|  |
| --- |
| Time Series Analysis |
| ARMA模型分析 |
|  |

|  |
| --- |
| 易芊然  2023-1-7 |

# 引言

AR模型有偏自相关函数截尾性质；MA模型有相关函数截尾性质。有些因果线性时间序列有与AR和MA类似的表现，但是不能在低阶实现偏自相关函数截尾或者相关函数截尾。ARMA模型结合了AR和MA模型， 在对数据拟合优度相近的情况下往往可以得到更简单的模型，而且不要求偏自相关函数截尾也不要求相关函数截尾。

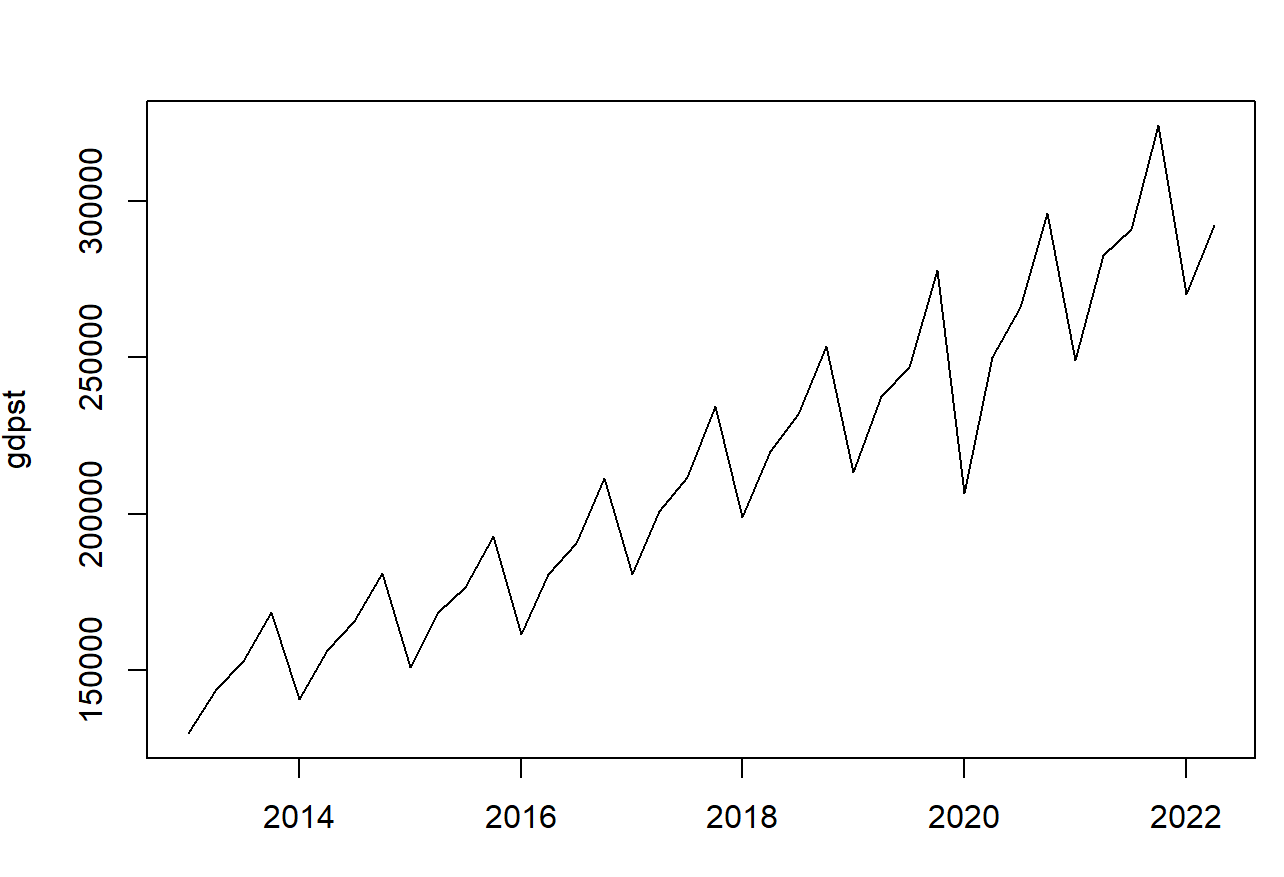
任何平稳的时间序列都可以建立ARMA模型，建立模型的基本思想是：将某个时间序列的SACF和SPACF的行为与各种理论ACF和PACF的行为匹配起来，挑选最佳匹配（或一组匹配的集合），估计模型的未知参数，并检查从模型拟合到的残差，发现可能的模型错误。

