Solutions: Assignment 2 Due 10/7 at Midnight

Julia Schedler

NOTE: I forgot to include a relevant detail for Part 2, number 7. The change is **bolded.** Sorry!!! Part 2 number 8 should be easier to answer now.

Part 1: Math

In class, we have worked with "Signal plus noise Model" (equation 1.5)

$$\text{Model: } x_t = 2\cos(2\pi\frac{t+15}{50}) + w_t$$
$$\text{Mean function: } \mathbb{E}(x_t) = 2\cos(2\pi\frac{t+15}{50})$$

1. [5 points] The mean function is derived in Example 2.4. Describe what happens in each step of the computation [3 points], and provide a "math stress" rating (1 = effortless, 100 = nightmare) and 3 emojis[2 points]. This is personal and there is no right answer.
2. [5 points] Is the signal plus noise model stationary in the mean?
3. [5 points] Write down $\gamma_x(s,t)$, the autocovariance function of x_t [3 points]. You may accomplish this in any way, including asking me personally in office hours or asking a classmate. Just make sure you cite the source![2 points]
4. [6 points] Consider the model:

$$y_t = x_t - 2\cos(2\pi\frac{t+15}{50})$$

   Compute the mean function of y_t [3 points]. Is y_t stationary in the mean?[1 point] How do you know?[2 points]

Part 2: Code

Note: I have set the code chunks here to have `eval: false` in the code chunk. Change that to `true` so that I can run your code easily.

0. [5 points] All your code runs without errors (unless that's the point), and if there is a message, explain what it means. (Bonus: to be nice to me, submit a rendered pdf)
1. [5 points] Simulate from an AR(1) process with coefficient 0.7 and 10 data points.

```r
library(astsa)

# your code here

w <- rnorm(10)
x_t <- stats::filter(x = w, filter = 0.7, method = "recursive")
```

2. [6 points] Look at the documentation for the `stats::lag` function (run `?lag` in the console). State what package the function is in and what the function does[4 points]. Using `k = 1` compute a lag(1) version of `x_t` that you simulated above[2 points].

```r
x_t_lag1 <- stats::lag(x_t, k = 1)
```

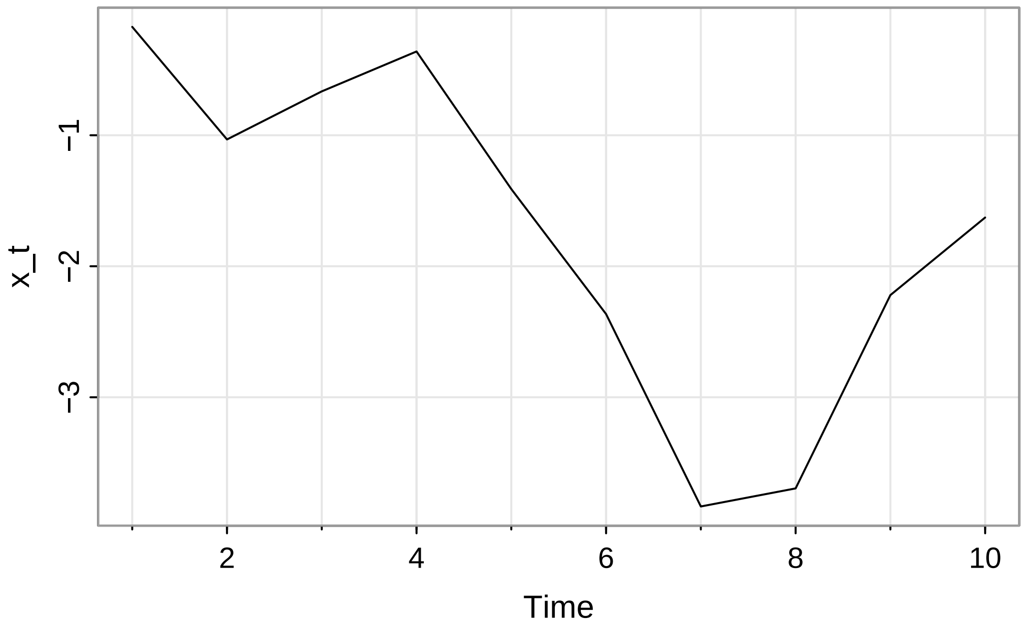
3. [3 points] Run the following code and compare `x_t` and `x_t_lag1`.

```r
cbind(x_t, x_t_lag1)
```

