**ARIMA model building – Example**

The goal of this section is to show how one can go through the process of choosing the best model for a time series data set.

Example: ARIMA(1,1,1) with ϕ1 = 0.7, θ1 = 0.4,  = 9, n=200 (arima111\_model\_build.R, arima111.csv, examine.mod.R)

1. Construct plots of xt vs. t and the estimated ACF to determine if the time series data is stationary

> arima111 <- read.csv(file = "arima111.csv")

> head(arima111)

time x

1 1 -143.2118

2 2 -142.8908

3 3 -138.0634

4 4 -133.5038

5 5 -132.7496

6 6 -132.2910

> x <- arima111$x

> #########################################################

> # Step #1 and #2

> #Plot of the data

> par(mfrow = c(1,1))

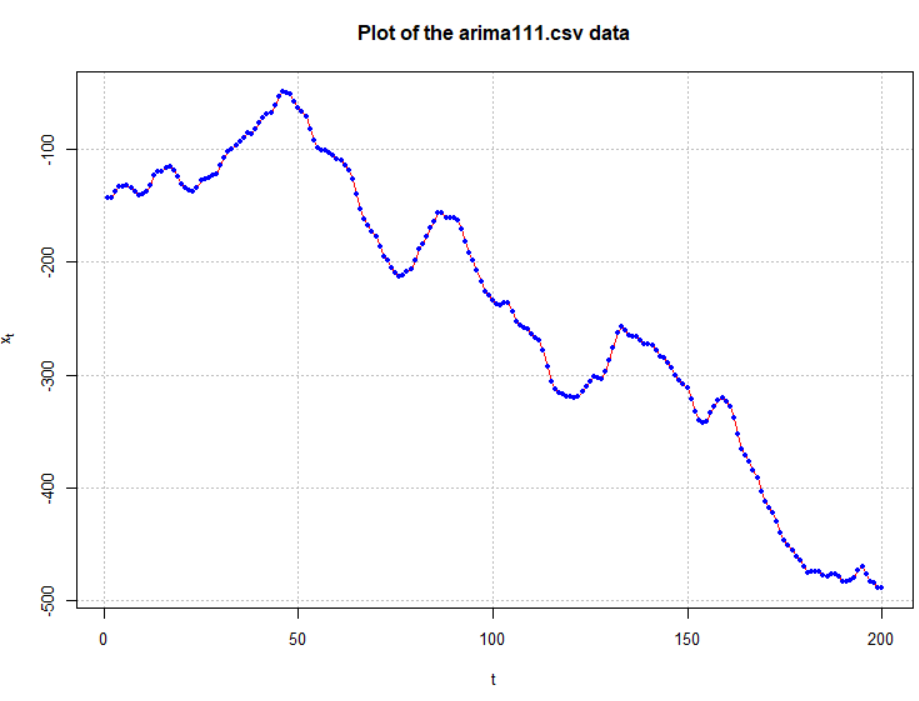
> plot(x = x, ylab = expression(x[t]), xlab = "t", type =

"l", col = "red", main = "Plot of the

arima111.csv data", panel.first = grid(col = "gray",

lty = "dotted"))

> points(x = x, pch = 20, col = "blue")



> #ACF and PACF of x\_t

> par(mfrow = c(1,2))

