

Homework 9

Joshua Ingram

11/9/2020

Problem 1

Given

$$Y_t = \delta t + \epsilon_t, \epsilon_t \sim_{i.i.d.} N(0, \sigma^2)$$

1.

$$Y'_t = Y_t - Y_{t-1}$$

a.

$$\begin{aligned} E[Y'_t] &= E[Y_t - Y_{t-1}] = E[Y_t] - E[Y_{t-1}] = E[\delta t + \epsilon_t] - E[\delta(t-1) + \epsilon_{t-1}] = E[\delta t] + 0 - E[\delta(t-1)] - 0 \\ &= E[\delta t] - E[\delta t] + E[\delta] = \delta \end{aligned}$$

Thus

$$E[Y'_t] = \delta$$

b.

$$Var(Y'_t) = Var(Y_t - Y_{t-1}) = Var(Y_t) + Var(Y_{t-1}) = Var(\delta t + \epsilon_t) + Var(\delta(t-1) + \epsilon_{t-1})$$

$$Var(\delta t) + Var(\epsilon_t) + Var(\delta(t-1)) + Var(\epsilon_{t-1}) = 0 + \sigma^2 + 0 + \sigma^2 = 2\sigma^2$$

Thus

$$Var(Y'_t) = 2\sigma^2$$

c.

Given

$$Y'_t = Y_t - Y_{t-1} = \delta t + \epsilon_t - \delta(t-1) - \epsilon_{t-1} = \delta t - \delta t + \delta + \epsilon_t - \epsilon_{t-1} = \delta + \epsilon_t - \epsilon_{t-1}$$

and

$$Y'_{t-1} = Y_{t-1} - Y_{t-2} = \delta(t-1) + \epsilon_{t-1} - \delta(t-2) - \epsilon_{t-2} = \delta t - \delta t - \delta + 2\delta + \epsilon_{t-1} - \epsilon_{t-2} = \delta + \epsilon_{t-1} - \epsilon_{t-2}$$

Then

$$\text{cov}(Y'_t, Y'_{t-1}) = \text{cov}(Y_t - Y_{t-1}, Y_{t-1} - Y_{t-2}) = \text{cov}(\delta + \epsilon_t - \epsilon_{t-1}, \delta + \epsilon_{t-1} - \epsilon_{t-2})$$

Given δ is a constant and each ϵ_t is i.i.d.

$$= \text{cov}(\epsilon_t - \epsilon_{t-1}, \epsilon_{t-1} - \epsilon_{t-2}) = \text{cov}(-\epsilon_{t-1}, \epsilon_{t-1}) = -\sigma^2$$

d.

Given

$$Y'_t = \delta + \epsilon_t - \epsilon_{t-1}$$

and

$$Y'_{t-k} = Y_{t-k} - Y_{t-(k+1)} = \delta(t-k) + \epsilon_{t-k} - \delta(t-(k+1)) - \epsilon_{t-(k+1)} = \delta t - \delta t - k\delta + (k+1)\delta + \epsilon_{t-k} - \epsilon_{t-(k+1)} = \delta + \epsilon_{t-k} - \epsilon_{t-(k+1)}$$

Such that $k > 1$, then

$$\text{cov}(Y_t, Y_{t-k}) = \text{cov}(Y_t - Y_{t-1}, Y_{t-k} - Y_{t-(k+1)}) = \text{cov}(\delta + \epsilon_t - \epsilon_{t-1}, \delta + \epsilon_{t-k} - \epsilon_{t-(k+1)})$$

Given δ is a constant and each ϵ_t is i.i.d.

$$= \text{cov}(\epsilon_t - \epsilon_{t-1}, \epsilon_{t-k} - \epsilon_{t-(k+1)}) = 0$$

Thus

$$\text{cov}(Y_t, Y_{t-k}) = 0$$

for all t where $k > 1$.

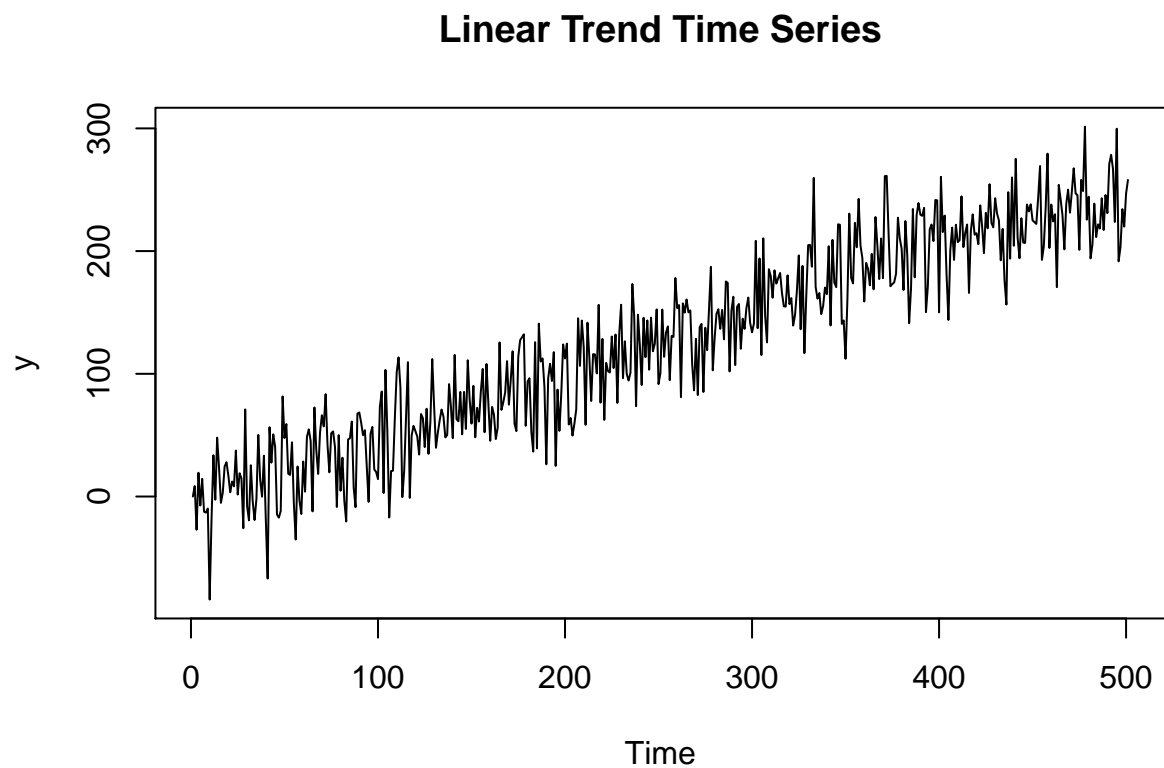
Given the preceding properties, differencing a linear trend makes the data stationary.