# 基于ARMA模型研究GDP

## 模型的识别

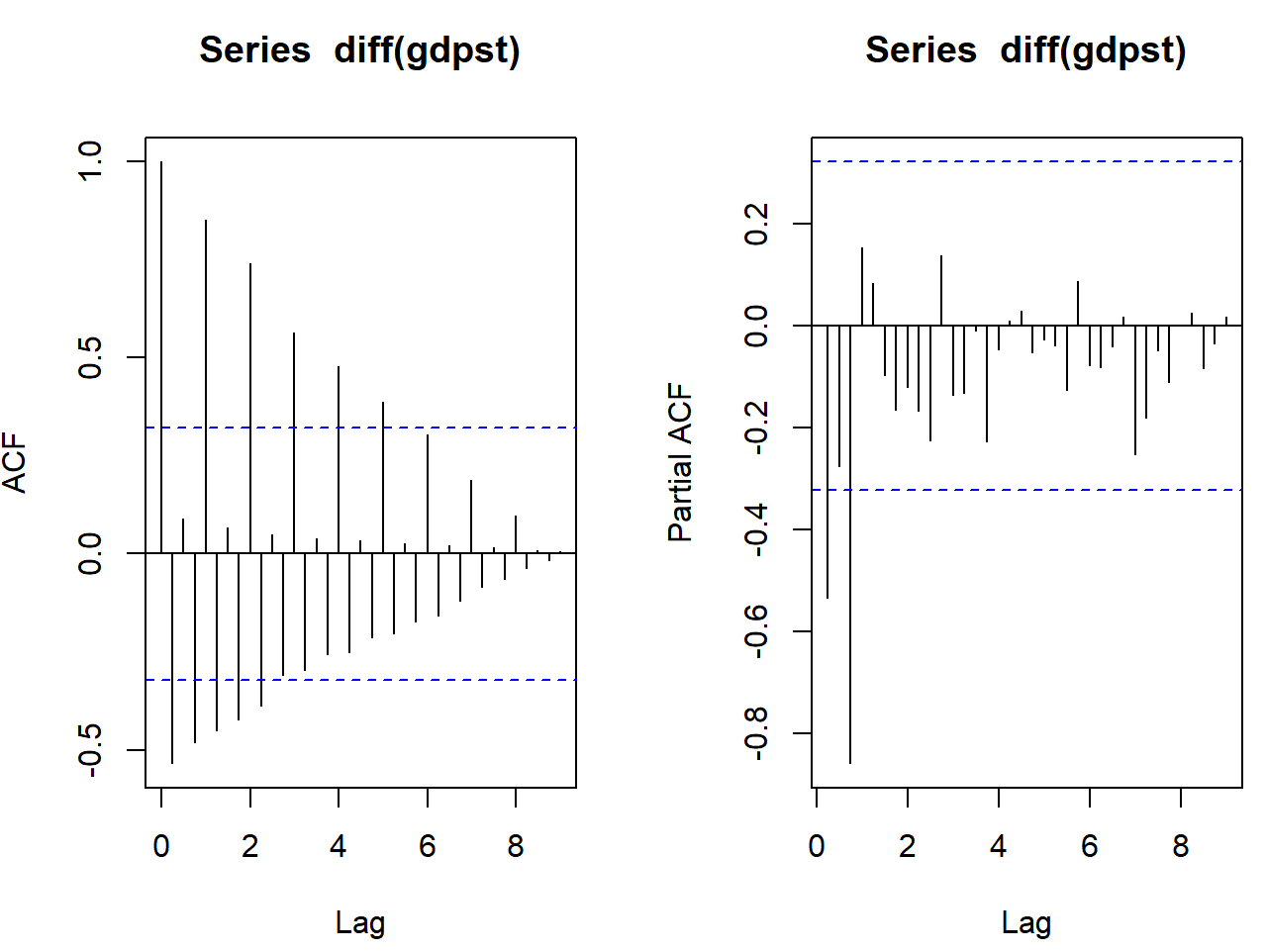
模型识别的基本工具是相关分析。

1. library(readxl)
2. test\_2 <- read\_excel("D:/Users/HUAWEI/Desktop/test.2.xls")
3. gdpst <- ts(test\_2$gdp00, start = c(2013,1), end=c(2022,2),freq = 4)
4. plot(gdpst) *#输出时序图*



可以看出gdp呈现很明显的季节性和趋势性。接着要通过作图法确定最适合的阶数，反应数据的动态特征。

1. acf(gdp1,lag=50)
2. pacf(gdp1,lag=50)
3. par(mfrow=c(1,2))
4. acf(diff(gdp1),lag=50)
5. pacf(diff(gdp1),lag=50)



观察SACF和SPACF的截尾情况，然后选择自回归阶数和移动平均阶数（显著不为0的自相关个数）

1. gdp.ar<-ar(diff(gdp1),method="mle")
2. gdp.ar$order
3. gdp.ar$ar
4. Mod(polyroot(c(1,-gdp.ar$ar)))
5. acf(gdp.ar$res[-(1:gdp.ar$order)],lag=50)