> acf(x = x, type = "correlation", lag.max = 20, xlim =

c(1,20), ylim = c(-1,1), main = "Estimated ACF of the

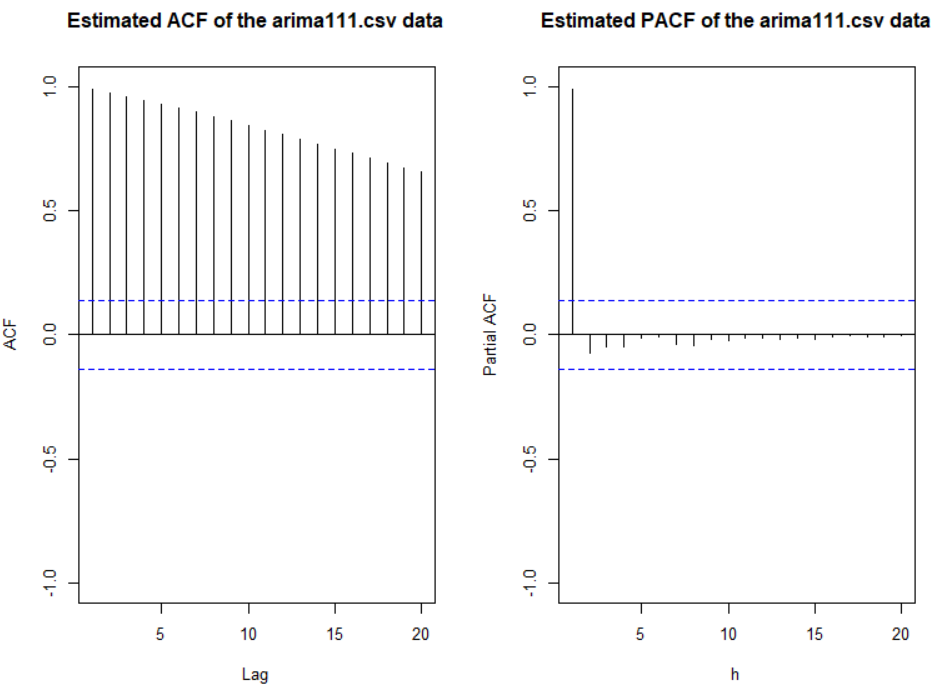
arima111.csv data")

> pacf(x = x, lag.max = 20, xlim = c(1,20), ylim =

c(-1,1), xlab = "h", main = "Estimated PACF of the

arima111.csv data")

> par(mfrow = c(1,1))



The data appears to be nonstationary in the mean because there is a linear trend in the xt vs. t plot and the autocorrelations are going to 0 very slowly. Below is what happens after the first differences are taken.

> #Examine the first differences

> plot(x = diff(x = x, lag = 1, differences = 1), ylab =

expression(x[t]-x[t-1]), xlab = "t", type = "l", col

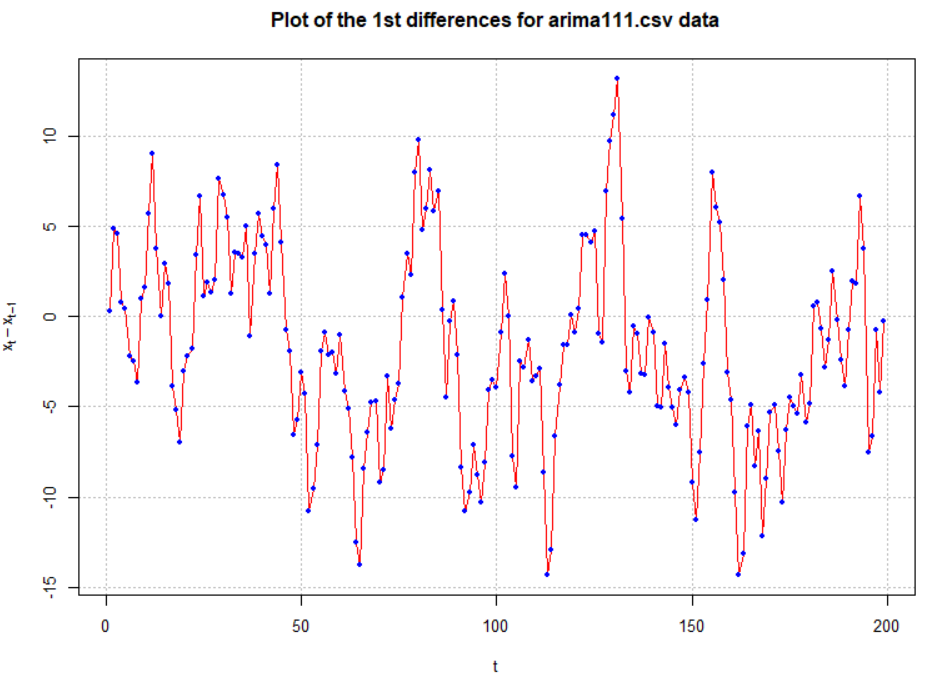
= "red", main = "Plot of the 1st differences for

arima111.csv data", panel.first = grid(col = "gray",

lty = "dotted"))

> points(x = diff(x = x, lag = 1, differences = 1), pch =

20, col = "blue")



> #ACF and PACF of x\_t - x\_t-1

> par(mfrow = c(1,2))

