324 次得到该时间序列。**

**(a) 对robot 序列建立AR(1) 模型，并估计其参数；**

library(TSA)

##   
## Attaching package: 'TSA'

## The following objects are masked from 'package:stats':  
##   
## acf, arima

## The following object is masked from 'package:utils':  
##   
## tar

data("robot")  
plot(robot, type = "o")  
robot.ar <- arima(robot, order = c(1, 0, 0), method = "ML")  
robot.ar

##   
## Call:  
## arima(x = robot, order = c(1, 0, 0), method = "ML")  
##   
## Coefficients:  
## ar1 intercept  
## 0.3076 0.0015  
## s.e. 0.0528 0.0002  
##   
## sigma^2 estimated as 6.482e-06: log likelihood = 1475.54, aic = -2947.08

通过建立AR(1)模型并根据ML进行参数估计，估计结果：

并且由上述估计结果可知，截距项与phi均在5%的水平下显著不为零.

**(b) 对robot 序列建立IMA(1,1) 模型，并估计其参数；**

robot.ima <- arima(robot, order = c(0, 1, 1), method = "ML")  
robot.ima

##   
## Call:  
## arima(x = robot, order = c(0, 1, 1), method = "ML")  
##   
## Coefficients:  
## ma1  
## -0.8713  
## s.e. 0.0389  
##   
## sigma^2 estimated as 6.069e-06: log likelihood = 1480.95, aic = -2959.9

通过建立IMA(1,1)模型并根据ML进行参数估计，估计结果：

并且由上述估计结果可知，theta均在5%的水平下显著不为零.

**(c) 应用AIC 准则比较(a) 和(b) 的结果；**

参考两次模型的输出结果， aic\_1 = -2947.08，aic\_2 = -2959.9，对比可知两模型的AIC较为接近，相对而言IMA模型略小，在该信息准则下表现效果略优于AR模型.

**(d) 通过Ljung-Box 检验，对robot 序列拟合的AR(1) 模型和IMA(1,1) 模型进行诊断， 并比较两个模型的诊断结果(设定最大滞后阶数为12)。**

Box.test(rstandard(robot.ar), lag = 6, type = "Ljung-Box", fitdf = 1)  
Box.test(rstandard(robot.ar), lag = 12, type = "Ljung-Box", fitdf = 1)  
Box.test(rstandard(robot.ar), lag = 18, type = "Ljung-Box", fitdf = 1)  
  
Box.test(rstandard(robot.ima), lag = 6, type = "Ljung-Box", fitdf = 1)  
Box.test(rstandard(robot.ima), lag = 12, type = "Ljung-Box", fitdf = 1)  
Box.test(rstandard(robot.ima), lag = 18, type = "Ljung-Box", fitdf = 1)

##   
## Box-Ljung test  
##   
## data: rstandard(robot.ar)  
## X-squared = 26.894, df = 5, p-value = 5.983e-05  
##   
##   
## Box-Ljung test  
##   
## data: rstandard(robot.ar)  
## X-squared = 52.477, df = 11, p-value = 2.234e-07  
##   
##   
## Box-Ljung test  
##   
## data: rstandard(robot.ar)  
## X-squared = 66.12, df = 17, p-value = 9.93e-08  
##   
##   
## Box-Ljung test  
##   
## data: rstandard(robot.ima)  
## X-squared = 5.0925, df = 5, p-value = 0.4047  
##   
##   
## Box-Ljung test  
##   
## data: rstandard(robot.ima)  
## X-squared = 17.131, df = 11, p-value = 0.1041  
##   
##   
## Box-Ljung test  
##   
## data: rstandard(robot.ima)  
## X-squared = 21.386, df = 17, p-value = 0.2095

根据上述输出内容可知，AR(1)的Ljung—Box检验在5%的水平下需要拒绝原假设，说明拟合效果较差；IMA(1)的Ljung-Box检验结果p值均大于10%，且大多数保证了在20%水平之上，可以认为拟合效果比较充分。

