

# STATS 531 Homework 4

Huy Le

February 2026

## Question 4.1

A. Consider the stationary AR(2) model

$$X_n = 1.5 X_{n-1} - 0.8 X_{n-2} + \epsilon_n, \quad (1)$$

where  $\{\epsilon_n\}$  is white noise with  $\text{Var}(\epsilon_n) = \sigma^2$ .

We first rewrite this in terms of the backshift operator  $B$ :

$$\phi(B) X_n = \epsilon_n, \quad \phi(B) = 1 - 1.5B + 0.8B^2.$$

Recall that the spectral density of a causal ARMA process is given by (1; 2)

$$f(\omega) = \sigma^2 \left| \frac{\theta(e^{-2\pi i \omega})}{\phi(e^{-2\pi i \omega})} \right|^2, \quad (2)$$

where  $\theta(z) = 1$  for a pure AR model. Since we have a pure AR(2), this reduces to

$$f(\omega) = \frac{\sigma^2}{|1 - 1.5 e^{-2\pi i \omega} + 0.8 e^{-4\pi i \omega}|^2}. \quad (3)$$

To evaluate the denominator, we let  $z = e^{-2\pi i \omega}$  so that

$$\phi(z) = 1 - 1.5z + 0.8z^2.$$

We then compute  $|\phi(z)|^2 = \phi(z)\overline{\phi(z)}$ . Since  $z\bar{z} = 1$ , we can expand this as

$$|\phi(e^{-2\pi i \omega})|^2 = 1 + 1.5^2 + 0.8^2 - 2(1.5) \cos(2\pi\omega) + 2(1.5)(0.8) \cos(2\pi\omega) - 2(0.8) \cos(4\pi\omega).$$

After simplifying, we obtain

$$|\phi(e^{-2\pi i \omega})|^2 = 2.89 - 0.6 \cos(2\pi\omega) - 1.6 \cos(4\pi\omega).$$

We verify this analytic expression numerically below and plot both the spectral density and the autocovariance function.

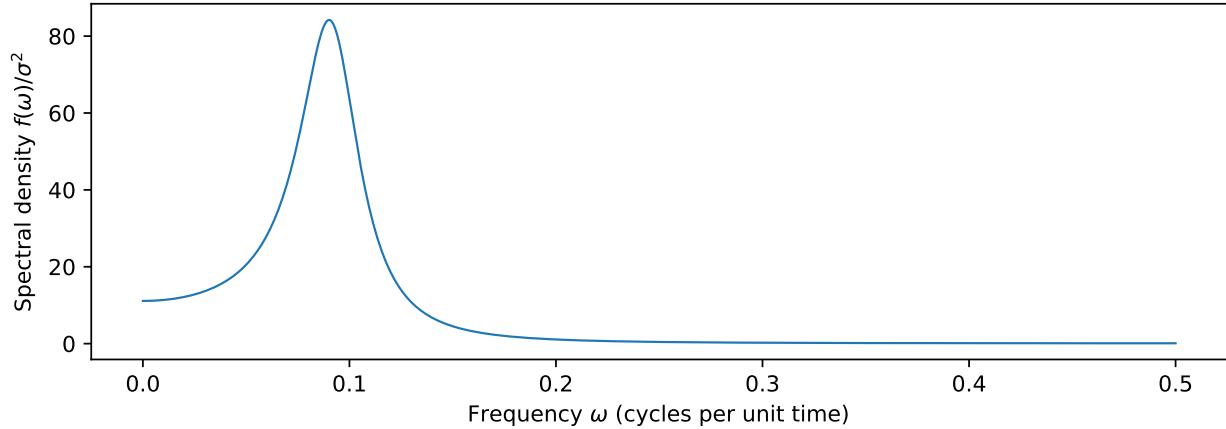


Figure 1: Spectral density of the AR(2) model  $X_n = 1.5X_{n-1} - 0.8X_{n-2} + \epsilon_n$ .

