Module 3 Assignment - ARMA Models

Filipp Krasovsky

11/15/2021

4.2

Let Wt be a white noise process with variance sigmaˆ2, let abs(Phi) < 1 be a constant. Consider X0 = w0 and Xt = PhiX(t-1) + Wt, t = 1,2…

(a) Show that Xt = sigma(j=0,t,Phi(j)*W(t-j)) for any t

(b) find E(xt)

b)

Show that $\mathcal{E}(X_t) = 0$.

$$\mathcal{E}(X_t) = \sum_{j=0}^{\infty} \mathcal{E}(w_{t-j}) \cdot \varphi^{j}$$

$$\mathcal{E}(w_{t-j}) = 0$$

$$\therefore$$

$$\mathcal{E}(X_t) = 0$$

(c) show that var(xt) = (sigmaˆ2 / (1-phiˆ2)) * (1-phiˆ2(t+1))

1. let $X_t = \varphi X_{t-1} + w_t$ where $|\varphi| < 1$

and $w_t \sim (0, \sigma^2)$

$X_t = \varphi(B) X_t + \theta(B) W_t$ where $\theta(B) = (1 + \theta_1 B + \theta_2 B^2 \dots)$

where $\varphi(B) = (1 - \varphi_1 B - \varphi_2 B^2 \dots)$

and $\theta(B) = 1$

$$x_t(1 - \varphi(B)) = \theta(B) W_t$$

$$\therefore X_t = \frac{\theta(B)}{1 - \varphi(B)} W_t = \varphi^{-1}(B)\, \theta(B) w_t = \psi(B) w_t$$

$$= \sum_{j=0}^{\infty} \psi_j w_{t-j}$$

$$X_t = \frac{\theta(B) w_t}{\varphi(B)} = \frac{w_t}{\varphi(B)}$$

because $|\varphi| < 1$, $\frac{1}{\varphi(z)} = \frac{1}{1 - \varphi z} = \sum_{j=0}^{\infty} \varphi^j z^j$

so $\psi_j = \varphi_j$

$$\therefore \quad X_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \boxed{\sum_{j=0}^{\infty} \varphi_j w_{t-j}}$$

Figure 1: part a

Prove that $var(x_t) = \frac{\sigma^2_w}{1-\varphi_2} \cdot \left(1-\varphi^{2(t+1)}\right)$

Induction:

$$Var(X_{t=0}) = \cancel{\frac{\sigma^2_w}{1-\varphi^2} \cdot (1-\varphi^2)} = \sigma^2_w$$

where $X_t = w_t$, $var(X_t) = var(w_t) = \sigma^2_w$

<u>for $X_{t\geq 1}$</u>

$$var(X_{t=1}) = \frac{\sigma^2_w}{1-\varphi_2} \cdot 1-\varphi^{2(2)} = \sigma^2 \cdot \frac{(1-\varphi^4)}{(1-\varphi 2)}$$

$$= \frac{\sigma^2_w(1-\cancel{\varphi^2})(1+\varphi^2)}{(1-\cancel{\varphi^2})} = (1+\varphi^2)\cdot\sigma^2_w$$

$$Var(x_{t=1}) = \varphi^2\, Var(x_0) + v(w_1)$$
$$= \varphi^2 \sigma^2_w + \sigma^2_w = (1+\varphi^2)\sigma^2_w$$

$$Var(x_{t=2}) = \frac{\sigma^2_w}{1-\varphi_2} \cdot 1-\varphi^6$$

$$Var(x_2) = \varphi^2\, var(x_1) + \sigma^2_w = \varphi^2(1+\varphi^2)\sigma^2_w + \sigma^2_w$$
$$\frac{\sigma^2_w(1-\varphi^6)}{(1-\varphi^2)} \qquad = (\varphi^2+\varphi^4+1)\sigma^2_w$$

$$(1-\varphi^6) = (\varphi^2+\varphi^4+1)(1-\varphi^2) = \cancel{\varphi^2}-\cancel{\varphi^4}+\cancel{\varphi^4}-\varphi^6+1-\cancel{\varphi^2} =$$
$$(1-\varphi^6)$$

etc.

Figure 2: part c

d) show that $h \geq 0$, $cov(x_{t+h}, x_t) = \varphi^h var(x_t)$

$$cov(X_{t+h}, x_t) = cov\left(\sum \varphi^j w_{t+h-j}, \sum \varphi^k w_{t-k}\right)$$

$$= cov\left(\left[w_{t+h} + \ldots + \varphi^h w_t + \varphi^{h+1} w_{t-1}\right], \left[\varphi^0 w_t + \varphi w_{t-1} + \ldots\right]\right)$$

$$= \sigma_w^2 \sum_{j=0}^{\infty} \varphi^{h+j} \varphi^j = \sigma_w^2 \varphi^h \sum_{j=0}^{\infty} \varphi^{2j} = \boxed{\frac{\sigma_w^2 \varphi^h}{1-\varphi^2}}$$

e) is X_t stationary? yes.

f) as $\lim_{t \to \infty}$, X_t becomes stationary:

$$\lim_{t \to \infty} Var(x_t) = \frac{\sigma^2}{1-\varphi^2}\left(1 - \varphi^{2(t+1)}\right)$$

given that $|\varphi| < 1$, $\lim_{t \to \infty} \varphi^{2(t+1)} \to 0$

$$\therefore \lim_{t \to \infty} Var(x_t) = \frac{\sigma^2}{1-\varphi^2} \cdot (1-0) = \frac{\sigma^2}{1-\varphi^2}$$

Figure 3: part a

(d) show that h>=0, cov(xt+h,xt) = phi(h)*var(xt)

(e) is Xt stationary?

(f) argue that as t->inf the process becomes stationary

(g) comment on how you could use the results to siumulate

n observations of a stationary gaussian AR(1) model from

simulated iid N(0,1) values

As in the examples provided in the textbook, we could generate an excess amount of values and discard the first X (usually ~50 values).

(h) supposed X0 = W0/sqrt(1-phi^2). is this process stationary? show var(xt) is constant.

X_0 is stationary.

$$X_0 = \frac{W_0}{\sqrt{1-\varphi^2}}$$

$$var(X_t) = Var\left(\frac{W_0}{\sqrt{1-\varphi^2}}\right)$$ were φ is constant

let $\alpha = \frac{1}{\sqrt{1-\varphi^2}}$, $var(\alpha W_0) = \alpha^2 var(W_0)$

$$Var(X_t) = \frac{1}{1-\varphi^2} \cdot \sigma^2_w$$

$$\left[\begin{array}{l} E(X_t) = 0 \\ Var(X_0) = \frac{\sigma^2_w}{1-\varphi^2} \text{ is constant.} \end{array}\right]$$

4.3 (8 Points) ## Consider the following models: ## Xt = .8Xt-1 - .15Xt-2 + wt - .3Wt-1 ## Xt =

Xt-1 - .5Xt-2 + Wt + Wt-1

(a) check the models for parameter redundancy, find

the reduced form

model 1: rewrite: (1 - .8B + .15B^2)Xt = (1-.3B)Wt (1-.5B)(1-.3B)Xt = (1-.3B)Wt (1-.5B)Xt = Wt Xt = .5Xt-1 + Wt Model 1 is an AR(1) series.

Model 2: Model 2 does not reduce, it is an ARMA(2,1) series.

(b) check if the models are causal and or invertible

in their reduced forms, where applicable.

(c) For each of the reduced models find the first 50

coefficients and see if their ARMAtoMA and ARMAtoAR

transformations converge to zero.

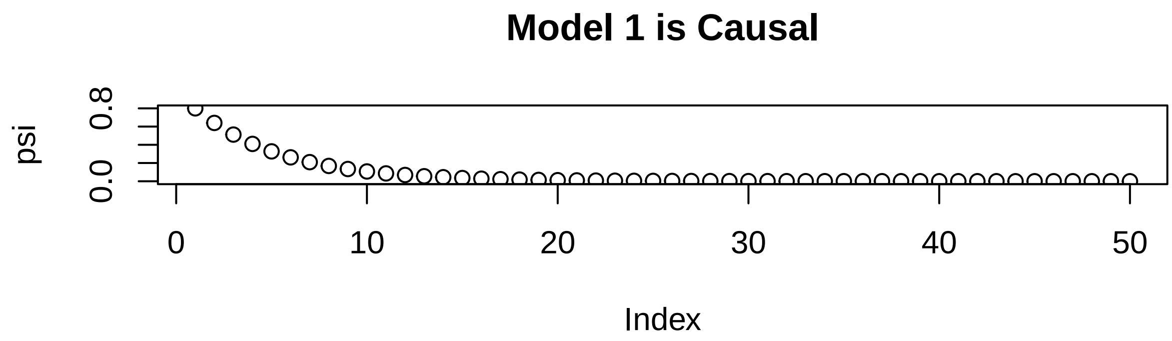
(both answered concurrently.)

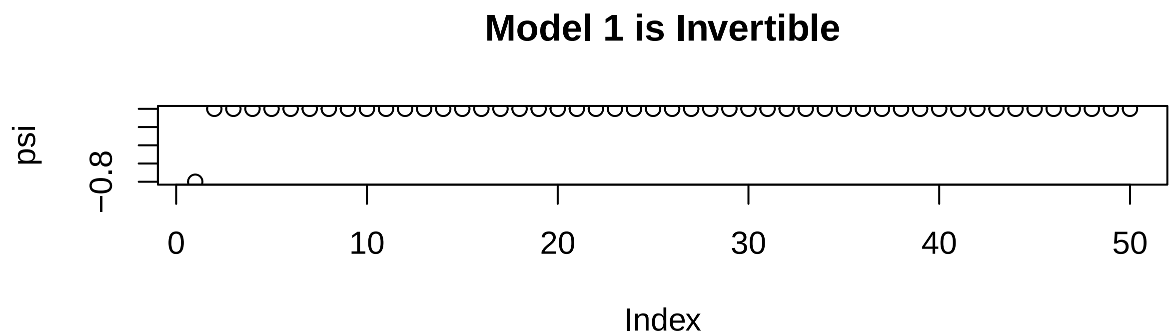
Models 1 and 2

```
require(astsa)
```

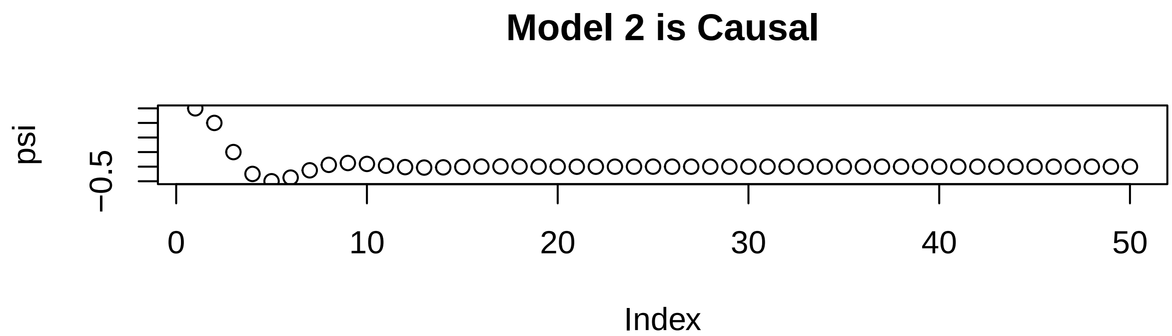
```
## Loading required package: astsa
```

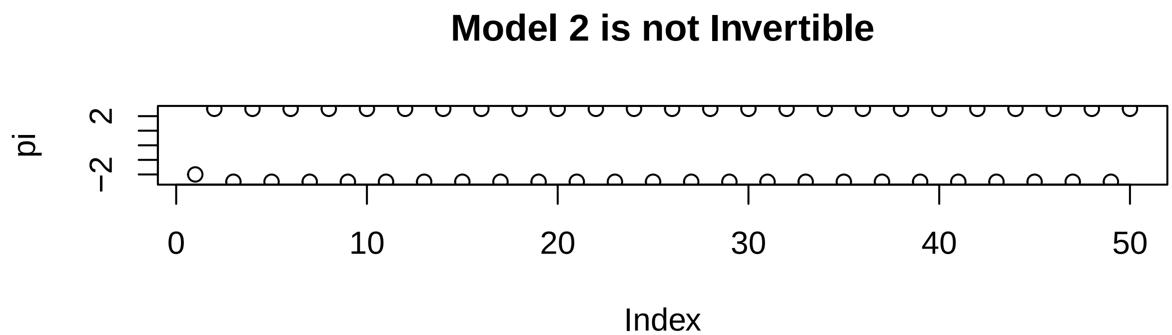
```
library(Metrics)
psi = ARMAtoMA(ar = c(0.8),ma=0,50)
par(mfrow=c(2,1))
plot(psi,main="Model 1 is Causal")

psi = ARMAtoAR(ar=c(0.8),ma=0,50)
plot(psi,main = "Model 1 is Invertible")
```