## 第五题

**对文件名为 hare 的序列建立AR(3)模型，并运用极大似然估计方法拟合模型。**

**(a) 绘制残差序列的样本ACF 图(设定最大滞后阶数为10)，并评述残差序列的相关性；**

# library(TSA)  
data("hare")  
plot(hare, type = "o")  
hare.ar3 <- arima(hare, order = c(3, 0, 0), method = "ML")  
print("AR(3) MLE Model:")  
hare.ar3  
hare.ar3.resid <- hare.ar3$residuals  
acf(hare.ar3.resid, lag.max = 10, main = "")

## [1] "AR(3) MLE Model:"  
##   
## Call:  
## arima(x = hare, order = c(3, 0, 0), method = "ML")  
##   
## Coefficients:  
## ar1 ar2 ar3 intercept  
## 1.0956 -0.4321 -0.2778 38.0404  
## s.e. 0.1884 0.2846 0.1951 3.8842  
##   
## sigma^2 estimated as 164.5: log likelihood = -124.6, aic = 257.21

|  |  |
| --- | --- |
|  |  |

模型拟合结果如上所示。残差序列样本ACF图如上图所示。由于前10阶的ACf均落在区间范围内，故其相关性统计意义上不显著，可以认为残差近似不相关。

**(b) 设定Ljung-Box 检验的最大滞后阶数为9，计算Ljung-Box 统计量，并给出检验结 论；**

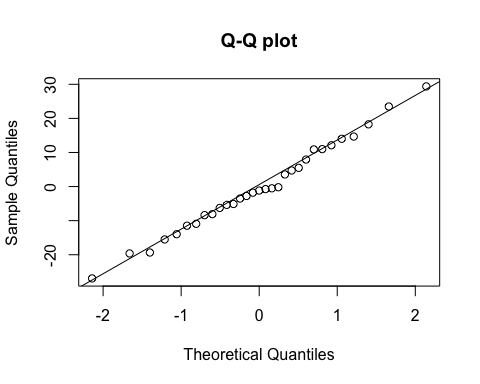
Box.test(rstandard(hare.ar3), lag = 3, type = "Ljung-Box", fitdf = 1)  
Box.test(rstandard(hare.ar3), lag = 6, type = "Ljung-Box", fitdf = 1)  
Box.test(rstandard(hare.ar3), lag = 9, type = "Ljung-Box", fitdf = 1)

##   
## Box-Ljung test  
##   
## data: rstandard(hare.ar3)  
## X-squared = 2.0847, df = 2, p-value = 0.3526  
##   
##   
## Box-Ljung test  
##   
## data: rstandard(hare.ar3)  
## X-squared = 5.1681, df = 5, p-value = 0.3957  
##   
##   
## Box-Ljung test  
##   
## data: rstandard(hare.ar3)  
## X-squared = 7.0685, df = 8, p-value = 0.5293

Ljung-Box 的检验结果如上图所示，由统计量可知其p值均显著大于20%，无法拒绝原假设，说明模型拟合充分。

**(c) 展示残差的正态Q-Q 图，并对图形进行评论；**

qqnorm(hare.ar3$residuals, main = "Q-Q plot")  
qqline(hare.ar3$residuals)



