Stat 443: Time Series and Forecasting

Lab 5

Caden Hewlett

February 09, 2024

Question 1

Given a time series, we can fit possible ARIMA models in R using the arima command. Look at the help page on this function before attempting the following activities.

Suppose $\{Z_t\}_{t\in\mathbb{N}} \sim \text{WN}(0, (0.8)^2)$ and

Consider stochastic process $\{X_t\}_{t\in\mathbb{N}}$ with:

$$X_t = 0.8X_{t-1} - \frac{1}{3}X_{t-2} + \frac{0.6}{\sqrt{3}}X_{t-3} + Z_t$$

Part 1

Name the process defined in equation (1), specifying its order.

Solution This is an AR process with $p = 3$ (i.e., $\{X_t\}_{t\in\mathbb{N}} = \text{AR}(3)$.)

Part 2

Based on an observed time series, the first step to identifying this model is to inspect the sample *acf*.

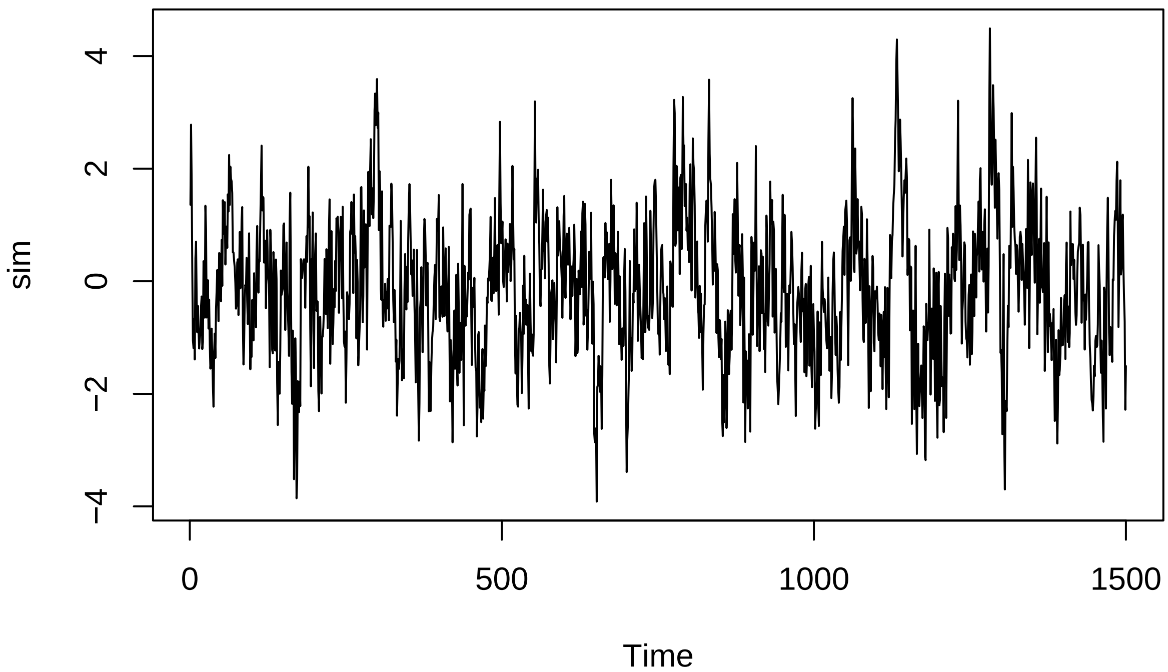
In the sample *acf* of an autoregressive time series, you would expect either exponential decay (slow, exponential decay) for positive α dominating the series, or, alternatively, a dampened (decreasing) sinusoidal curve in the sample acf.

Once you have decided that the process is autoregressive, the order is then determined from the *pacf*, and you would look for the greatest k such that $\hat{\alpha}_{kk}$ is significant (i.e. greater than $\pm 2/\sqrt{n}$ or by a corresponding test against the theoretical normal describing the partial autocorrelations), and $\forall p > k$, the *pacf* values of p are insignificant. Then, under this setting, k is the order of the AR process.