As we can see in Figure 1, the spectral density shows a prominent peak near  $\omega \approx 0.1$ , which indicates a quasi-periodic oscillation. We can understand this by examining the roots of the AR(2) polynomial:  $\phi(z) = 1 - 1.5z + 0.8z^2$  has complex roots since the discriminant is  $1.5^2 - 4(0.8) = -0.95 < 0$ . Specifically, the roots are at  $z = \frac{1.5 \pm i\sqrt{0.95}}{1.6}$ , which have modulus  $|z| = 1/\sqrt{0.8} \approx 1.118 > 1$ , confirming stationarity. Moreover, the frequency of the spectral peak corresponds to the argument of these complex roots.

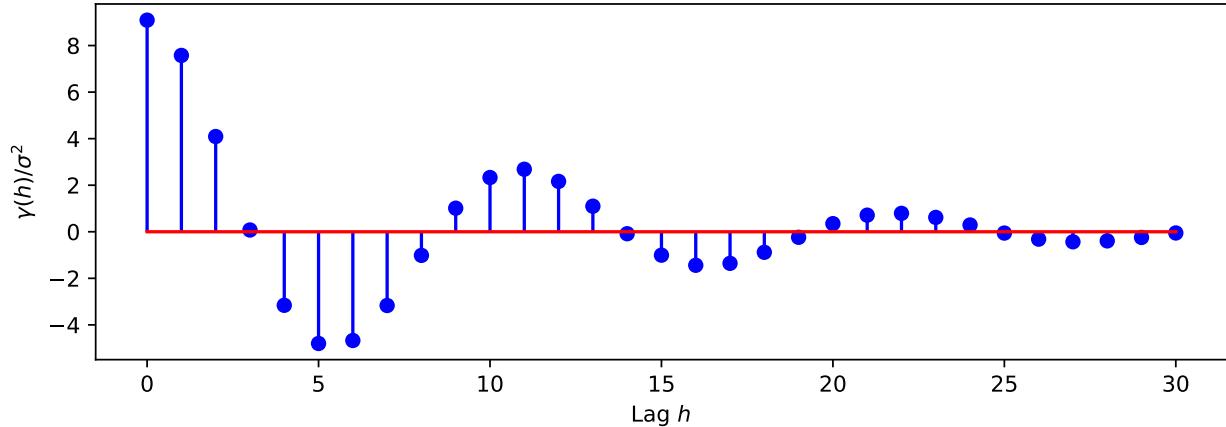


Figure 2: Autocovariance function of the AR(2) model (in units of  $\sigma^2$ ), computed from the AR coefficients.

In Figure 2, we observe that the autocovariance function displays damped oscillations. This is the time-domain signature of the complex AR roots, and the oscillatory decay directly reflects the quasi-periodic behavior we identified in the spectral peak above.

**B.** Consider the MA(2) moving mean

$$X_n = \epsilon_{n-2} + \epsilon_{n-1} + \epsilon_n, \quad (4)$$

where  $\{\epsilon_n\}$  is white noise with  $\text{Var}(\epsilon_n) = \sigma^2$ .

We write this in operator notation as  $X_n = \theta(B)\epsilon_n$  with  $\theta(B) = 1 + B + B^2$ . Accordingly, the spectral density is

$$f(\omega) = \sigma^2 |\theta(e^{-2\pi i \omega})|^2 = \sigma^2 |1 + e^{-2\pi i \omega} + e^{-4\pi i \omega}|^2. \quad (5)$$