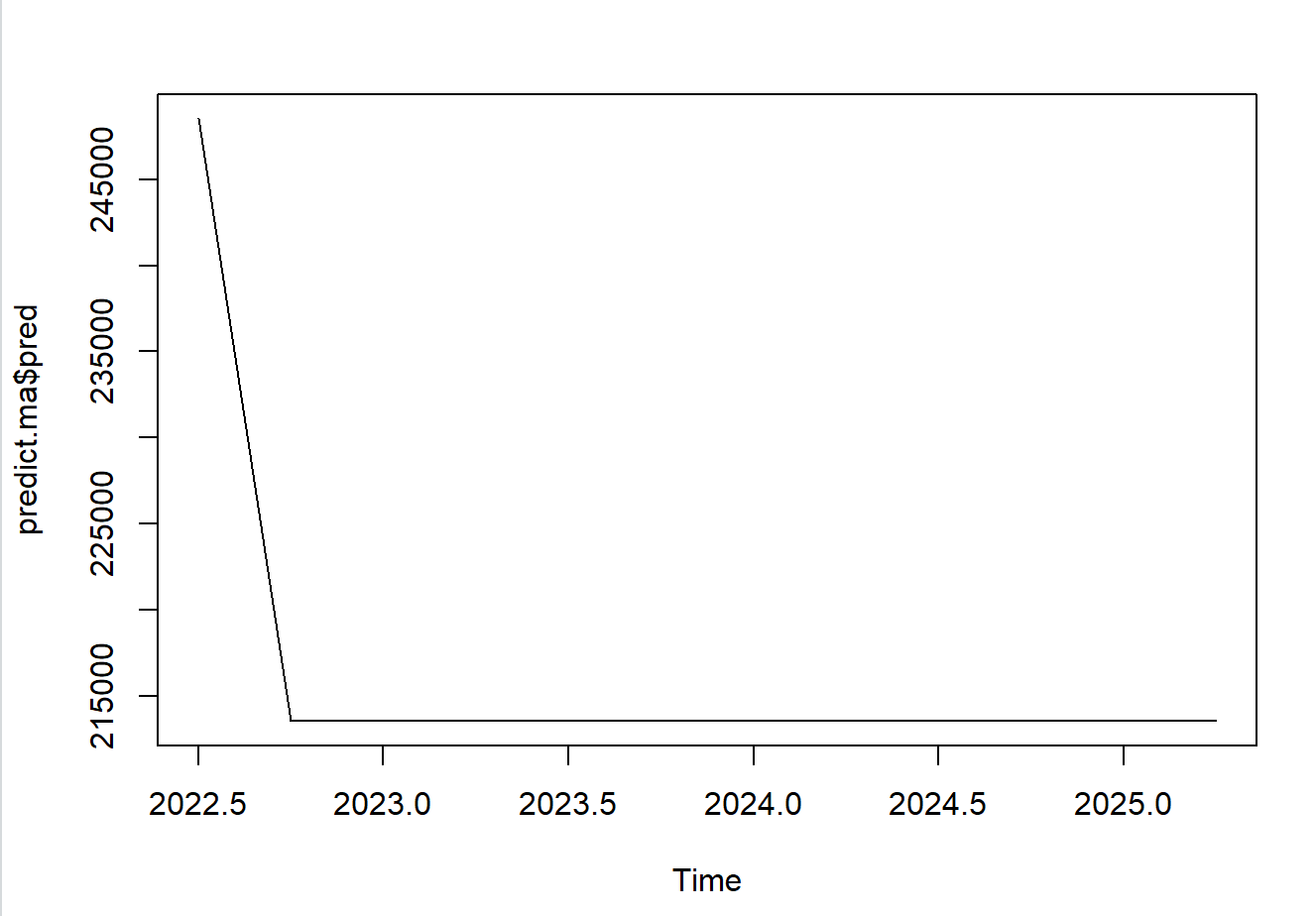