Part 3

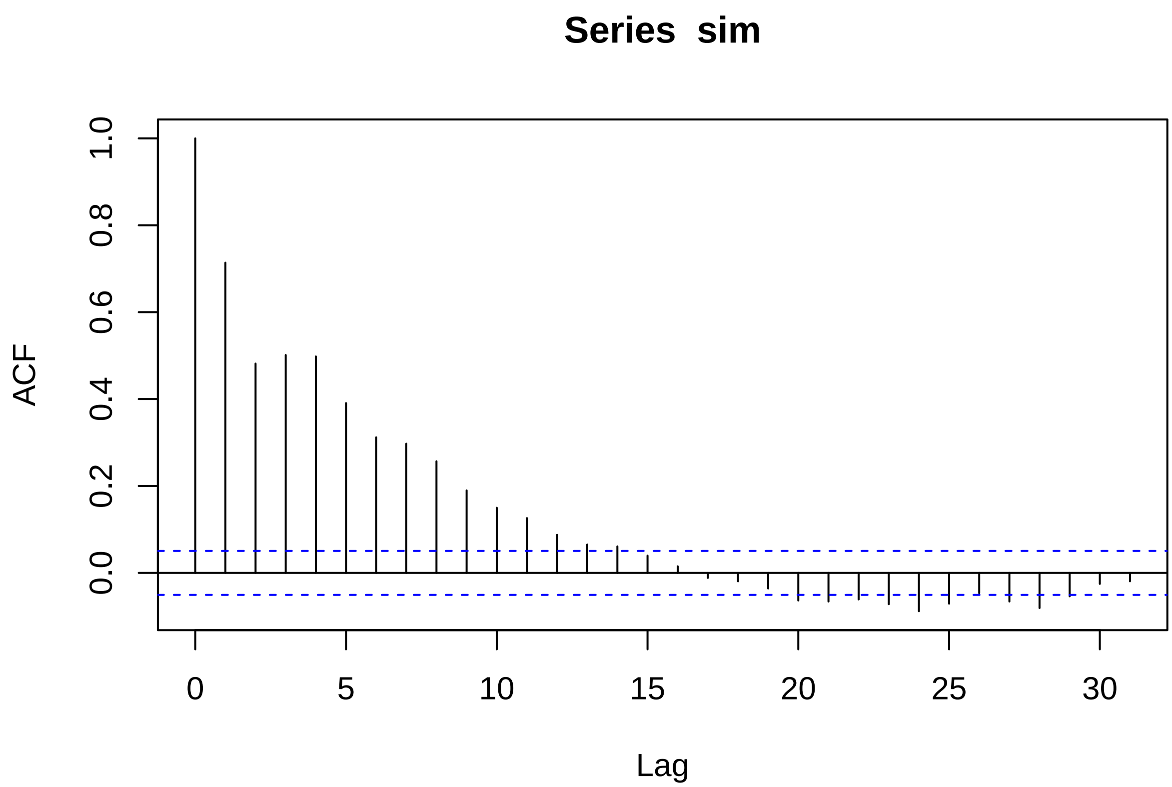
Use the command `set.seed(23456)` to set the random seed for reproducibility and then use function `arima.sim()` to generate 1500 observations from the model in (1). Plot the simulated time series.

```
set.seed(23456)
coefs = c(0.8, -1/3, (0.6)/sqrt(3))
sim = arima.sim(n = 1500, list(ar = coefs), sd = 0.8)
plot(sim)
```