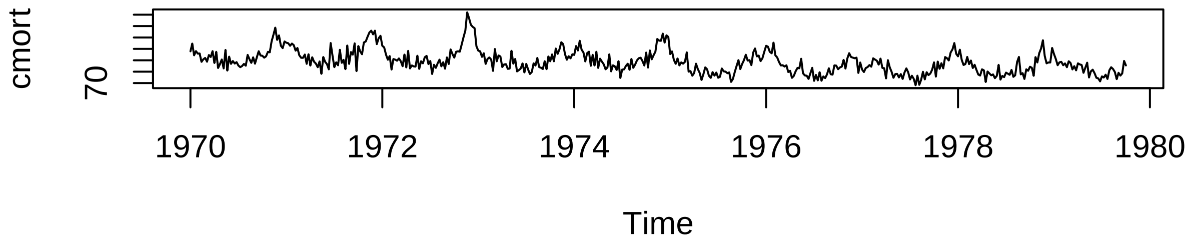

```
psi = ARMAtoMA(ar=c(1,-.5),ma=1,50)
par(mfrow=c(2,1))
plot(psi,main="Model 2 is Causal")

pi = ARMAtoAR(ar=c(1,-.5),ma=1,50)
plot(pi,main = "Model 2 is not Invertible")
```

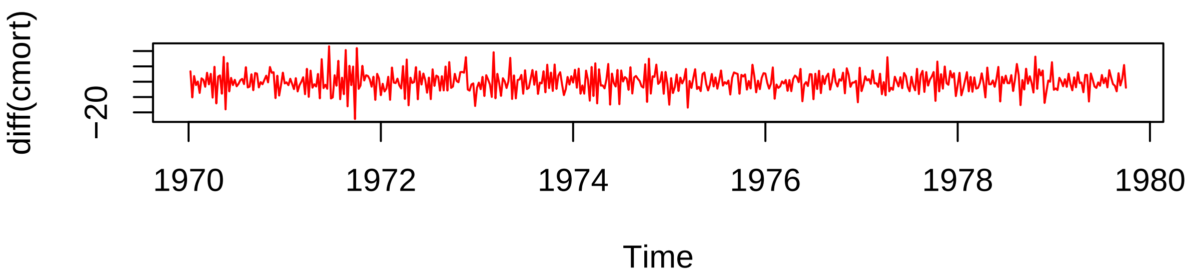




Let Ct be the cardiovascular mortality series discussed in 3.5 and let Xt = diff(ct) be the differenced data. ## (a) plot xt and compare it to the actual data plotted in 3.2. Why does differencing seem reasonable?

```
par(mfrow=c(2,1))
plot(cmort,col='black')
plot(diff(cmort),col='red')
```

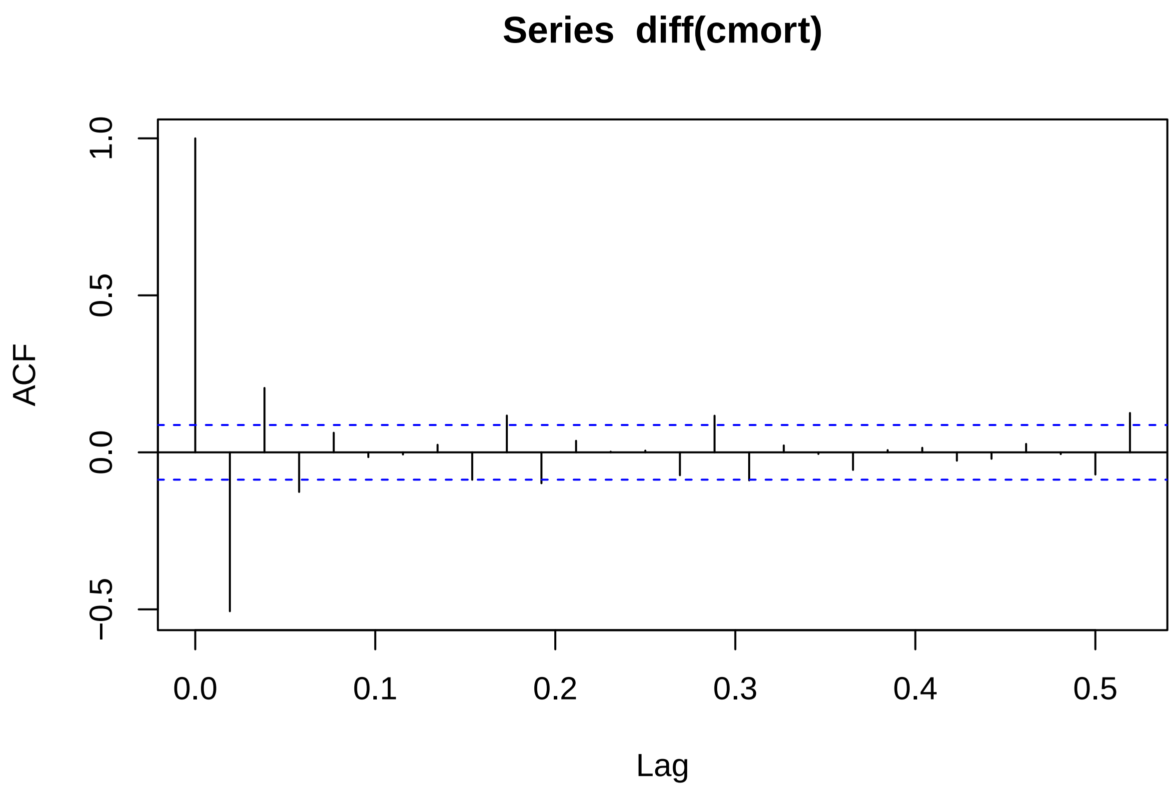




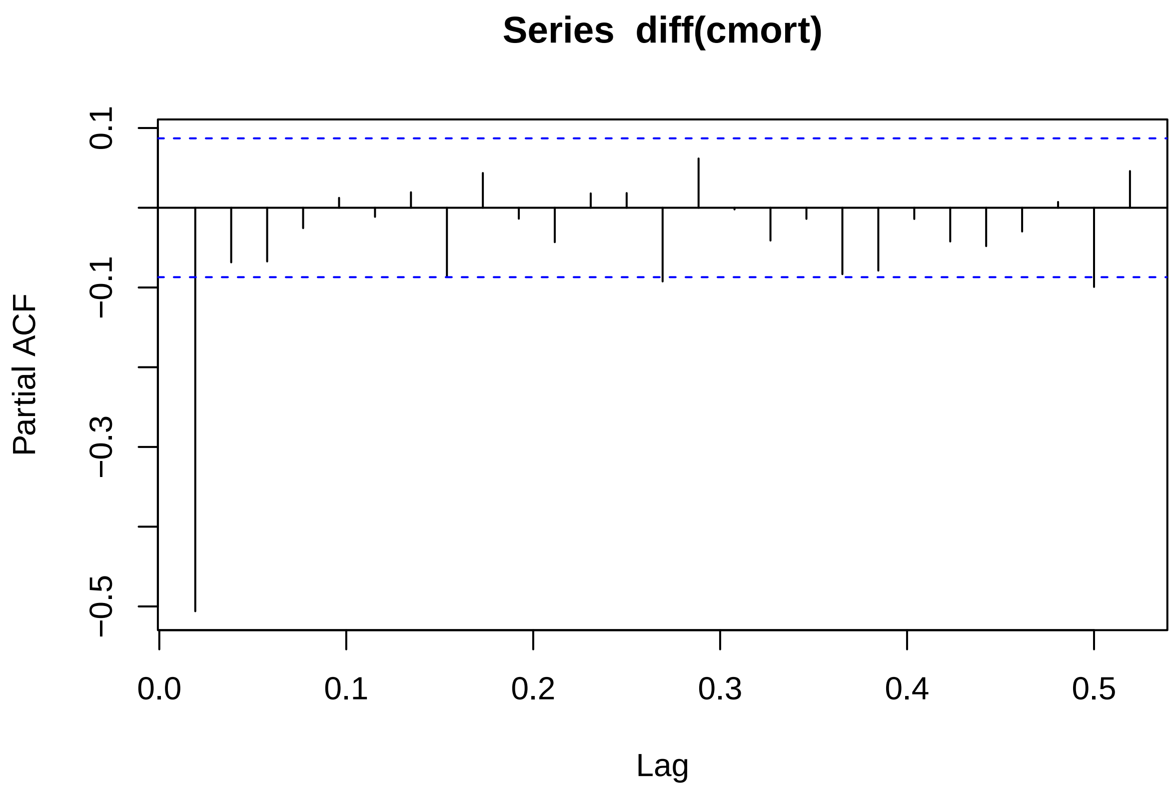
On face value, the original dataset seems to be non-stationary and difficult to model, while the differenced data has less drift and appears stationary.

(b) calculate and plot the sample acf and pacf of xt using table 4.1, argue that an AR(1) is appropriate.

```
acf(diff(cmort))
```




```
pacf(diff(cmort))
```



The ACF does not seem to cut off at any particular lag, gradually decaying, while the PACF has a strong spike at lag = 1, thus making a reasonable case to fit an AR(1) model.

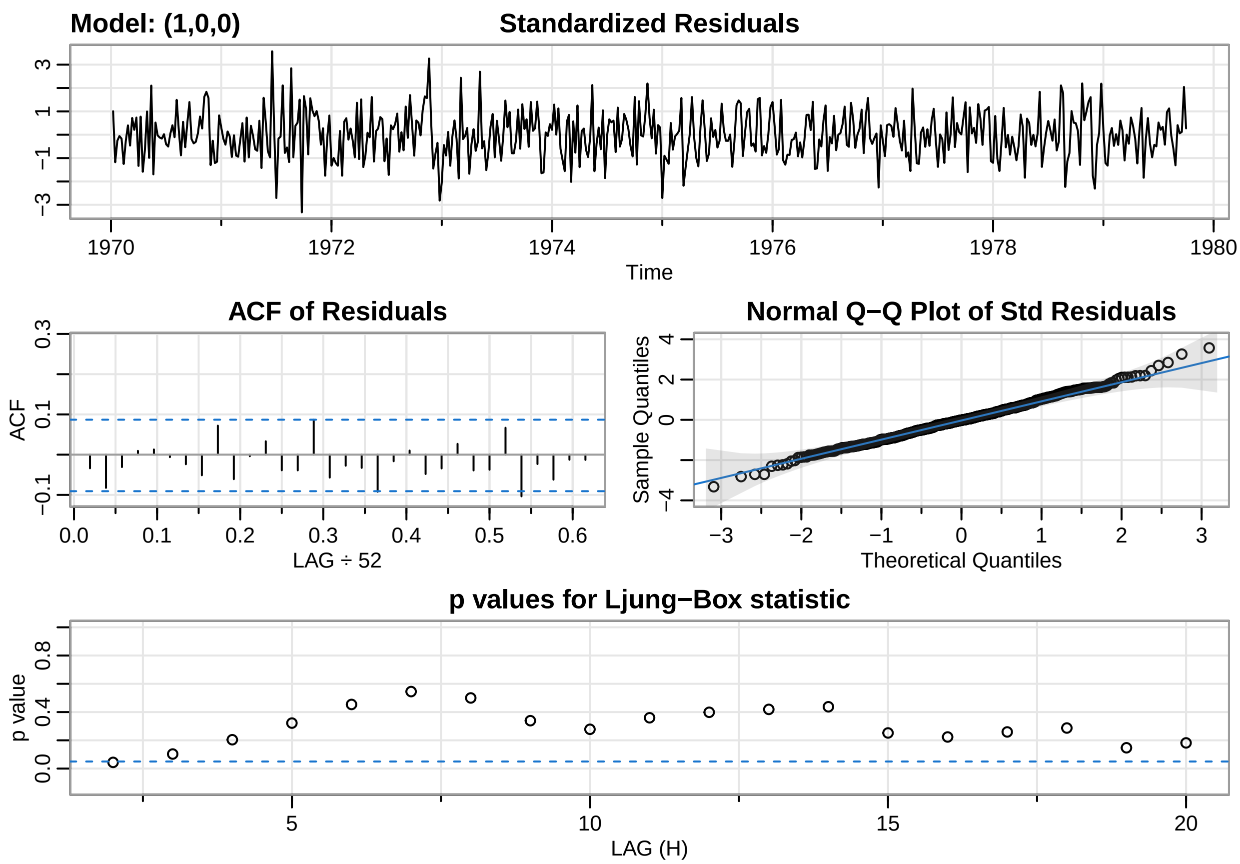
(c) fit an AR1 using maximum likelihood. comment on the significance of the regression parameters.

what is the estimate of the white noise variance?