> acf(x = diff(x = x, lag = 1, differences = 1), type =

"correlation", lag.max = 20, xlim = c(1,20), ylim =

c(-1,1), main = "Est. ACF 1st diff. for arima111.csv

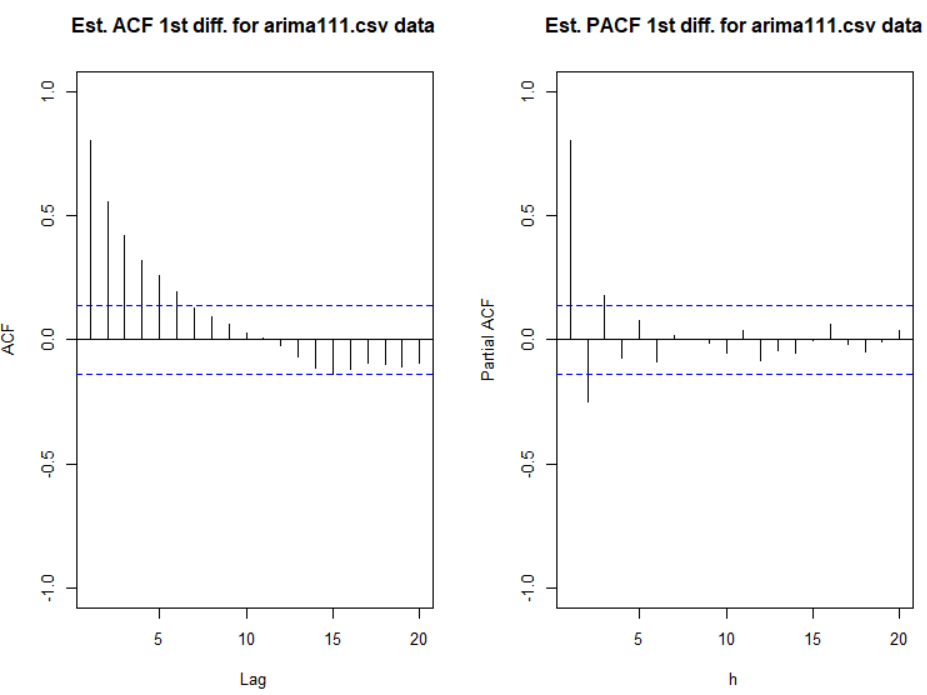
data")

> pacf(x = diff(x = x, lag = 1, differences = 1), lag.max

= 20, xlim = c(1,20), ylim = c(-1,1), xlab = "h",

main = "Est. PACF 1st diff. for arima111.csv data")

> par(mfrow = c(1,1))



The data now appears to be stationary in the mean. Also, there appears to be no evidence against constant variance.

1. Construct plots of the estimated ACF and PACF of the stationary series.

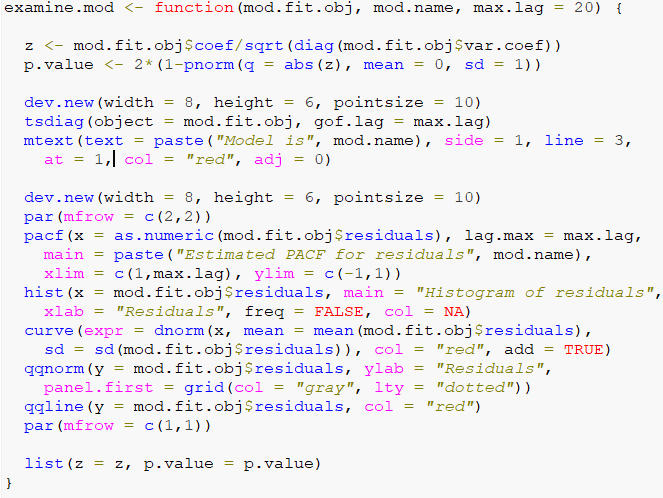
The plots are above. The ACF is tailing off toward 0 after q = 1 lags. The PACF may also be tailing off toward 0 after p = 1 lags. This would be indicative of an ARMA(1,1,1) model.

Other possible models include:

* ARIMA(2,1,0) or ARIMA(3,1,0) because the PACF possibly cuts off to 0 after lag 2 or 3.
* Try simple models such as ARIMA(1,1,0) or ARIMA(0,1,1) to see if ACF and PACF plots of the model residuals help determine what types of changes should be made to the model. This is often a very good strategy to follow when it is not obvious from an ACF or PACF what the model should be.

1. Find the estimated models using maximum likelihood estimation AND
2. For each model chosen, investigate diagnostic measures.