2.

```
set.seed(420)
y <- 0
for(i in 1:500){
  y <- c(y, 0.5*i+ rnorm(1,0,30))
}
```

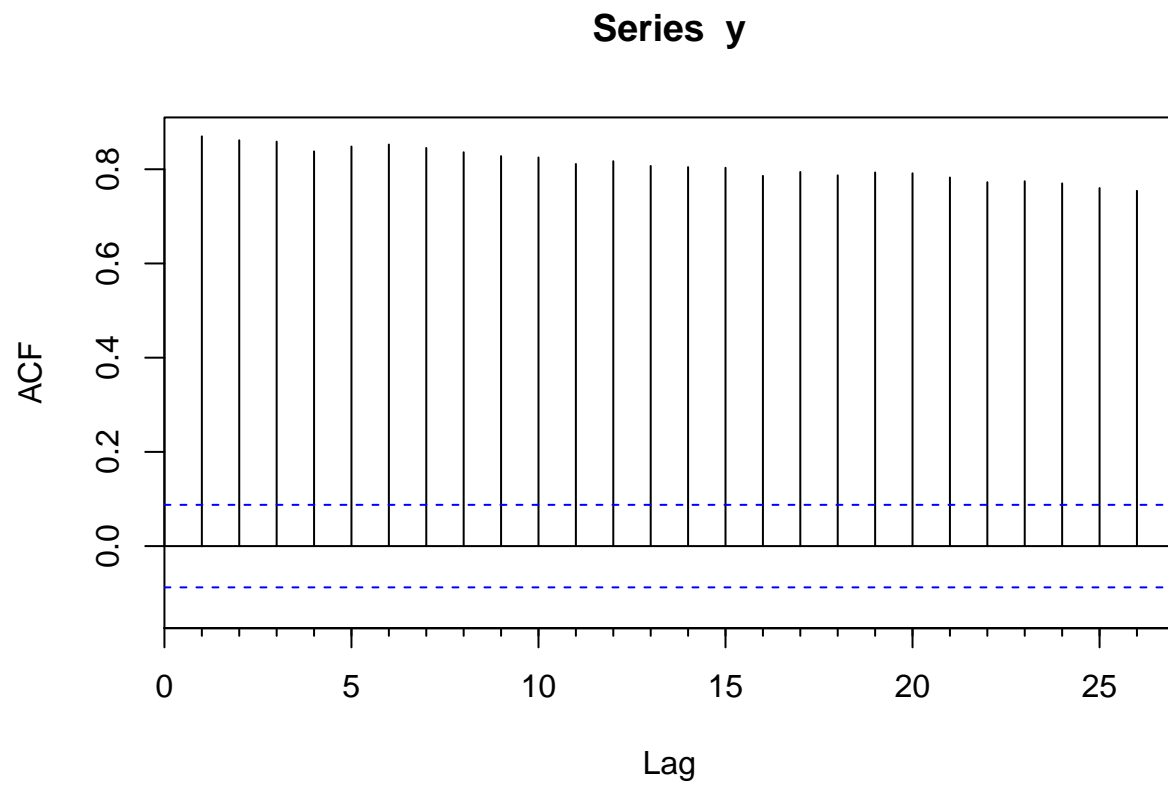
a.

```
plot(y, type="l", xlab="Time")
title("Linear Trend Time Series")
```



This time series data is clearly not stationary, as the mean is non-constant. I would expect a slowly decreasing ACF because of the increasing trend in the data with no seasonality.

```
Acf(y)
```

