502_HW6

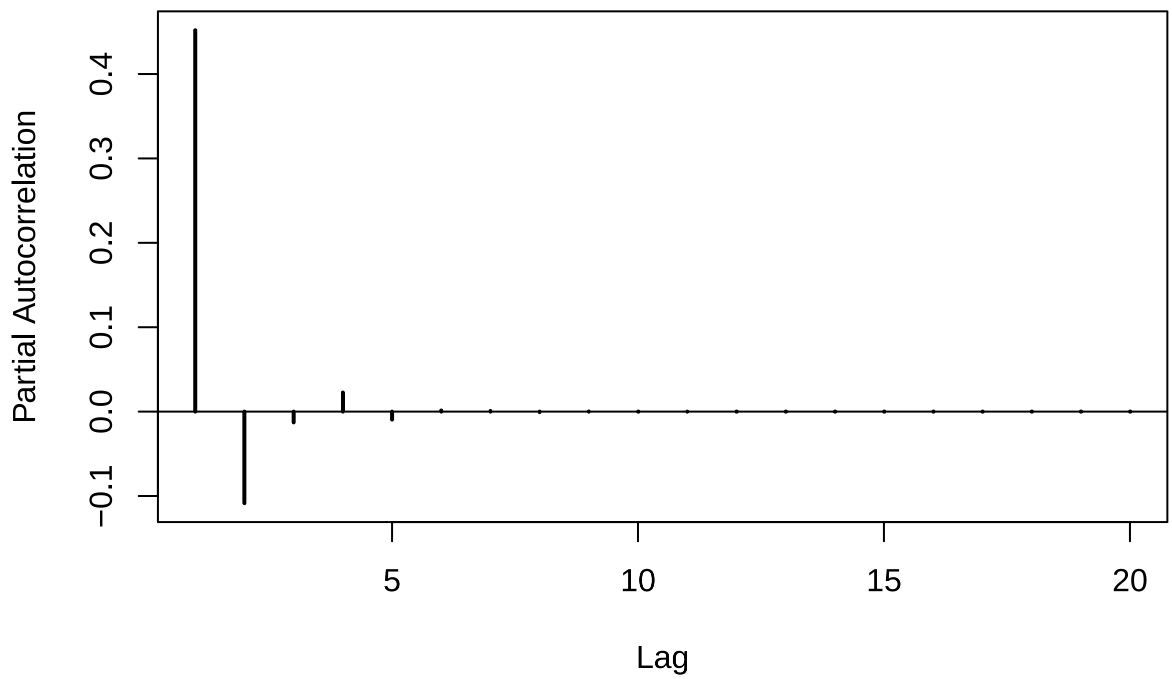
Shengbo Jin

2/26/2022

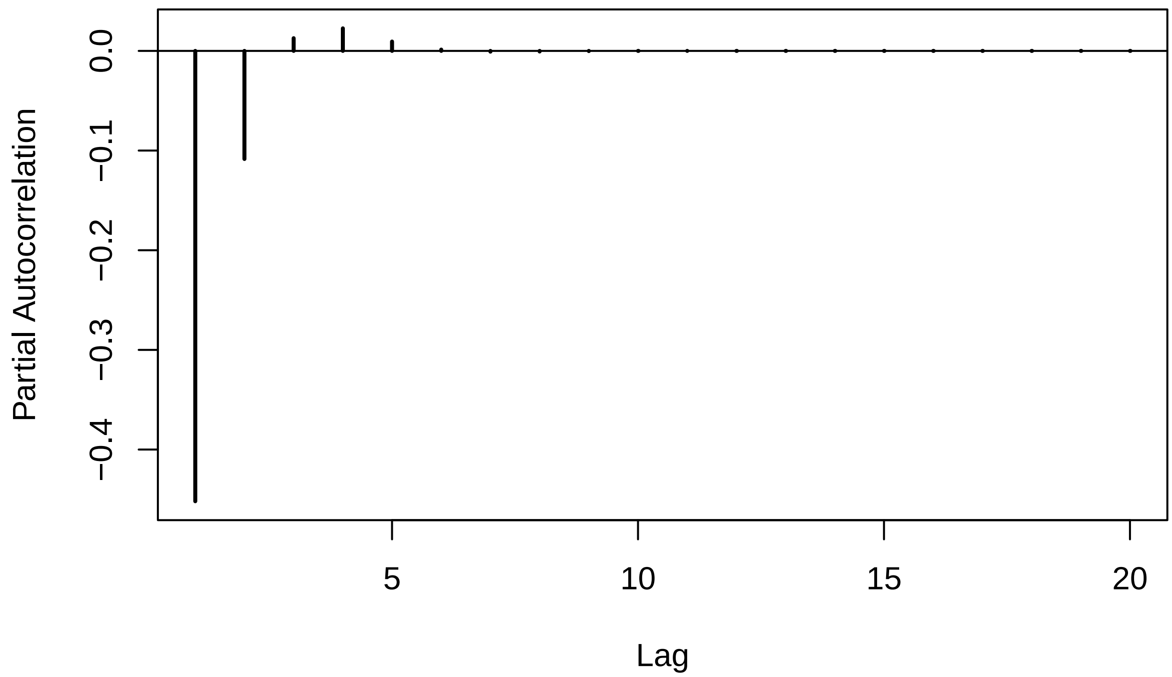
Q1

(a)

```r
plot(1:20, ARMAacf(ma=c(0.5,0.15), lag.max=20, pacf=TRUE), type="h",
     xlab="Lag", ylab="Partial Autocorrelation", lwd=2)
abline(h=0)
```




```r
plot(1:20, ARMAacf(ma=c(-0.5,0.15), lag.max=20, pacf=TRUE), type="h",