To help perform #4 correctly, I wrote a function called examine.mod()which automatically does most of R parts needed. The function puts together the code that was used in the AR(1) data example. You will need to run the entire function first before using it! Below is a screen capture of the function as it appears in a programming editor.



ARIMA(1,1,1):

> source("examine.mod.R")

> mod.fit111 <- arima(x = x, order = c(1, 1, 1))

> mod.fit111

Call:

arima(x = x, order = c(1, 1, 1))

Coefficients:

ar1 ma1

0.6720 0.4681

s.e. 0.0637 0.0904

sigma^2 estimated as 9.558: log likelihood = -507.68, aic = 1021.36

> save.it <- examine.mod(mod.fit.obj = mod.fit111, mod.name

= "ARIMA(1,1,1)", max.lag = 20)

> save.it

$z

ar1 ma1

10.545076 5.177294

$p.value

ar1 ma1

0.000000e+00 2.251268e-07





ARIMA(2,1,0):

> #ARIMA(2,1,0)

> mod.fit210 <- arima(x = x, order = c(2, 1, 0))

> mod.fit210

Call:

arima(x = x, order = c(2, 1, 0))

Coefficients:

ar1 ar2

1.0137 -0.2371

s.e. 0.0690 0.0691

sigma^2 estimated as 9.97: log likelihood = -511.8, aic = 1029.59

> # Need to run function in examine.mod.R file first

> examine.mod(mod.fit.obj = mod.fit210, mod.name

= "ARIMA(2,1,0)", max.lag = 20)

$z

ar1 ar2

14.701106 -3.431004

$p.value

ar1 ar2

0.0000000000 0.0006013512

ARIMA(3,1,0):

> #ARIMA(3,1,0)

> mod.fit310 <- arima(x = x, order = c(3, 1, 0))

> mod.fit310

Call:

arima(x = x, order = c(3, 1, 0))

Coefficients:

ar1 ar2 ar3

1.0627 -0.4436 0.2010

s.e. 0.0697 0.0985 0.0697

sigma^2 estimated as 9.565: log likelihood = -507.73, aic = 1023.46

> examine.mod(mod.fit.obj = mod.fit310, mod.name =

"ARIMA(3,1,0)", max.lag = 20)

$z

ar1 ar2 ar3

15.251131 -4.501187 2.883806

$p.value

ar1 ar2 ar3

0.000000e+00 6.757494e-06 3.929004e-03



Model notes:

* ARIMA(1,1,1)
  + - Ljung-Box-Pierce test does not indicate any significant groups of autocorrelations.
    - ACF and PACF plots do not exhibit any significant values.
    - There are a few standardized residuals outside of the -1.96 and 1.96 range, but still < 3.
    - The normality assumption appears to be satisfied because the normal distribution appears to fit the histogram of residuals well.
    - The hypothesis test for ϕ1=0 and θ1=0 results in

|  |  |  |
| --- | --- | --- |
|  | Test statistic | P-value |
| ϕ1 | 10.54 | <0.0001 |
| θ1 | 5.18 | 2.3×10-7 |

Because the p-values are small, there is sufficient evidence to indicate that ϕ1 ≠ 0 and θ1 ≠ 0.

* ARIMA(2,1,0)
  + - Ljung-Box-Pierce test indicates marginally significant groups of autocorrelations for H = 2 to 6.
    - The ACF plot has a marginally significant value at lag 2. The PACF plot has a marginally significant value at lag 2. This indicates that an additional AR or MA term may need to be added to the model.
    - There are a few standardized residuals outside of the -1.96 and 1.96 range with one close to -3.
    - The normality assumption appears to be satisfied because the normal distribution appears to fit the histogram of residuals well.
    - The test for ϕi = 0 results in

|  |  |  |
| --- | --- | --- |
|  | Test statistic | P-value |
| ϕ1 | 14.70 | <0.0001 |
| ϕ2 | -3.43 | 0.0006 |

Because the p-values are small, there is sufficient evidence to indicate that ϕ1 ≠ 0 and ϕ2 ≠ 0.