由该Q-Q图可知大多数数据都落在基准线上，整体散点可以近似认为排列成一条直线，说明残差数据的正态性较好。

**(d) 请用Shapiro-Wilk 检验考察残差序列的正态性。**

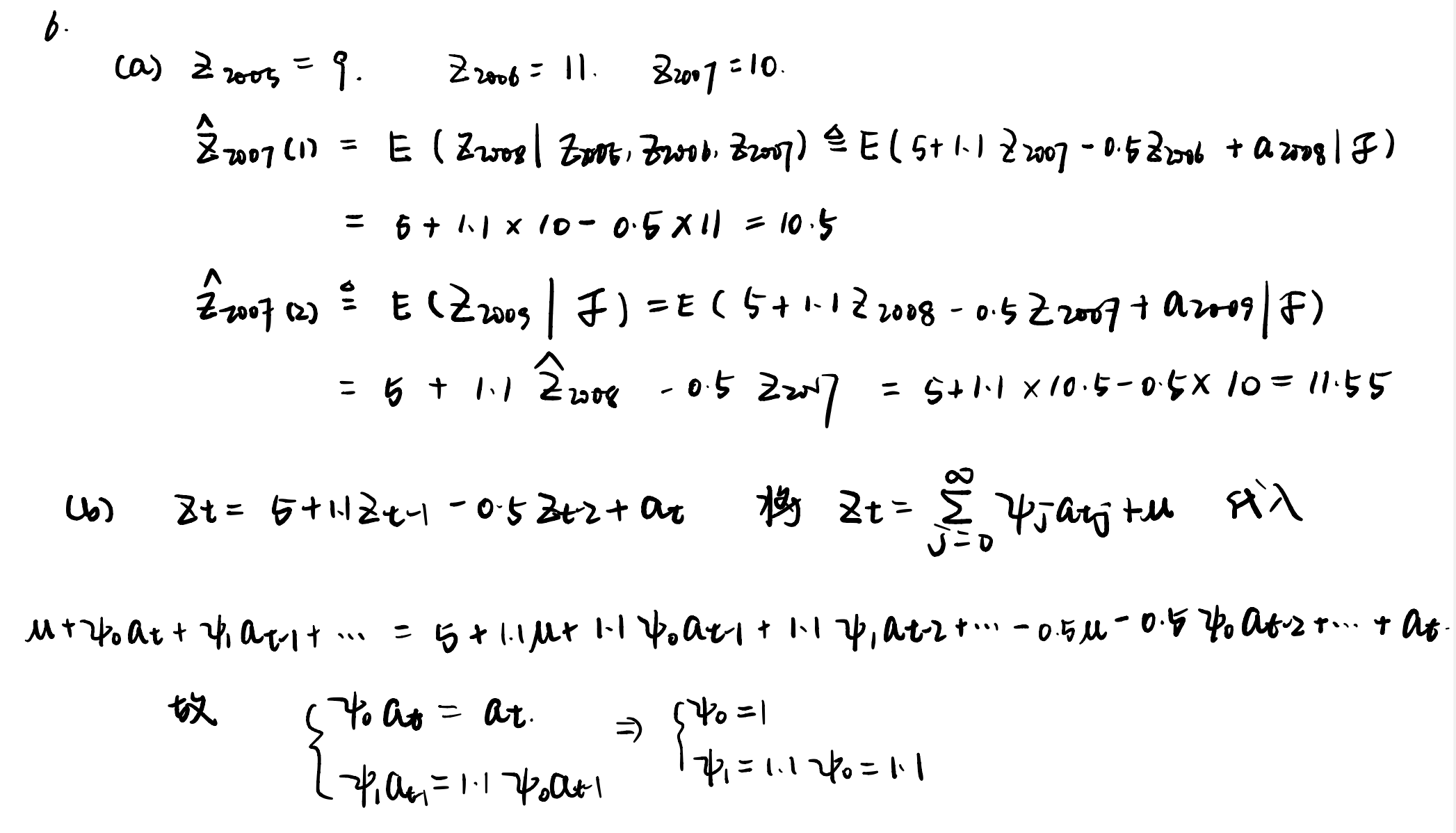
shapiro.test(hare.ar3$residuals)