(d)comment on the residuals.

```
xt = diff(cmort)
sarima(xdata=xt,p=1,d=0,q=0,no.constant=TRUE)
```

```
## initial  value 1.908720
## iter   2 value 1.760373
## iter   3 value 1.760373
## iter   4 value 1.760373
## iter   4 value 1.760373
## final  value 1.760373
## converged
## initial  value 1.760679
## iter   1 value 1.760679
## final  value 1.760679
## converged
```



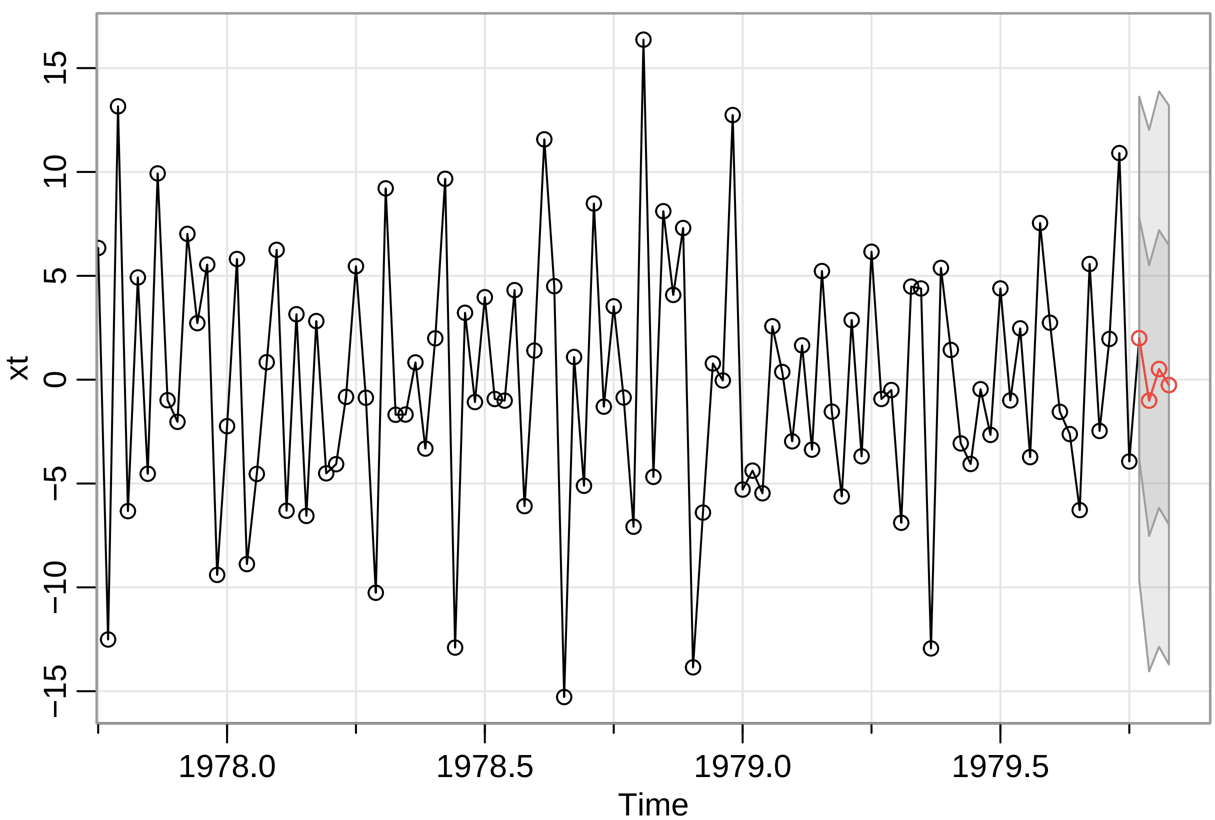

```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     xreg = xmean, include.mean = FALSE, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##           ar1
##       -0.5064
## s.e.   0.0383
##
## sigma^2 estimated as 33.81:  log likelihood = -1612.07,  aic = 3228.13
##
## $degrees_of_freedom
## [1] 506
##
## $ttable
##     Estimate     SE  t.value p.value
## ar1  -0.5064 0.0383 -13.2224       0
##
## $AIC
## [1] 6.367124
##
## $AICc
## [1] 6.36714
```

```
##
## $BIC
## [1] 6.383805
```

The white noise variance is estimated at 33.81, while the AR1 coefficient has been estimated at -0.5, suggesting that moments in time are negatively correlated with the data point one unit of time prior, and positively correlated with moments in time two units behind (and forward). The residuals seem to be a white noise series with an ACF that has no significant spikes. this is also supported by the Ljung-Box p-values all being significant enough to fail to reject the null hypothesis that the data are indipendently distributed.

e) Assuming the fitted model is the true model, find the forecasts over a four-week horizon, for m=1,2,3,4, and the corresponding 95% prediction intervals, n=508 here. The easiest way to do this is to use sarima.for from astsa.

```
m4pred = sarima.for(xt,n.ahead=4,p=1,d=0,q=0,no.constant=TRUE)
```




```
print(m4pred)
```

```
## $pred
## Time Series:
## Start = c(1979, 41)
## End = c(1979, 44)
```

```
## Frequency = 52
## [1]  1.9950506 -1.0102099  0.5115279 -0.2590162
##
## $se
## Time Series:
## Start = c(1979, 41)
## End = c(1979, 44)
## Frequency = 52
## [1] 5.814686 6.517634 6.685975 6.728459
```

(f) show how the values were calculated

presumably using the formula xt = -0.5064 * x(t-1) as the point estimate with a confidence interval using alpha=.025 and .05. mu* St.Dev *(sqrt(1+(1/n))), where t is a tabled value from the t distribution which depends on the confidence level and sample size.

(g) what is the one step ahead forecast value value?

```
print(m4pred$pred[1])
```

```
## [1] 1.995051
```

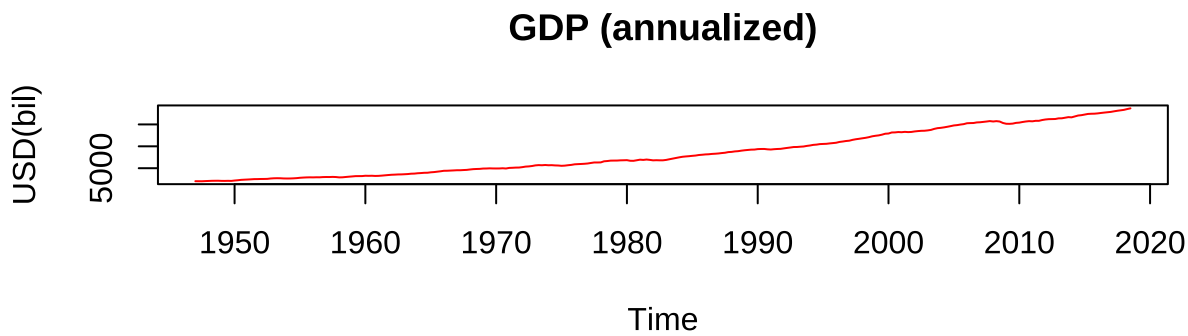
5.2

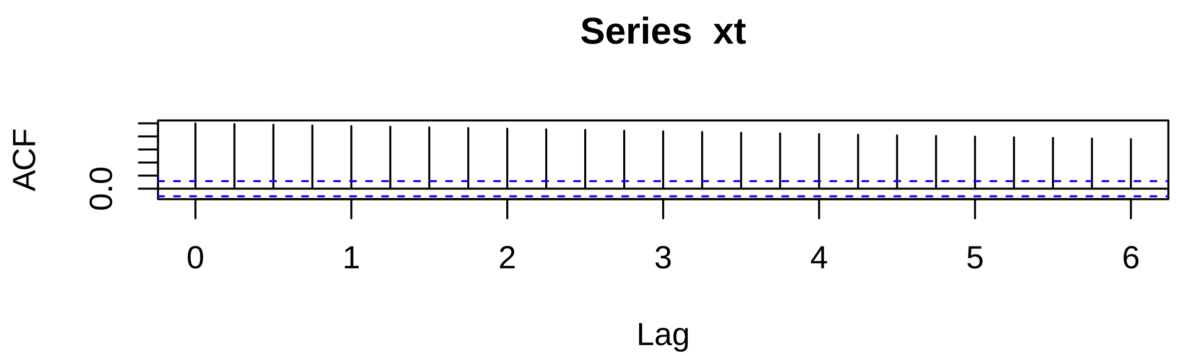
in example 5.6 we fit an ARIMA model. repeat the analysis for the US GDP series in GDP.

discuss all aspects of the fit as specified in the beginning of 5.2 from plotting the data to diagnostics

and model choice.

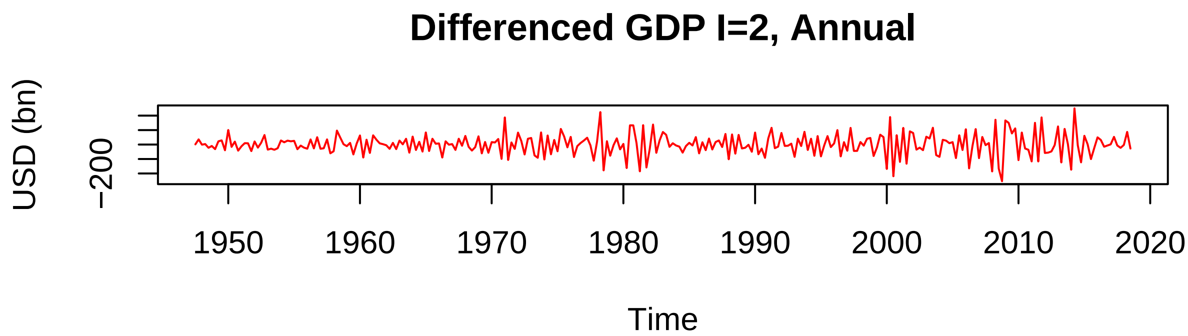
```
par(mfrow=c(2,1))
xt = astsa::gdp;
#1. plot the data
plot(xt,col='red',main="GDP (annualized)",ylab="USD(bil)")
#2. we can also observe the ACF - a slow decay in the ACF suggests differencing.
acf(xt)
```

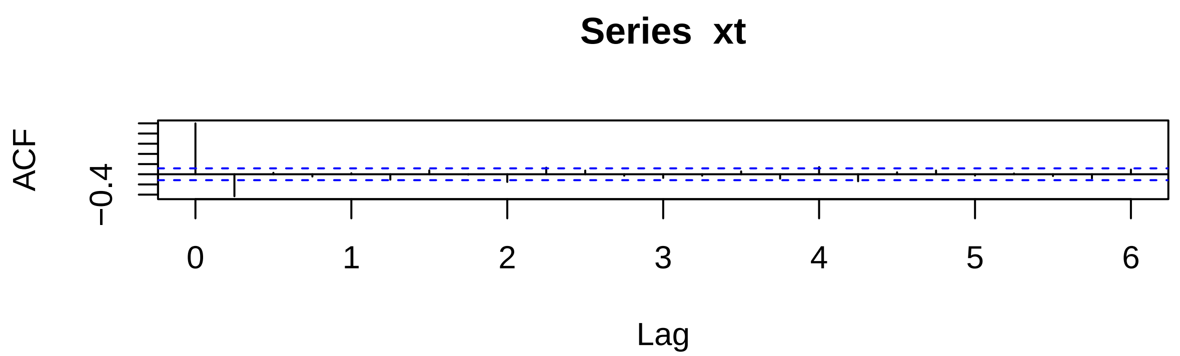




This series is not stationary and therefore needs to be transformed based on the data itself and the acf.

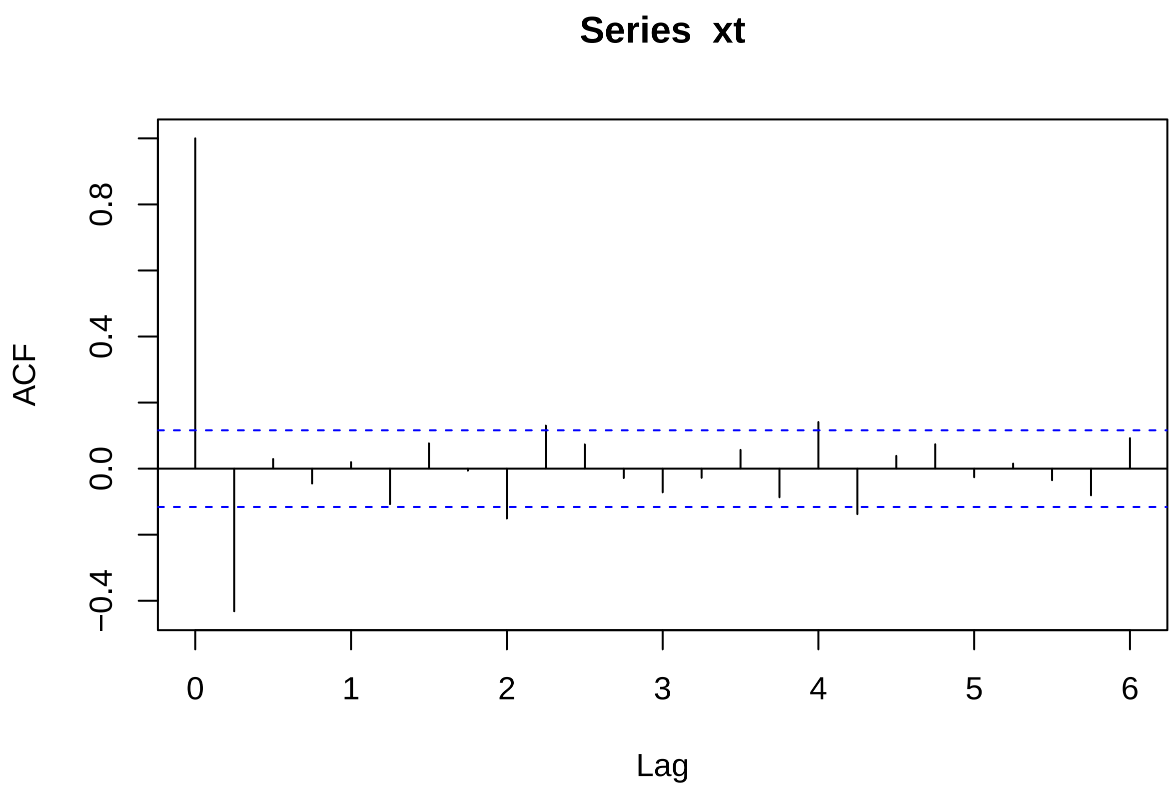
```
#2. transform the data
par(mfrow=c(2,1))
xt = diff(diff(gdp))
plot(xt,col='red',main="Differenced GDP I=2, Annual",ylab='USD (bn)')
acf(xt)
```