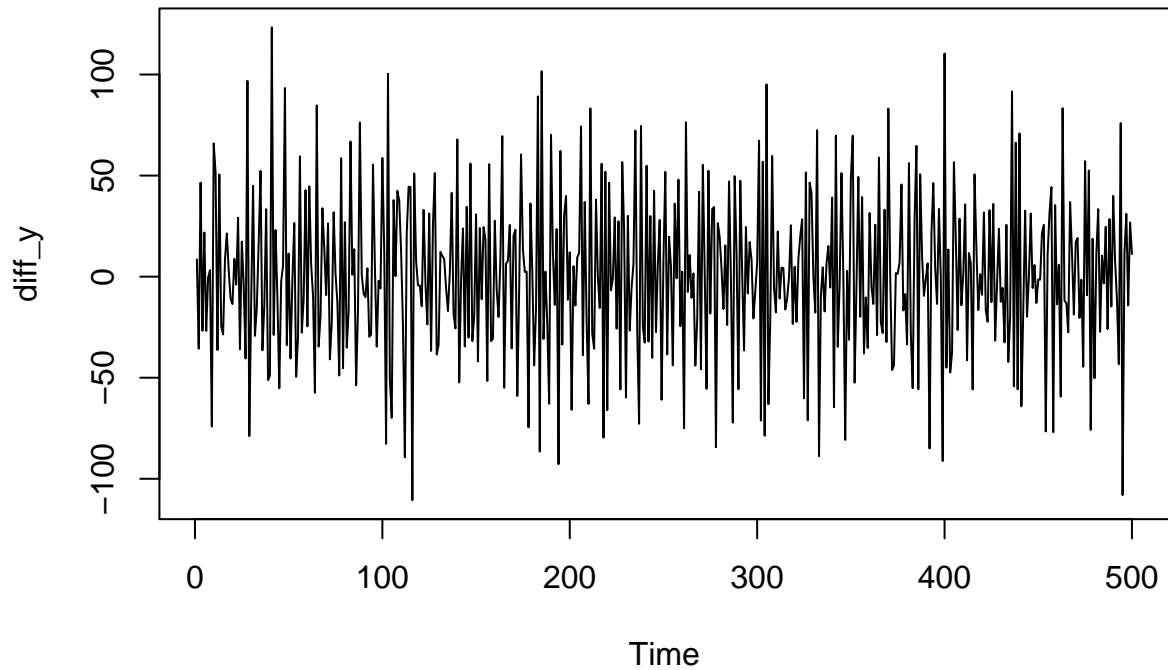


b.

```
diff_y <- diff(y)

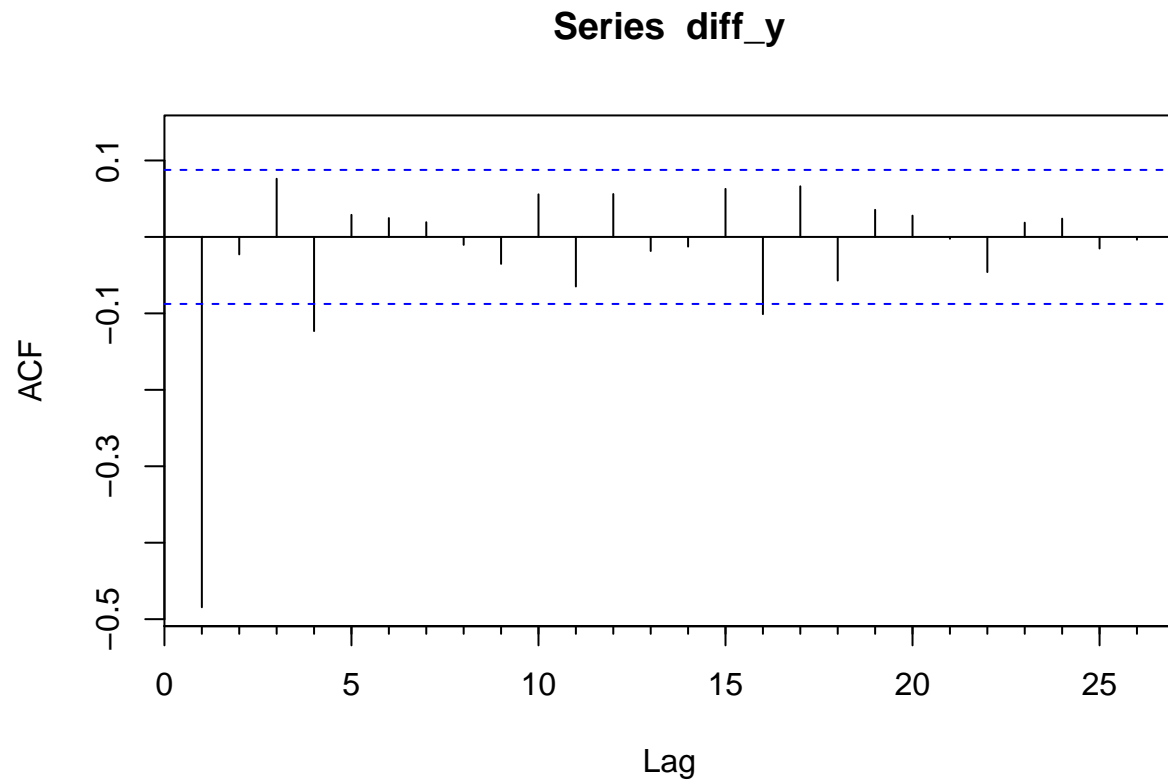
plot(diff_y, type="l", xlab="Time")
title("Differenced Linear Trend Time Series")
```

Differenced Linear Trend Time Series



This does look like a stationary time series because the mean is constant and same for the variance. I would expect an ACF where all autocorrelations are within the blue lines with no clear pattern. There would likely be a high autocorrelation for lag 1.

```
Acf(diff_y)
```



3.

```
# original time series
kpss.test(y)
```

```
## Warning in kpss.test(y): p-value smaller than printed p-value
```

```
##
## KPSS Test for Level Stationarity
##
## data: y
## KPSS Level = 8.3389, Truncation lag parameter = 5, p-value = 0.01
```

```
# differenced time series
kpss.test(diff_y)
```

```
## Warning in kpss.test(diff_y): p-value greater than printed p-value
```

```
##
## KPSS Test for Level Stationarity
##
## data: diff_y
## KPSS Level = 0.0062142, Truncation lag parameter = 5, p-value = 0.1
```

We use the KPSS test to determine whether the series is stationary. For the original time series data, we receive a small p-value which gives us significant evidence to reject the null, that the series is stationary, in favor of the alternative, that the series is not stationary.

For the differenced time series, we receive a p-value of 0.1, so we retain the null, that the time series is stationary. These tests do confirm the visual intuitions.

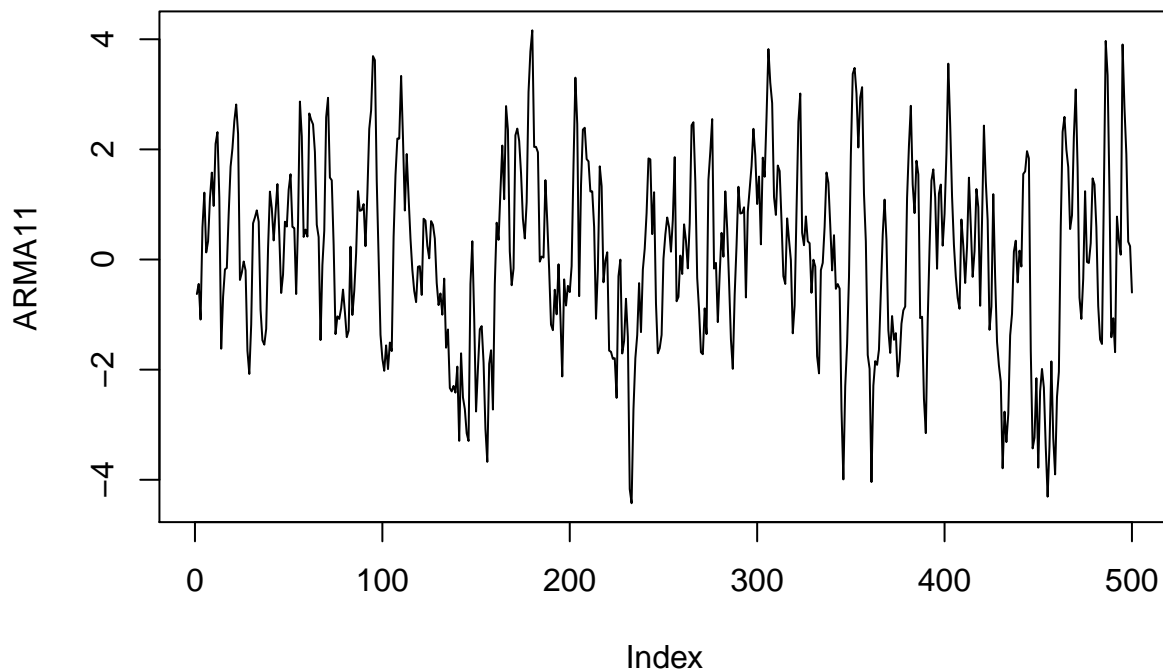
Problem 2

1.

```
ts.length <- 500
sigma <- 1

set.seed(1)
err <- rnorm(ts.length, sd=1)
ARMA11 <- err[1]
for (j in 2:ts.length){
  ARMA11 <- c(ARMA11, 0.7*ARMA11[j-1]+0.3*err[j-1]+err[j])
}

plot(ARMA11, type="l")
```



