```r
library(TSA)
library(itsmr)
```

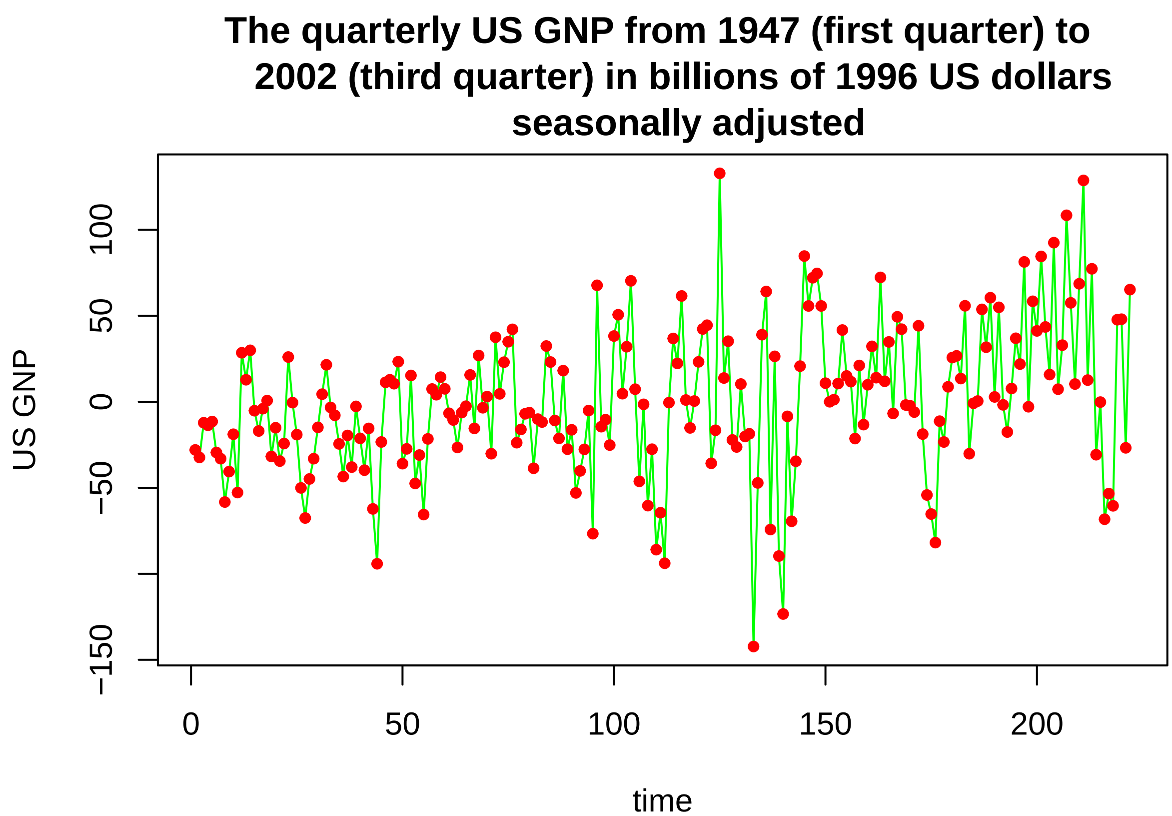
1. US GNP

(a) MA(2) model

1. Read in the data.

(a) Plot the data.

```r
GNP = read.table('gnp-47-02D.txt')
plot(GNP$V1, type = 'l', col = 'green',
     xlab = 'time', ylab = 'US GNP',
     main = 'The quarterly US GNP from 1947 (first quarter) to
     2002 (third quarter) in billions of 1996 US dollars
     seasonally adjusted')
points(GNP$V1, pch = 20, col = 'red')
```



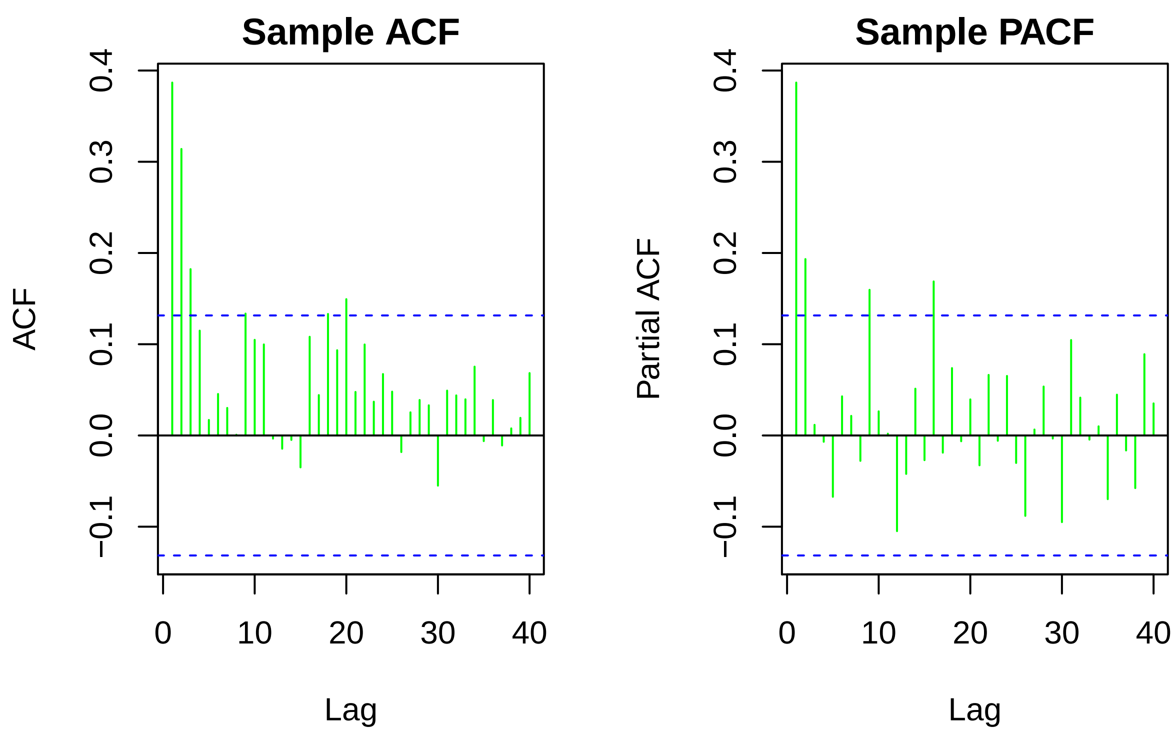
(b) Use the plot to make a preliminary assessment of stationary.

The time series seems have a slightly upward trend, and the variance is constant except for time 100 to 150. Overall, i would still consider this time series as stationary.

(c) Plot the sample acf/pacf.

```
par(mfrow = c(1,2))
sampleacf = acf(GNP$V1, lag.max = 40, plot = F)
plot(sampleacf, main = "", col = 'green')
title("Sample ACF",line=0.5)

samplepacf = pacf(GNP$V1, lag.max = 40, plot = F)
plot(samplepacf, main = "", col = 'green')
title("Sample PACF", line=0.5)
```



(d) Use the plot to make observations about possible orders of dependency.

The pacf is significantly nonzero at lag 2 and insignificant after lag 2 while there is an exponential decay in acf values for small increasing lags, So we might try an AR(2) model.

2. Estimate the parameters in the specified model using the specified method.

(a) Give the estimated coefficients.

```
ma2fit = stats::arima(GNP$V1, order = c(0,0,2), method = 'ML')
ma2fit$coef
```

```
##       ma1       ma2 intercept
## 0.2937883 0.2221881 0.0364831
```

(b) Give the value of the AIC or AICC.

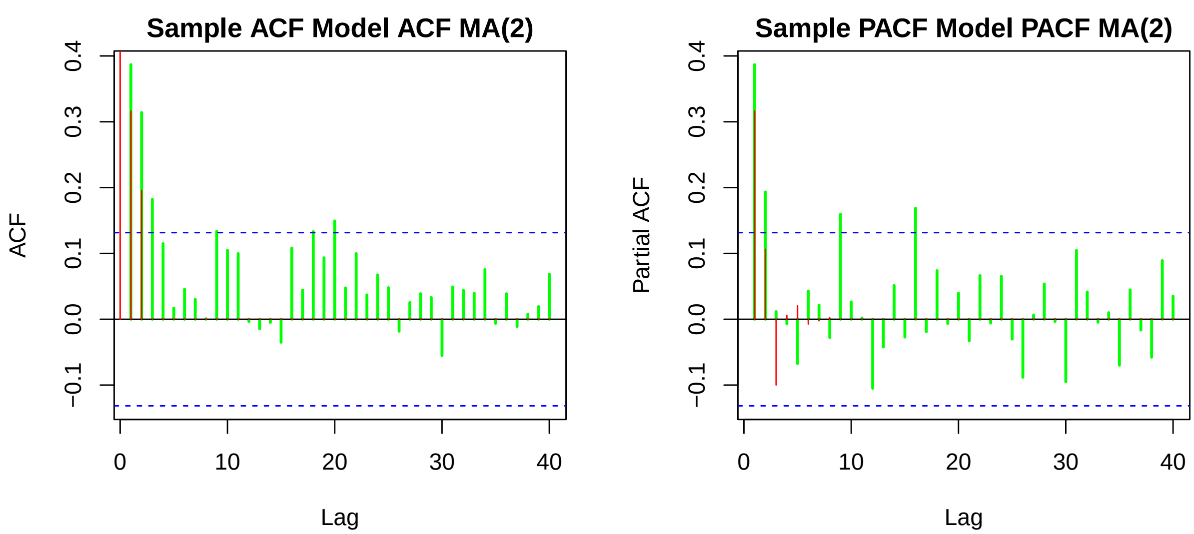
```
ma2fit$aic
```

```
## [1] 2262.106
```

(c) Plot the model and sample acf/pacf.

```
par(mfrow = c(1,2))
plot(sampleacf, main = "", col = 'green', lwd = 2)
lines(x = 0:40, ARMAacf(ma = c(0.2937883,0.2221881), lag.max = 40),
      type = 'h', col = 'red')
title("Sample ACF Model ACF MA(2)",line=0.5)

plot(samplepacf, main = "", col = 'green', lwd = 2)
lines(ARMAacf(ma = c(0.2937883,0.2221881), lag.max = 40, pacf = TRUE),
      type = 'h', col = 'red')
title("Sample PACF Model PACF MA(2)", line=0.5)
```



(d) Use the plot from (c) to assess the quality of the model fit.

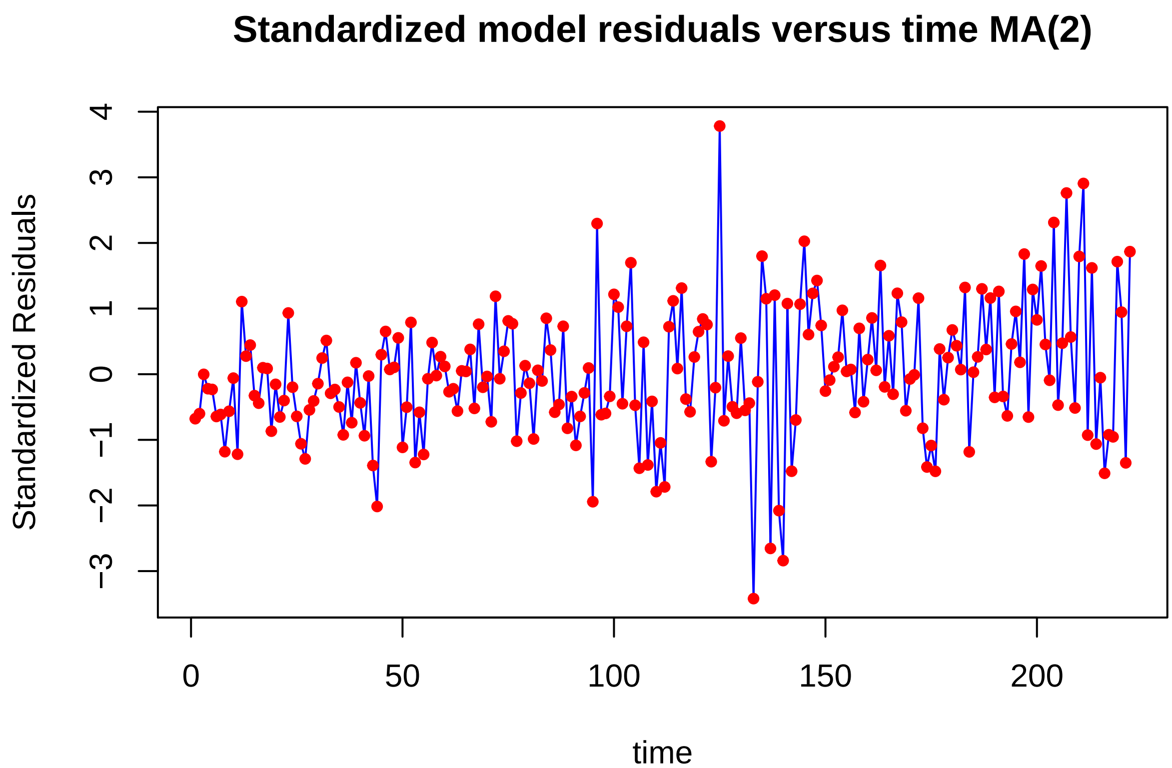
We did not observe a good agreement in both acf and pacf values.