* ARIMA(3,1,0)
  + - Ljung-Box-Pierce test does not indicate any significant groups of autocorrelations.
    - ACF and PACF plots do not exhibit any significant values.
    - There are a few standardized residuals outside of the -1.96 and 1.96 range with two close to -3.
    - The normality assumption appears to be satisfied because the normal distribution appears to fit the histogram of residuals well.
    - The test for ϕi = 0 results in

|  |  |  |
| --- | --- | --- |
|  | Test statistic | P-value |
| ϕ1 | 15.25 | <0.0001 |
| ϕ2 | -4.50 | 6.8×10-6 |
| ϕ3 | 2.88 | 0.0039 |

Because the p-values are small, there is sufficient evidence to indicate that ϕ1≠0, ϕ2≠0, and ϕ3≠0.

Summary: Both the ARIMA(1,1,1) and ARIMA(3,1,0) models appear to be better than the ARIMA(2,1,0) model. The ARIMA(1,1,1) model may be a “little” better than the ARIMA(3,1,0) model due to the examination of the standardized residuals.

Suppose I tried an ARIMA(1,1,0) model with the hope that the residual ACF and PACF plots help suggest changes to the model. Below are the plots and other summary measures.

> #ARIMA(1,1,0)

> mod.fit110 <- arima(x = x, order = c(1, 1, 0))

> mod.fit110

all:

arima(x = x, order = c(1, 1, 0))

Coefficients:

ar1

0.8180

s.e. 0.0398

sigma^2 estimated as 10.57: log likelihood = -517.51, aic = 1039.02

> examine.mod(mod.fit.obj = mod.fit110, mod.name =

"ARIMA(1,1,0)", max.lag = 20)

$z

ar1

20.52674

$p.value

ar1

0





Both the ACF and PACF plots have possibly significant values at lag = 1 and 2. The problem is knowing if these are signs of cutting off or tailing off and where does the tailing off start. If the PACF was tailing off to 0 starting at lag 1, the significant ACF value(s) would indicate correctly that a MA term was needed.

Suppose I tried an ARIMA(0,1,1) model with the hope that the residual ACF and PACF plots help suggest changes to the model. Below are the plots and other summary measures.

> #ARIMA(0,1,1)

> mod.fit011 <- arima(x = x, order = c(0, 1, 1))

> mod.fit011

Call:

arima(x = x, order = c(0, 1, 1))

Coefficients:

ma1

0.8371

s.e. 0.0321

sigma^2 estimated as 13.6: log likelihood = -542.67, aic = 1089.35

> examine.mod(mod.fit.obj = mod.fit011, mod.name =

"ARIMA(0,1,1)", max.lag = 20)

$z

ma1

26.05175

$p.value

ma1

0





Notice the pattern among the standardized residuals indicating dependence still exists among them. There are significant partial autocorrelations at lag 1 and 2 and they are non-significant after lag 2. The ACF appears to be tailing off to 0 (maybe, although lag 2 is greater than lag 1). This suggests that AR terms may need to be added to the model. If ϕ1 is added, the results shown for ARIMA(1,1,1) earlier are found. Normally, I would add only one term at a time. However, given the PACF here, it may be reasonable to add both ϕ1 and ϕ2 to the model:

> mod.fit211 <- arima(x = x, order = c(2, 1, 1))

> mod.fit211

Call:

arima(x = x, order = c(2, 1, 1))

Coefficients:

ar1 ar2 ma1

0.3636 0.2926 0.7394

s.e. 0.1431 0.1329 0.1107

sigma^2 estimated as 9.365: log likelihood = -505.69, aic = 1019.38

> examine.mod(mod.fit.obj = mod.fit211, mod.name =

"ARIMA(2,1,1)", max.lag = 20)

$z

ar1 ar2 ma1

2.541125 2.202128 6.676700

$p.value

ar1 ar2 ma1

1.104964e-02 2.765624e-02 2.443823e-11



The tests for ϕ1 and ϕ2 = 0 result in somewhat marginally significant p-values. Thus, it is difficult to decide whether or not to include both in the model. Generally, one will build the models in a hierarchical manner so that if ϕ2 is in the model then ϕ1 will also be in the model. The main time when this generality is not done is in the case of seasonal dependence (for example, data is collected quarterly) and models for that situation will be discussed later in the course.

Note that the Ljung-Box-Pierce test results in no reject Ho’s.

1. Using each model that gets through 4), pick the best model based upon model parsimony and the AIC. The best model corresponds to the one with the smallest number of AR and MA terms and the smallest AIC.

* Akaike’s information criterion (AIC)

Although all of the models in the table below did not satisfy the model assumptions, I decided to present all of the AIC values for illustrative purposes. Normally, I would have only examined the AIC for the ARIMA(1,1,1), ARIMA(3,1,0), and ARIMA(2,1,1) models.