     xlab="Lag", ylab="Partial Autocorrelation", lwd=2)
abline(h=0)
```



The partial autocorrelation functions of both MA(2) tail off. For each lag, both pacf have the same absolute values. They are exactly the same in even lags, but opposite in odd lags.

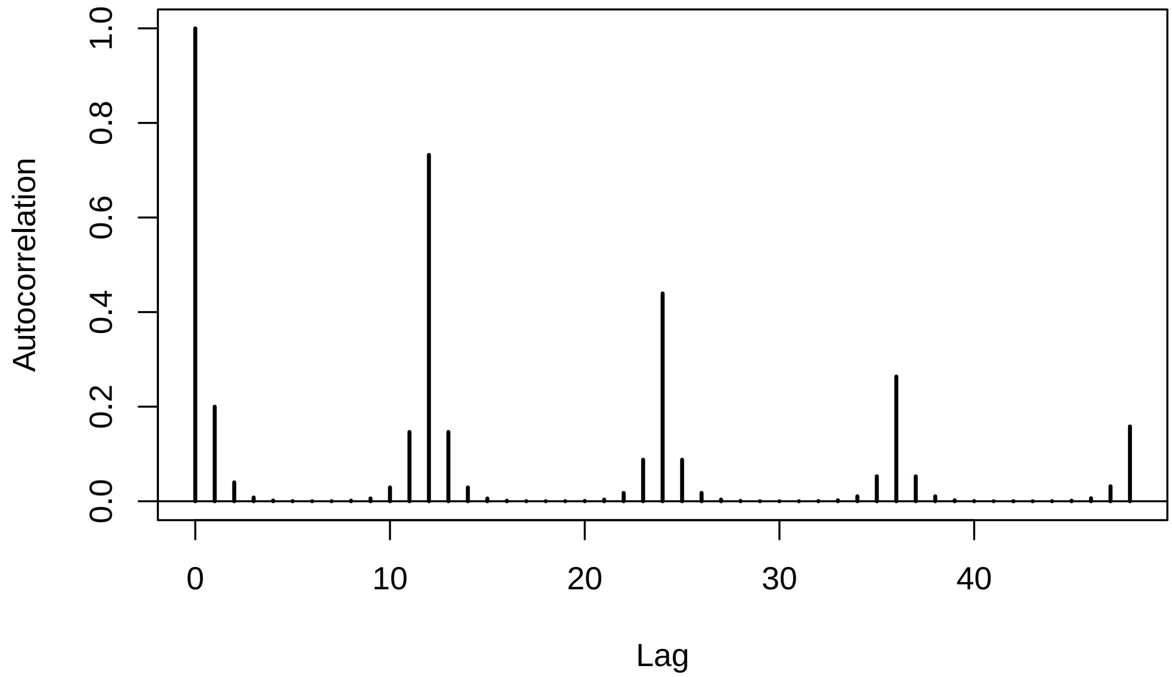
(b)