3. Analyze the standardized residuals.

(a) Plot the standardized residuals.

```
ma2resid = rstandard(ma2fit)
plot(ma2resid, type = 'l', col = 'blue',
```

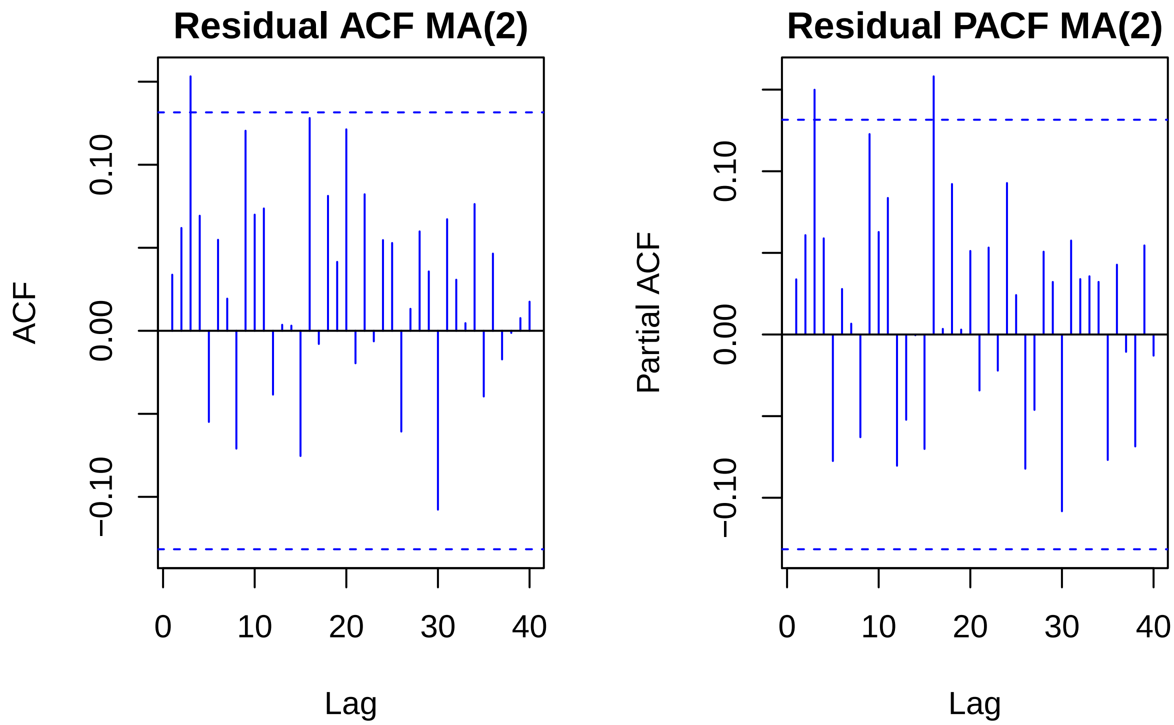
```
     xlab = 'time', ylab = 'Standardized Residuals',
     main = 'Standardized model residuals versus time MA(2)')
points(ma2resid, pch = 20, col = 'red')
```



(b) Plot the sample acf/pacf for the standardized residuals.

```
par(mfrow = c(1,2))
ma2residacf = acf(ma2resid, lag.max = 40, plot = F)
plot(ma2residacf, main = "", col = 'blue')
title("Residual ACF MA(2)",line=0.5)

ma2residpacf = pacf(ma2resid, lag.max = 40, plot = F)
plot(ma2residpacf, main = "", col = 'blue')
title("Residual PACF MA(2)", line=0.5)
```



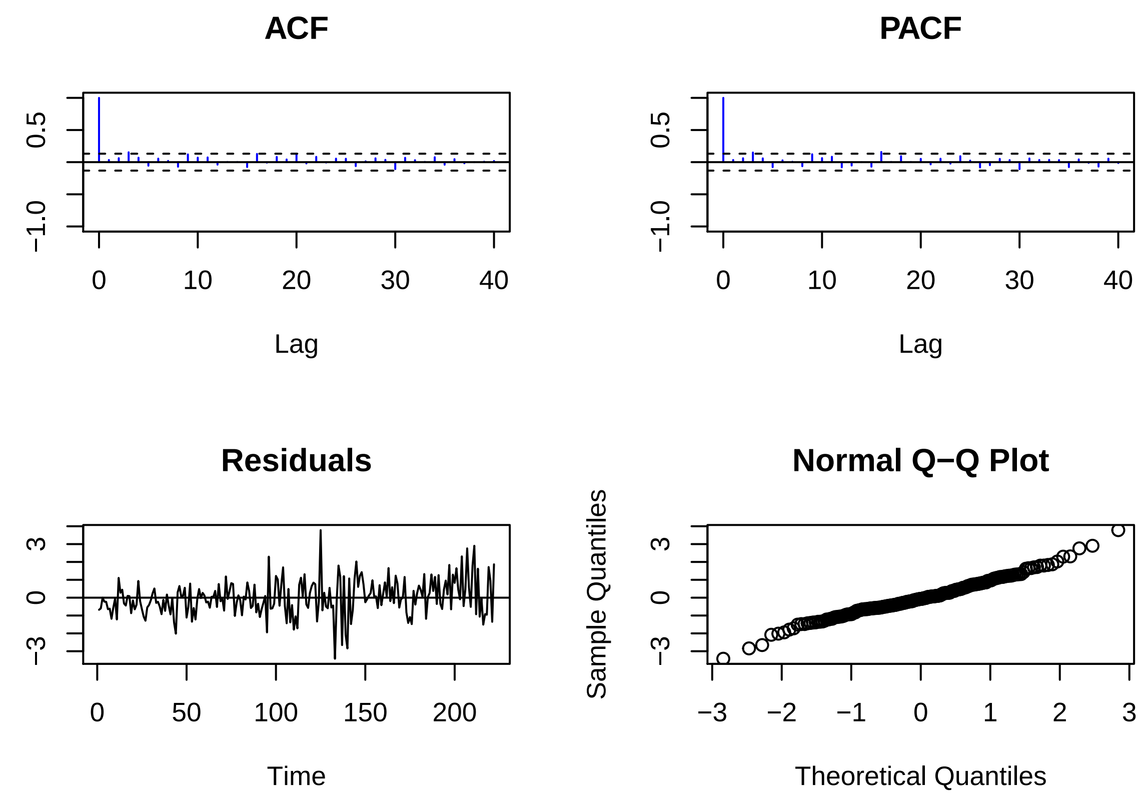
(c) Assess the plots for the hypothesis that the residuals are iid.

The plots of the residual acf and pacf do not support the rejection of iid noise hypothesis.

(d) Evaluate the Ljung-Box and McLeod-Li statistics and indicate if they support rejection of the iid hypothesis.

```
test(ma2resid)
```

```
## Null hypothesis: Residuals are iid noise.
## Test                          Distribution Statistic   p-value
## Ljung-Box Q                  Q ~ chisq(20)     27.36    0.1254
## McLeod-Li Q                  Q ~ chisq(20)     62.07         0 *
## Turning points T  (T-146.7)/6.3 ~ N(0,1)       152     0.394
## Diff signs S      (S-110.5)/4.3 ~ N(0,1)       116     0.202
## Rank P        (P-12265.5)/553.1 ~ N(0,1)     14335     2e-04 *
```



The Ljung-Box statistic $Q_{LB} = 27.36$ with p-value 0.1254. This does not provide sufficient evidence to reject the iid hypothesis.

The McLeod-Li statistic $Q_{ML} = 62.07$ with p-value almost 0. This provide sufficient evidence to reject the iid hypothesis.

(e) Give a final assessment on the validity of the iid hypothesis.

According to the residual acf/pacf plot and statistical tests, My conclusion is that the residuals are behaving like iid noise.

4. Use the results to give a summary evaluation about the quality of the fitted model.

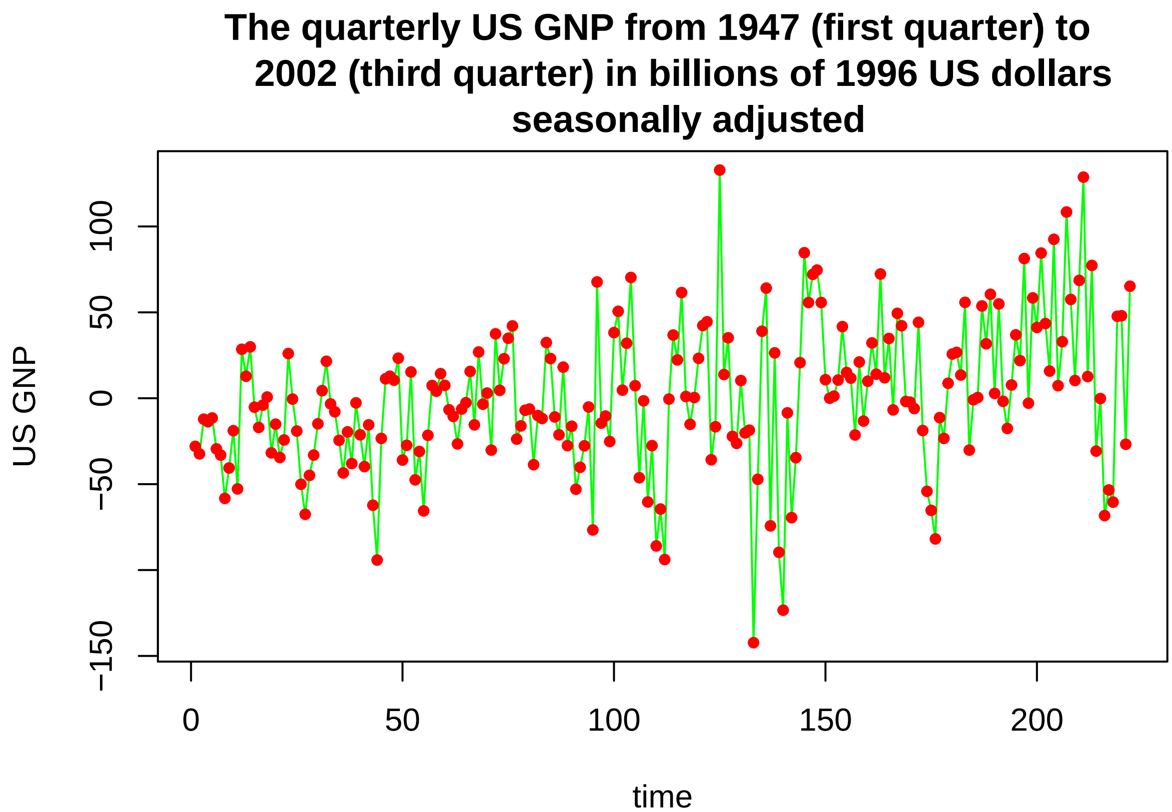
Even though the residuals are behaving like iid noise, the model acf and pacf do not have a good agreement on sample acf and pacf. I do not think MA(2) model is a good fit. We might try some other models. Next we will try to fit an AR(2) model.

(b) AR(2) model

1. Read in the data.

(a) Plot the data.

```
GNP = read.table('gnp-47-02D.txt')
plot(GNP$V1, type = 'l', col = 'green',
     xlab = 'time', ylab = 'US GNP',
     main = 'The quarterly US GNP from 1947 (first quarter) to
     2002 (third quarter) in billions of 1996 US dollars
     seasonally adjusted')
points(GNP$V1, pch = 20, col = 'red')
```



(b) Use the plot to make a preliminary assessment of stationary.

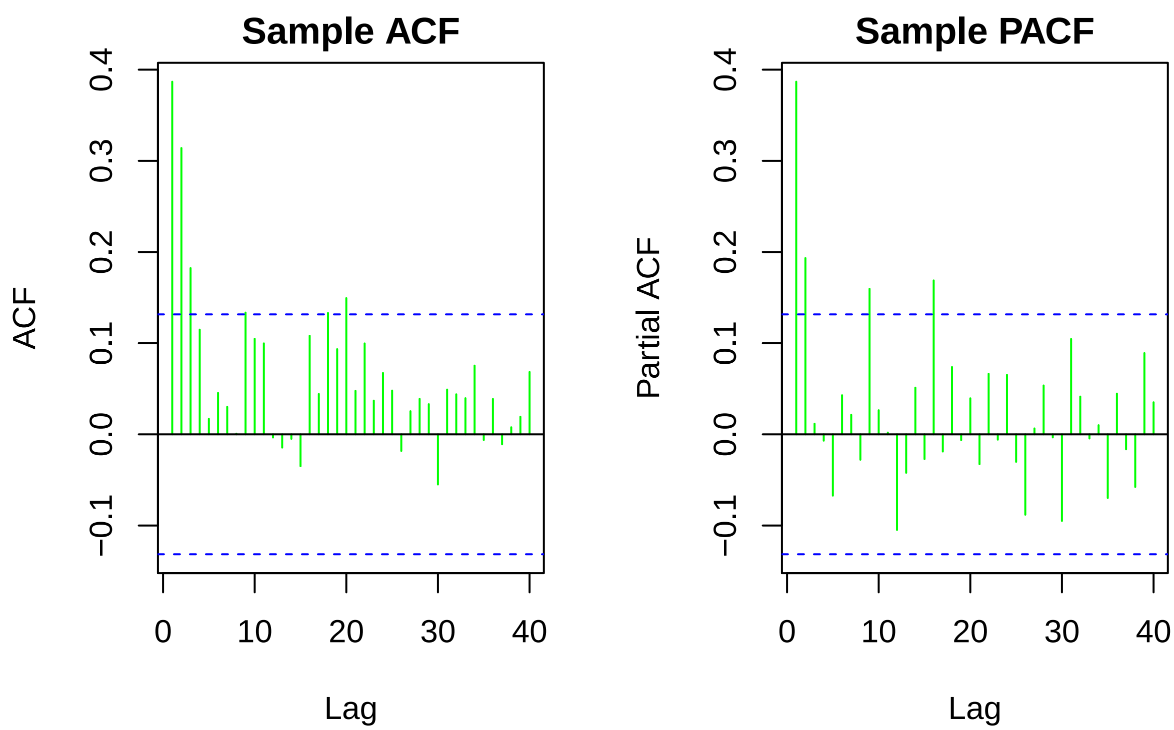
The time series seems have a slightly upward trend, and the variance is constant except for time 100 to 150. Overall, i would still consider this time series as stationary.