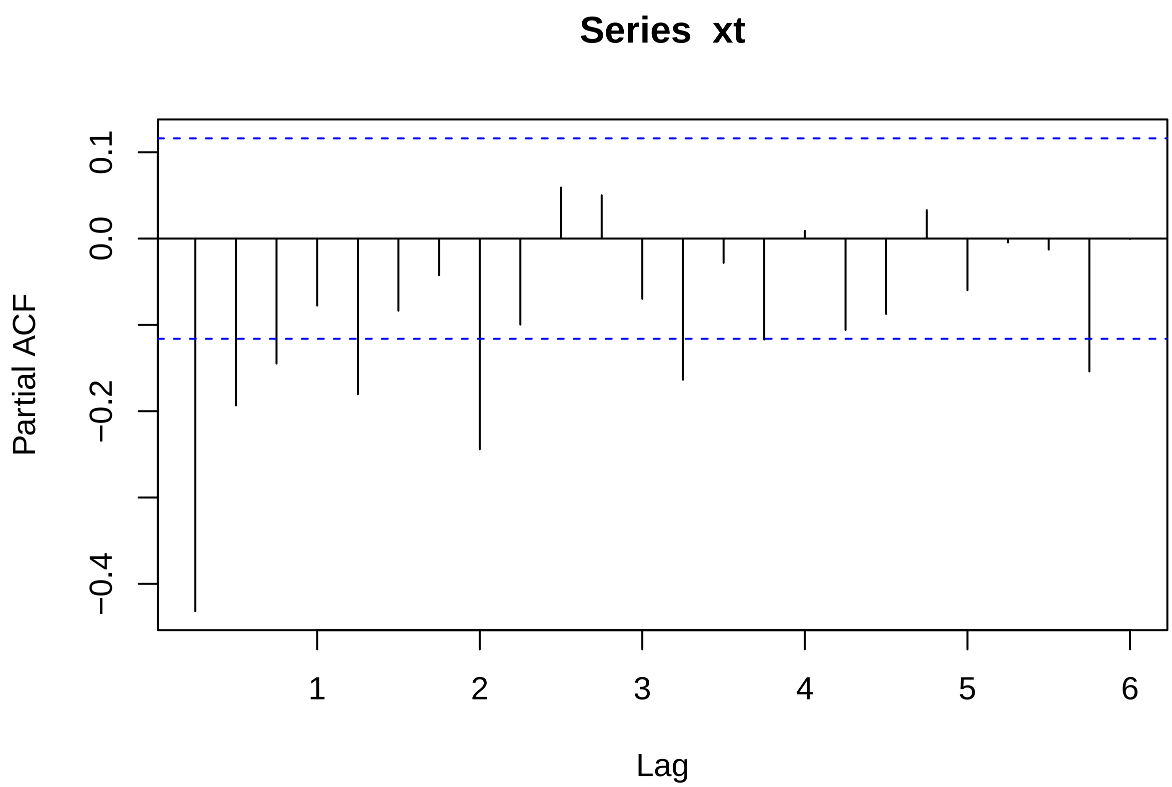


This model is mostly stationary, but had a significant drop in 2008, around the time of the financial recession. A more robust model would take this into consideration, but we will proceed with observing the dependence orders. The ACF suggests a quick gradual decay, so additional differencing is not necessary.

```
acf(xt)
```




```
pacf(xt)
```

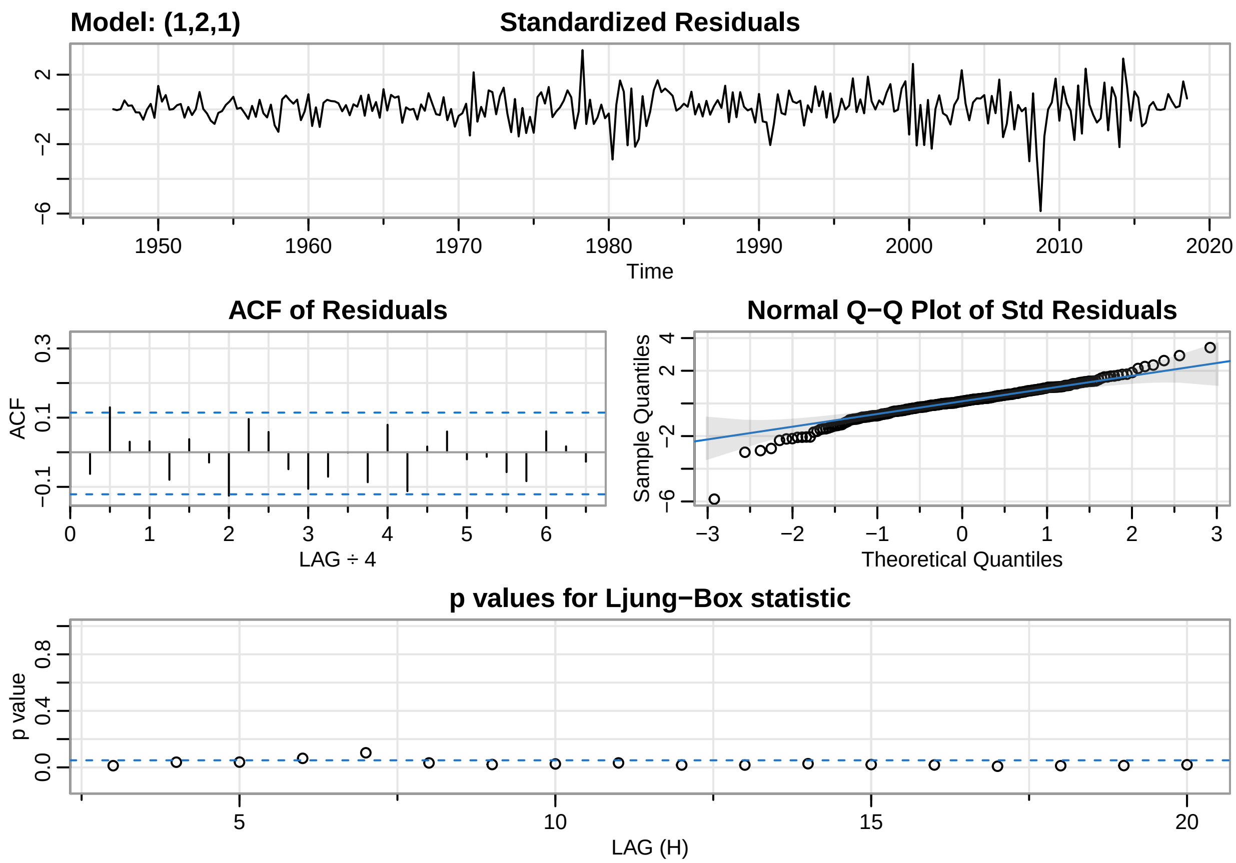


Initial behavior suggests an MA model based on the prominent ACF spikes at lags 1 and 2, with gradually decaying PACF spikes.

```
#iterate over all  possibilities
for (i in 1:3){
  print(sarima(xdata=gdp,p=1,d=2,q=i,no.constant=TRUE))
}
```

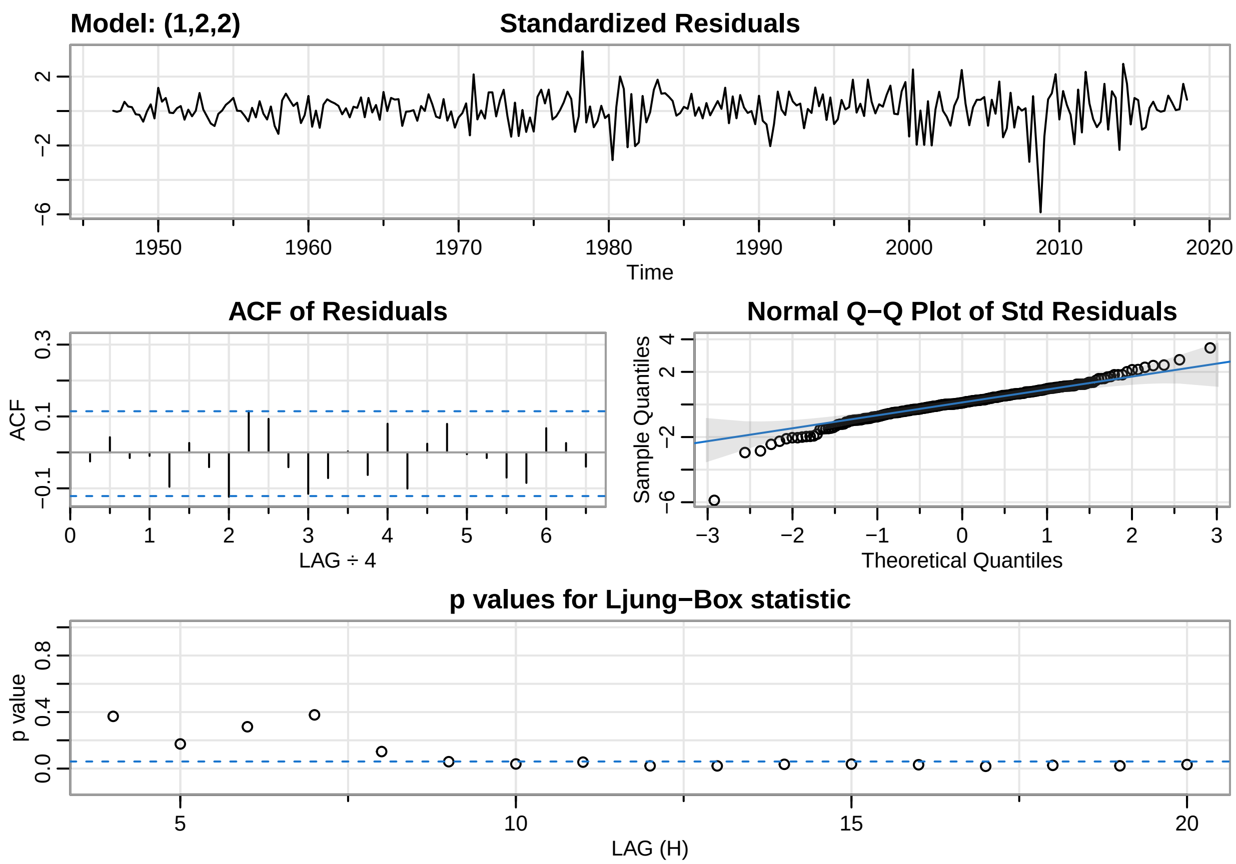
```
## initial  value 4.306246
## iter   2 value 4.217085
## iter   3 value 4.178588
## iter   4 value 4.174787
## iter   5 value 4.153340
## iter   6 value 4.137937
## iter   7 value 4.132149
## iter   8 value 4.127528
## iter   9 value 4.126766
## iter  10 value 4.125878
## iter  11 value 4.125756
## iter  12 value 4.125692
## iter  13 value 4.125576
## iter  14 value 4.125548
## iter  14 value 4.125548
## iter  14 value 4.125548
## final  value 4.125548
## converged
## initial  value 4.125025
```

```
## iter   2 value 4.124928
## iter   3 value 4.124882
## iter   4 value 4.124875
## iter   5 value 4.124875
## iter   5 value 4.124875
## iter   5 value 4.124875
## final  value 4.124875
## converged
```