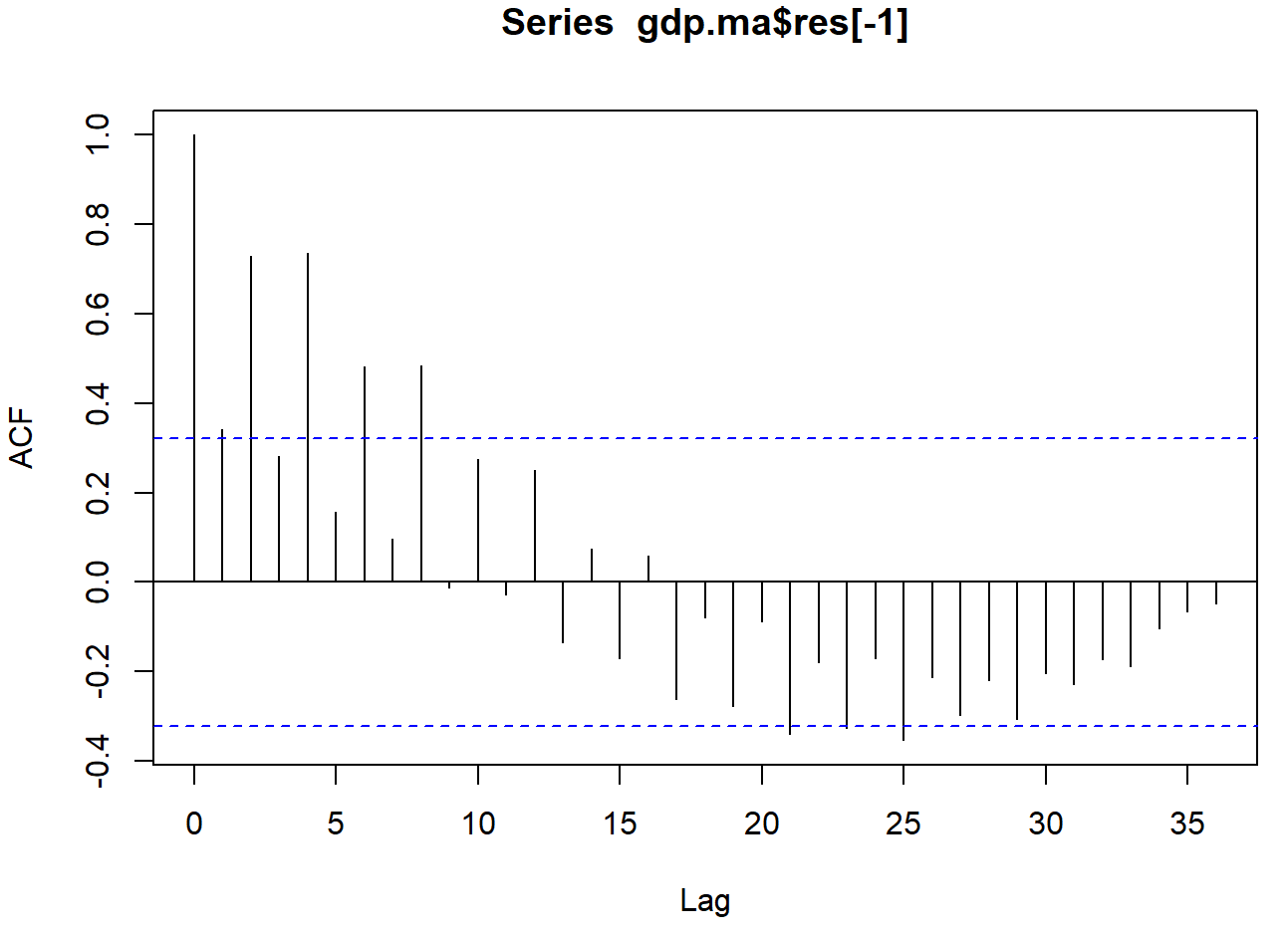


