STAT 443 Lab 3

Caden Hewlett

2024-01-26

Question 1

We'll first simulate the data for $\{Z_t\}_{t \in \mathbb{Z}}$.

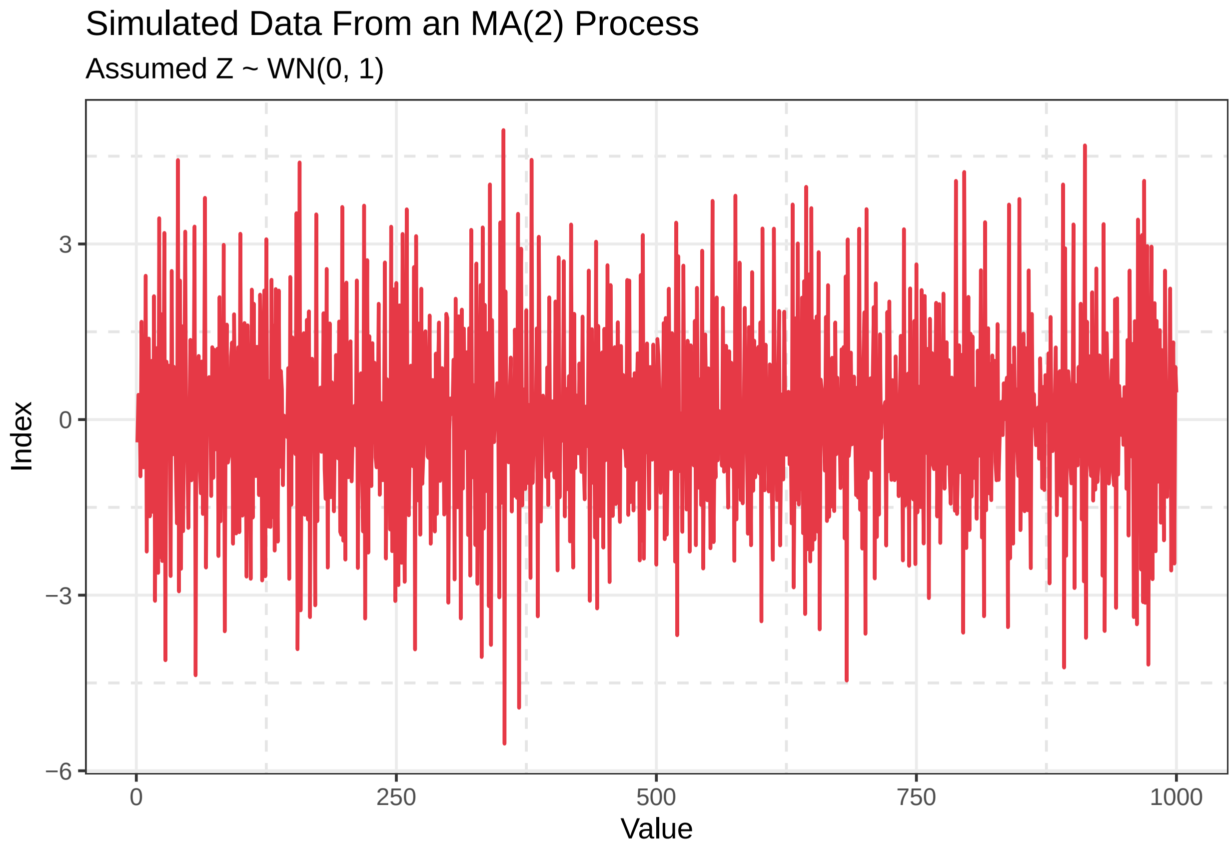
```r
coefs = c(-1.3, 0.4)

# simulate the MA(2) process
simulated_data <- arima.sim(n = 1000,
                            model = list(ma = coefs))
```

Part 1

Create a time series plot of the data.

```r
p1data = fortify.zoo(simulated_data)
p1 <- ggplot(p1data, aes(x = Index, y = simulated_data)) +
  geom_line(color = "#E63946", linewidth = 0.65) +
  labs(
    title = "Simulated Data From an MA(2) Process",
    subtitle = "Assumed Z ~ WN(0, 1)",
    x = "Value",
    y = "Index"
  ) + theme_bw() +
  theme(panel.grid.minor = element_line(
    color = "grey90",
    linetype = "dashed",
    linewidth = 0.5
  ))
print(p1)
```



What would we expect the acf of $\{X_t\}$ to look like?