(c) Plot the sample acf/pacf.

```
par(mfrow = c(1,2))
sampleacf = acf(GNP$V1, lag.max = 40, plot = F)
plot(sampleacf, main = "", col = 'green')
title("Sample ACF",line=0.5)

samplepacf = pacf(GNP$V1, lag.max = 40, plot = F)
```

```
plot(samplepacf, main = "", col = 'green')
title("Sample PACF", line=0.5)
```



(d) Use the plot to make observations about possible orders of dependency.

The pacf is significantly nonzero at lag 2 and insignificant after lag 2 while there is an exponential decay in acf values for small increasing lags, So we might try an AR(2) model.

2. Estimate the parameters in the specified model using the specified method.

(a) Give the estimated coefficients.

```
ar2fit = stats::arima(GNP$V1, order = c(2,0,0), method = 'ML')
ar2fit$coef
```

```
##        ar1        ar2  intercept
## 0.31355139 0.19309695 0.05500756
```

(b) Give the value of the AIC or AICC.

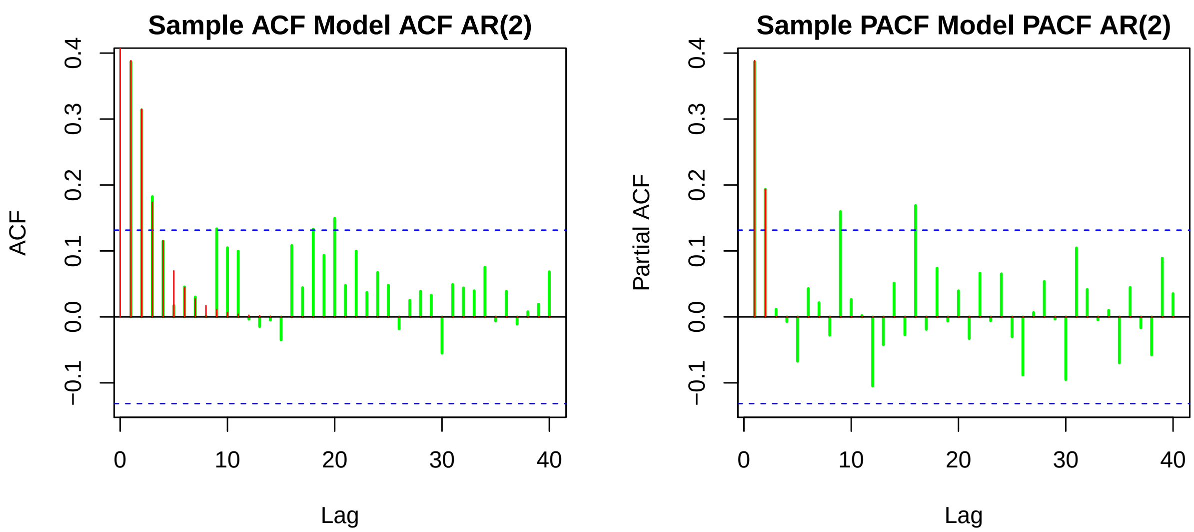
```
ar2fit$aic
```

```
## [1] 2255.34
```

(c) Plot the model and sample acf/pacf.

```
par(mfrow = c(1,2))
plot(sampleacf, main = "", col = 'green', lwd = 2)
lines(x = 0:40, ARMAacf(ar = c(0.31355139, 0.19309695), lag.max = 40),
      type = 'h', col = 'red')
title("Sample ACF Model ACF AR(2)",line=0.5)

plot(samplepacf, main = "", col = 'green', lwd = 2)
lines(ARMAacf(ar = c(0.31355139, 0.19309695), lag.max = 40, pacf = TRUE),
      type = 'h', col = 'red')
title("Sample PACF Model PACF AR(2)", line=0.5)
```



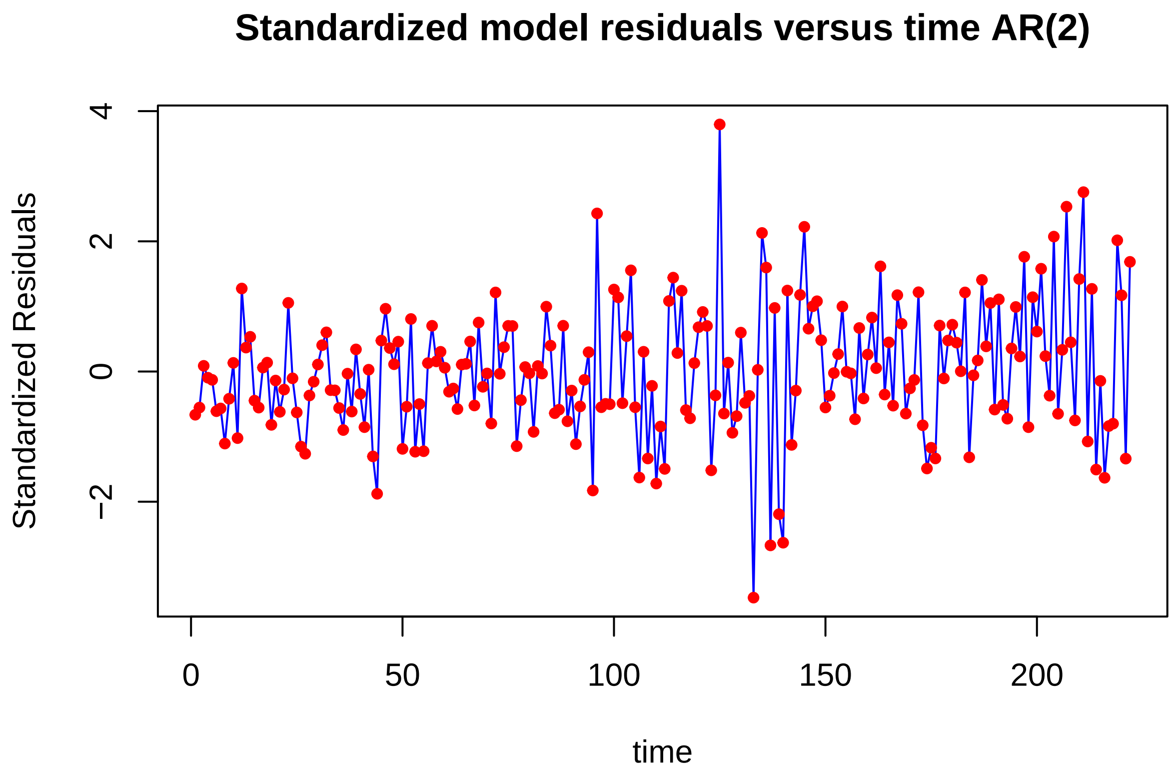
(d) Use the plot from (c) to assess the quality of the model fit.

We see good agreement in the acf values up to lag 5 and good agreement in the significant pacf values.

3. Analyze the standardized residuals.

(a) Plot the standardized residuals.

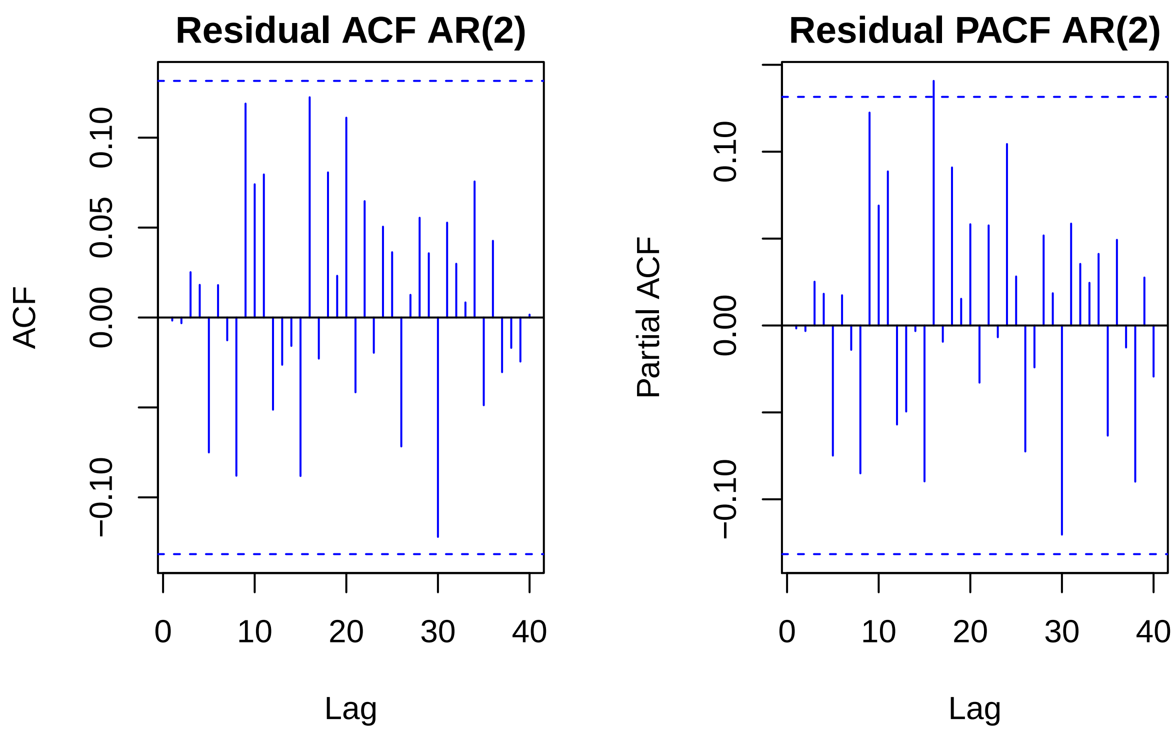
```
ar2resid = rstandard(ar2fit)
plot(ar2resid, type = 'l', col = 'blue',
     xlab = 'time', ylab = 'Standardized Residuals',
     main = 'Standardized model residuals versus time AR(2)')
points(ar2resid, pch = 20, col = 'red')
```



(b) Plot the sample acf/pacf for the standardized residuals.

```
par(mfrow = c(1,2))
ar2residacf = acf(ar2resid, lag.max = 40, plot = F)
plot(ar2residacf, main = "", col = 'blue')
title("Residual ACF AR(2)",line=0.5)

ar2residpacf = pacf(ar2resid, lag.max = 40, plot = F)
plot(ar2residpacf, main = "", col = 'blue')
title("Residual PACF AR(2)", line=0.5)
```



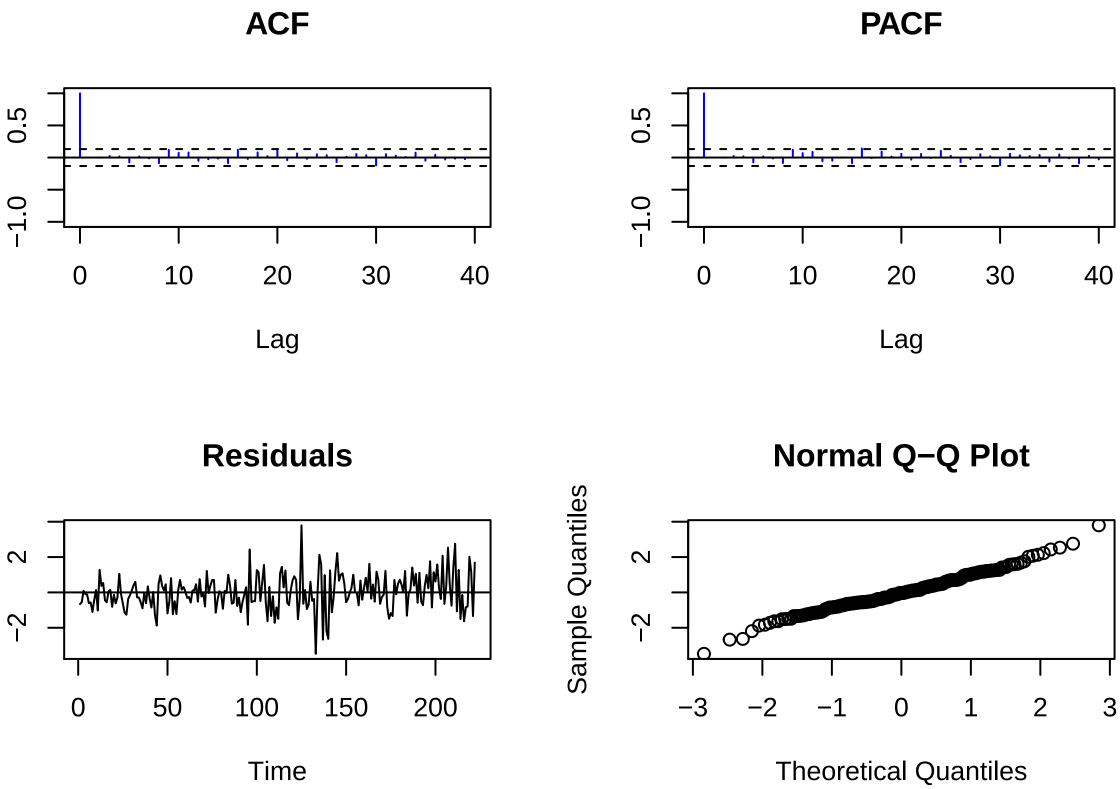
(c) Assess the plots for the hypothesis that the residuals are iid.

The plots of the residual acf and pacf do not support the rejection of iid noise hypothesis.