可以看出根都在单位圆外，故AR模型满足平稳性。

1. gdp.ma<-arima(gdp1,order=c(0,0,1))
2. gdp.ma
3. tstats<-coef(gdp.ma)/sqrt(diag(vcov(gdp.ma)))
4. tstats

对初选的模型采用信息准则进行筛选。

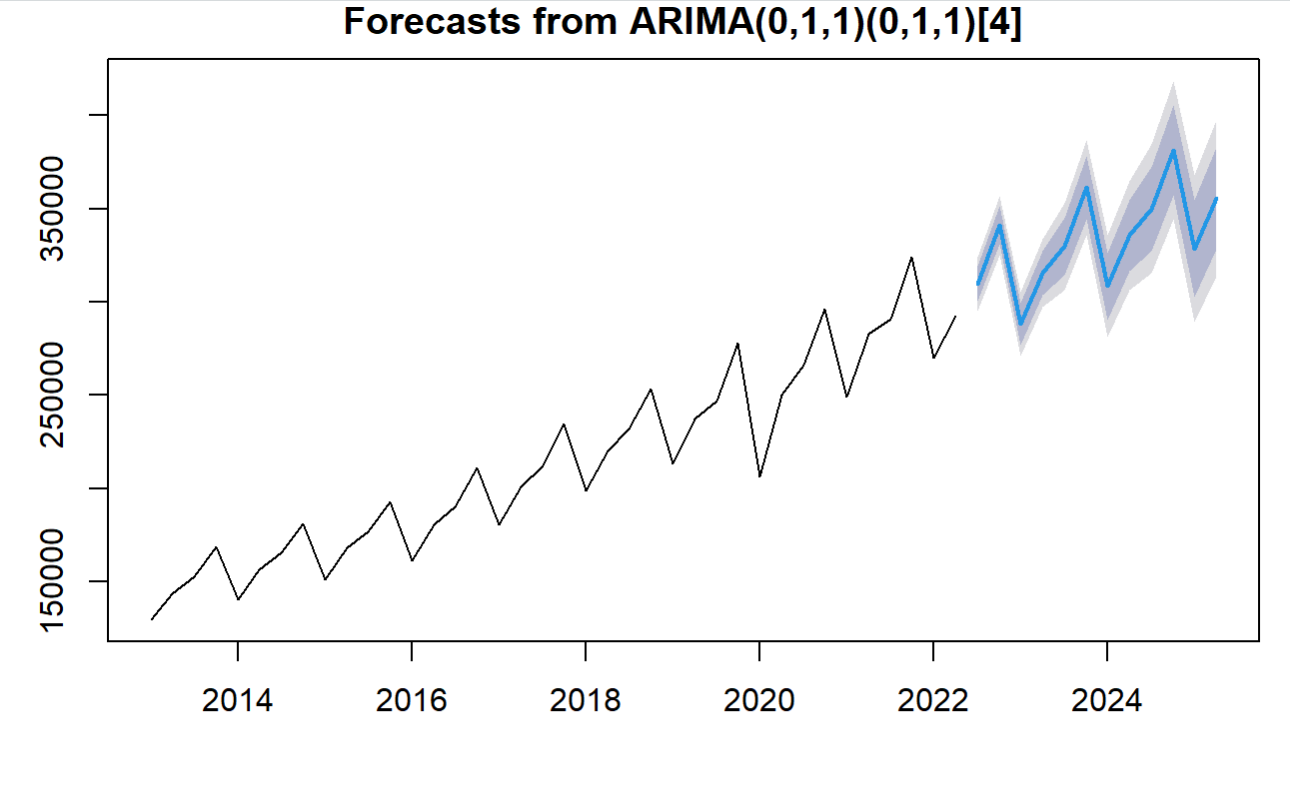
1. acf(gdp.ma$res[-1],lag=50)
2. AIC(gdp.ma)
3. BIC(gdp.ma)
4. jarque.bera.test(gdp.ma$residuals)
5. predict.ma<-predict(gdp.ma,n.ahead=36)
6. plot(predict.ma$pred)