$$\phi(B)\Phi(B^{12})\Delta_{12} Y_t = \theta(B)\Theta(B^{12})\epsilon_t$$

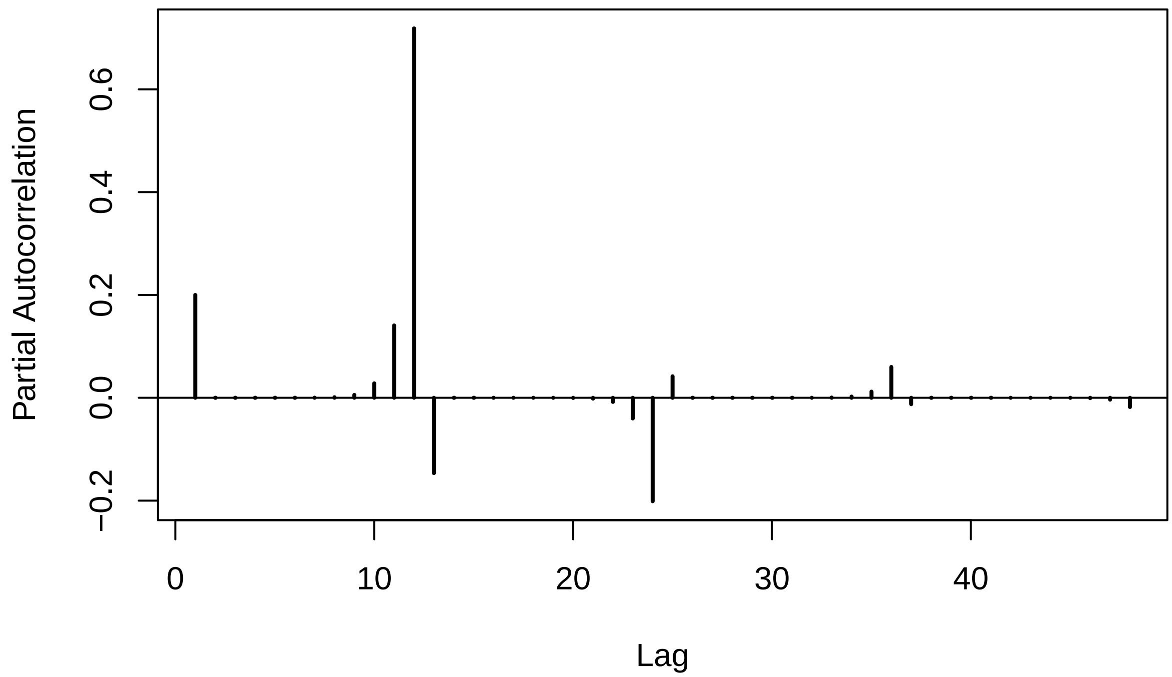
$$(1-\phi_1 B)(1-\Phi_1 B^{12}) Y_t = (1+\Theta_1 B^{12})\epsilon_t$$

$$Y_t - \phi_1 Y_{t-1} - \Phi_1 Y_{t-12} + \phi_1 \Phi_1 Y_{t-13} = \epsilon_t + \Theta_1 \epsilon_{t-12}$$

```
plot(0:48, ARMAacf(ar=c(0.2,rep(0,10),0.6,-0.12), ma=c(rep(0,11),0.3),
                   lag.max=48), type="h", xlab="Lag",
                   ylab="Autocorrelation", lwd=2)
abline(h=0)
```




```
plot(1:48, ARMAacf(ar=c(0.2,rep(0,10),0.6,-0.12), ma=c(rep(0,11),0.3),
                   lag.max=48, pacf=TRUE), type="h", xlab="Lag",
                   ylab="Partial Autocorrelation", lwd=2)
abline(h=0)
```



At each period of 12, both acf and pacf tails off after 1, so the seasonal order should be (1,1). At lag 1, acf tails off but pacf cuts off, so the regular order should be (1,0).