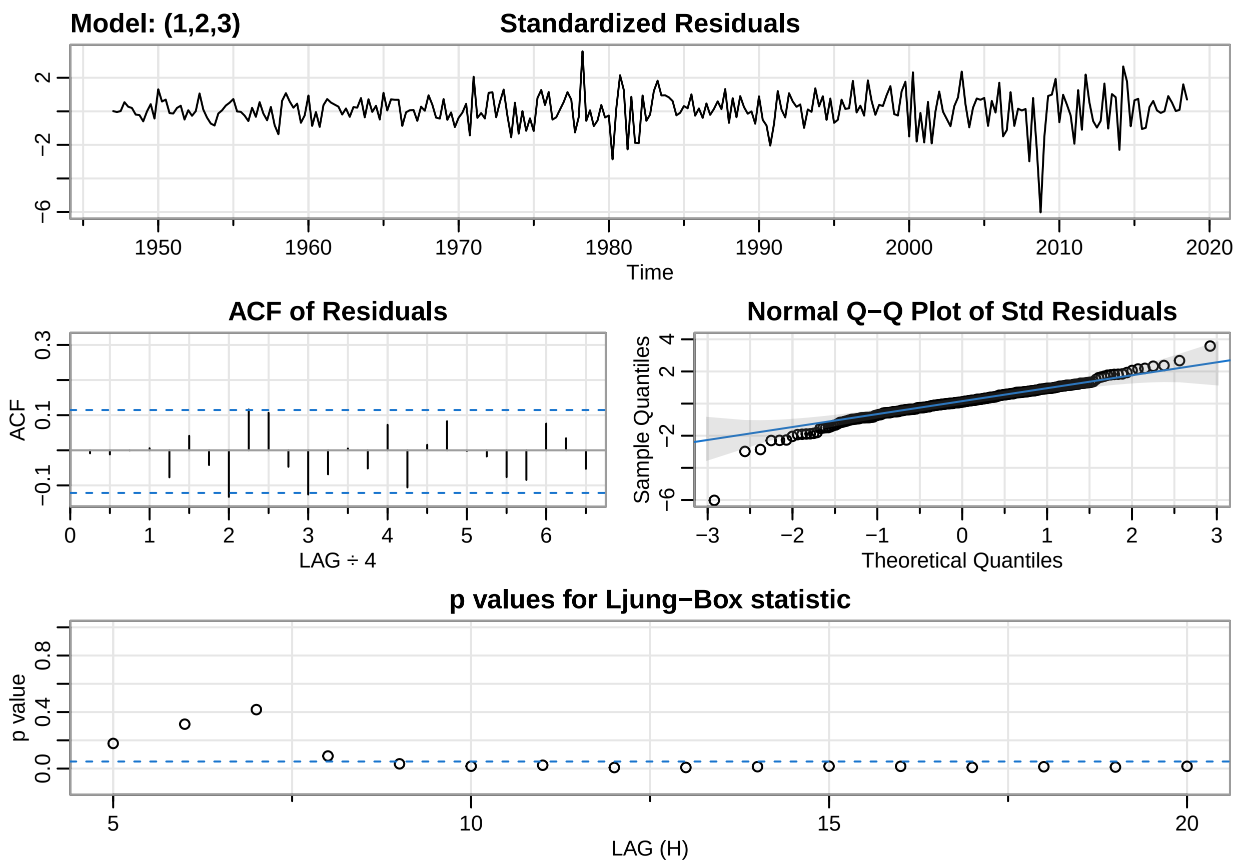

```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     include.mean = !no.constant, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1      ma1
##       0.3452  -0.9693
## s.e.  0.0591   0.0165
##
## sigma^2 estimated as 3798:  log likelihood = -1579.99,  aic = 3165.97
##
## $degrees_of_freedom
## [1] 283
```

```
##
## $ttable
##      Estimate     SE  t.value p.value
## ar1    0.3452 0.0591   5.8412       0
## ma1   -0.9693 0.0165 -58.7617       0
##
## $AIC
## [1] 11.10868
##
## $AICc
## [1] 11.10883
##
## $BIC
## [1] 11.14713
##
## initial  value 4.306246
## iter   2 value 4.224153
## iter   3 value 4.160797
## iter   4 value 4.150560
## iter   5 value 4.149246
## iter   6 value 4.148795
## iter   7 value 4.147655
## iter   8 value 4.147156
## iter   9 value 4.146189
## iter  10 value 4.144975
## iter  11 value 4.141673
## iter  12 value 4.130080
## iter  13 value 4.125154
## iter  14 value 4.119628
## iter  15 value 4.118340
## iter  16 value 4.117077
## iter  17 value 4.116734
## iter  18 value 4.115877
## iter  19 value 4.115359
## iter  20 value 4.115321
## iter  21 value 4.115148
## iter  22 value 4.115106
## iter  23 value 4.115105
## iter  24 value 4.115018
## iter  25 value 4.114936
## iter  26 value 4.114935
## iter  27 value 4.114934
## iter  27 value 4.114934
## iter  27 value 4.114934
## final  value 4.114934
## converged
## initial  value 4.114484
## iter   2 value 4.114338
## iter   3 value 4.114328
## iter   4 value 4.114327
## iter   4 value 4.114327
## iter   4 value 4.114327
## final  value 4.114327
## converged
```




```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     include.mean = !no.constant, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1      ma1     ma2
##       0.6560  -1.3181  0.3330
## s.e.  0.1034   0.1223  0.1177
##
## sigma^2 estimated as 3720:  log likelihood = -1576.98,  aic = 3161.96
##
## $degrees_of_freedom
## [1] 282
##
## $ttable
##     Estimate     SE  t.value p.value
## ar1   0.6560 0.1034   6.3450   0.000
## ma1  -1.3181 0.1223 -10.7813   0.000
## ma2   0.3330 0.1177   2.8306   0.005
##
## $AIC
## [1] 11.0946
##
```

```
## $AICc
## [1] 11.0949
##
## $BIC
## [1] 11.14586
##
## initial  value 4.306246
## iter   2 value 4.214362
## iter   3 value 4.142711
## iter   4 value 4.135559
## iter   5 value 4.129135
## iter   6 value 4.127703
## iter   7 value 4.127151
## iter   8 value 4.125891
## iter   9 value 4.125784
## iter  10 value 4.125574
## iter  11 value 4.125538
## iter  12 value 4.125386
## iter  13 value 4.119562
## iter  14 value 4.116961
## iter  15 value 4.115132
## iter  16 value 4.113646
## iter  17 value 4.113048
## iter  18 value 4.112874
## iter  19 value 4.112772
## iter  20 value 4.112764
## iter  21 value 4.112758
## iter  22 value 4.112756
## iter  22 value 4.112755
## iter  22 value 4.112755
## final  value 4.112755
## converged
## initial  value 4.112254
## iter   2 value 4.112109
## iter   3 value 4.112086
## iter   4 value 4.112081
## iter   5 value 4.112081
## iter   6 value 4.112081
## iter   6 value 4.112081
## iter   6 value 4.112081
## final  value 4.112081
## converged
```




```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     include.mean = !no.constant, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1      ma1     ma2      ma3
##       0.5342  -1.2136  0.3261  -0.0921
## s.e.  0.1690   0.1730  0.1387   0.0787
##
## sigma^2 estimated as 3703:  log likelihood = -1576.34,  aic = 3162.68
##
## $degrees_of_freedom
## [1] 281
##
## $ttable
##     Estimate     SE t.value p.value
## ar1   0.5342 0.1690  3.1613  0.0017
## ma1  -1.2136 0.1730 -7.0144  0.0000
## ma2   0.3261 0.1387  2.3507  0.0194
## ma3  -0.0921 0.0787 -1.1700  0.2430
##
## $AIC
## [1] 11.09713
```

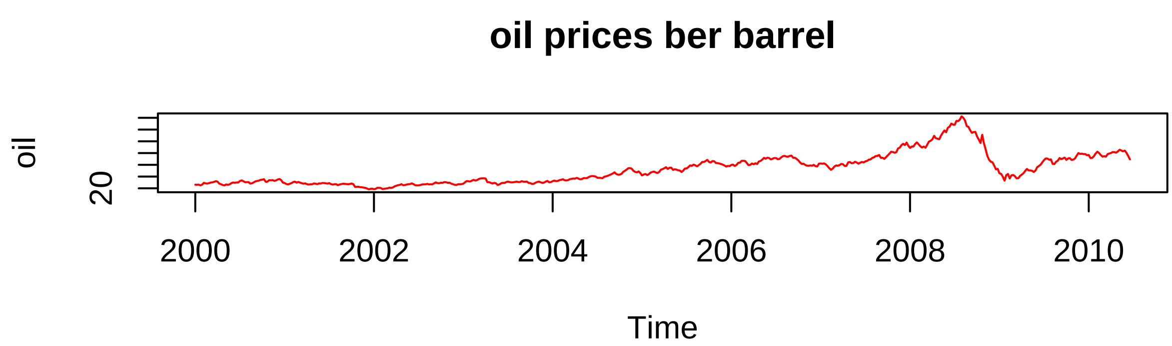
```
##
## $AICc
## [1] 11.09763
##
## $BIC
## [1] 11.1612
```

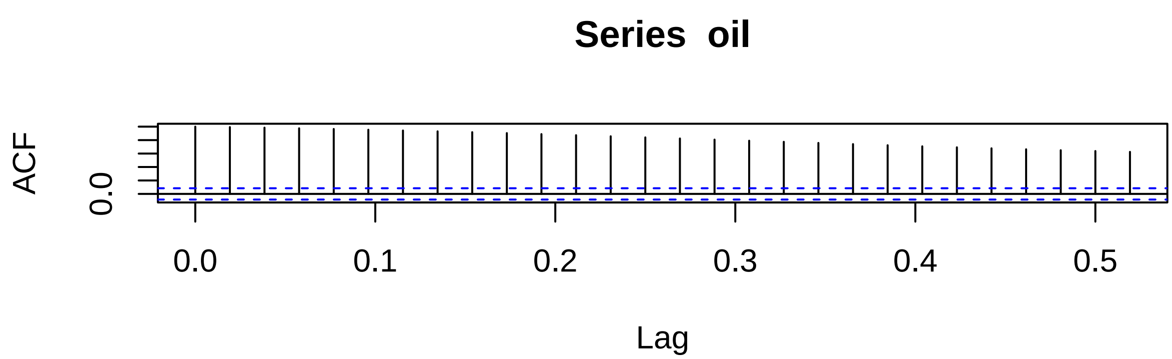
Of these, it appears that an ARIMA(1,2,2) model performs best, where I=2 because of the differenced dataset when looking at the AIC and the distribution of the residuals.

Example 5.3 [8 points].

##Crude oil prices in dollars per barrel are in oil. Fit an ARIMA(p,d,q) model to the growth rate performing ##all necessary diagnostics. Comment.

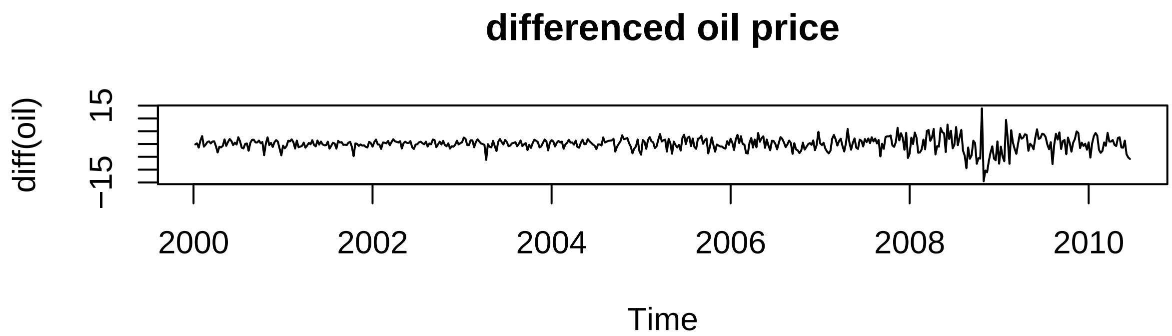
```
par(mfrow=c(2,1))
plot(oil,main="oil prices ber barrel",col='red')
acf(oil)
```

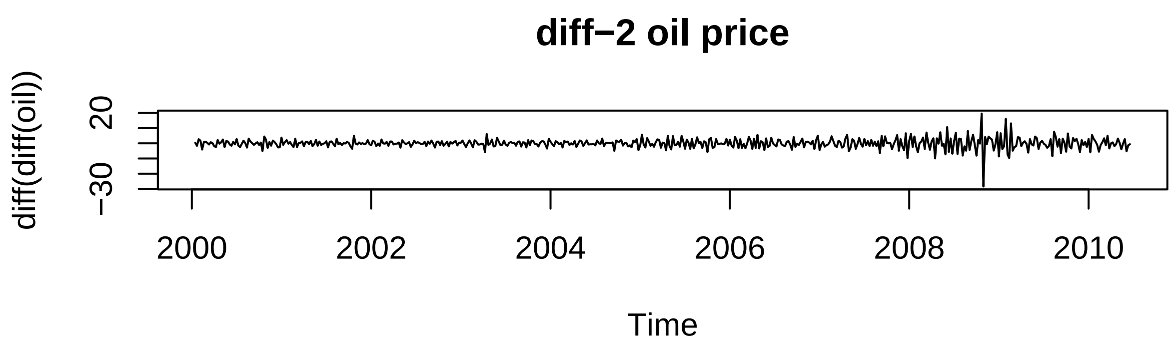




This series is non-stationary and this is evident in the slowly decaying acf.