We now expand this analytically. Letting  $z = e^{-2\pi i \omega}$ , we have

$$|1 + z + z^2|^2 = (1 + z + z^2)(1 + \bar{z} + \bar{z}^2).$$

We then use the identities  $z\bar{z} = 1$  and  $z + \bar{z} = 2 \cos(2\pi\omega)$  to obtain

$$\begin{aligned} |1 + z + z^2|^2 &= 3 + 4 \cos(2\pi\omega) + 2 \cos(4\pi\omega) \\ &= 3 + 4 \cos(2\pi\omega) + 2(2 \cos^2(2\pi\omega) - 1) \\ &= 1 + 4 \cos(2\pi\omega) + 4 \cos^2(2\pi\omega) \\ &= (1 + 2 \cos(2\pi\omega))^2. \end{aligned}$$

We therefore arrive at the closed-form expression

$$f(\omega) = \sigma^2 (1 + 2 \cos(2\pi\omega))^2. \quad (6)$$

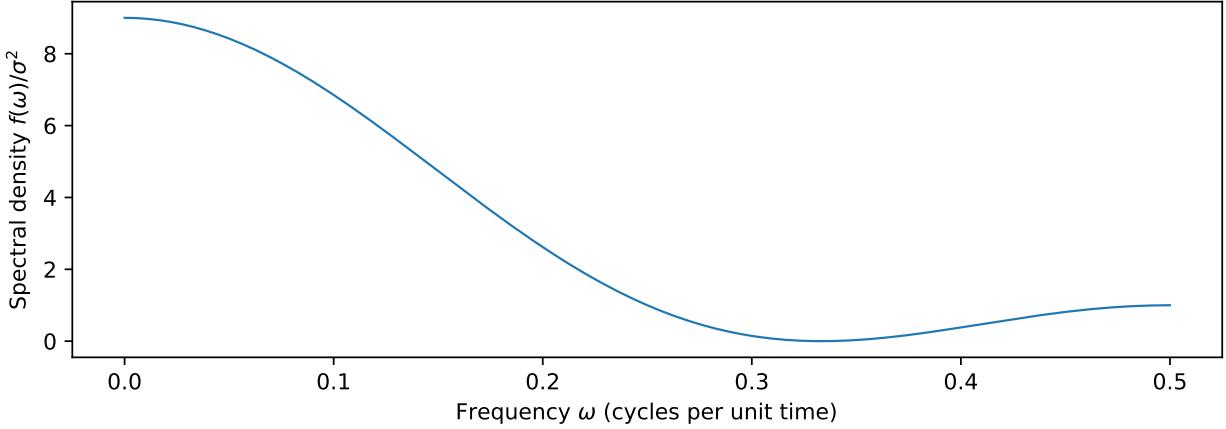


Figure 3: Spectral density of the MA(2) moving mean  $X_n = \epsilon_{n-2} + \epsilon_{n-1} + \epsilon_n$ .

As Figure 3 confirms, the numerical and analytic formulas agree exactly. We note that the spectral density equals zero at  $\omega = 1/3$  (where  $\cos(2\pi/3) = -1/2$ , so  $1 + 2(-1/2) = 0$ ). Meanwhile, the maximum occurs at  $\omega = 0$  where  $f(0) = 9\sigma^2$ .