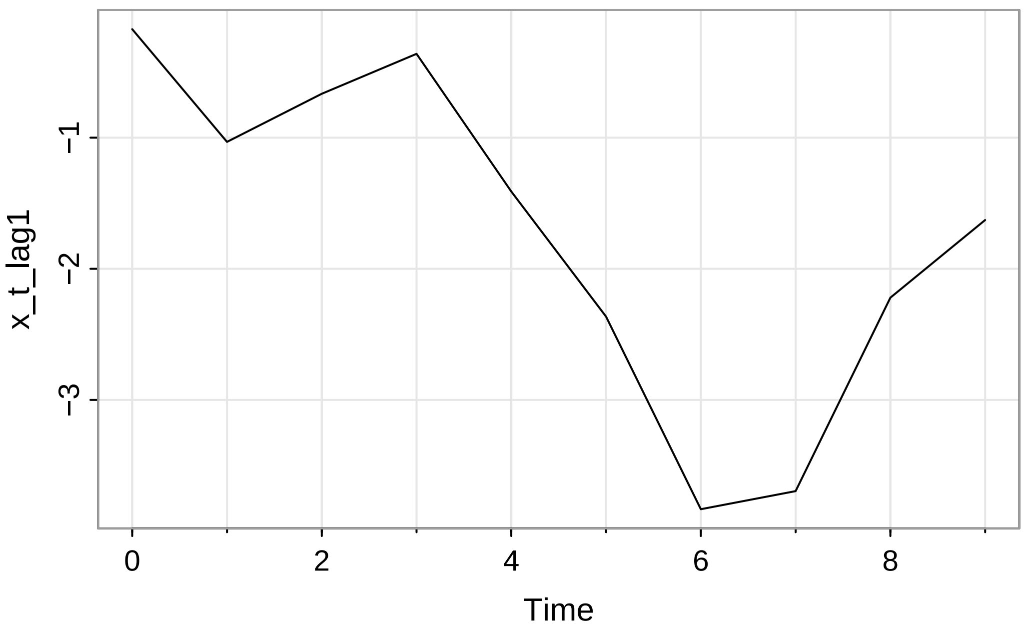
```
Time Series:
Start = 0
End = 10
Frequency = 1
          x_t   x_t_lag1
 0         NA -0.1726278
 1 -0.1726278 -1.0320927
 2 -1.0320927 -0.6651790
 3 -0.6651790 -0.3606563
 4 -0.3606563 -1.4105394
 5 -1.4105394 -2.3649387
 6 -2.3649387 -3.8335712
 7 -3.8335712 -3.6957562
 8 -3.6957562 -2.2203032
 9 -2.2203032 -1.6280108
10 -1.6280108         NA
```

4. Make a time series plot of `x_t` and `x_t_1`. Do you notice the same features as when in the previous question?

```r
tsplot(x_t)
```




```r
tsplot(x_t_lag1)
```