Q2

```
dax.logprice <- log(EuStockMarkets[1:1000,1])
```

(a)

```
library(forecast)
d1q1.fit <- Arima(dax.logprice, order=c(0,1,1))
(d1q1.fc <- forecast(d1q1.fit, h=1, level=0.95))
```

```
##      Point Forecast    Lo 95   Hi 95
## 1001       7.609808 7.590796 7.62882
```

(b)

$$X_{t+1} = \varepsilon_{t+1} + \theta\varepsilon_t$$

$$\gamma(h) = \begin{cases} (1+\theta^2)\sigma^2 & h=0 \\ \theta\sigma^2 & h=1 \\ 0 & h=2,3,\cdots \end{cases}$$

$$\hat{X}_{n+1} = a_1X_1 + a_2X_2 + \cdots\ a_nX_n$$

$$\Gamma_n = (\gamma(i-j))_{i,j=1}^{n} \in R^{n\times n} \quad c = (\gamma(n),\dots,\gamma(1)) \in R^n$$

$$a = \Gamma_n^{-1}c$$

$$\hat{X}_{n+1} = a^T X$$

$$\hat{Y}_{n+1} = Y_n + \hat{X}_{n+1}$$

```r
theta <- d1q1.fit$coef
sigma2 <- d1q1.fit$sigma2
Yn <- dax.logprice[1000]

n <- 1000
c <- numeric(n)
c[n] <- sigma2*theta

G = matrix(nrow=n,ncol=n,0)

for (i in 1:n) {
    for (j in i:n) {
      if (i-j == 0 ) {
        G[i,j] <- (1+theta^2)*sigma2
      }
      if (abs(i-j) == 1){
        G[i,j] <- theta*sigma2
      }
      if (abs(i-j) > 1){
        G[i,j] <- 0
      }
    }
}

a <- solve(G) %*% c
```

```r
X <- c(0, diff(dax.logprice))

(Yhat <- Yn + t(a) %*% X)
```

```
##          [,1]
## [1,] 7.609808
```

(c)

$$
\begin{aligned}
& Var(Y_{n+1} - \hat{Y}_{n+1}) \\
&= Var(Y_n + X_{n+1} - Y_n - \hat{X}_{n+1}) \\
&= Var(X_{n+1} - \hat{X}_{n+1}) \\
&= Var(\hat{\varepsilon}_{n+1}) = \sigma^2 \\
& E(Y_{n+1} - \hat{Y}_{n+1}) = 0 \\
& P\left(-z_{0.975} < \frac{Y_{n+1} - \hat{Y}_{n+1}}{\sigma} < z_{0.975}\right) = 0.95 \\
& (\hat{Y}_{n+1} - z_{0.975}\sigma \,,\, \hat{Y}_{n+1} + z_{0.975}\sigma)
\end{aligned}
$$

```r
low <- Yhat - qnorm(0.975)*sqrt(d1q1.fit$sigma2)
high <- Yhat + qnorm(0.975)*sqrt(d1q1.fit$sigma2)
(PI <- data.frame(Yhat, low, high))
```

```
##       Yhat      low    high
## 1 7.609808 7.590796 7.62882
```

Q3

$$\Delta Y_t = (1-B)Y_t = Y_t - Y_{t-1} = \varepsilon_t$$

$\because \{\varepsilon_t\} \sim WN(0, \sigma^2)$

$\therefore \{\Delta Y_t\}$ is stationary

$$\Delta^2 Y_t = (1-B)\Delta Y_t = (1-B)\varepsilon_t$$
$$= \varepsilon_t - \varepsilon_{t-1}$$

$$\theta(z) = 1-z = 0$$
$$z = 1$$

$\therefore \{\Delta^2 Y_t\}$ is a $MA(1)$ process where the moving average polynomial has a unit root