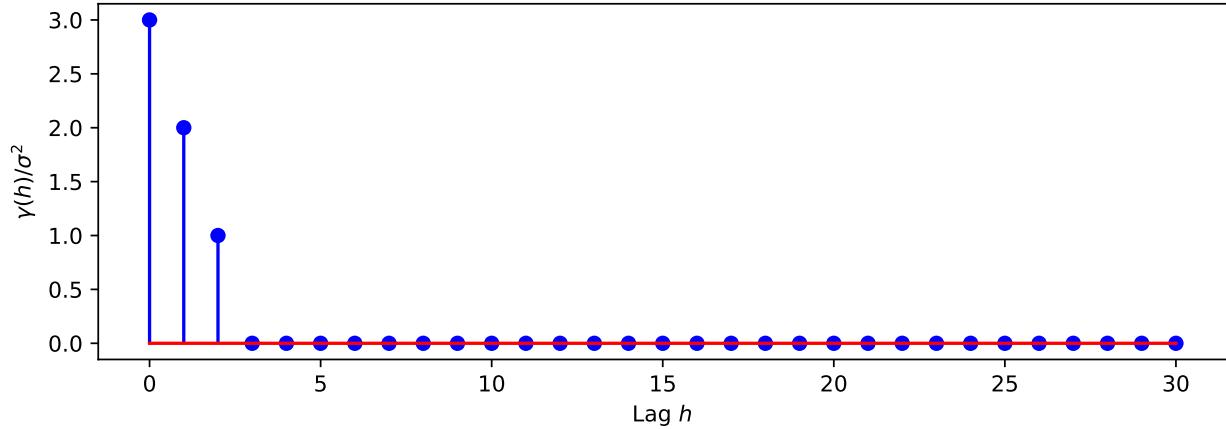


Figure 4: Autocovariance function of the MA(2) moving mean (in units of  $\sigma^2$ ).

We can compute the autocovariance of this MA(2) directly:  $\gamma(0) = 3\sigma^2$ ,  $\gamma(1) = 2\sigma^2$ ,  $\gamma(2) = \sigma^2$ , and  $\gamma(h) = 0$  for  $|h| > 2$ . This sharp cutoff is expected, since an MA( $q$ ) process always has autocovariance that vanishes beyond lag  $q$ .

C. The AR(2) has a sharp spectral peak and its autocovariance oscillates and decays slowly, while the MA(2) puts most of its power at low frequencies and its autocovariance cuts off after just 2 lags. Basically, the AR model is good at capturing narrow peaks in the spectrum, whereas the MA model is better at representing zeros.

## Question 4.2

The sunspot data from the WDC-SILSO (3) contains 3325 monthly mean sunspot numbers from 1749.0 to 2026.0. Figure 5 shows the well-known approximately 11-year solar cycle.

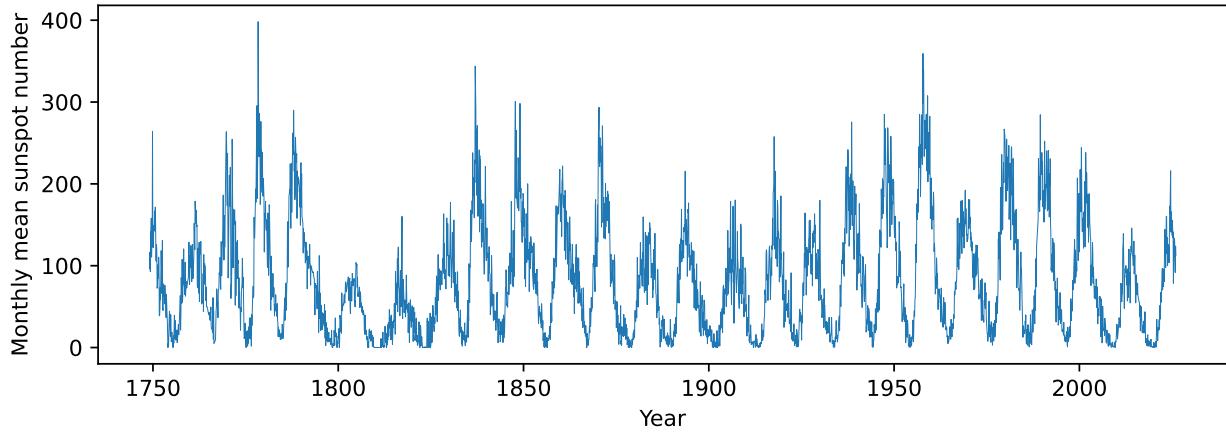


Figure 5: Monthly mean sunspot number, 1749–2026.