I would expect the ACF to (obviously) have a value of 1 for $h = 0$. Further, I would expect the autocorrelation for lag values $h \in \{1, 2\}$ to be negative. The reason is this: if an innovation occurs at time t, we can consider our model forward-shifted by two to think about what is going to happen two time steps from now.

$$X_{t+2} = Z_{t+2} - 1.3Z_{t+1} + 0.4Z_t$$

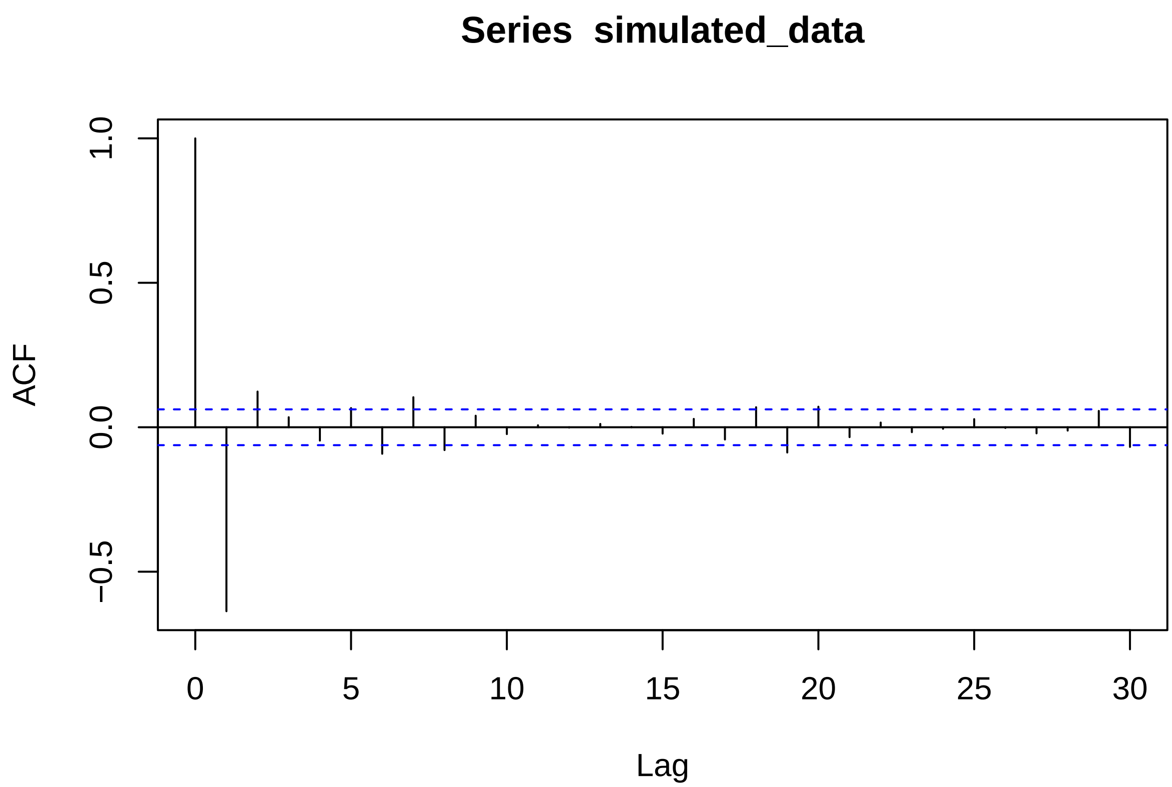
.

So the innovation is going to "ripple" forward, but be multiplied by a factor of -1.3 and become negatively-valued. So I would *roughly* expect a negative correlation for lags of 1 and 2, descending into noise for lags greater than 2.