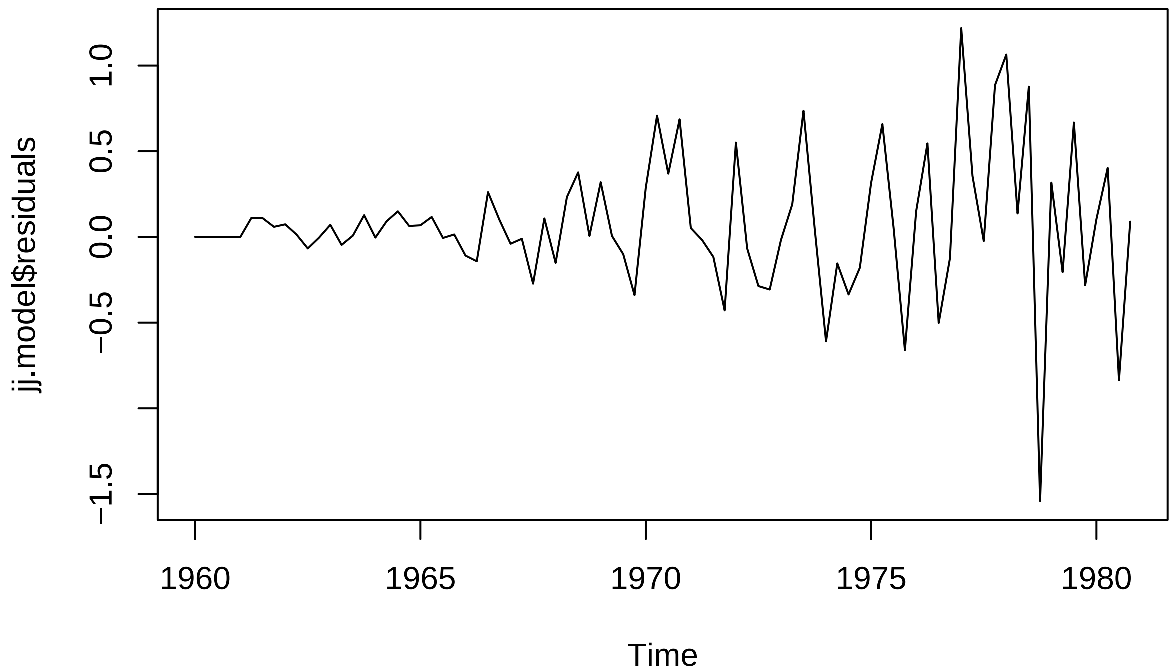
Q4

```r
data(jj, package="astsa")
head(jj, 8)
```

```
##      Qtr1 Qtr2 Qtr3 Qtr4
## 1960 0.71 0.63 0.85 0.44
## 1961 0.61 0.69 0.92 0.55
```

(a)

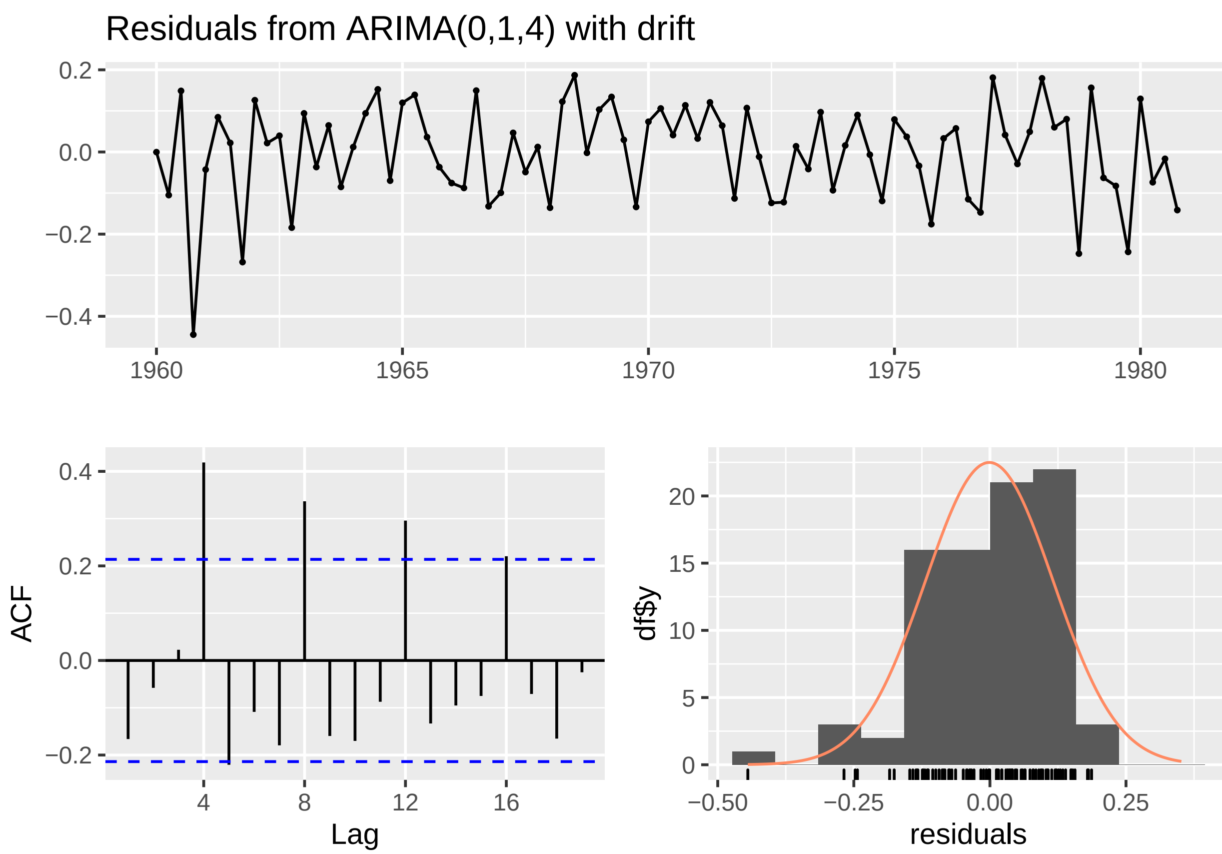
```r
jj.model <- auto.arima(jj)
plot(jj.model$residuals)
```