a.

```
kpss.test(ARMA11)
```

```
## Warning in kpss.test(ARMA11): p-value greater than printed p-value
```

```
##
```

```
## KPSS Test for Level Stationarity
```

```
##
```

```
## data: ARMA11
```

```
## KPSS Level = 0.11731, Truncation lag parameter = 5, p-value = 0.1
```

With a p-value of 0.1 after conducting the KPSS test, we do not have significant evidence to reject the null, that the time series is stationary. This does confirm my intuition from HW 8.

b.

```
ndiffs(ARMA11)
```

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

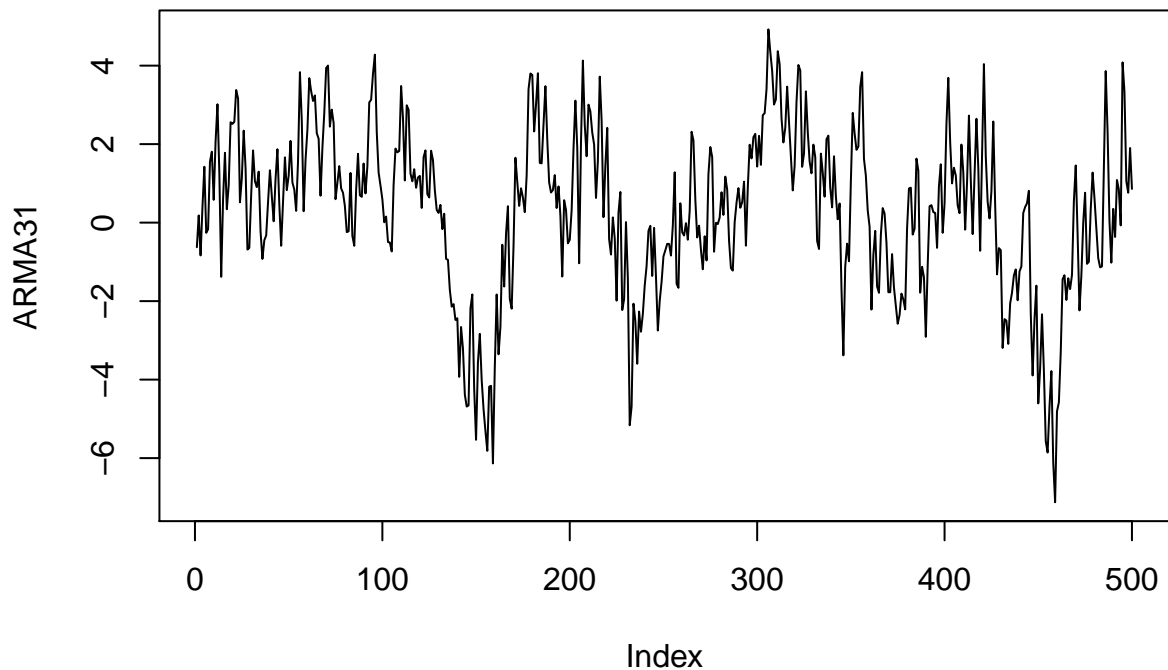
No differencing needed, so it corresponds to my conclusion from part a.

c.

ndiffs yielded a 0.

2.

```
set.seed(1)
err <- rnorm(ts.length, sd=1)
ARMA31 <- err[1:3]
for(j in 4:ts.length){
  ARMA31 <- c(ARMA31, sum(c(0.7, -0.3, 0.5)*ARMA31[c(j-1,j-2,j-3)])+0.3*err[j-1]+err[j])
}
plot(ARMA31, type="l")
```

```
kpss.test(ARMA31)
```

```
##
## KPSS Test for Level Stationarity
##
## data: ARMA31
## KPSS Level = 0.4863, Truncation lag parameter = 5, p-value = 0.04475
```

With a p-value of 0.04475, we have significant evidence to reject the null in favor of the alternative, that the time series is not stationary. These results were in line with my intuition in the last homework.

b.