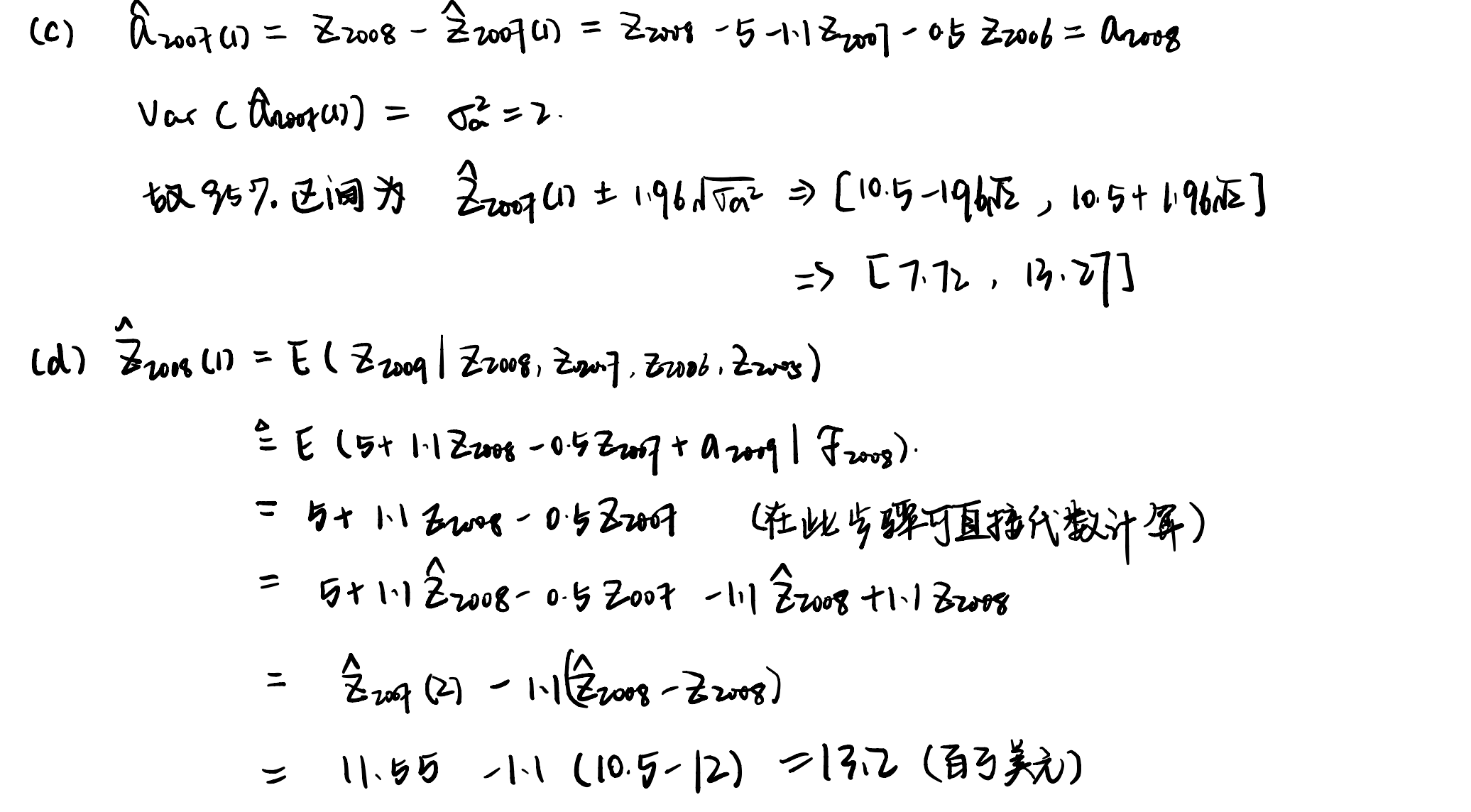
##   
## Shapiro-Wilk normality test  
##   
## data: hare.ar3$residuals  
## W = 0.99068, p-value = 0.9932

该检验的输出结果如上所示，其p值显著大于5%，说明无法拒绝原假设，即可以认为残差序列是服从正态分布的。

## 第六题

**假定某公司的年销售额(单位:百万美元)符合 AR(2) 模型 Zt = 5 + 1.1Zt−1 − 0.5Zt−2 + at, i.i.d. 其中，at ~ N (0, 2).   
(a) 如果 2005 年、2006 年和 2007 年的销售额分别是 9、11 和 10(百万美元)，预测 2008 年和 2009 年的销售额;   
(b) 证明该模型的 MA 展式中对应 ψ1 = 1.1;   
(c) 计算问题 (a) 中 2008 年的 95% 预测区间;   
(d) 如果 2008 年的销售额为 1200 万美元，更新对 2009 年的预测。**





## 第七题

**考虑模型: Zt = β0 + β1t + Xt, 其中，Xt = φXt−1 +at，at ~WN(0,σa2). 假定β0，β1 和φ已知，求证:前置l期的最 小均方误差预测是 Zˆt(l) = β0 +β1(t+l)+φl (Zt −β0 −β1t).**