There is nonconstant variance.

(b)

```
logjj.model <- auto.arima(log(jj), seasonal=FALSE)
checkresiduals(logjj.model)
```

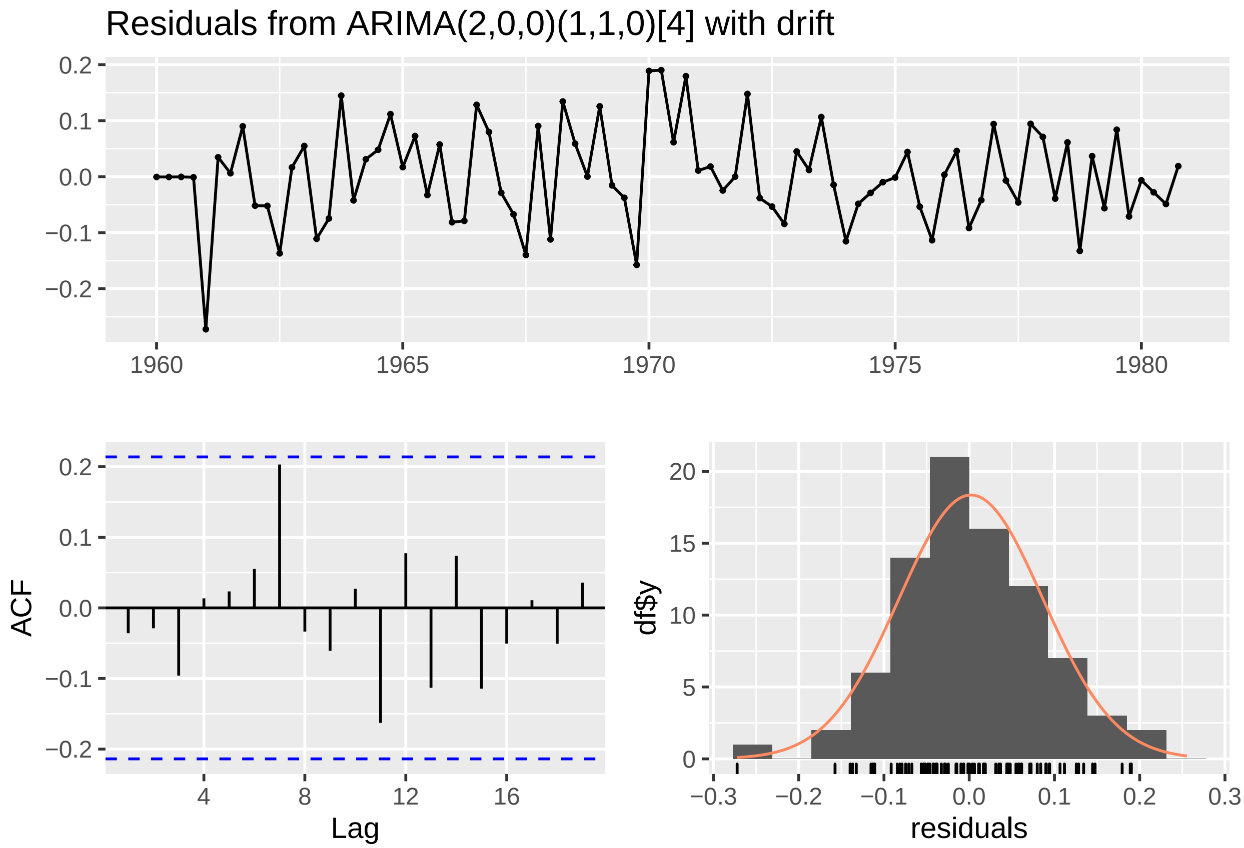



```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,1,4) with drift
## Q* = 37.911, df = 3, p-value = 2.952e-08
##
## Model df: 5.   Total lags used: 8
```