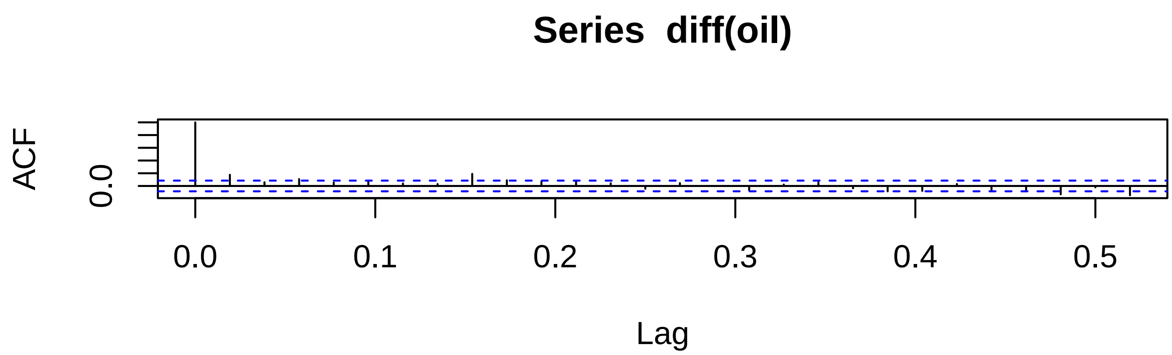
```
#transformations
par(mfrow=c(2,1))
plot(diff(oil),main='differenced oil price')
plot(diff(diff(oil)),main='diff-2 oil price')
```

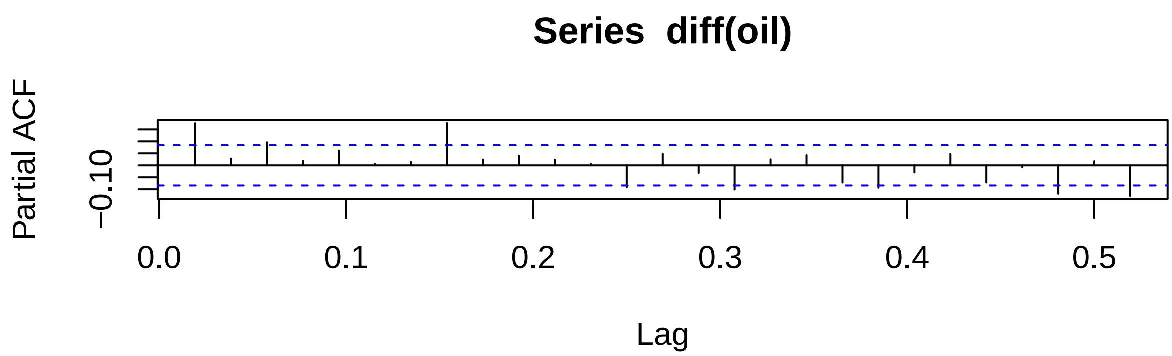




We begin with a diff-1 acf/pacf analysis:

```
par(mfrow=c(2,1))
acf(diff(oil))
pacf(diff(oil))
```



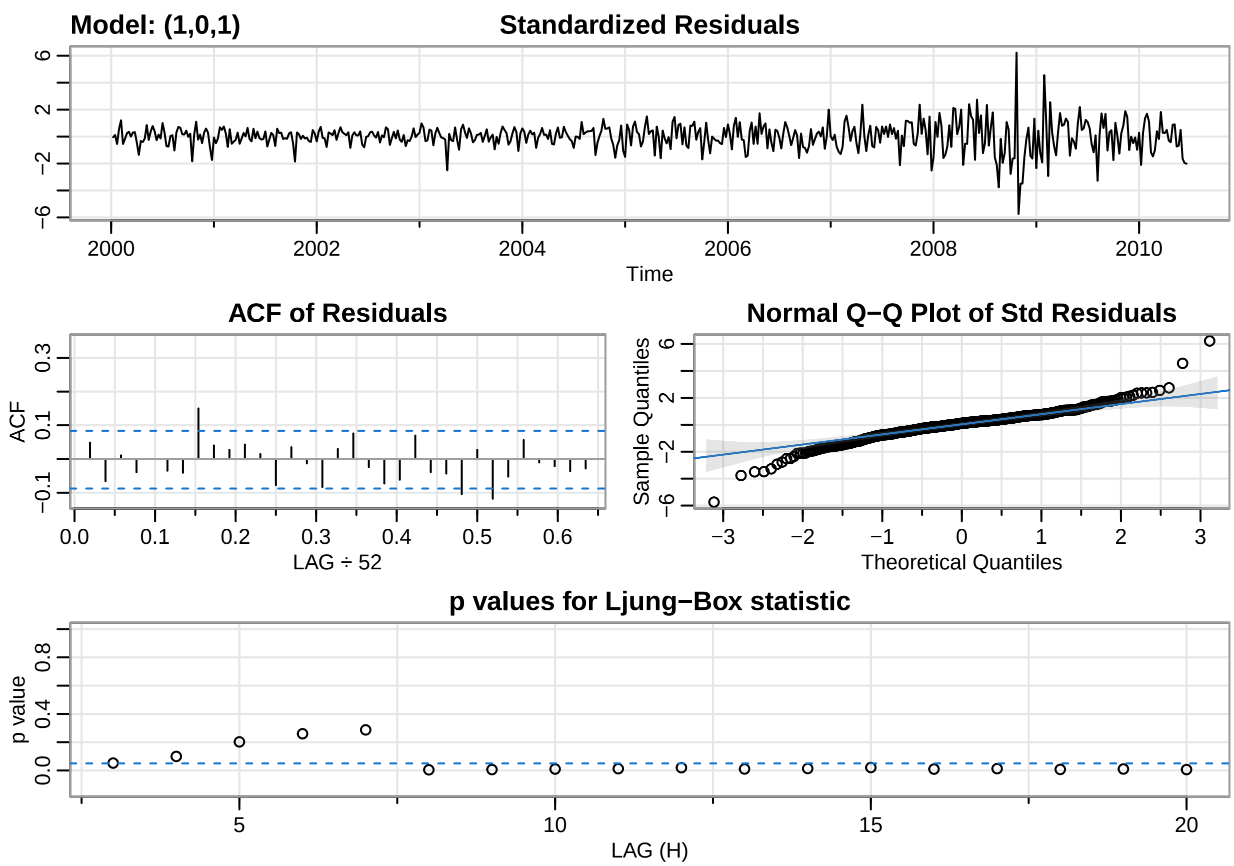


Behavior suggests the possibility of an ARMA(1,1) or an ARMA(2,1) process.

```
#parameter estimation
#iterate over all  possibilities
for (i in 1:3){
  print(sarima(xdata=diff(oil),p=i,d=0,q=1,no.constant=TRUE))
}
```

```
## initial  value 0.953039
## iter   2 value 0.943050
## iter   3 value 0.937488
## iter   4 value 0.937475
## iter   5 value 0.937347
## iter   6 value 0.937056
## iter   7 value 0.936921
## iter   8 value 0.936832
## iter   9 value 0.936814
## iter  10 value 0.936538
## iter  11 value 0.936130
## iter  12 value 0.935719
## iter  13 value 0.935510
## iter  14 value 0.935485
## iter  15 value 0.931710
## iter  16 value 0.931403
## iter  17 value 0.931337
## iter  18 value 0.931115
## iter  19 value 0.931054
```

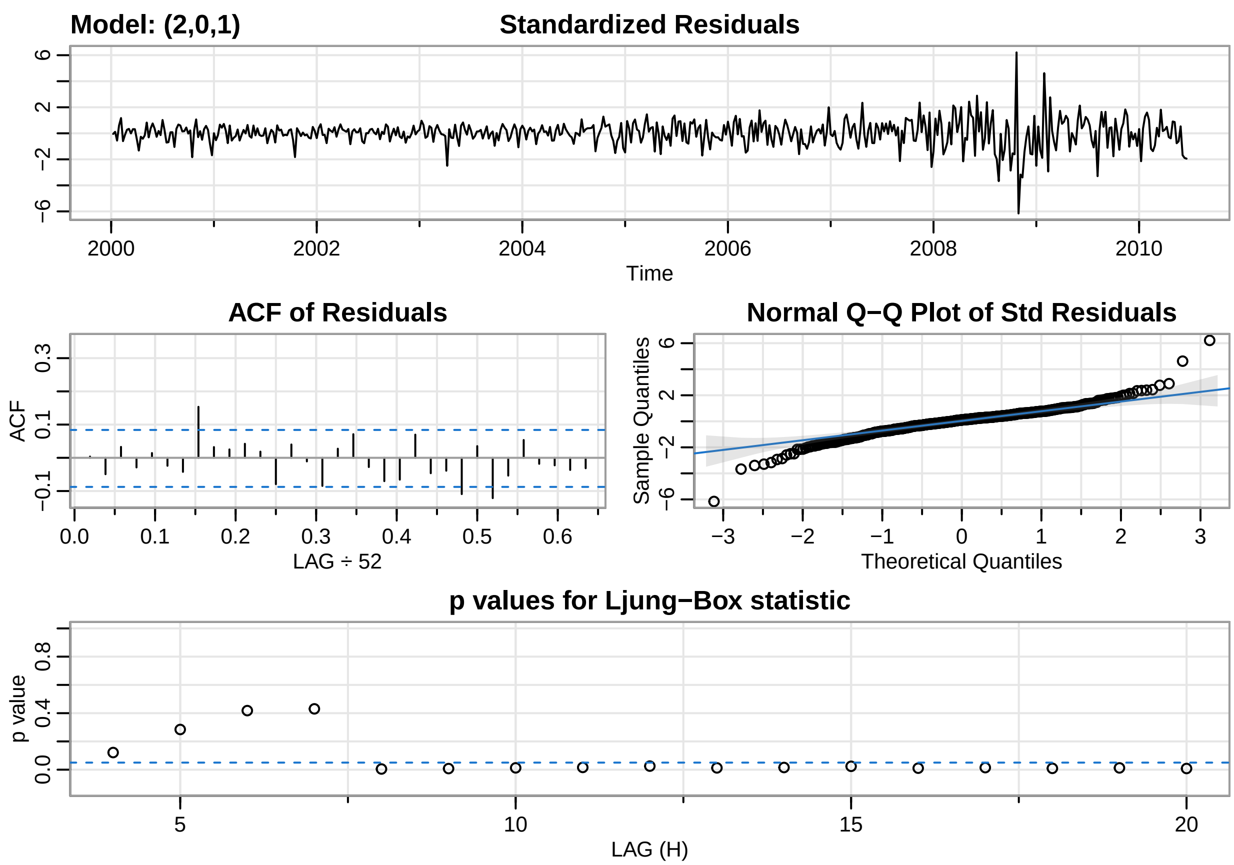
```
## iter  20 value 0.931020
## iter  21 value 0.931019
## iter  22 value 0.931014
## iter  23 value 0.931013
## iter  23 value 0.931013
## iter  23 value 0.931013
## final  value 0.931013
## converged
## initial  value 0.930160
## iter   2 value 0.930159
## iter   3 value 0.930159
## iter   3 value 0.930159
## iter   3 value 0.930159
## final  value 0.930159
## converged
```




```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     xreg = xmean, include.mean = FALSE, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1      ma1
```

```
##       0.8760  -0.7723
## s.e.  0.0585   0.0762
##
## sigma^2 estimated as 6.425:  log likelihood = -1277.91,  aic = 2561.82
##
## $degrees_of_freedom
## [1] 542
##
## $ttable
##     Estimate     SE  t.value p.value
## ar1   0.8760 0.0585  14.9744       0
## ma1  -0.7723 0.0762 -10.1288       0
##
## $AIC
## [1] 4.709225
##
## $AICc
## [1] 4.709265
##
## $BIC
## [1] 4.732932
##
## initial  value 0.953950
## iter   2 value 0.953316
## iter   3 value 0.937850
## iter   4 value 0.937711
## iter   5 value 0.937707
## iter   6 value 0.937703
## iter   7 value 0.937685
## iter   8 value 0.937639
## iter   9 value 0.937631
## iter  10 value 0.937587
## iter  11 value 0.937584
## iter  12 value 0.937330
## iter  13 value 0.937087
## iter  14 value 0.936729
## iter  15 value 0.936582
## iter  16 value 0.935104
## iter  17 value 0.934732
## iter  18 value 0.934670
## iter  19 value 0.934063
## iter  20 value 0.933458
## iter  21 value 0.933074
## iter  22 value 0.932855
## iter  23 value 0.932487
## iter  24 value 0.932217
## iter  25 value 0.931335
## iter  26 value 0.931104
## iter  27 value 0.930521
## iter  28 value 0.930215
## iter  29 value 0.930047
## iter  30 value 0.929939
## iter  31 value 0.929923
## iter  32 value 0.929900
```