(d) Evaluate the Ljung-Box and McLeod-Li statistics and indicate if they support rejection of the iid hypothesis.

```
test(ar2resid)
```

```
## Null hypothesis: Residuals are iid noise.
## Test                          Distribution Statistic   p-value
## Ljung-Box Q                  Q ~ chisq(20)     20.72    0.4141
## McLeod-Li Q                  Q ~ chisq(20)     72.17         0 *
## Turning points T  (T-146.7)/6.3 ~ N(0,1)        146    0.9151
## Diff signs S      (S-110.5)/4.3 ~ N(0,1)        116     0.202
## Rank P        (P-12265.5)/553.1 ~ N(0,1)      13778    0.0062 *
```



The Ljung-Box statistic $Q_{LB} = 20.72$ with p-value 0.4141. This does not provide sufficient evidence to reject the iid hypothesis.

The McLeod-Li statistic $Q_{ML} = 72.17$ with p-value almost 0. This provide sufficient evidence to reject the iid hypothesis.

(e) Give a final assessment on the validity of the iid hypothesis.

According to the residual acf/pacf plots and statistical tests, My conclusion is that the residuals are behaving like iid noise.

4. Use the results to give a summary evaluation about the quality of the fitted model.

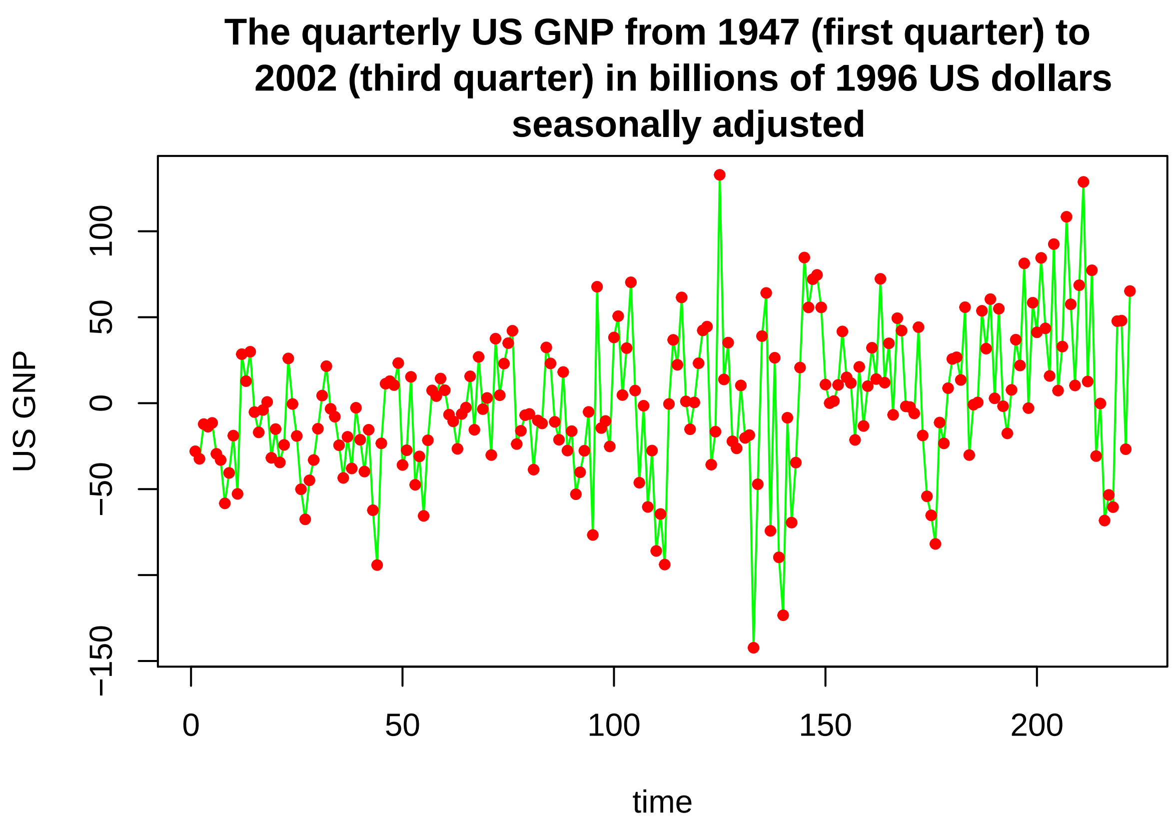
The residuals are behaving like iid noise, and the model acf and pacf have a good agreement on sample acf and pacf. I would conclude that AR(2) model is a good fit, and this is what i guessed in question 1(d). It is a good thing to try more models and compare them and select the best one, so next we will fit an ARMA(1,1) model.

(c) ARMA(1,1) model

1. Read in the data.

(a) Plot the data.

```
GNP = read.table('gnp-47-02D.txt')
plot(GNP$V1, type = 'l', col = 'green',
     xlab = 'time', ylab = 'US GNP',
     main = 'The quarterly US GNP from 1947 (first quarter) to
     2002 (third quarter) in billions of 1996 US dollars
     seasonally adjusted')
points(GNP$V1, pch = 20, col = 'red')
```



(b) Use the plot to make a preliminary assessment of stationary.

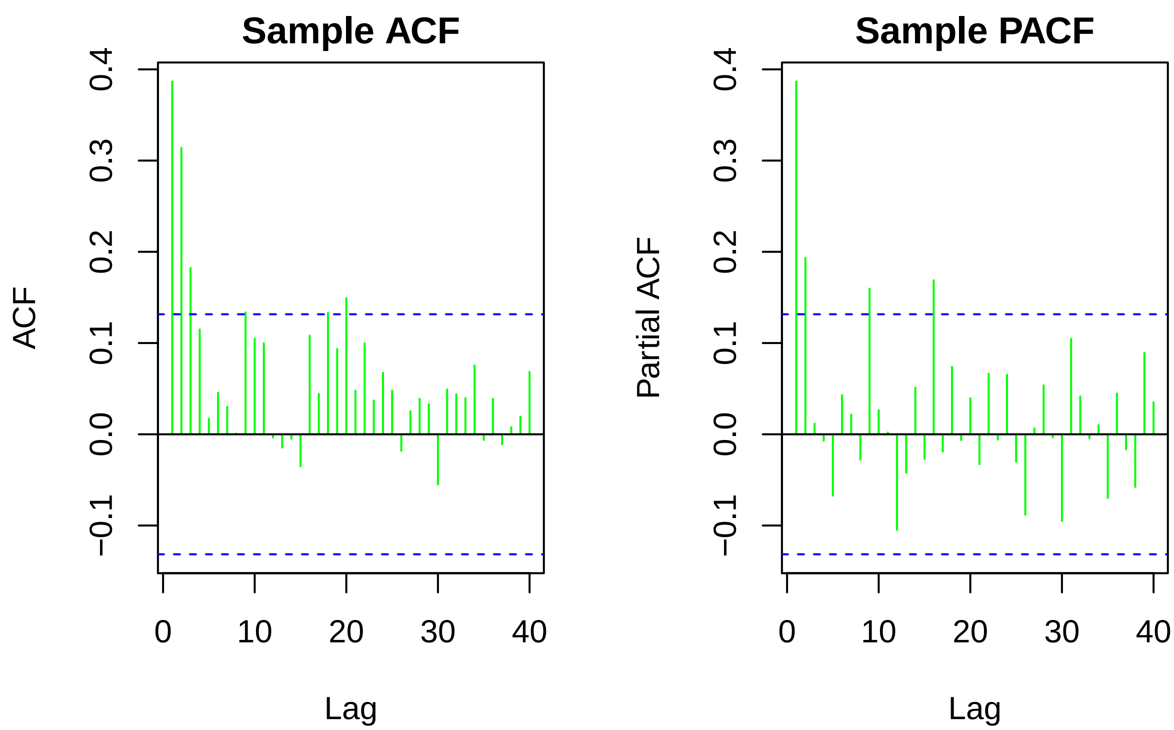
The time series seems have a slightly upward trend, and the variance is constant except for time 100 to 150. Overall, i would still consider this time series as stationary.

(c) Plot the sample acf/pacf.

```
par(mfrow = c(1,2))
sampleacf = acf(GNP$V1, lag.max = 40, plot = F)
plot(sampleacf, main = "", col = 'green')
```

```
title("Sample ACF",line=0.5)

samplepacf = pacf(GNP$V1, lag.max = 40, plot = F)
plot(samplepacf, main = "", col = 'green')
title("Sample PACF", line=0.5)
```



(d) Use the plot to make observations about possible orders of dependency.

The pacf is significantly nonzero at lag 2 and insignificant after lag 2 while there is an exponential decay in acf values for small increasing lags, So we might try an AR(2) model.

2. Estimate the parameters in the specified model using the specified method.

(a) Give the estimated coefficients.

```
arma11fit = stats::arima(GNP$V1, order = c(1,0,1), method = 'ML')
arma11fit$coef
```

```
##        ar1         ma1   intercept
##  0.70187278 -0.36973530  0.08457291
```

(b) Give the value of the AIC or AICC.

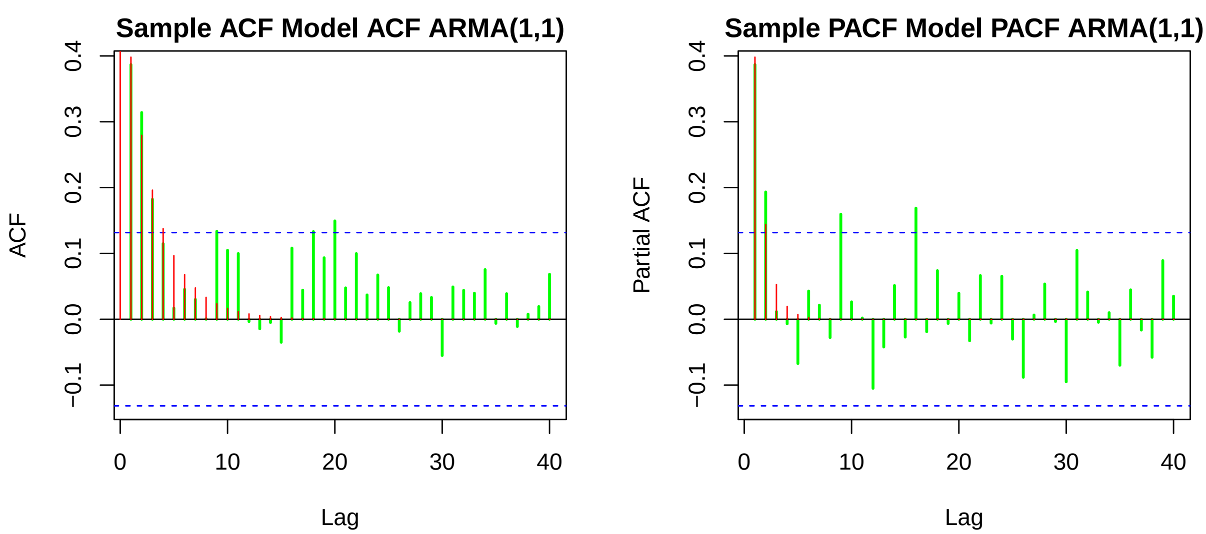
```
arma11fit$aic
```

```
## [1] 2256.534
```

(c) Plot the model and sample acf/pacf.

```
par(mfrow = c(1,2))
plot(sampleacf, main = "", col = 'green', lwd = 2)
lines(x = 0:40, ARMAacf(ar = 0.70187278, ma = -0.36973530, lag.max = 40),
      type = 'h', col = 'red')
title("Sample ACF Model ACF ARMA(1,1)",line=0.5)

plot(samplepacf, main = "", col = 'green', lwd = 2)
lines(ARMAacf(ar = 0.70187278, ma = -0.36973530, lag.max = 40, pacf = TRUE),
      type = 'h', col = 'red')
title("Sample PACF Model PACF ARMA(1,1)", line=0.5)
```



(d) Use the plot from (c) to assess the quality of the model fit.

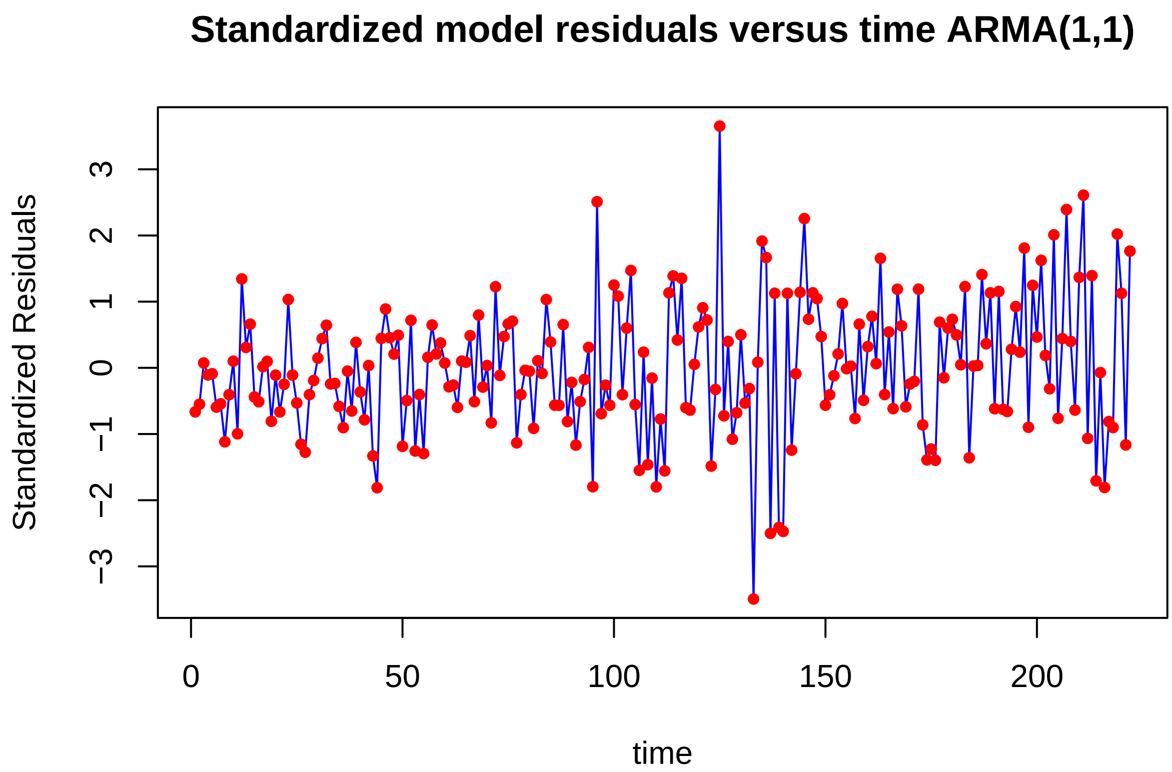
We see good agreement in the acf values up to lag 1 and good agreement in the pacf values up to lag 1.