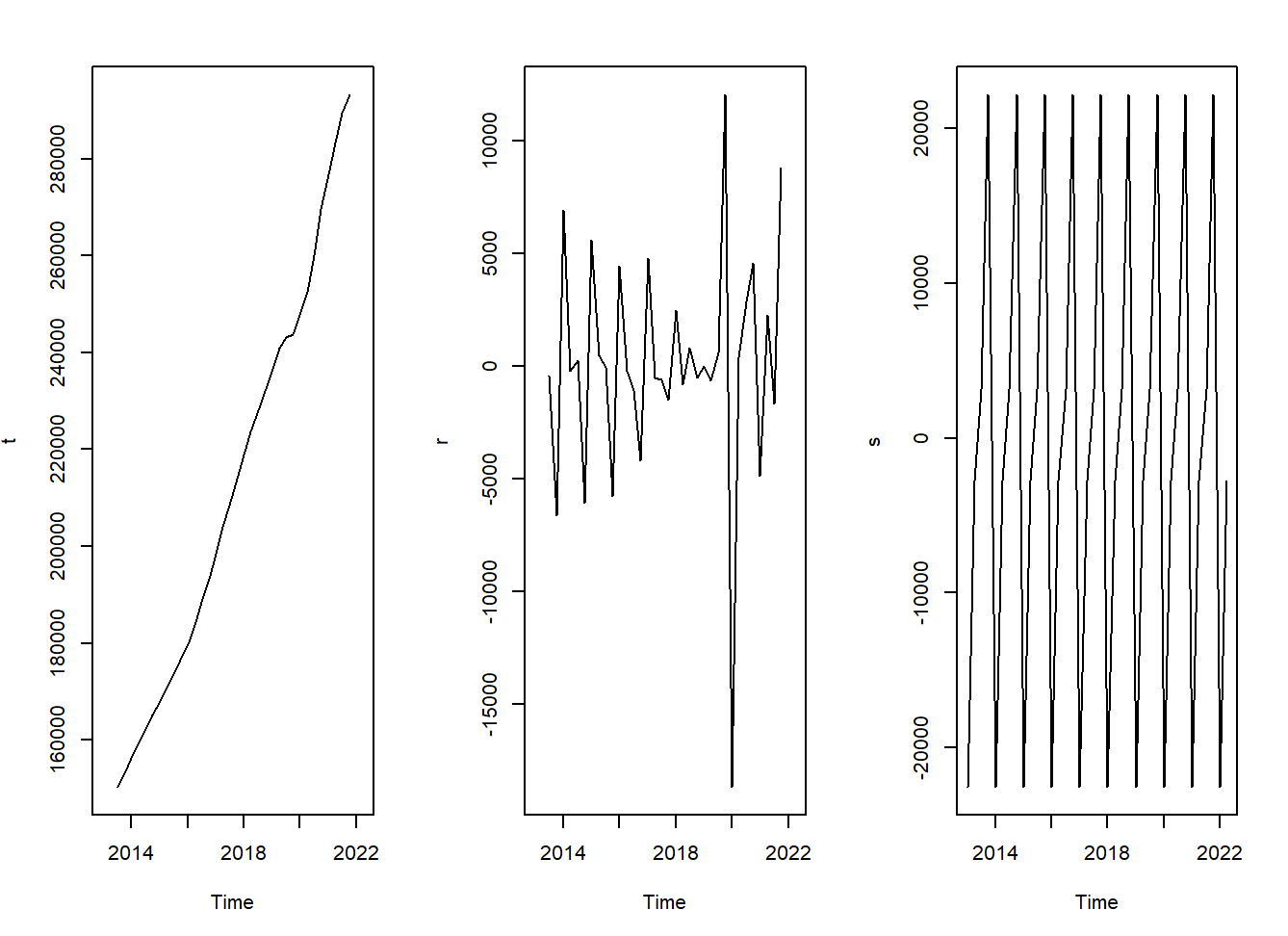
可以发现MA模型的滞后阶数为1阶。

## 模型的参数估计及模型预测

1. 在没有考虑gdp的季节性时估计最佳模型
2. auto.arima(gdpst,max.p=8,max.q=8,d=0,D=0,seasonal=FALSE)
3. bestmodel<-auto.arima(gdpst)
4. gg=forecast(bestmodel,h=12)
5. plot(gg)
6. lines(gdpst)



1. 考虑季节性
2. t<-decompose(gdpst)$trend
3. s<-decompose(gdpst)$seasonal
4. r<-decompose(gdpst)$random
5. par(mfrow=c(1,3))
6. plot(t)
7. plot(r)
8. plot(s)
9. ss<-ma(gdpst,3)
10. ss



为了平滑季节性，MA模型选择2q+1个观测量

1. ss<-ma(gdpst,3)
2. ss
3. plot(ss)plot(gdpst)
4. par(mfrow=c(1,1))
5. acf(diff(na.omit(ss)),lag=40)
6. gdp.arima1<-arima(gdpst,order=c(1,1,0),seas=list(order=c(1,0,0),4),method = 'ML')
7. gdp.arima2<-arima(gdpst,order=c(1,1,0),seas=list(order=c(1,1,0),4),method = 'ML')
8. BIC(gdp.arima1)
9. BIC(gdp.arima2)
10. acf(resid(gdp.arima1))
11. acf(resid(gdp.arima2))

最终得到关于gdp包含季节性的最佳模型

1. bestmodel2<-auto.arima(gdpst,max.p=8,max.q=8,d=1,D=1,seasonal=TRUE)
2. #consider the non-stationary and seasonal effect
3. bestmodel2$arma
4. bestmodel2
5. tstats<-coef(bestmodel2)/sqrt(diag(bestmodel2$var.coef))
6. tstats
7. par(mfrow=c(1,3))
8. acf(resid(bestmodel2))
9. acf(resid(bestmodel2),ci=0.99)
10. #show the selection of lag order
11. gg=forecast(bestmodel2,h=12)
12. plot(gg)
13. lines(gdpst)