To estimate the spectral density, I tried three approaches. First, we compute the raw periodogram, defined as  $I(\omega_j) = \frac{1}{N} |\sum_{n=1}^N X_n e^{-2\pi i \omega_j n}|^2$ . Second, we apply a smoothed periodogram using the Daniell kernel (1), which averages the raw periodogram over a window of neighboring frequencies to reduce variance at the cost of frequency resolution. We try several different bandwidths to assess sensitivity. Third, we fit an AR model with order selected by AIC and compute the implied parametric spectral density.

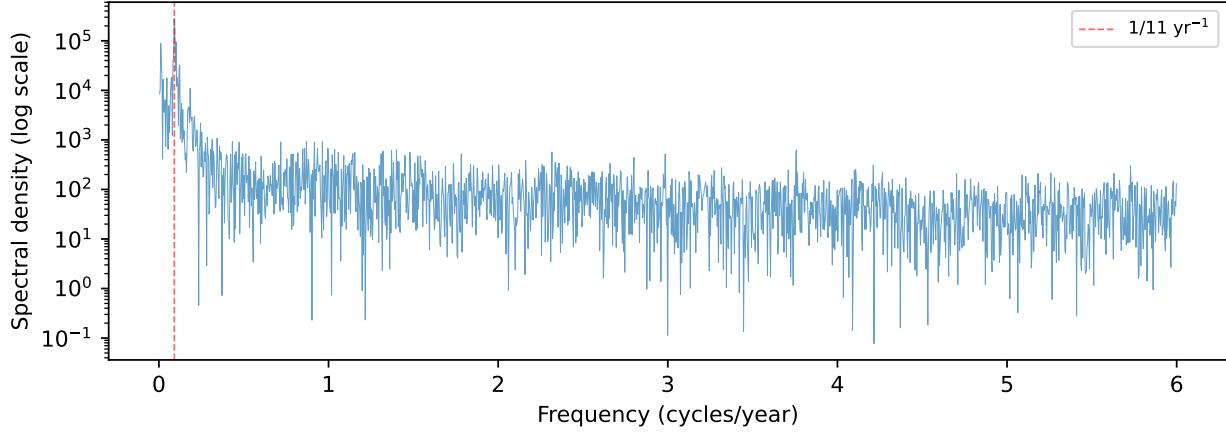


Figure 6: Raw periodogram of the sunspot series (log scale).

As we can see in Figure 6, the raw periodogram is quite noisy, which is expected since it is an inconsistent estimator. Nevertheless, it clearly reveals a dominant peak near  $1/11 \approx 0.091$  cycles/year, corresponding to the well-known  $\sim 11$ -year solar cycle.