3. Analyze the standardized residuals.

(a) Plot the standardized residuals.

```
arma11resid = rstandard(arma11fit)
plot(arma11resid, type = 'l', col = 'blue',
```

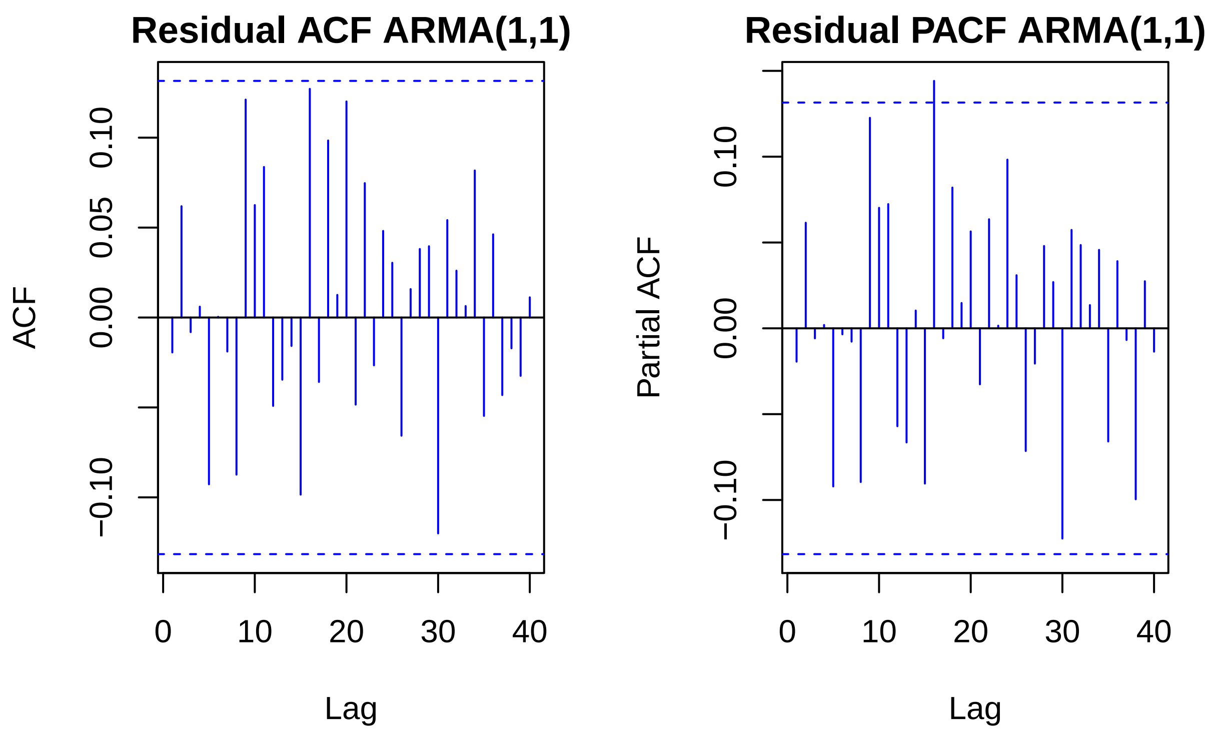
```
     xlab = 'time', ylab = 'Standardized Residuals',
     main = 'Standardized model residuals versus time ARMA(1,1)')
points(arma11resid, pch = 20, col = 'red')
```



(b) Plot the sample acf/pacf for the standardized residuals.

```
par(mfrow = c(1,2))
arma11residacf = acf(arma11resid, lag.max = 40, plot = F)
plot(arma11residacf, main = "", col = 'blue')
title("Residual ACF ARMA(1,1)",line=0.5)

arma11residpacf = pacf(arma11resid, lag.max = 40, plot = F)
plot(arma11residpacf, main = "", col = 'blue')
title("Residual PACF ARMA(1,1)", line=0.5)
```



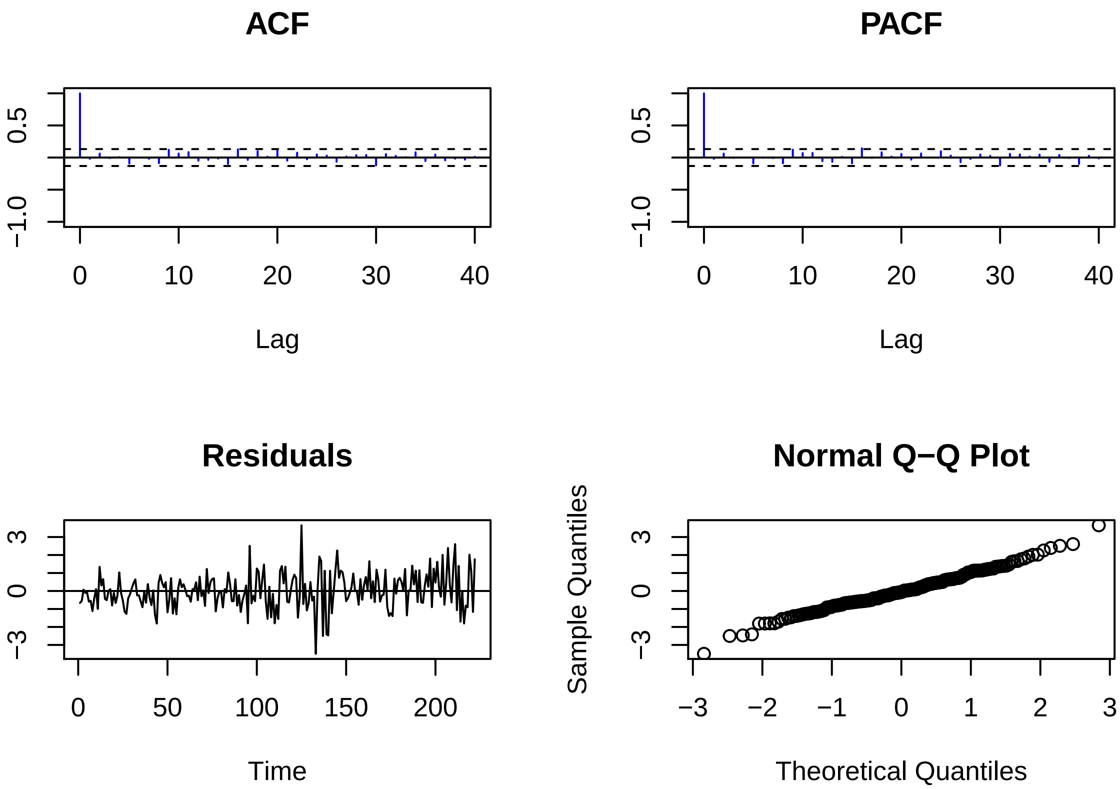
(c) Assess the plots for the hypothesis that the residuals are iid.

The plots of the residual acf and pacf do not support the rejection of iid noise hypothesis.

(d) Evaluate the Ljung-Box and McLeod-Li statistics and indicate if they support rejection of the iid hypothesis.

```
test(arma11resid)
```

```
## Null hypothesis: Residuals are iid noise.
## Test                        Distribution Statistic   p-value
## Ljung-Box Q                Q ~ chisq(20)      24.2    0.2337
## McLeod-Li Q                Q ~ chisq(20)     72.62         0 *
## Turning points T  (T-146.7)/6.3 ~ N(0,1)       156    0.1358
## Diff signs S      (S-110.5)/4.3 ~ N(0,1)       116     0.202
## Rank P        (P-12265.5)/553.1 ~ N(0,1)     13748    0.0074 *
```



The Ljung-Box statistic $Q_{LB} = 24.2$ with p-value 0.2337. This does not provide sufficient evidence to reject the iid hypothesis.

The McLeod-Li statistic $Q_{ML} = 72.62$ with p-value almost 0. This provide sufficient evidence to reject the iid hypothesis.

(e) Give a final assessment on the validity of the iid hypothesis.

According to the residual acf/pacf plots and statistical tests, My conclusion is that the residuals are behaving like iid noise.

4. Use the results to give a summary evaluation about the quality of the fitted model.

The residuals are behaving like iid noise, but the model acf and pacf do not have a good agreement on sample acf and pacf. I think ARMA(1,1) is not a good fit. It is worse than AR(2) model but have a better performance than MA(2) model.

(d) Use the results from (a) - (c) to decide which model is best and provide a reason for your choice.

We fitted MA(2), AR(2), and ARMA(1,1) models. The residuals plot for these three models are all behaving like iid noise.

```
tab <- matrix(c(ma2fit$aic, ar2fit$aic, arma11fit$aic), ncol=1, byrow=TRUE)
colnames(tab) <- c('AIC')
rownames(tab) <- c('MA(2) model','AR(2) model','ARMA(1,1) model')
tab <- as.table(tab)
tab
```

```
##                      AIC
## MA(2) model     2262.106
## AR(2) model     2255.340
## ARMA(1,1) model 2256.534
```

The AIC has a minimum value for AR(2) model. The AIC suggests that the AR(2) model is the best fit followed by ARMA(1,1) model and this is similar to the analysis of model and sample acf/pacf we did earlier. We observe a good agreement for AR(2) model. In conclusion, I would conclude that AR(2) is the best model.

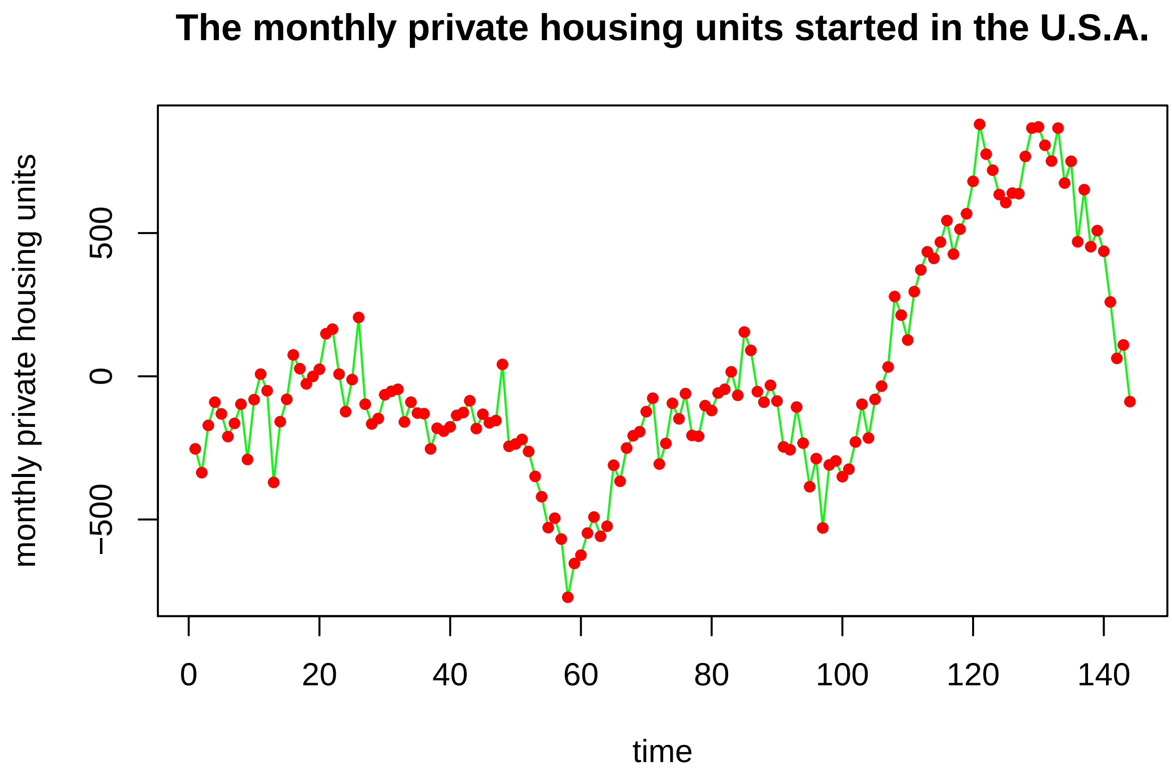
2. monthly private housing units started in the U.S.A.