At a 5% significance level, the test rejects null hypothesis that the autocorrelations of the residuals up to lag 8 are zero. Thus, to a large extent, the residuals are not white noise. The ACF plot of residuals shows a period of 4.

(c)

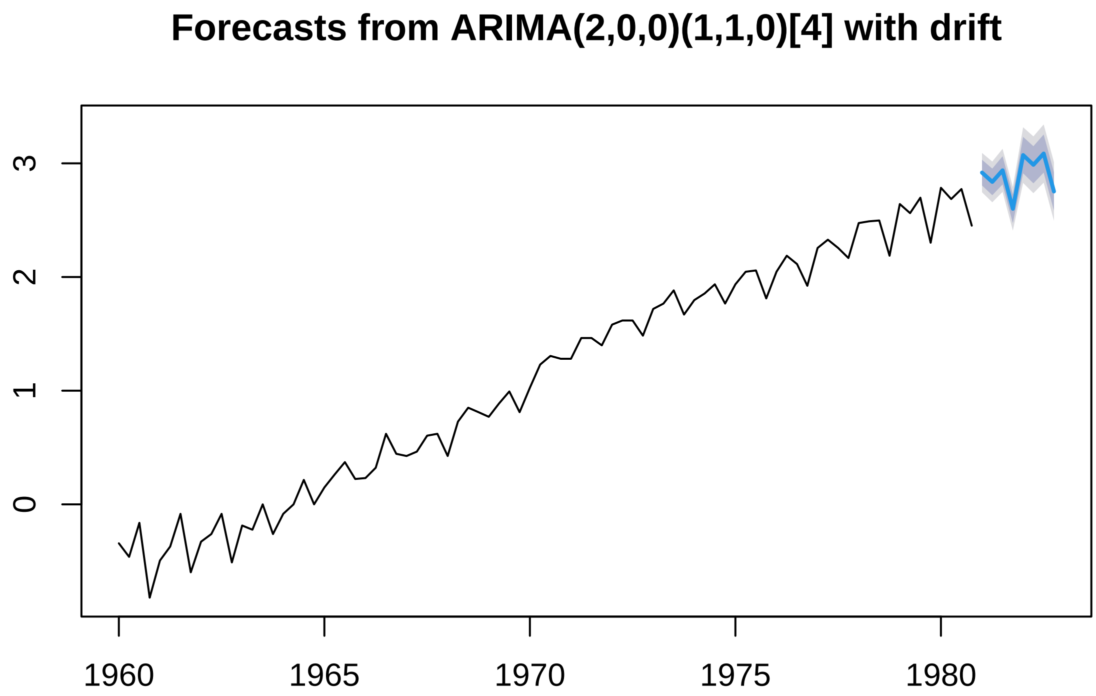
```
logjjs.model <- auto.arima(log(jj))
checkresiduals(logjjs.model)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,0,0)(1,1,0)[4] with drift
## Q* = 5.3331, df = 4, p-value = 0.2548
##
## Model df: 4.   Total lags used: 8
```

The first plot shows residuals are stationary and constant variance. At a 5% significance level, the test fails to reject the null that the autocorrelations of the residuals up to lag 8 are zero. The third plot shows the residuals are normally distributed. In conclusion, the residuals are consistent with white noise, so this model is a good fit.

```
logjjs.fc <- forecast(logjjs.model, h=8)
plot(logjjs.fc)
```


Forecasts from ARIMA(2,0,0)(1,1,0)[4] with drift
0
1
2
3
1960
1965
1970
1975
1980