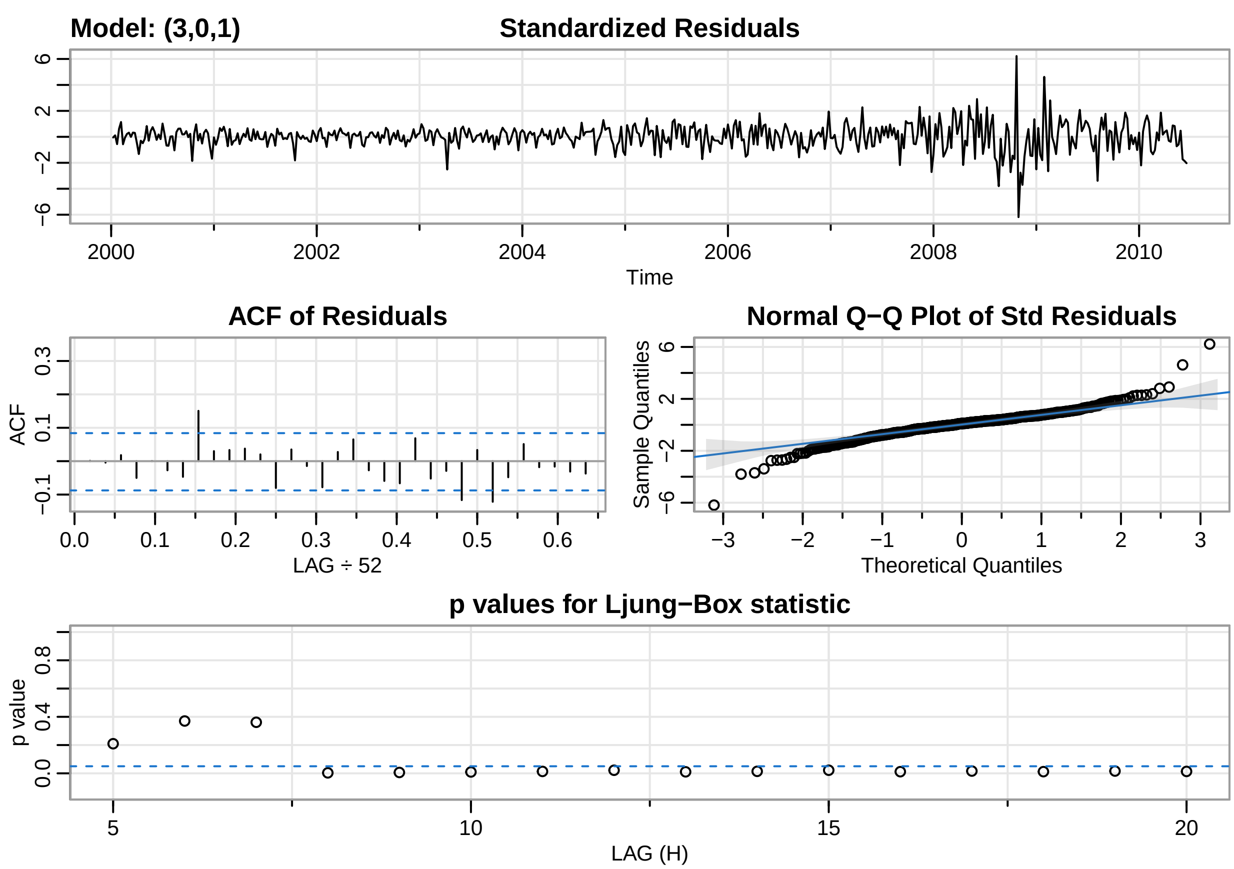
```
## iter  33 value 0.929899
## iter  34 value 0.929899
## iter  35 value 0.929899
## iter  35 value 0.929899
## iter  35 value 0.929899
## final  value 0.929899
## converged
## initial  value 0.928214
## iter   2 value 0.928214
## iter   3 value 0.928213
## iter   4 value 0.928213
## iter   5 value 0.928213
## iter   6 value 0.928213
## iter   7 value 0.928213
## iter   7 value 0.928213
## iter   7 value 0.928213
## final  value 0.928213
## converged
```




```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     xreg = xmean, include.mean = FALSE, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
```

```
##
## Coefficients:
##          ar1      ar2      ma1
##       0.9718  -0.0702  -0.8195
## s.e.  0.0782   0.0485   0.0646
##
## sigma^2 estimated as 6.4:  log likelihood = -1276.85,  aic = 2561.7
##
## $degrees_of_freedom
## [1] 541
##
## $ttable
##     Estimate     SE  t.value p.value
## ar1   0.9718 0.0782  12.4346  0.0000
## ar2  -0.0702 0.0485  -1.4459  0.1488
## ma1  -0.8195 0.0646 -12.6877  0.0000
##
## $AIC
## [1] 4.709009
##
## $AICc
## [1] 4.709091
##
## $BIC
## [1] 4.740619
##
## initial  value 0.954609
## iter   2 value 0.949897
## iter   3 value 0.933641
## iter   4 value 0.933315
## iter   5 value 0.933304
## iter   6 value 0.933303
## iter   7 value 0.933286
## iter   8 value 0.933256
## iter   9 value 0.933173
## iter  10 value 0.933018
## iter  11 value 0.932869
## iter  12 value 0.932755
## iter  13 value 0.932747
## iter  14 value 0.932745
## iter  15 value 0.932738
## iter  16 value 0.932718
## iter  17 value 0.932657
## iter  18 value 0.932381
## iter  19 value 0.932323
## iter  20 value 0.932192
## iter  21 value 0.932074
## iter  22 value 0.932019
## iter  23 value 0.931989
## iter  24 value 0.931913
## iter  25 value 0.931631
## iter  26 value 0.931286
## iter  27 value 0.931190
## iter  28 value 0.930873
```

```
## iter  29 value 0.930644
## iter  30 value 0.930610
## iter  31 value 0.930386
## iter  32 value 0.930367
## iter  33 value 0.930360
## iter  34 value 0.930352
## iter  35 value 0.930347
## iter  36 value 0.930346
## iter  37 value 0.930346
## iter  37 value 0.930346
## final  value 0.930346
## converged
## initial  value 0.926600
## iter   2 value 0.926594
## iter   3 value 0.926591
## iter   4 value 0.926589
## iter   5 value 0.926577
## iter   6 value 0.926569
## iter   7 value 0.926564
## iter   8 value 0.926563
## iter   8 value 0.926563
## iter   8 value 0.926563
## final  value 0.926563
## converged
```




```
## $fit
```

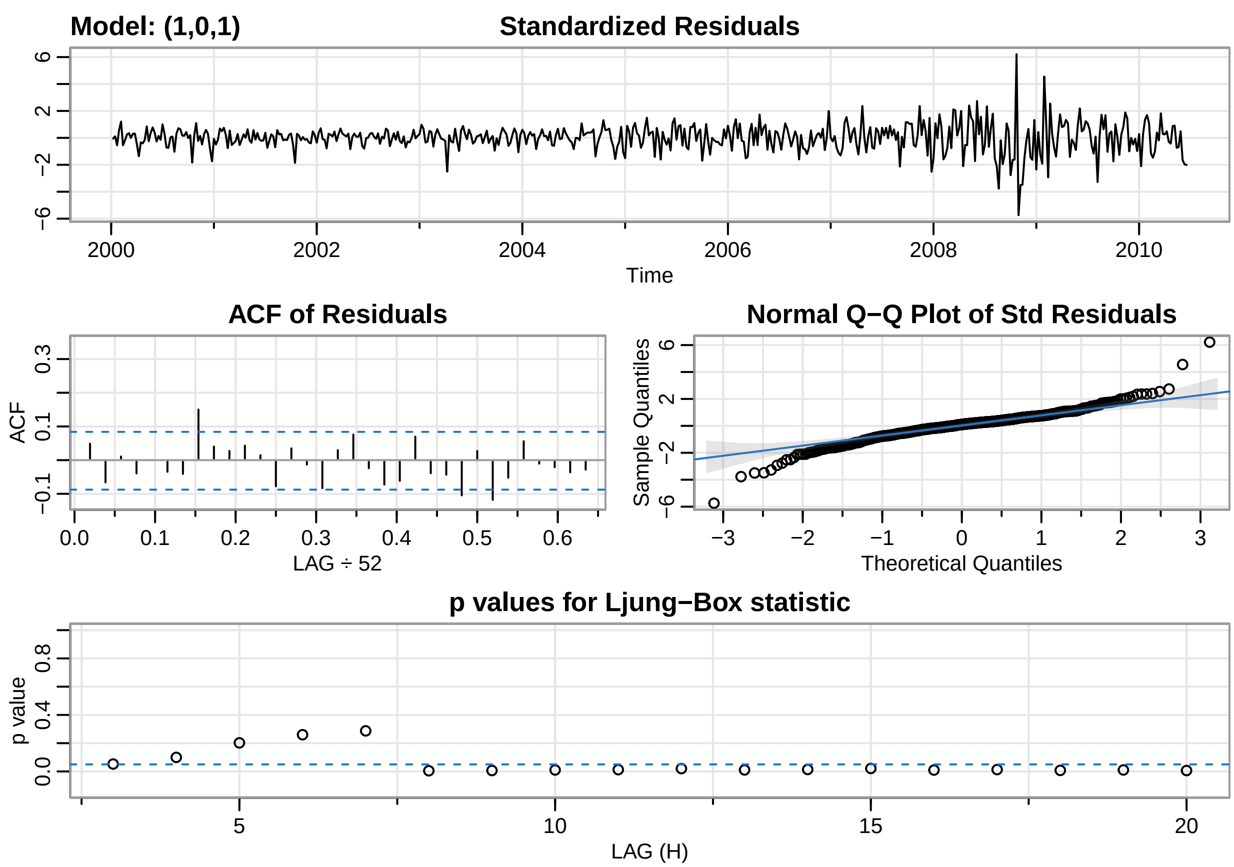
```
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     xreg = xmean, include.mean = FALSE, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1      ar2     ar3      ma1
##       0.9368  -0.1160  0.0620  -0.7811
## s.e.  0.0908   0.0597  0.0458   0.0813
##
## sigma^2 estimated as 6.378:  log likelihood = -1275.95,  aic = 2561.91
##
## $degrees_of_freedom
## [1] 540
##
## $ttable
##     Estimate     SE t.value p.value
## ar1   0.9368 0.0908 10.3200  0.0000
## ar2  -0.1160 0.0597 -1.9432  0.0525
## ar3   0.0620 0.0458  1.3547  0.1761
## ma1  -0.7811 0.0813 -9.6072  0.0000
##
## $AIC
## [1] 4.709385
##
## $AICc
## [1] 4.709522
##
## $BIC
## [1] 4.748898
```

Results suggest that an ar2 and ar3 component are not significant, so we retain the AR-1 portion and can move on to tuning for MA.

```
for (i in 1:2){
  print(sarima(xdata=diff(oil),p=1,d=0,q=i,no.constant=TRUE))
}
```

```
## initial  value 0.953039
## iter   2 value 0.943050
## iter   3 value 0.937488
## iter   4 value 0.937475
## iter   5 value 0.937347
## iter   6 value 0.937056
## iter   7 value 0.936921
## iter   8 value 0.936832
## iter   9 value 0.936814
## iter  10 value 0.936538
## iter  11 value 0.936130
## iter  12 value 0.935719
## iter  13 value 0.935510
## iter  14 value 0.935485
```

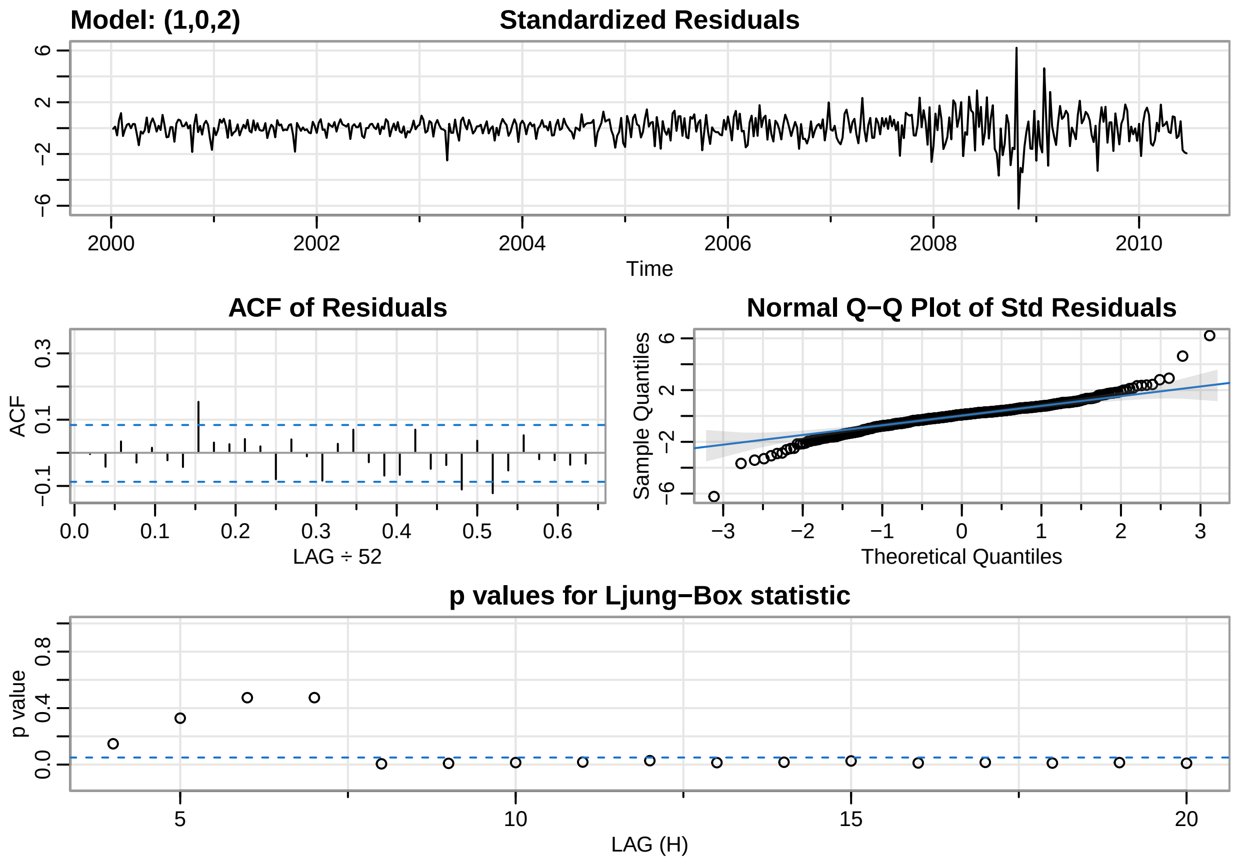
```
## iter  15 value 0.931710
## iter  16 value 0.931403
## iter  17 value 0.931337
## iter  18 value 0.931115
## iter  19 value 0.931054
## iter  20 value 0.931020
## iter  21 value 0.931019
## iter  22 value 0.931014
## iter  23 value 0.931013
## iter  23 value 0.931013
## iter  23 value 0.931013
## final  value 0.931013
## converged
## initial  value 0.930160
## iter   2 value 0.930159
## iter   3 value 0.930159
## iter   3 value 0.930159
## iter   3 value 0.930159
## final  value 0.930159
## converged
```




```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
```

```
##     xreg = xmean, include.mean = FALSE, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1      ma1
##       0.8760  -0.7723
## s.e.  0.0585   0.0762
##
## sigma^2 estimated as 6.425:  log likelihood = -1277.91,  aic = 2561.82
##
## $degrees_of_freedom
## [1] 542
##
## $ttable
##     Estimate     SE  t.value p.value
## ar1   0.8760 0.0585  14.9744       0
## ma1  -0.7723 0.0762 -10.1288       0
##
## $AIC
## [1] 4.709225
##
## $AICc
## [1] 4.709265
##
## $BIC
## [1] 4.732932
##
## initial  value 0.953039
## iter   2 value 0.942654
## iter   3 value 0.937328
## iter   4 value 0.937291
## iter   5 value 0.937284
## iter   6 value 0.937227
## iter   7 value 0.937095
## iter   8 value 0.936622
## iter   9 value 0.936019
## iter  10 value 0.935373
## iter  11 value 0.935340
## iter  12 value 0.934242
## iter  13 value 0.934165
## iter  14 value 0.933967
## iter  15 value 0.933865
## iter  16 value 0.932545
## iter  17 value 0.932350
## iter  18 value 0.932158
## iter  19 value 0.930898
## iter  20 value 0.930088
## iter  21 value 0.929661
## iter  22 value 0.929225
## iter  23 value 0.928792
## iter  24 value 0.928784
## iter  25 value 0.928768
## iter  26 value 0.928761
## iter  27 value 0.928761
```

```
## iter  27 value 0.928761
## iter  27 value 0.928761
## final  value 0.928761
## converged
## initial  value 0.927894
## iter   2 value 0.927893
## iter   3 value 0.927893
## iter   4 value 0.927893
## iter   5 value 0.927893
## iter   6 value 0.927893
## iter   7 value 0.927893
## iter   7 value 0.927893
## iter   7 value 0.927893
## final  value 0.927893
## converged
```