(a) ARMA(3,2) model

1. Read in the data.

(a) Plot the data.

```
apc = read.table('apc.txt')
# remove mean
apc$V1 = apc$V1 - mean(apc$V1)
plot(apc$V1, type = 'l', col = 'green',
     xlab = 'time', ylab = 'monthly private housing units',
     main = 'The monthly private housing units started in the U.S.A.')
points(apc$V1, pch = 20, col = 'red')
```



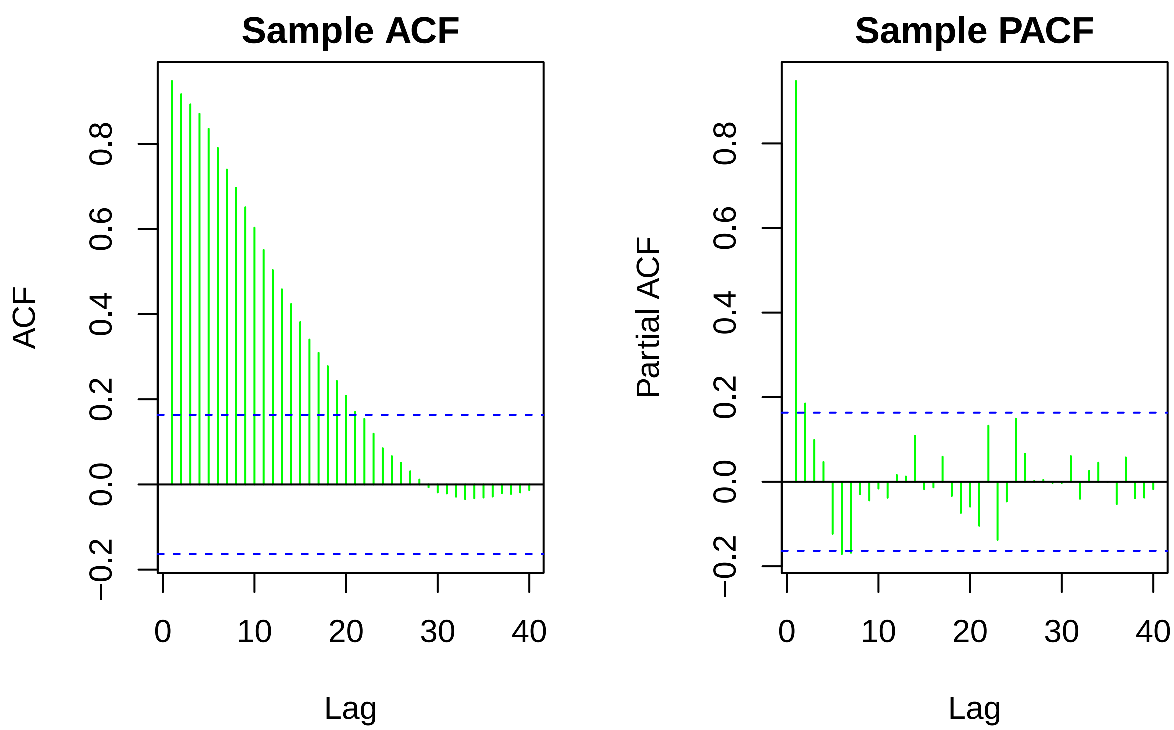
(b) Use the plot to make a preliminary assessment of stationary.

The time series have a clear trend and increase variance, so i think this is not a stationary time series.

(c) Plot the sample acf/pacf.

```
par(mfrow = c(1,2))
sampleacf2 = acf(apc$V1, lag.max = 40, plot = F)
plot(sampleacf2, main = "", col = 'green')
title("Sample ACF",line=0.5)

samplepacf2 = pacf(apc$V1, lag.max = 40, plot = F)
plot(samplepacf2, main = "", col = 'green')
title("Sample PACF", line=0.5)
```



(d) Use the plot to make observations about possible orders of dependency.

There is a slow decay in acf values for small increasing lags instead of exponential decay, this suggests that the time series is not stationary. However, MLE method produces estimates of the ϕ and θ parameters by assuming the data are observations of a stationary gaussian time series and maximizing the likelihood with respect to the parameters. So i would consider remove the trend first and plot acf and pacf again, then make observations about possible orders of dependency.

2. Estimate the parameters in the specified model using the specified method.

(a) Give the estimated coefficients.

```
arma32fit = stats::arima(apc$V1, order = c(3,0,2), method = 'ML')
arma32fit$coef
```

```
##         ar1         ar2         ar3         ma1         ma2   intercept
##  0.07570715  0.93538904 -0.07104240  0.66428593 -0.26983521 -42.07872059
```

(b) Give the value of the AIC or AICC.

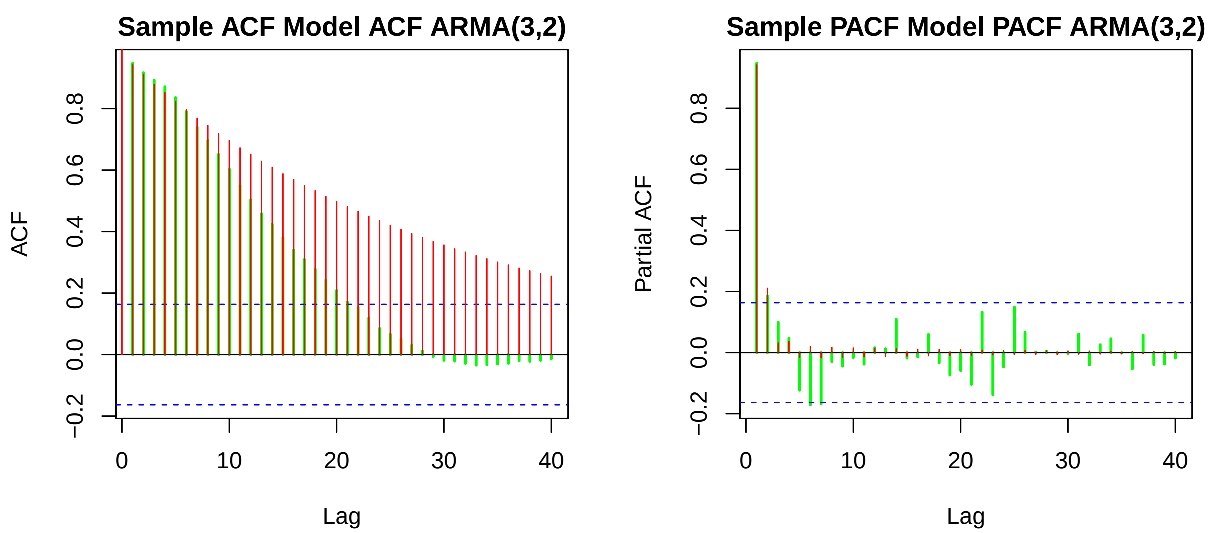
```
arma32fit$aic
```

```
## [1] 1786.285
```

(c) Plot the model and sample acf/pacf.

```
par(mfrow = c(1,2))
plot(sampleacf2, main = "", col = 'green', lwd = 2)
lines(x = 0:40, ARMAacf(ar = c(0.07570715, 0.93538904, -0.07104240),
                        ma = c(0.66428593, -0.26983521), lag.max = 40),
      type = 'h', col = 'red')
title("Sample ACF Model ACF ARMA(3,2)",line=0.5)

plot(samplepacf2, main = "", col = 'green', lwd = 2)
lines(ARMAacf(ar = c(0.07570715, 0.93538904, -0.07104240),
              ma = c(0.66428593, -0.26983521), lag.max = 40, pacf = TRUE),
      type = 'h', col = 'red')
title("Sample PACF Model PACF ARMA(3,2)", line=0.5)
```



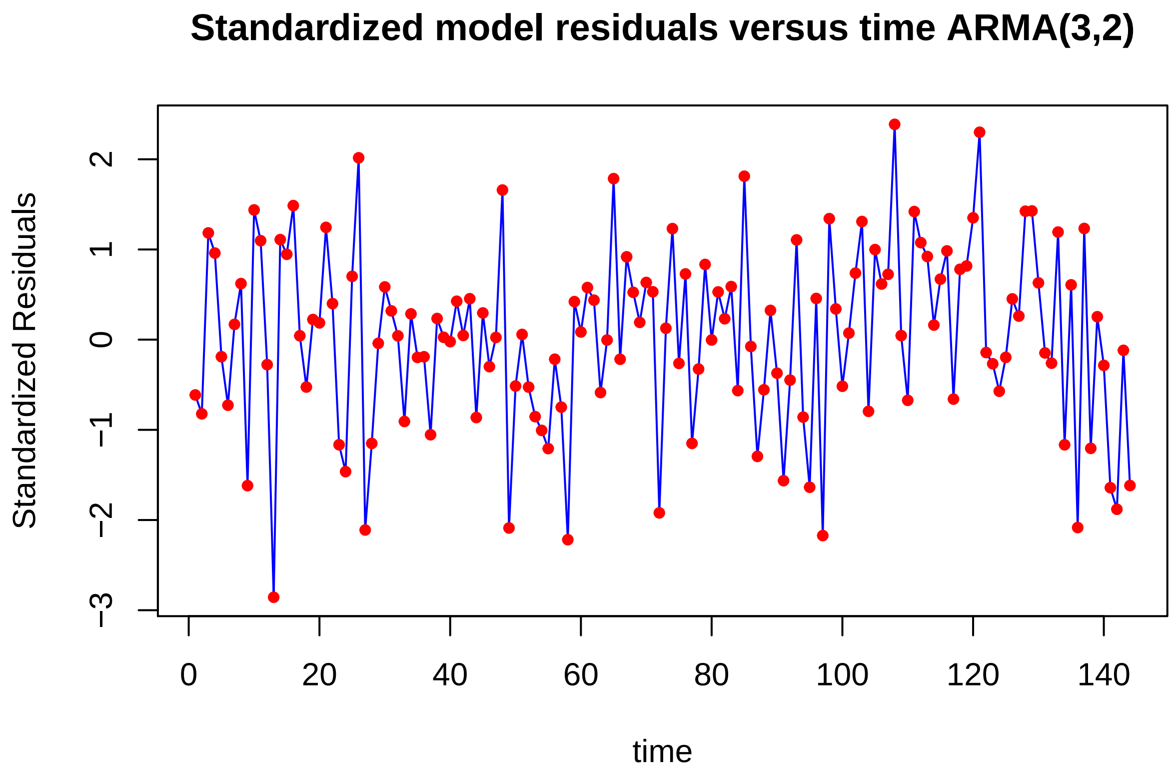
(d) Use the plot from (c) to assess the quality of the model fit.

We see good agreement in the significant pacf values but not a good agreement in the acf values.

3. Analyze the standardized residuals.

(a) Plot the standardized residuals.

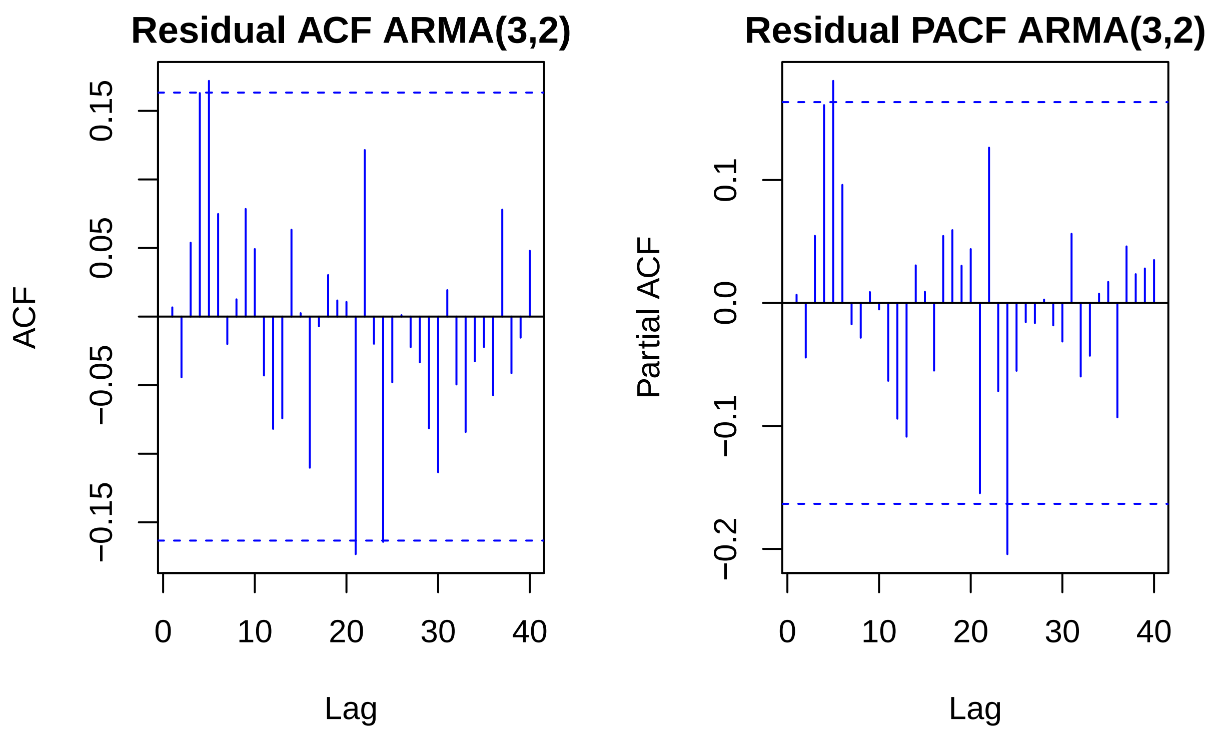
```
arma32resid = rstandard(arma32fit)
plot(arma32resid, type = 'l', col = 'blue',
     xlab = 'time', ylab = 'Standardized Residuals',
     main = 'Standardized model residuals versus time ARMA(3,2)')
points(arma32resid, pch = 20, col = 'red')
```



(b) Plot the sample acf/pacf for the standardized residuals.

```
par(mfrow = c(1,2))
arma32residacf = acf(arma32resid, lag.max = 40, plot = F)
plot(arma32residacf, main = "", col = 'blue')
title("Residual ACF ARMA(3,2)",line=0.5)

arma32residpacf = pacf(arma32resid, lag.max = 40, plot = F)
plot(arma32residpacf, main = "", col = 'blue')
title("Residual PACF ARMA(3,2)", line=0.5)
```