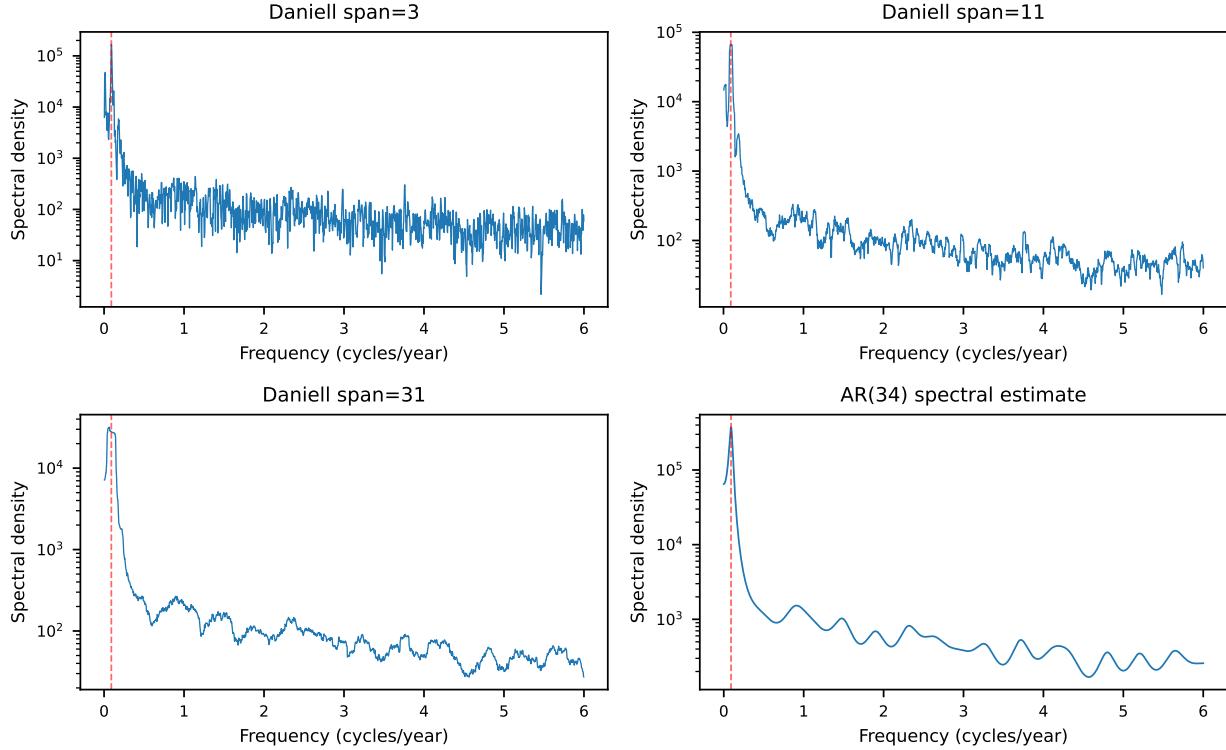


Figure 7: Smoothed periodogram estimates using different bandwidths, compared with AR-based spectral estimate.

Looking at Figure 7, the smoothed periodogram with a moderate bandwidth (span around 11 to 31) seems to work best. When the span is too small (like 3), the plot is still pretty noisy and hard to read. But when it gets too large, the peaks start getting smeared out and you lose detail. The span=11 and span=31 versions both pick up the main cycle clearly while being smooth enough to interpret.

I also tried a parametric approach by fitting an AR model with order chosen by AIC, which gave AR(34). This one gives the cleanest looking estimate since it is smooth by construction, but it does assume the data actually follows an AR process, which might not be perfectly true. Still, for data like this with a strong periodic signal, AR models tend to do a good job at picking up sharp spectral peaks (1).

The biggest peak in the spectrum sits at about 0.0902 cycles/year, so the period is roughly 11.1 years. This lines up with the Schwabe cycle, which is the well-known ~11-year sunspot cycle caused by the flipping of the sun's magnetic field (3).

There is also some power showing up near  $2/11 \approx 0.18$  cycles/year, which is the second harmonic. This makes sense because the sunspot cycle is not a perfect sine wave — sunspot numbers tend to rise quickly and fall off slowly, so you need higher harmonics to capture that asymmetry.

At the low-frequency end (below 1/11 cycles/year), there is broad power that could be related to longer-term changes in how strong each solar cycle is. Some people call this the Gleissberg cycle, which has a period of around 80–100 years. You can kind of see this in the time series itself: some cycles have much higher peaks than others, and this variation happens over many decades.

So overall, the spectrum is dominated by the ~11-year cycle and its harmonics, which is consistent with what solar physicists have known for a long time.

## Acknowledgments

AI (Claude, Anthropic) was used for indenting contents and calculation for a better format, and help integrate code from the Jupyter notebook into the Quarto (.qmd) report format

## Bibliography

- [1] Robert H. Shumway and David S. Stoffer. *Time Series Analysis and Its Applications: With R Examples*. 2017.
- [2] Edward L. Ionides. Stats 531 (winter 2026) course notes: Analysis of time series. <https://ionides.github.io/531w26/>, 2026.
- [3] Royal Observatory of Belgium. World data center – sunspot index and long-term solar observations (wdc-silso). <https://www.sidc.be/SILSO/home>, 2026.