Part 4

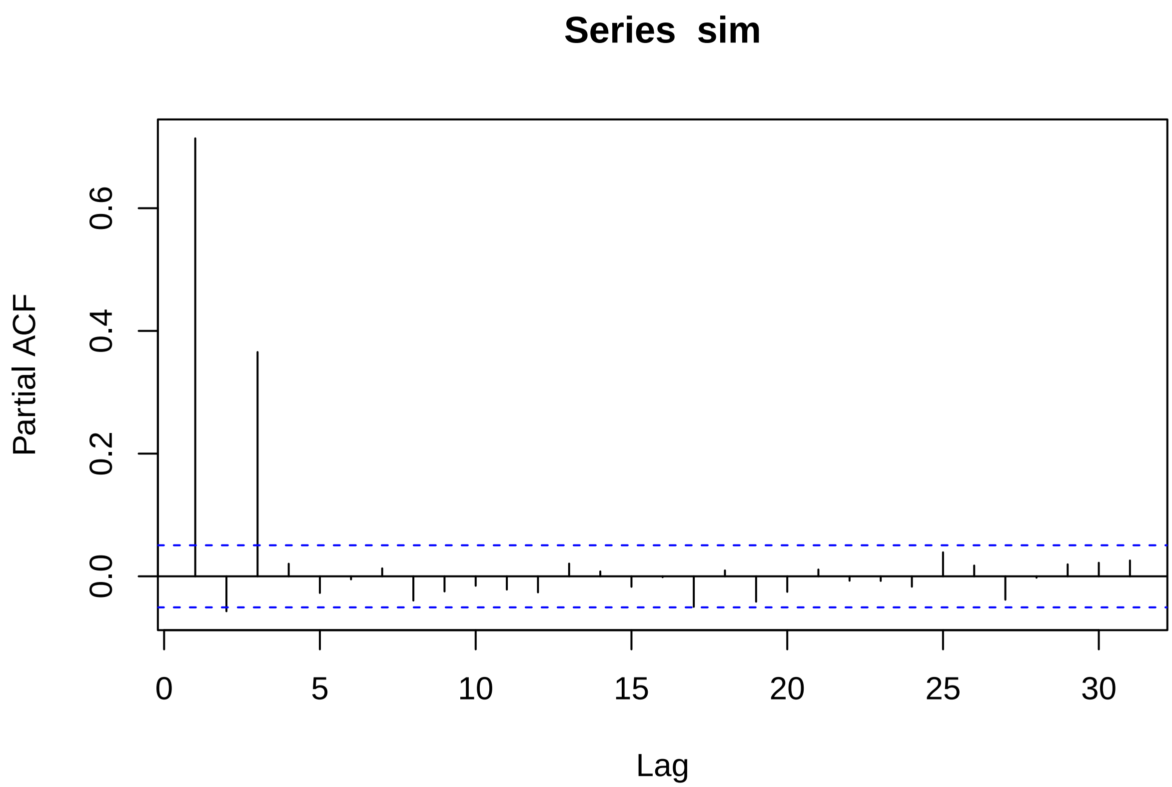
```
acf(sim)
```



Comments: We see a slow, exponential decay of sample *acf* values. This is what we would expect from an autoregressive process dominated by positive α terms. What I mean by this is that since $\alpha_2 < 0$, we should observe a sinusoidal pattern; however, since $|\alpha_1| > |\alpha_3| > |\alpha_2|$ that effect would be more mild than if α_2 is greatest in magnitude. We see this reflected in the *acf*. There is slow exponential decay as we'd expect, as well as an eventual low-magnitude sinusoidal pattern.

Part 5

```
pacf(sim)
```



This is exactly what I would expect of a *pacf* of an order 3 autoregressive process. We see that $\hat{\alpha}_{11}$ and $\hat{\alpha}_{33}$ are well above the threshold for significance (with $\hat{\alpha}_{22}$ also slightly above this threshold.) Importantly, though, in this *pacf* we have the following relationship:

$$\forall k > 3, |\hat{\alpha}_{kk}| < \frac{2}{\sqrt{n}}$$

Which is the exact definition of order $p = 3$ as we saw in lecture. This matches the process well. ## Part 6

Since this is an $AR(p = 3)$ process, we can apply $ARIMA(p = 3, d = 0, q = 0)$. We let the mean equal to zero, setting `include.mean = F`.

```
order_x = c(3, 0, 0)
method = c("CSS-ML", "ML", "CSS")
# iterate through all
results = (sapply(method, function(M){
  fit = arima(sim, order = order_x, include.mean = F, method = M)
  c(fit$coef[1], fit$coef[2], fit$coef[3], fit$sigma2)
  }))
# prepare table
truth = c(coefs, 0.8^2)
df = data.frame(t(results))
colnames(df) = c("a1", "a2", "a3", "sigma_2")
kable(rbind(Truth = round(truth,3), round(df, 3)))
```

	a1	a2	a3	sigma_2
Truth	0.800	-0.333	0.346	0.640
CSS-ML	0.775	-0.330	0.368	0.637
ML	0.775	-0.330	0.368	0.637
CSS	0.774	-0.330	0.368	0.636