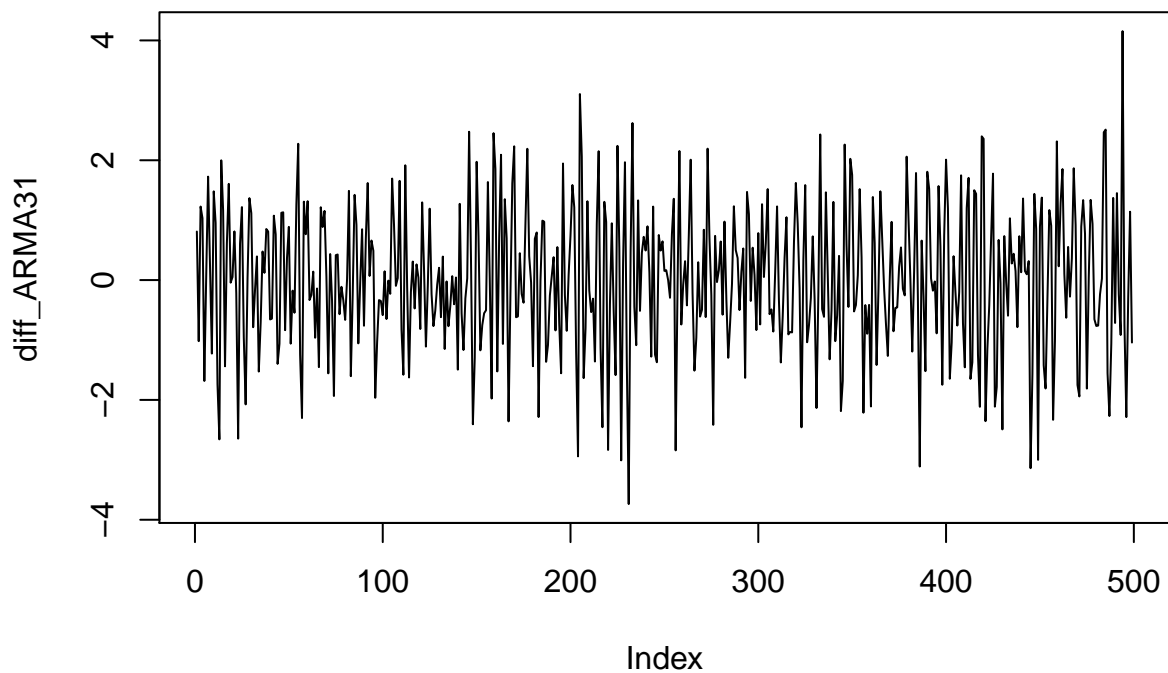
```
ndiffs(ARMA31)
```

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

This result is consistent with the results in part a.

c.

```
diff_ARMA31 <- diff(ARMA31)
plot(diff_ARMA31, type="l")
```



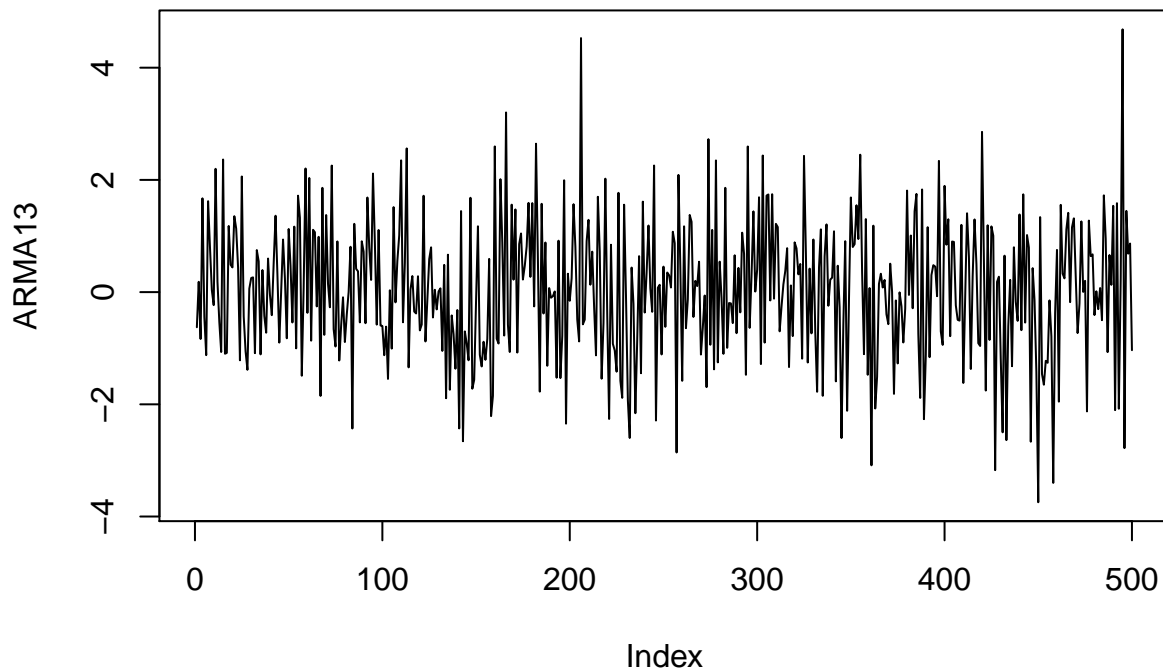
```
kpss.test(diff_ARMA31)
```

```
## Warning in kpss.test(diff_ARMA31): p-value greater than printed p-value
```

```
##
## KPSS Test for Level Stationarity
##
## data: diff_ARMA31
## KPSS Level = 0.014646, Truncation lag parameter = 5, p-value = 0.1
```

3.

```
set.seed(1)
err <- rnorm(ts.length, sd=1)
ARMA13 <- err[1:3]
for(j in 4:ts.length){
  ARMA13 <- c(ARMA13, 0.3*tail(ARMA13,1)+ sum(c(-0.7, 0.3, 0.5)*err[c(j-1, j-2,j-3)])+err[j])
}
plot(ARMA13, type="l")
```



```
kpss.test(ARMA13)
```

```
## Warning in kpss.test(ARMA13): p-value greater than printed p-value
```

```
##
```

```
## KPSS Test for Level Stationarity
```

```
##
```

```
## data: ARMA13
```

```
## KPSS Level = 0.097715, Truncation lag parameter = 5, p-value = 0.1
```

With a p-value of 0.10, we do not have significant evidence to reject the null hypothesis, so the time series is stationary. This is consistent with my intuition on the last homework.

b.

```
ndiffs(ARMA13)
```

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

This is consistent with my results from the last part.