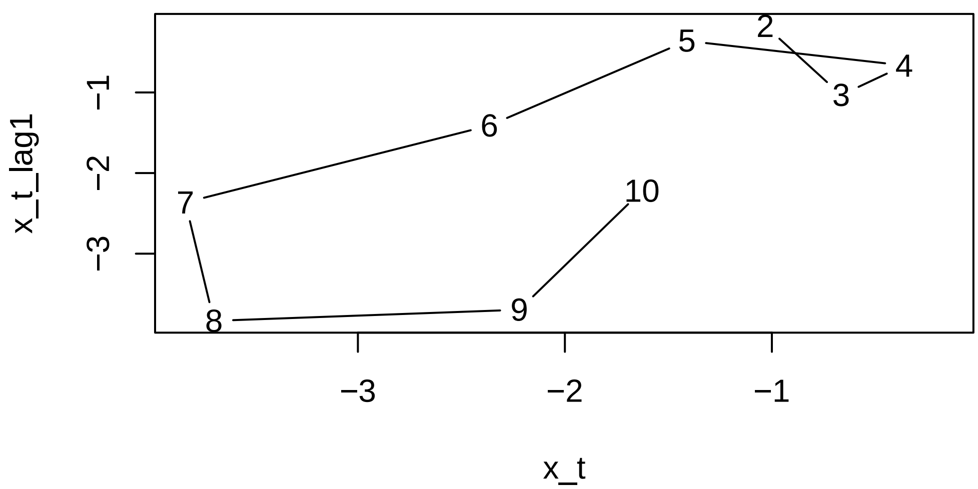


5. Run the below code. Why are the plots different? Are either particularly useful?

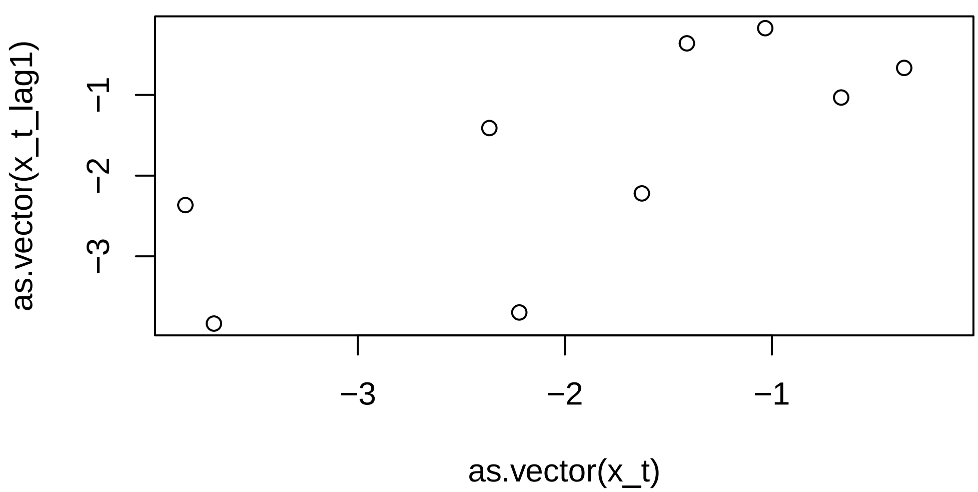
```
plot(x_t, x_t_lag1)
plot(as.vector(x_t), as.vector(x_t_lag1))
```

6. Instead of using `stats::lag`, use `dplyr::lag` to create a new version of `x_t_lag`. Repeat the code from steps 2-5. Describe how the output has changed.

```
x_t_lag1 <- dplyr::lag(as.vector(x_t), n = 1)
plot(x_t, x_t_lag1)
```




```r
plot(as.vector(x_t), as.vector(x_t_lag1))
```



7. **Re-simulate an AR(1) process as in number 1, but this time with 100 observations. Also recompute `x_t_lag1`.** Fit an intercept-free regression model **to**

predict `x_t` **from** `x_t_lag`. Provide the value of the slope estimate and interpret the value in the context of this simulation.

```r
w <- rnorm(100)
x_t <- stats::filter(x = w, filter = 0.7, method = "recursive")

x_t_lag1 <- dplyr::lag(as.vector(x_t), n = 1)

linear_model <- lm(x_t ~ -1 + x_t_lag1)
summary(linear_model)
```

```
Call:
lm(formula = x_t ~ -1 + x_t_lag1)

Residuals:
    Min      1Q  Median      3Q     Max
-2.2005 -0.7563 -0.1541  0.3911  2.0320

Coefficients:
         Estimate Std. Error t value Pr(>|t|)
x_t_lag1  0.69336    0.07289   9.512 1.38e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9661 on 98 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.4801,    Adjusted R-squared:  0.4748
F-statistic: 90.49 on 1 and 98 DF,  p-value: 1.384e-15
```

8. [11 points] Plot the `acf` of `x_t`[2 points] and the `acf` of the residuals from the regression model[4 points]. Which looks more like white noise?[2 points] What does this tell you about the temporal structure in `x_t` and its residuals?[3 points]

```r
acf(x_t)
acf(residuals(linear_model))
```

Part 3: Reading

[9 points] Read sections 2.8 and 2.9 from Forecasting Principles and Practice. Make 3 connections [3 points each] to content from the course textbook (equations or similar examples.).