> data.frame(mod.name = c("ARIMA(1,1,1)", "ARIMA(2,1,0)",

"ARIMA(3,1,0)", "ARIMA(1,1,0)", "ARIMA(0,1,1)",

"ARIMA(2,1,1)"),

AIC = c(mod.fit111$aic, mod.fit210$aic, mod.fit310$aic,

mod.fit110$aic, mod.fit011$aic, mod.fit211$aic))

mod.name AIC

1 ARIMA(1,1,1) 1021.360

2 ARIMA(2,1,0) 1029.591

3 ARIMA(3,1,0) 1023.459

4 ARIMA(1,1,0) 1039.015

5 ARIMA(0,1,1) 1089.346

6 ARIMA(2,1,1) 1019.383

Notice that the ARIMA(2,1,1) model has the lowest AIC! This model appeared to satisfy all of the model’s assumptions so this is the best model to use (according to AIC).

The ARIMA(1,1,1) model also has a low AIC. Because the data was simulated from this type of model, one would expect this AIC to be the lowest! Generally, this should happen.

There is some justification for using the ARIMA(1,1,1) model over the ARIMA(2,1,1) model. Because the ARIMA(1,1,1) has less parameters and the model assumptions were satisfied, this may be the better model to use. Also, remember the marginally significant result for the ϕ2 = 0 test.

Which model should you choose? There is sufficient justification for both models. One could also examine the MSE of forecasts for a number of the last observations to help make a decision as well.

Once you choose the model, state it as “ARIMA(p,d,q)” and written out with the estimated parameters.

What if included the δ term in the model?

I tried this with a few of the best models found above and received very similar results. The δ term is marginally significant. Please see the program for the actual code used to investigate the addition of δ term.

Note that the ARIMA(2,1,1) model with the δ term has the lowest AIC with a value of 1018.78! The p-value for the δ = 0 test is 0.0881.

1. Begin forecasting!

Model building process summarized

1. Construct plots of xt vs. t and the estimated ACF to determine if the time series data is stationary
   1. If it is not stationary in the variance, make the appropriate transformation.
   2. If it is not stationary in the mean, examine differences of xt.
   3. Use plots of xt vs. t and the estimated ACF to determine if what you did worked.
2. Construct plots of the estimated ACF and PACF of the stationary series (call it xt).
   1. Match patterns in these plots with those of ARMA models.
   2. Determine a few models to investigate further.
3. Find the estimated models using maximum likelihood estimation.
4. For each model chosen, investigate the diagnostic measures.
   1. Construct plots of the estimated residual ACF and PACF. If the plots do not look similar to corresponding plots from a white noise process, the model is not acceptable. Patterns in the plot may suggest what types of changes need to be made to the model.
   2. Examine the standardized residuals for outliers.
   3. Plot the standardized residuals versus time to look for autocorrelation.
   4. Investigate the normality assumption for wt by constructing histograms and Q-Q plots of the residuals. If the histogram is not symmetric, mound shape, or the Q-Q plot points do not fall on a straight line, there is evidence against normality. Investigate similar transformations as previously done to stabilize the variance.
   5. Perform the Ljung-Box-Pierce test on the residuals. If the test rejects the residual white noise hypothesis, other models need to be examined.
   6. Perform hypothesis tests for ϕi = 0 and θj = 0 parameters. If the test does not reject these hypotheses, then a different model may be needed.
   7. Examine the various diagnostic measures in an iterative manner until at least one model satisfies all of the model’s assumptions.
5. Using each model that gets through 4), pick the best model based upon model parsimony and the AIC. The best model corresponds to the one with the smallest number of AR and MA terms and the smallest AIC.
6. Choose one model and begin forecasting.

Final notes:

* Some models can be quite similar to one another when written out. Shumway and Stoffer’s GNP example provides an interesting case for how an AR(1) model is similar to a MA(2) model.
* There are automated procedures that choose the best ARIMA model. Here is a quote from Box, Jenkins, and Reinsel’s textbook regarding them:

Such procedures…have been useful, but they should be viewed as supplementary guidelines to assist in the model selection process. In particular they should not be used as a substitute for careful examination of characteristics of the estimated autocorrelation and partial autocorrelation function of the series, and critical examination of the residuals for the model inadequacies should always be included as a major aspect of the model selection process.