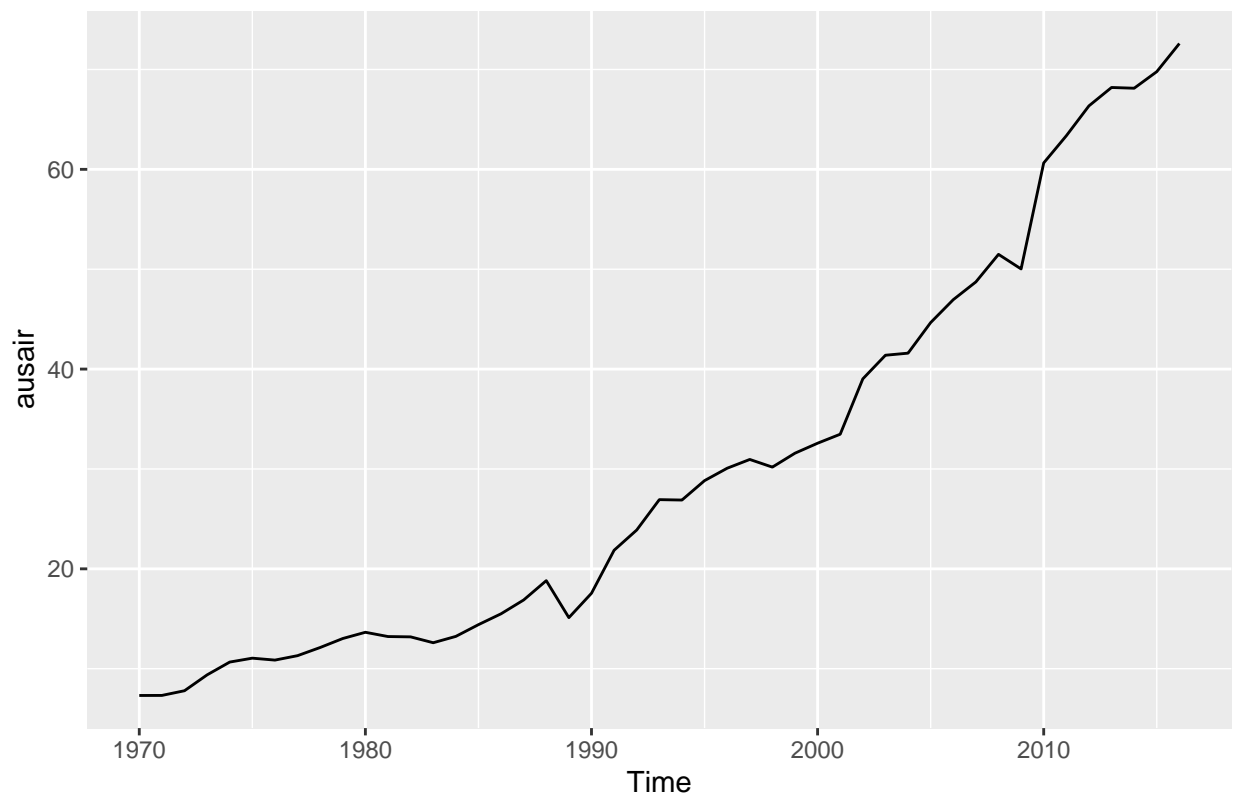
c.

Problem 3

1.

a.

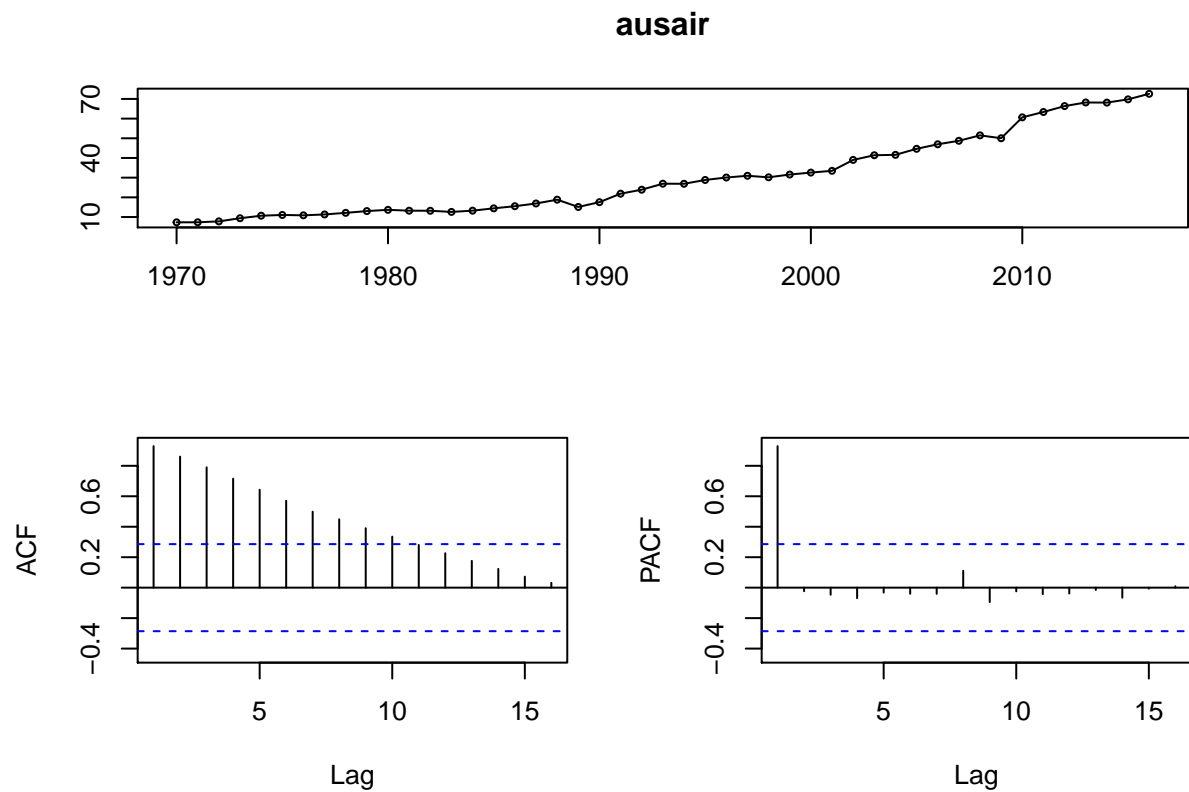
```
autoplot(ausair)
```



Variane looks fine.

b.