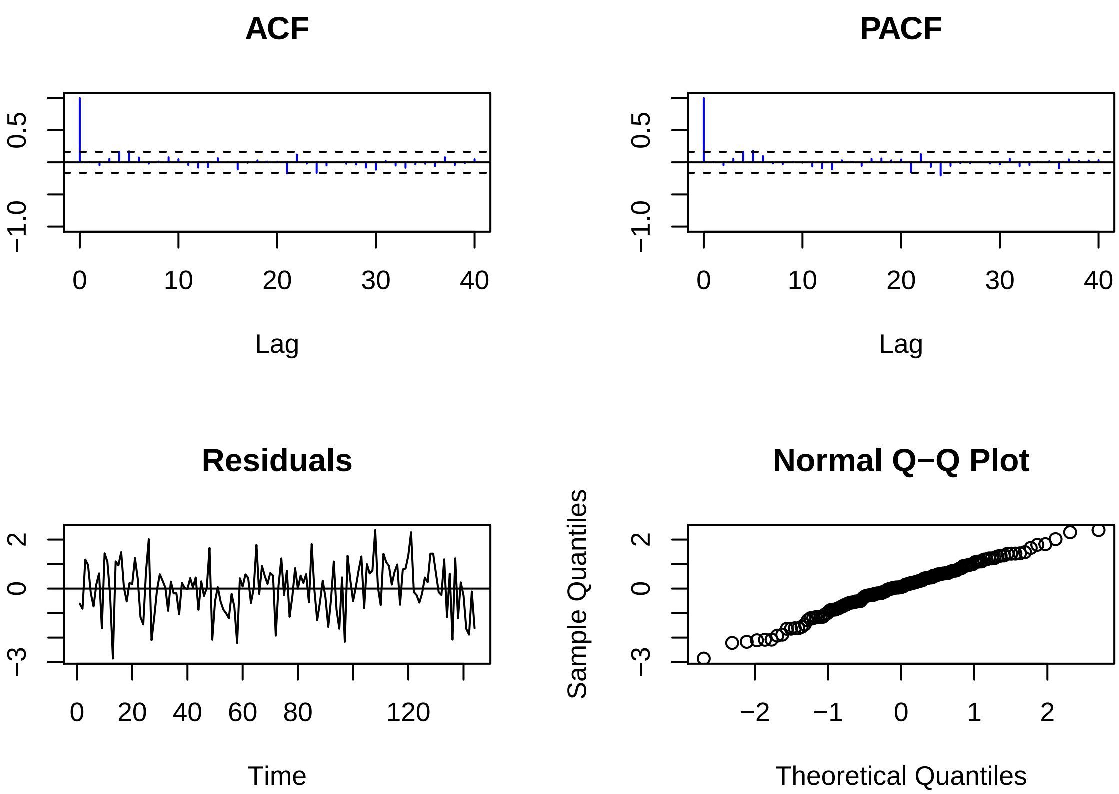
(c) Assess the plots for the hypothesis that the residuals are iid.

The plots of the residual acf and pacf do not support the rejection of iid noise hypothesis.

(d) Evaluate the Ljung-Box and McLeod-Li statistics and indicate if they support rejection of the iid hypothesis.

```
test(arma32resid)
```

```
## Null hypothesis: Residuals are iid noise.
## Test                        Distribution Statistic   p-value
## Ljung-Box Q                Q ~ chisq(20)     16.55     0.682
## McLeod-Li Q                Q ~ chisq(20)     15.23    0.7628
## Turning points T       (T-94.7)/5 ~ N(0,1)        92    0.5958
## Diff signs S         (S-71.5)/3.5 ~ N(0,1)        66    0.1136
## Rank P           (P-5148)/289.5 ~ N(0,1)      5356    0.4724
```



The Ljung-Box statistic $Q_{LB} = 16.55$ with p-value 0.685. This does not provide sufficient evidence to reject the iid hypothesis.

The McLeod-Li statistic $Q_{ML} = 15.23$ with p-value 0.7628. This does not provide sufficient evidence to reject the iid hypothesis.

(e) Give a final assessment on the validity of the iid hypothesis.

According to the residual acf/pacf plots and statistical tests, My conclusion is that the residuals are behaving like iid noise.

4. Use the results to give a summary evaluation about the quality of the fitted model.

The residuals are behaving like iid noise, but the model acf do not have a good agreement on sample acf. I do not think ARMA(3,2) is a good fit. Again, MLE assuming the data are observations of a stationary gaussian time series, but our data do not look like stationary time series. So next we will try to remove the trend first by differencing the data and plot the acf/pacf to determined p, q for ARMA(p,q) model.

(b) ARMA(1,1) model to the differenced data ∇x_t.

1. Read in the data.

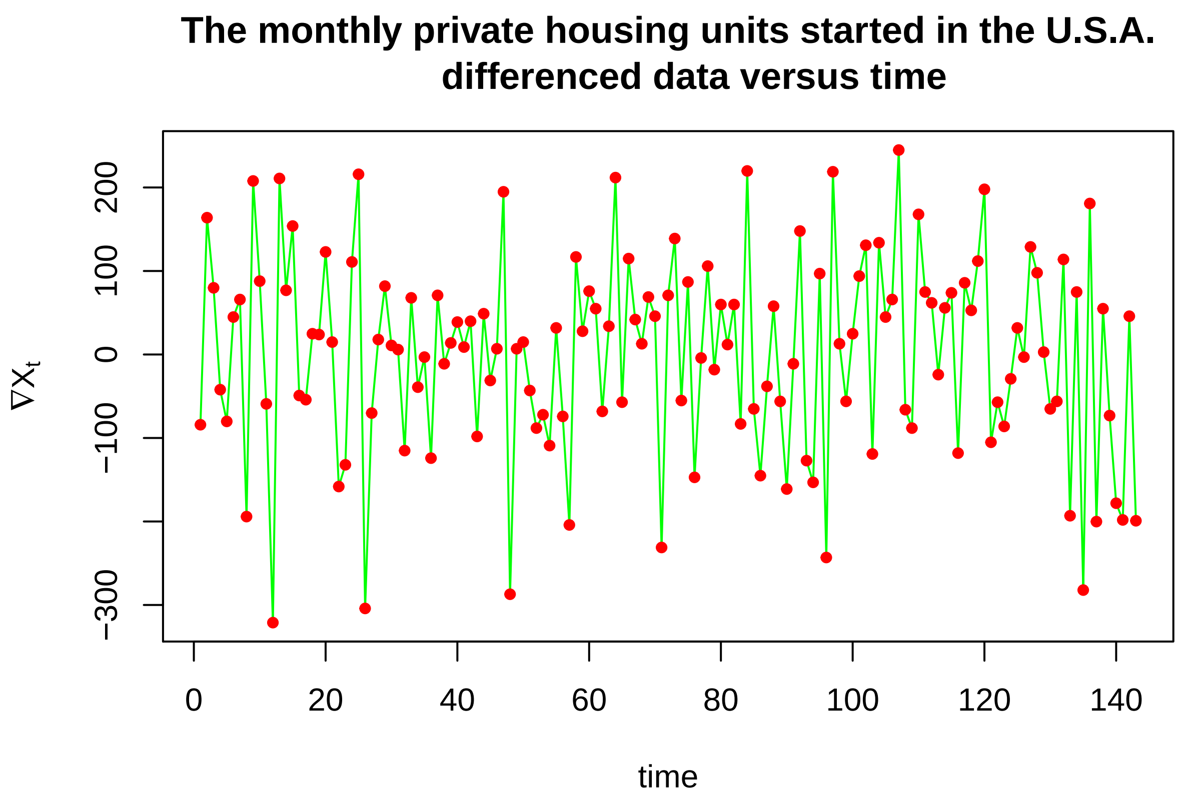
(a) Plot the data.

```r
apc = read.table('apc.txt')

# difference the data
apc = diff(apc$V1)

# remove mean
apc = apc - mean(apc)

plot(apc, type = 'l', col = 'green',
     xlab = 'time', ylab = expression(paste(nabla,X[t])),
     main = 'The monthly private housing units started in the U.S.A.
     differenced data versus time')
points(apc, pch = 20, col = 'red')
```



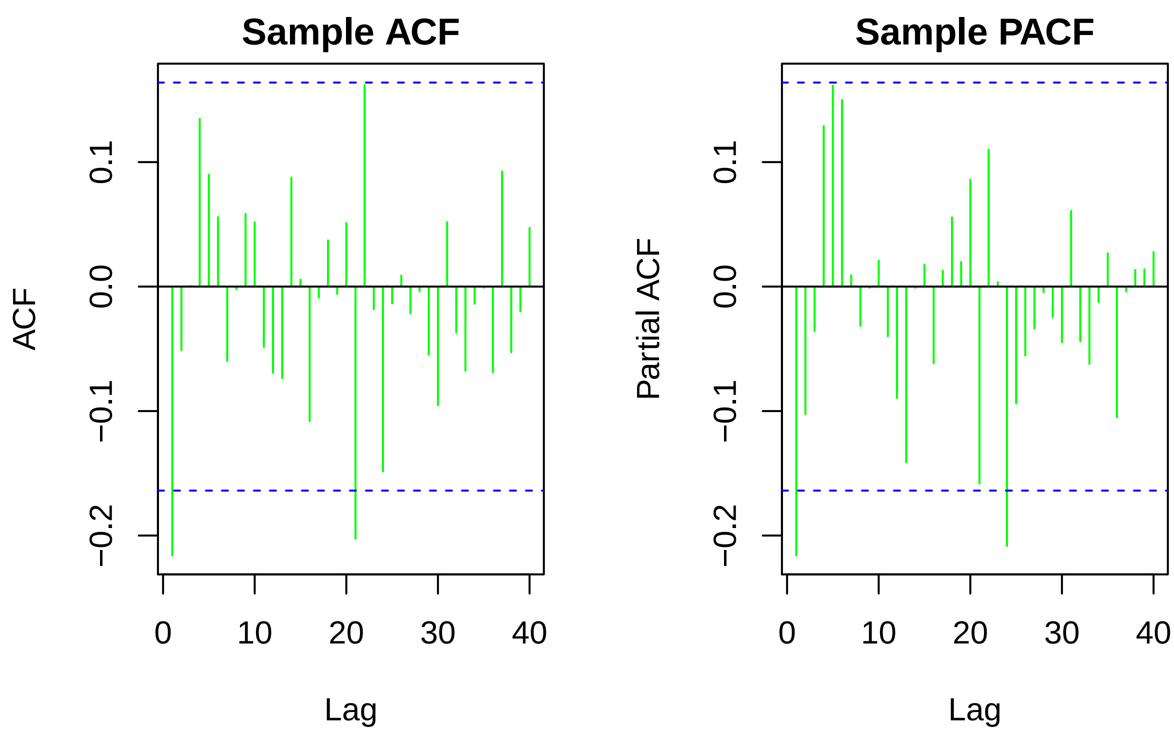
(b) Use the plot to make a preliminary assessment of stationary.

The time series do not have a clear trend and variance is constant, so i think this is a stationary time series.

(c) Plot the sample acf/pacf.

```
par(mfrow = c(1,2))
sampleacf3 = acf(apc, lag.max = 40, plot = F)
plot(sampleacf3, main = "", col = 'green')
title("Sample ACF",line=0.5)

samplepacf3 = pacf(apc, lag.max = 40, plot = F)
plot(samplepacf3, main = "", col = 'green')
title("Sample PACF", line=0.5)
```



(d) Use the plot to make observations about possible orders of dependency.

The pacf is significant for lags less than 1 and generally insignificant after lag 1 and the acf is significant for lags less than 1 and generally insignificant after 1, we might try an ARMA(1,1) model.

2. Estimate the parameters in the specified model using the specified method.

(a) Give the estimated coefficients.

```
arma11fit2 = stats::arima(apc, order = c(1,0,1), method = 'ML')
arma11fit2$coef
```

```
##        ar1        ma1  intercept
##  0.0429182 -0.2866843  0.4270416
```

(b) Give the value of the AIC or AICC.