Part 2

Here is the acf for the simulated data:

```
acf(simulated_data)
```



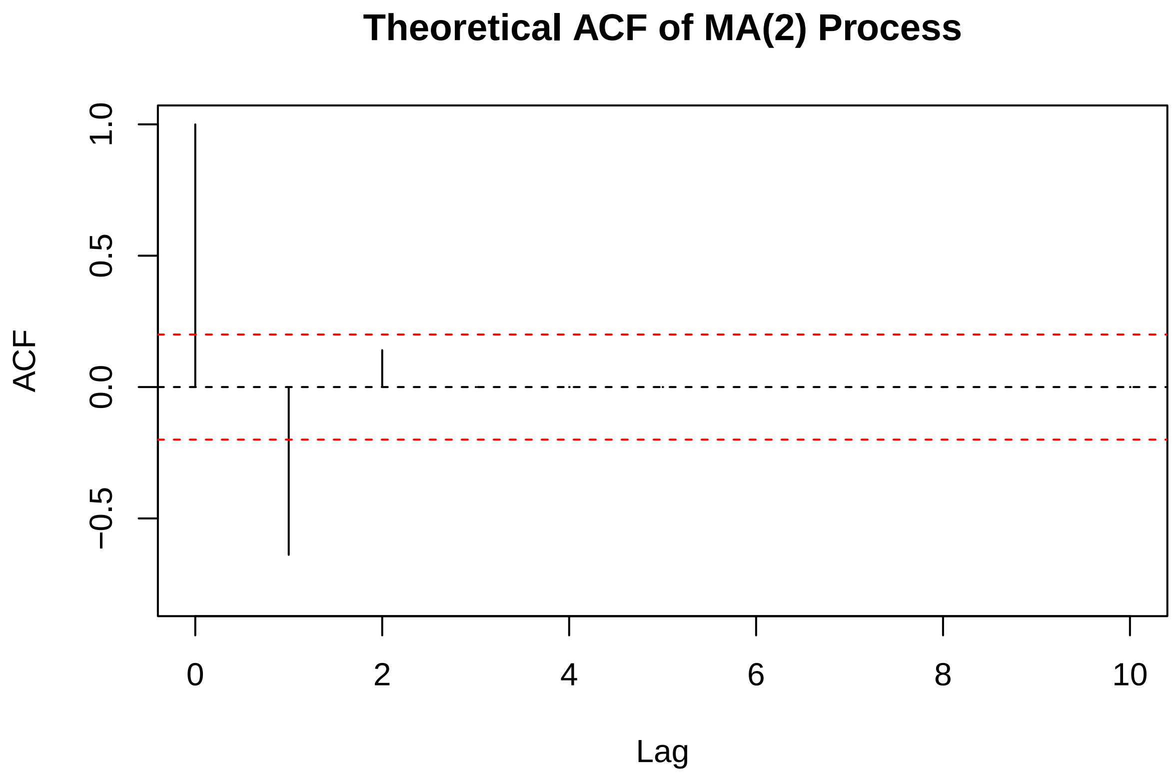
The behaviour of this ACF is sort of what I expected, but a little bit different. Interestingly, after running it a few times I began to notice that the aurocorrelation for lag $h = 2$ was "more negative" in repeated trials than for $h = 1$. Another interesting observation was that the observed ACF value for lag $h = 3$ was consistently moderately/weakly positive over repeated trials. I have a few ideas on why this is happening (to be discussed later.)

Part 3

Create the ACF Correlogram using `ARMAacf`.

```r
#  theoretical ACF for an MA(2) process
acf_values <- ARMAacf(ma = c(-1.3, 0.4), lag.max = 10)

# plotted theoretical ACF
plot(acf_values, x = 0:(length(acf_values)-1), type = "h",
     main = "Theoretical ACF of MA(2) Process",
     xlab = "Lag", ylab = "ACF", ylim = c(-0.8, 1))
abline(h = 0, lty = 'dashed')
abline(h = -0.2, lty = 'dashed', col = 'red')
abline(h = 0.2, lty = 'dashed', col = 'red')
```



Part 4

Explain the behaviour of the sample acf here.

To explain the AVF values, perhaps it is best to use an example.

Let's say there was an innovation with value 1 at time t, such that

$$X_t = Z_t - 1.3Z_{t-1} + 0.4Z_{t-2} = 1 + -1.3(0) + 0.4(0) = 1$$

And, for the sake of simplicity, Z_{t-1} and Z_{t-2} are both zero.

The effects of this innovation are going to "ripple forward." Let's say that time step $t+1$ has no innovation, and is a more regular value, e.g. $Z_{t+1} = 0.5$. In this case,

$$X_{t+1} = Z_{t+1} - 1.3Z_t + 0.4Z_{t-1} = 0.5 + -1.3(1) + 0.4(0) = -0.8$$

So, we can consider the pair $\{Z_t, Z_{t+1}\} = \{1, -0.8\}$. They're going to be negatively correlated on average (i.e. under this data production schema), as we can see in the ACF plot above for $h = 1$.

Let's continue with this example. Let's say that there is again no innovation at time step $t+2$, so the value of $Z_{t+2} = 0.3$. In this case,

$$X_{t+2} = Z_{t+2} - 1.3Z_{t+1} + 0.4Z_t = 0.3 + -1.3(-0.8) + 0.4(1) = 1.73$$

Now, there will be a weakly positive correlation between $\{Z_t, Z_{t+2}\} = \{1, 1.73\}$ for a lag of 2, which we also see in the ACF plot.

Then for further iterations, the process devolves into noise, which makes sense considering it is an $MA(2)$ process. This is reflected in the sample acf plot as zero autocorrelation.

Question 2

Part 1

Part 2

Part 3

Part 4

Question 3

Part 1

Part 2

Part 3

Part 4