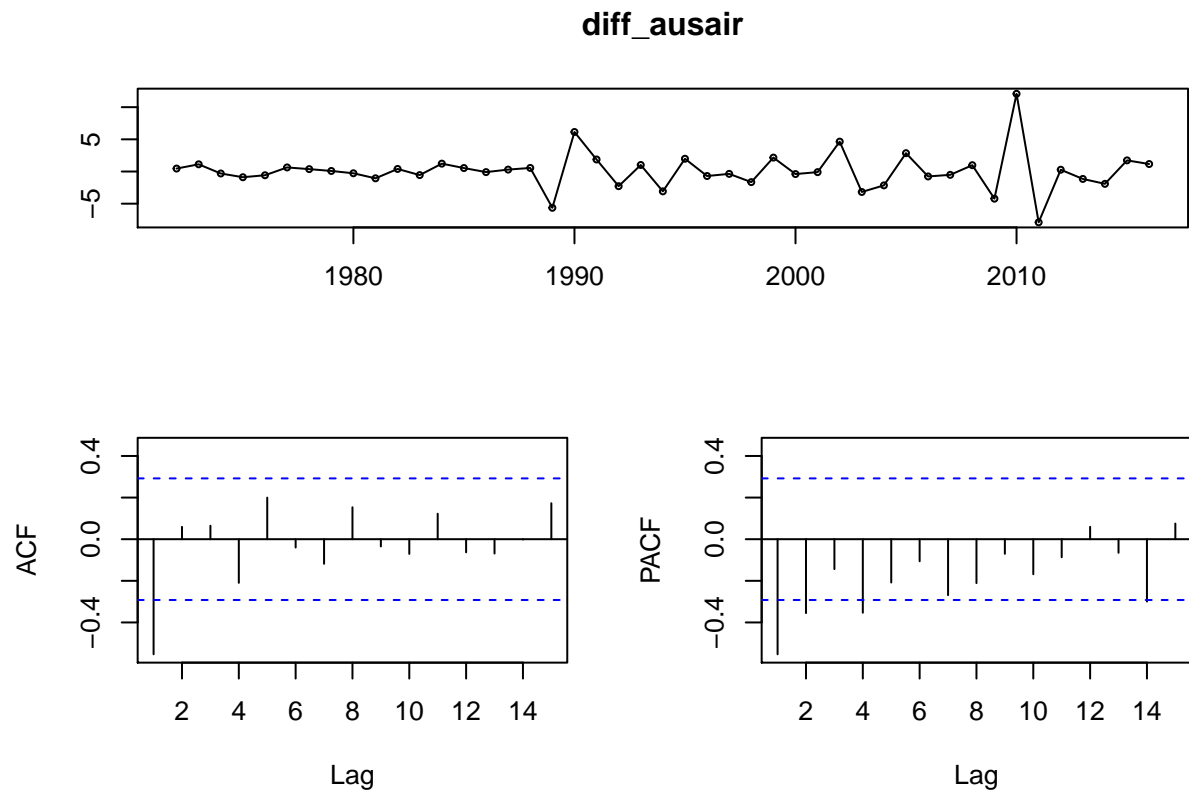
```
tsdisplay(ausair)
```



```
ndiffs(ausair)
```

```
## [1] 2
```

```
diff_ausair <- diff(ausair, differences = 2)
tsdisplay(diff_ausair)
```



c.

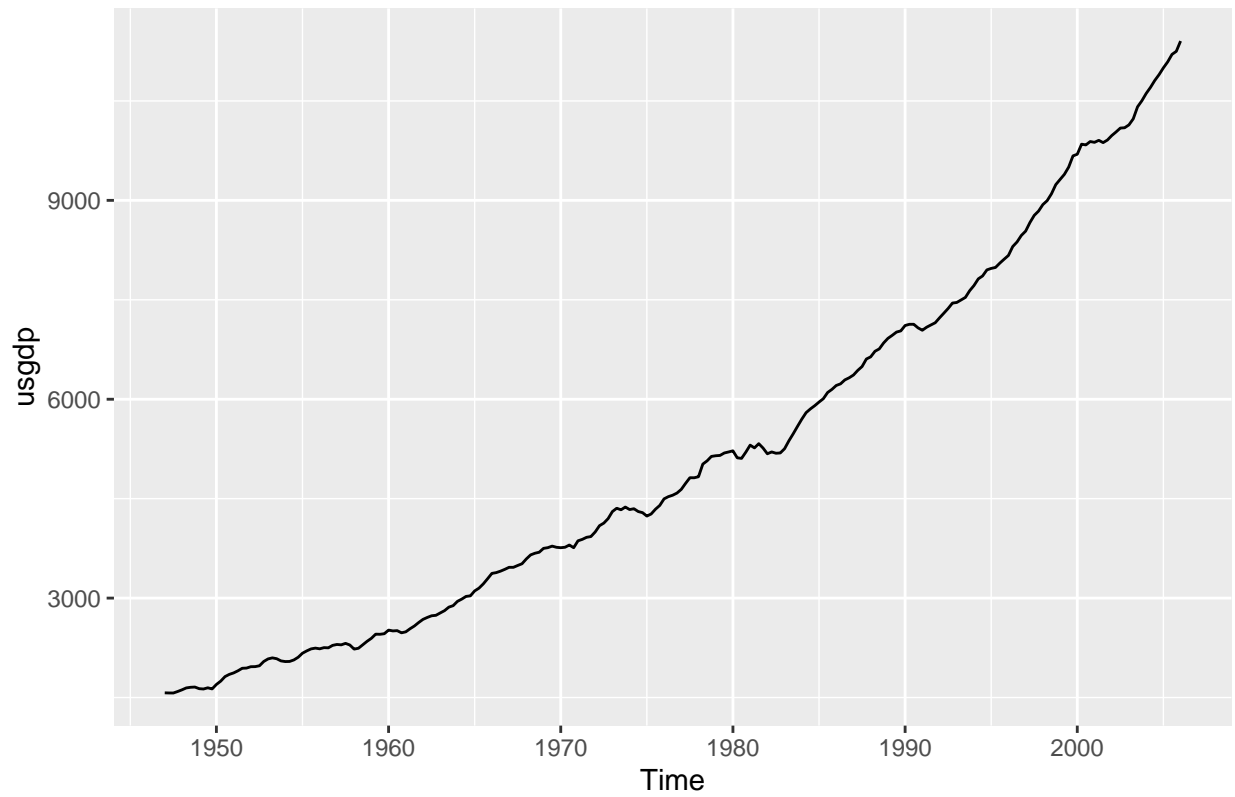
For this time series, ARIMA(0,2,1) would be appropriate. I chose this because of the spike in the ACF for lag 1 and the sudden drop off, as well as the quick decrease in the PACF plot. d is 2 because of the second order differencing.

$$y_t = \mu + \theta_1 \epsilon_{t-1} + \epsilon_t, \quad d = 2$$

2.

a.