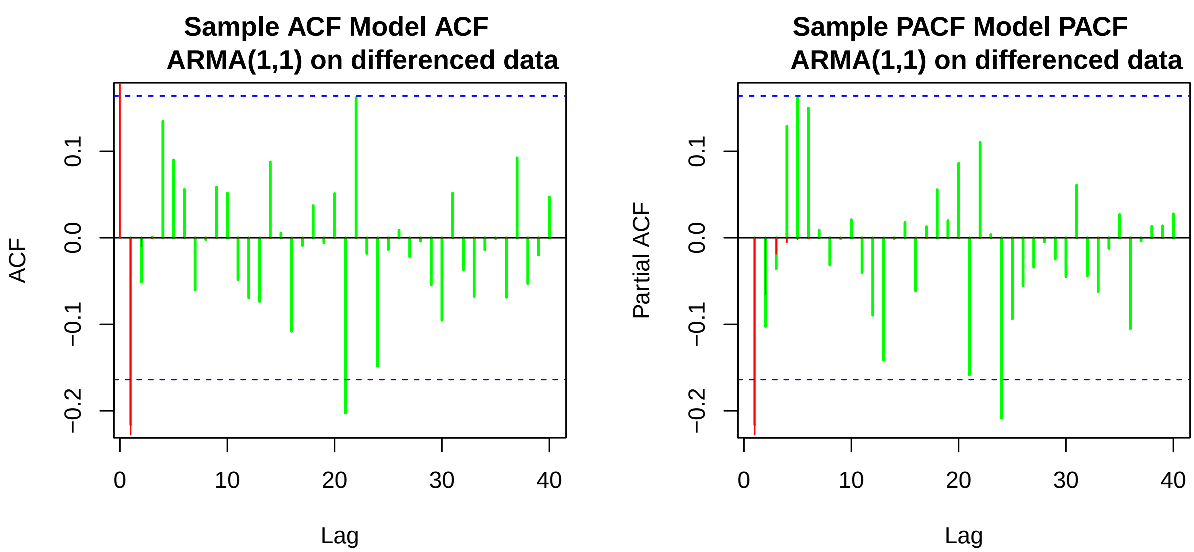
```
arma11fit2$aic
```

```
## [1] 1768.86
```

(c) Plot the model and sample acf/pacf.

```
par(mfrow = c(1,2))
par(mar = c(5, 4, 4.5, 2) + 0.1)
plot(sampleacf3, main = "", col = 'green', lwd = 2)
lines(x = 0:40, ARMAacf(ar = 0.0429182,
                        ma = -0.2866843, lag.max = 40),
      type = 'h', col = 'red')
title("Sample ACF Model ACF
      ARMA(1,1) on differenced data",line=0.5)

plot(samplepacf3, main = "", col = 'green', lwd = 2)
lines(ARMAacf(ar = 0.0429182,
              ma = -0.2866843, lag.max = 40, pacf = TRUE),
      type = 'h', col = 'red')
title("Sample PACF Model PACF
      ARMA(1,1) on differenced data", line=0.5)
```



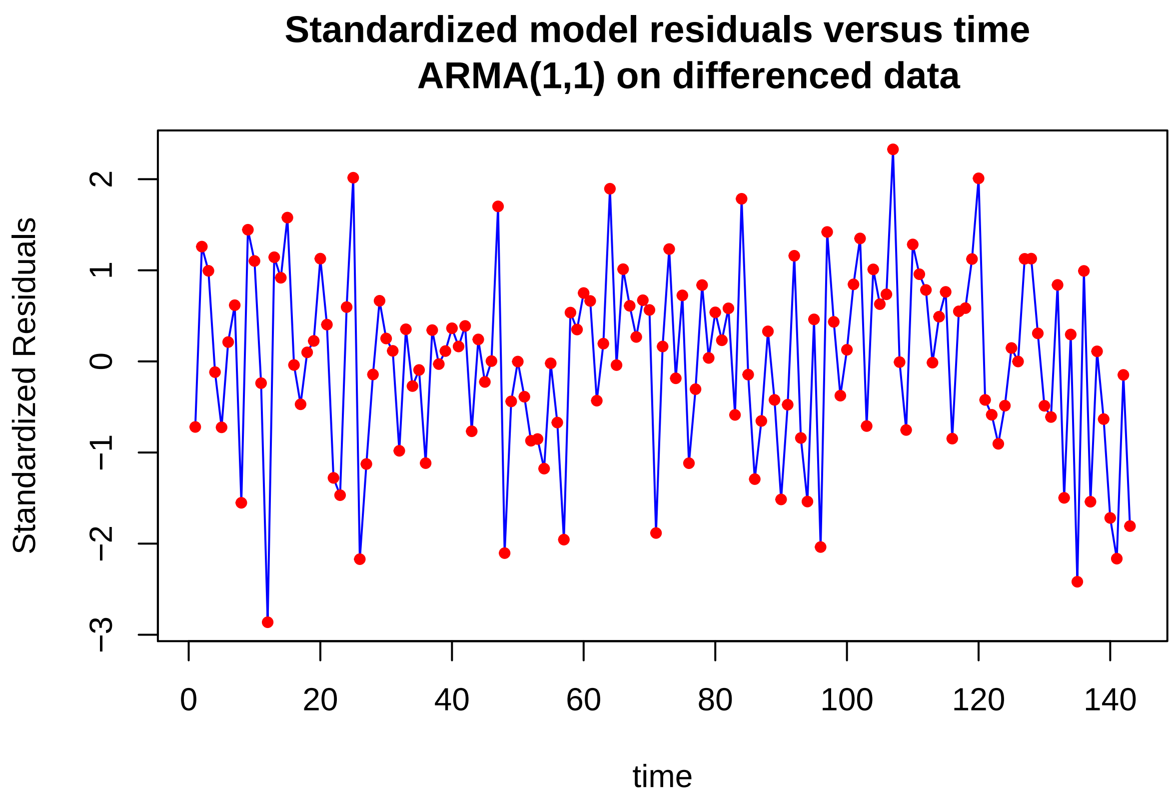
(d) Use the plot from (c) to assess the quality of the model fit.

We see good agreement in the significant acf and pacf values.

3. Analyze the standardized residuals.

(a) Plot the standardized residuals.

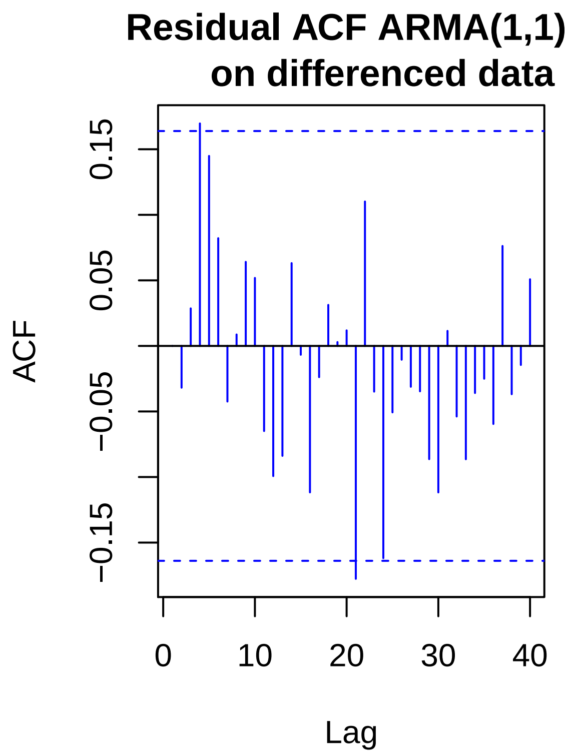
```
arma11diffresid = rstandard(arma11fit2)
plot(arma11diffresid, type = 'l', col = 'blue',
     xlab = 'time', ylab = 'Standardized Residuals',
     main = 'Standardized model residuals versus time
     ARMA(1,1) on differenced data')
points(arma11diffresid, pch = 20, col = 'red')
```

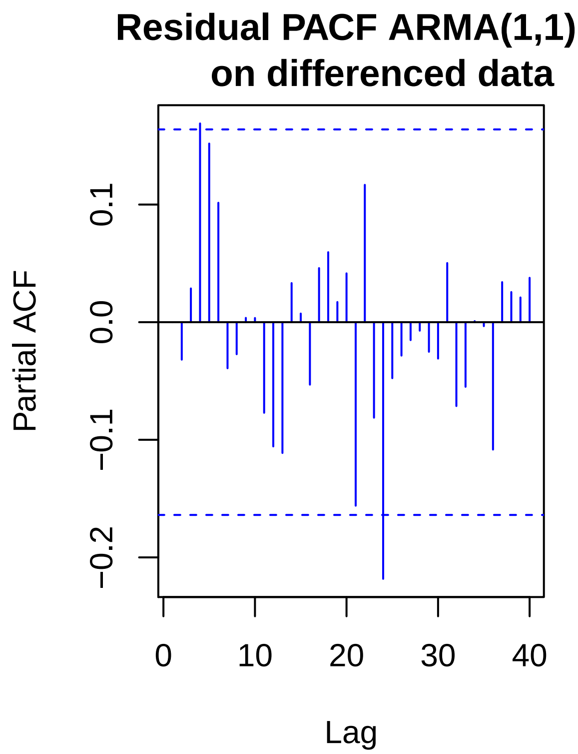


(b) Plot the sample acf/pacf for the standardized residuals.

```
par(mfrow = c(1,2))
par(mar = c(5, 4, 4.5, 2) + 0.1)
arma11diffresidacf = acf(arma11diffresid, lag.max = 40, plot = F)
plot(arma11diffresidacf, main = "", col = 'blue')
title("Residual ACF ARMA(1,1)
      on differenced data",line=0.5)

arma11diffresidpacf = pacf(arma11diffresid, lag.max = 40, plot = F)
plot(arma11diffresidpacf, main = "", col = 'blue')
title("Residual PACF ARMA(1,1)
      on differenced data", line=0.5)
```





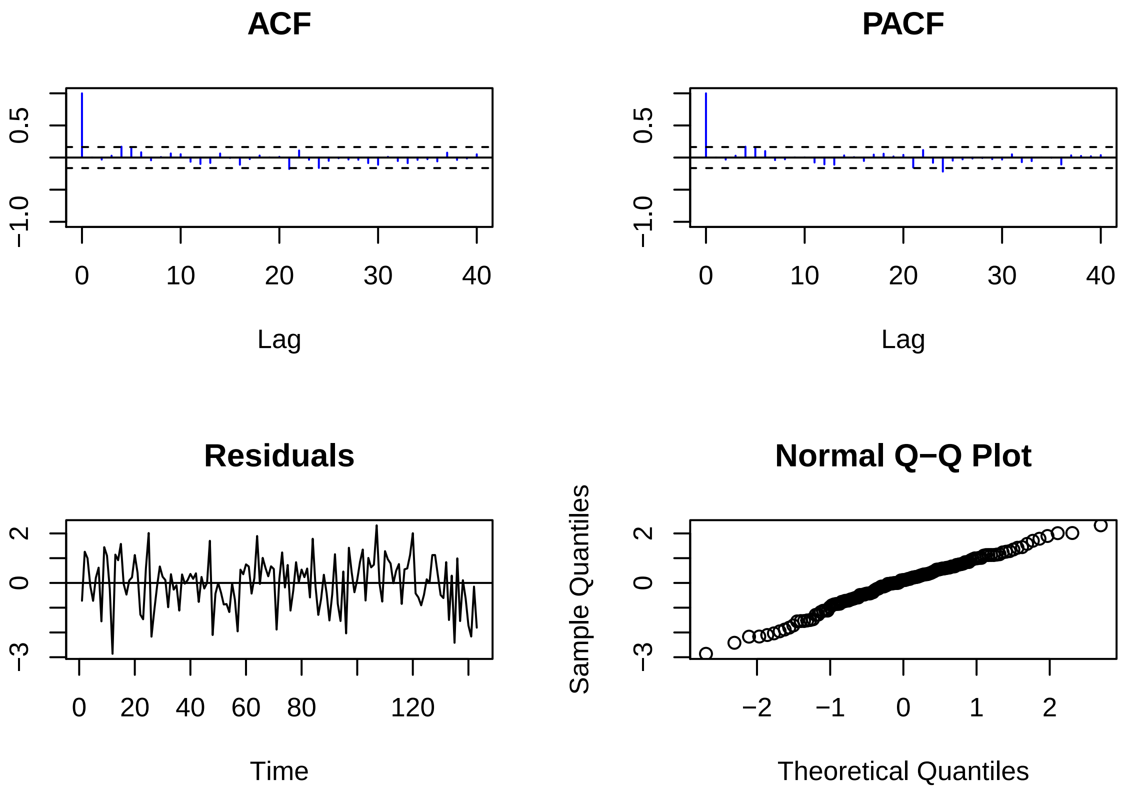
(c) Assess the plots for the hypothesis that the residuals are iid.

The plots of the residual acf and pacf do not support the rejection of iid noise hypothesis.

(d) Evaluate the Ljung-Box and McLeod-Li statistics and indicate if they support rejection of the iid hypothesis.

```
test(arma11diffresid)
```

```
## Null hypothesis: Residuals are iid noise.
## Test                        Distribution Statistic   p-value
## Ljung-Box Q                Q ~ chisq(20)     16.39    0.6919
## McLeod-Li Q                Q ~ chisq(20)     14.31    0.8144
## Turning points T       (T-94)/5 ~ N(0,1)        91    0.5493
## Diff signs S         (S-71)/3.5 ~ N(0,1)        69    0.5637
## Rank P         (P-5076.5)/286.5 ~ N(0,1)      4966    0.6997
```



The Ljung-Box statistic $Q_{LB} = 16.39$ with p-value 0.6919. This does not provide sufficient evidence to reject the iid hypothesis.

The McLeod-Li statistic $Q_{ML} = 14.31$ with p-value 0.8144. This does not provide sufficient evidence to reject the iid hypothesis.

(e) Give a final assessment on the validity of the iid hypothesis.

According to the residual acf/pacf plots and statistical tests, My conclusion is that the residuals are behaving like iid noise.

4. Use the results to give a summary evaluation about the quality of the fitted model.

The residuals are behaving like iid noise, and the model acf/pacf have a good agreement on sample acf/pacf. I would conclude that ARMA(1,1) on differenced data is a good fit.

(d) Use the results from (a) - (c) to decide which model is best and provide a reason for your choice.

We fitted ARMA(3,2) model, and ARMA(1,1) model on differenced data. The residuals plot for these two models are all behaving like iid noise.

```
tab <- matrix(c(arma32fit$aic, arma11fit2$aic), ncol=1, byrow=TRUE)
colnames(tab) <- c('AIC')
rownames(tab) <- c('ARMA(3,2) model', 'ARMA(1,1) model on differenced data')
tab <- as.table(tab)
tab
```

```
##                                          AIC
## ARMA(3,2) model                     1786.285
## ARMA(1,1) model on differenced data 1768.860
```

The AIC has a minimum value for ARMA(1,1) model on differenced data. The AIC suggests that the ARMA(1,1) model on differenced data is the best fit and this is similar to the analysis of model and sample acf/pacf we did earlier. We observe a good agreement for ARMA(1,1) model on differenced data. In conclusion, I would conclude that ARMA(1,1) on differenced data is the best model, and we should remove trend and seasonality first to get a stationary time series first and then fit the ARMA model.