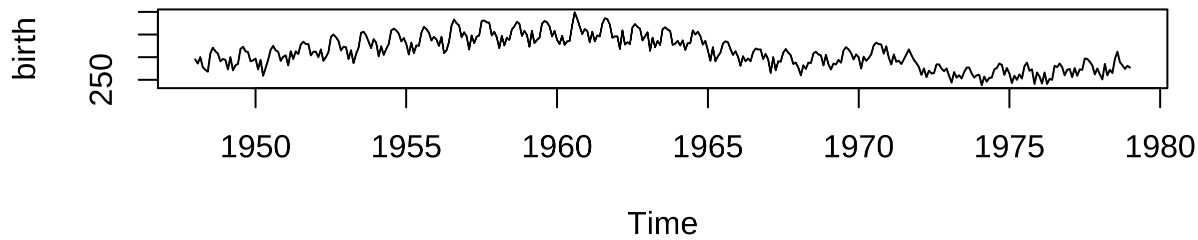

```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     xreg = xmean, include.mean = FALSE, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1      ma1      ma2
```

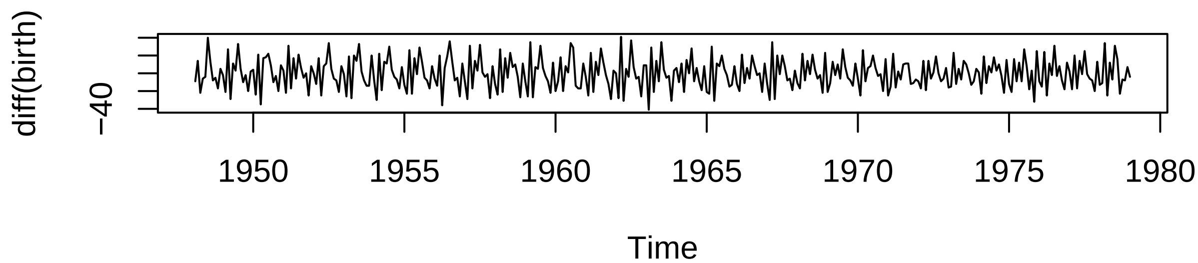
```
##       0.8943  -0.7337  -0.0724
## s.e.  0.0483   0.0658   0.0465
##
## sigma^2 estimated as 6.396:  log likelihood = -1276.68,  aic = 2561.35
##
## $degrees_of_freedom
## [1] 541
##
## $ttable
##     Estimate     SE  t.value p.value
## ar1   0.8943 0.0483  18.5204  0.0000
## ma1  -0.7337 0.0658 -11.1476  0.0000
## ma2  -0.0724 0.0465  -1.5567  0.1201
##
## $AIC
## [1] 4.70837
##
## $AICc
## [1] 4.708451
##
## $BIC
## [1] 4.739979
```

Results also confirm that an MA-2 component is not significant, so we proceed with an ARMA(1,1) model for the difference in oil prices, or an ARIMA(1,1,1) for the oil prices originally used.

#Example 5.11 [8 points] ##Fit a seasonal ARIMA model of your choice to the U.S. Live Birth Series, birth. Use the estimated model to forecast the next 12 months.

```r
par(mfrow=c(2,1))
plot(birth)
plot(diff(birth))
```

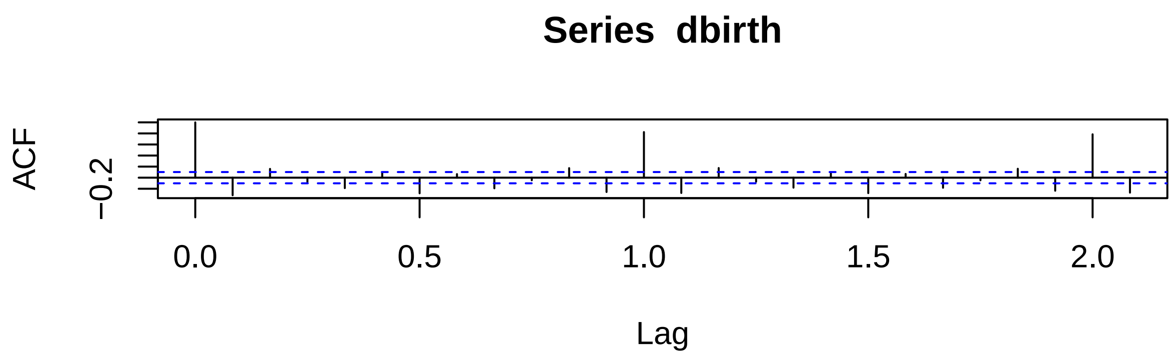


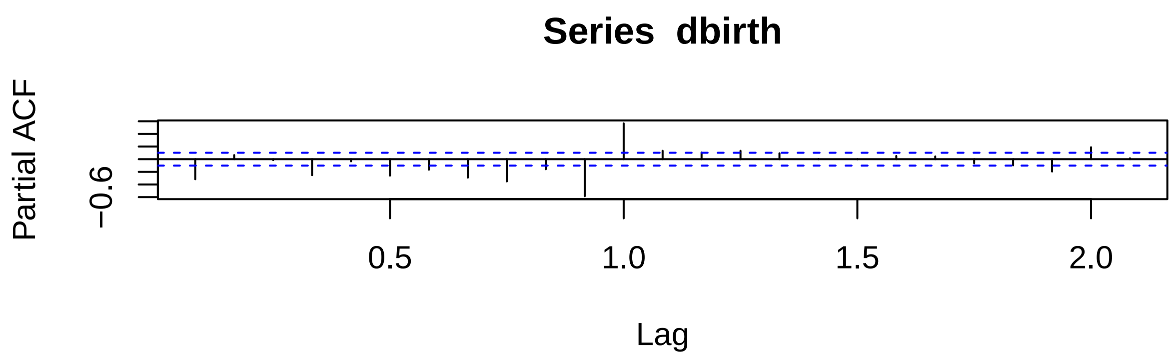



```
dbirth = diff(birth)
```

Differencing helps us remove the trend for the data, which leaves us with a very strong seasonal component.

```
par(mfrow=c(2,1))
acf(dbirth)
pacf(dbirth)
```





Three strong spikes in the ACF suggest an MA-3, however, the 12-lag difference between the prominent spikes in the ACF also suggests that this could be a seasonal MA(1) combined with an MA(1)/(2) process. However, one could make the argument that the PACF exhibits seasonality as well, so we may consider a SAR(1) component.

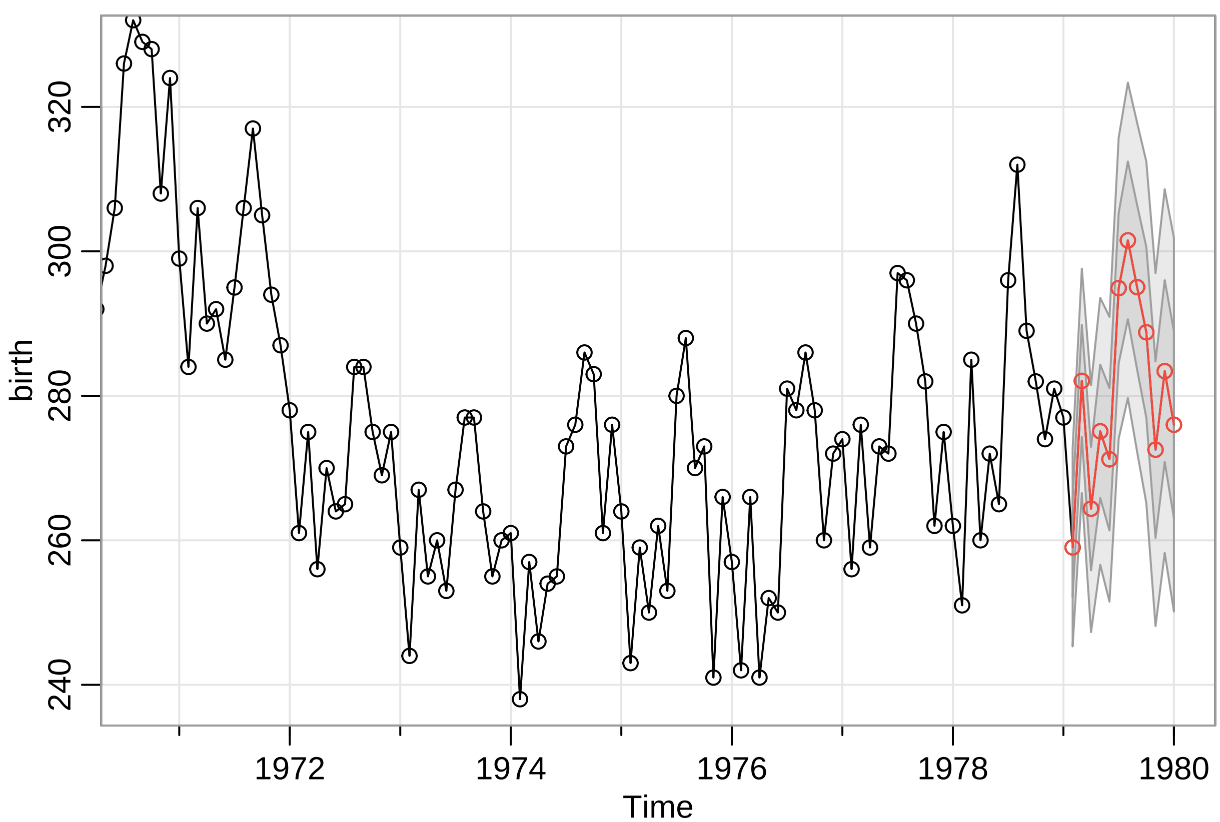
```
for (i in 1:2){
  for(j in 0:1){
    this.arima = arima(dbirth,order=c(i,0,i),seasonal=c(j,0,1),include.mean = FALSE)
    print(this.arima$coef)
    print(this.arima$aic)
    print(mean(this.arima$residuals))
    print(Box.test(this.arima$residuals)$p.value)
    #print(rmse(fitted(this.arima),dbirth))
    print("*****END OF MODEL*****")
  }
}
```

```
##        ar1        ma1       sma1
## -0.3459155  0.0364264  0.6026217
## [1] 2920.34
## [1] -0.00500309
## [1] 0.9804725
## [1] "*****END OF MODEL*****"
##        ar1        ma1       sar1       sma1
##  0.3145299 -0.7095535  0.9952291 -0.7922652
## [1] 2519.296
## [1] -0.03969498
```

```
## [1] 0.862053
## [1] "*****END OF MODEL*****"
##         ar1        ar2        ma1        ma2       sma1
##  0.1587327  0.3250529 -0.6047559 -0.3014571  0.6059992
## [1] 2885.804
## [1] -0.1109419
## [1] 0.6668825
## [1] "*****END OF MODEL*****"
##         ar1        ar2        ma1        ma2       sar1       sma1
##  1.2690218 -0.3850170 -1.6582345  0.7196604  0.9959411 -0.8070206
## [1] 2520.234
## [1] -0.03384715
## [1] 0.9818574
## [1] "*****END OF MODEL*****"
```

based on the AIC, parsimony, and the RMSE, we would consider an ARMA(1,1) * SARMA(1,1) model.

```
#forecast
sarima.for(xdata=birth,n.ahead=12, p = 1, q=1, d= 0, P=1, Q = 1,D=0,S=12)
```




```
## $pred
##           Jan      Feb      Mar      Apr      May      Jun      Jul      Aug
## 1979          259.0198 282.0706 264.3860 275.0874 271.2173 294.9100 301.5226
## 1980 276.0030
##           Sep      Oct      Nov      Dec
```

```
## 1979 295.0547 288.7981 272.5489 283.4228
## 1980
##
## $se
##            Jan       Feb       Mar       Apr       May       Jun       Jul
## 1979           6.853381  7.764838  8.548591  9.238379  9.855014 10.412463
## 1980 12.932289
##            Aug       Sep       Oct       Nov       Dec
## 1979 10.920674 11.387051 11.817296 12.215924 12.586585
## 1980
```