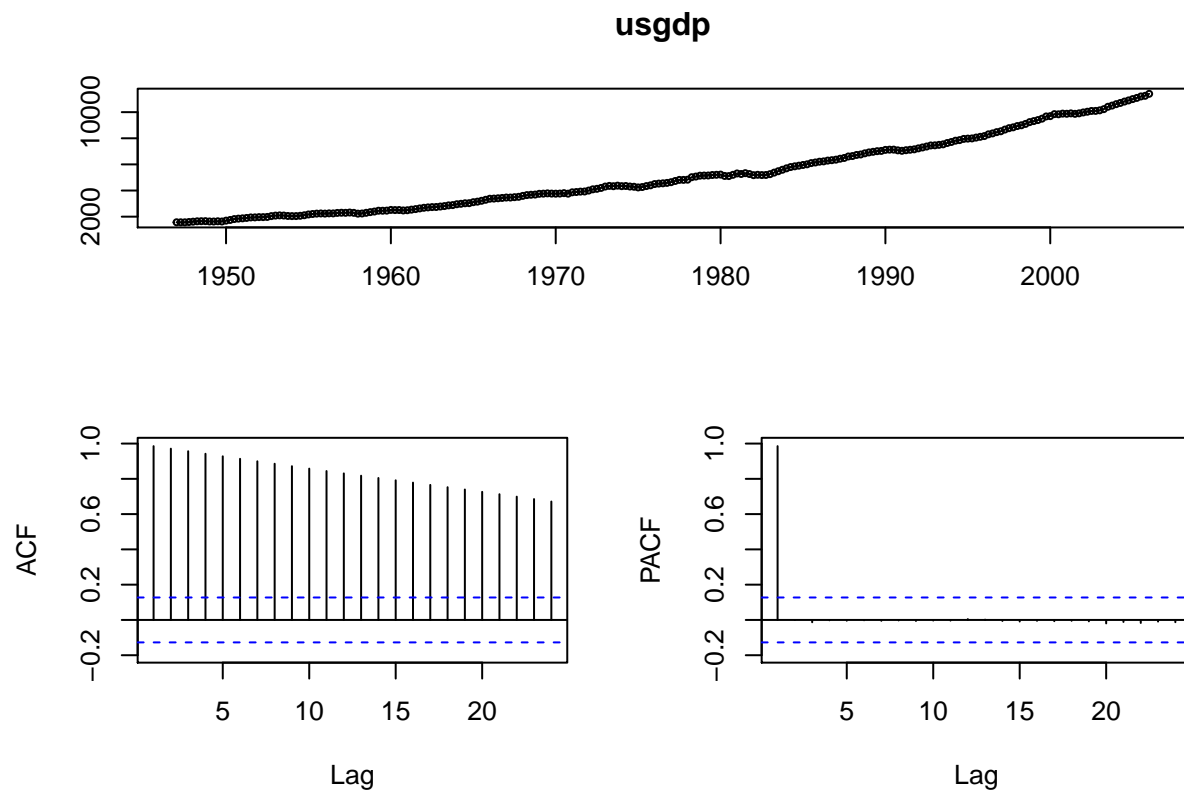
```
autoplot(usgdp)
```



Variance looks fine.

b.

```
tsdisplay(usgdp)
```



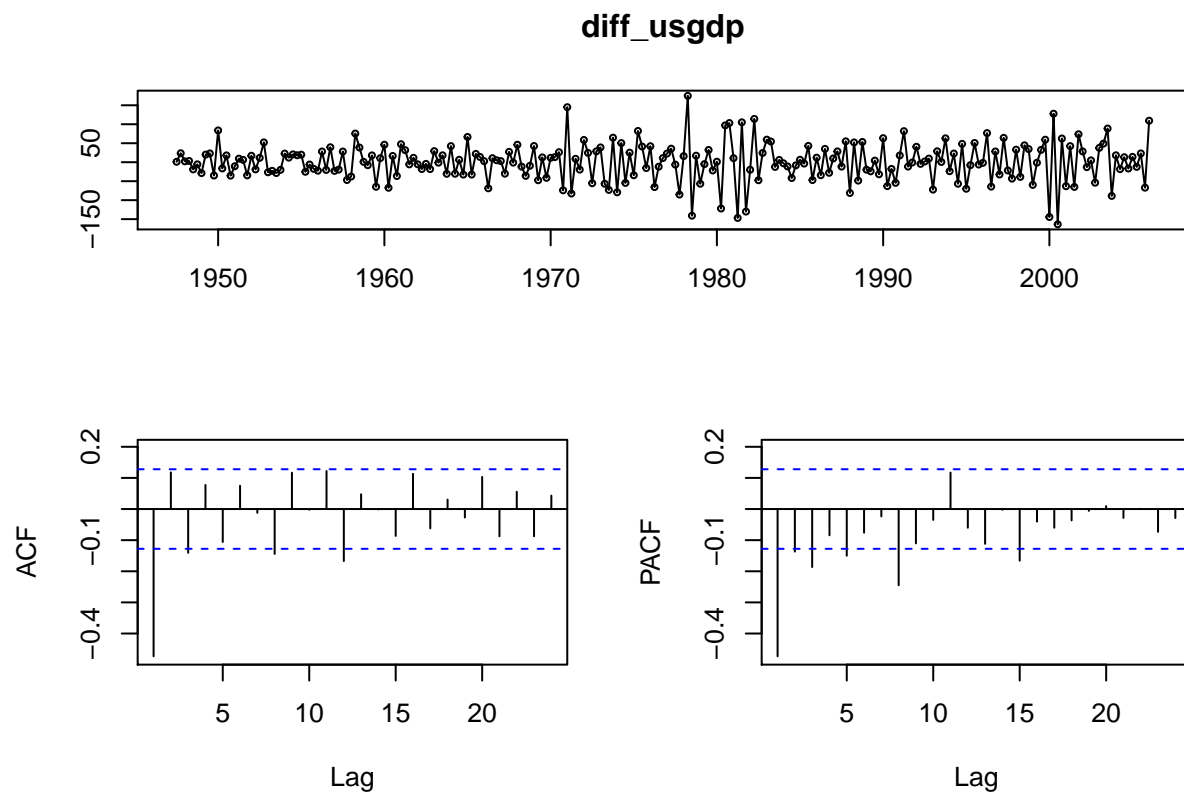
```
ndiffs(usgdp)
```

```
## [1] 2
```

```
nsdiffs(usgdp)
```

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

```
diff_usgdp <- diff(usgdp, differences = 2)
tsdisplay(diff_usgdp)
```

c.

For this, ARIMA(1,2,1) would be appropriate. I chose this because of the spike for lag 1 in the ACF and the spike in the PACF, followed by sudden drops in both plots. d is two because of the differencing being on the second order.

$$y_t = \mu + \phi_1 y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t, \quad d = 2$$