STAT 153 - Introduction to Time Series
Lecture Thirteen

Fall 2022, UC Berkeley

Aditya Guntuboyina

October 7, 2022

AutoRegressive (AR) Models are our next topic of study. We shall motivate their use through the sunspots dataset. This was how the AR models were originally invented by Yule [1].

1 The Yule Model for the Sunspots Data

For the sunspots dataset, we previously employed the model

$$Y_i = \mu + \alpha_1 \cos(\omega t_i) + \alpha_2 \sin(\omega t_i) + \epsilon_i \qquad \text{for } i = 1, \dots, n \tag{1}$$

We used a Bayesian method to infer the angular frequency parameter $\omega = 2\pi f$ (which is the main parameter of interest) and this led to an estimated period of close to 11 years (which is often cited as the period of the solar cycle). Note however that (1) is not ideal for the sunspots dataset for at least two reasons: (a) the fit to the data is not very good (some of the oscillations have a much higher amplitude than that explained by the single sinusoid), (b) data generated from the model (1) look much more "noisy" compared to the actual sunspots data. Starting with these observations, Yule [1] proposed an alternative model that is also based on a single sinuosoid. This alternative model is based on the idea of AR modeling.

Yule started with the following basic observation. Let s_t denote the sinusoid:

$$s_t = \mu + \alpha_1 \cos(\omega t) + \alpha_2 \sin(\omega t) \tag{2}$$

The same sinusoid can be understood as the solution to a specific difference equation. To derive the difference equation, let us first note that, in continuous time, $s(t)$ satisfies

$$s''(t) = -\omega^2 \left(\alpha_1 \cos(\omega t) + \alpha_2 \sin(\omega t)\right) = -\omega^2 \left(s(t) - \mu\right). \tag{3}$$

In discrete time (where $t \in \{\dots, -2, -1, 0, 1, 2, \dots\}$), the sequence (2) satisfies the following difference equation that is analogous to (3):

$$s_{t+2} - 2s_{t+1} + s_t = 2(\cos\omega - 1)\left(s_{t+1} - \mu\right). \tag{4}$$

To see this, note that

$$\begin{aligned} &s_{t+2} - 2s_{t+1} + s_t \\ &= \alpha_1 \left(\cos(\omega(t+2)) - 2\cos(\omega(t+1)) + \cos(\omega t)\right) + \alpha_2 \left(\sin(\omega(t+2)) - 2\sin(\omega(t+1)) + \sin(\omega t)\right) \end{aligned}$$

Writing $A = \omega(t+1)$ and $B = \omega$, we get

$$\begin{aligned}\cos(\omega(t+2)) - 2\cos(\omega(t+1)) + \cos(\omega t) &= \cos(A+B) - 2\cos A + \cos(A-B) \\ &= 2\cos A(\cos B - 1) \\ &= 2(\cos\omega - 1)\cos(w(t+1))\end{aligned}$$

and similarly

$$\sin(\omega(t+2)) - 2\sin(\omega(t+1)) + \sin(\omega t) = 2(\cos\omega - 1)\sin(\omega(t+1))$$

This proves

$$s_{t+2} - 2s_{t+1} + s_t = 2(\cos\omega - 1)\left(\alpha_1\cos(\omega(t+1)) + \alpha_2\sin(\omega(t+1))\right) = 2(\cos\omega - 1)(s_{t+1} - \mu)$$

thereby establishing (4).

The converse is also true in the sense that every solution $\{s_t\}$ to the difference equation (4) say, for $t = 0, 1, 2, \dots$, with given values of s_0 and s_1 (initial conditions) is of the form (2) for some α_1 and α_2. To see this, let $g_t = s_t - \mu$ and note that $\{g_t\}$ satisfies

$$g_{t+2} - 2g_{t+1} + g_t = 2(\cos\omega - 1)g_{t+1}.$$

We find α_1 and α_2 such that

$$h_t := \alpha_1\cos(\omega t) + \alpha_2\sin(\omega t)$$

matches g_t for $t = 0, 1$. Now if $g_t = h_t$ and $g_{t+1} = h_{t+1}$, then

$$\begin{aligned}g_{t+2} &= (2\cos\omega)g_{t+1} - g_t \\ &= (2\cos\omega)\left(\alpha_1\cos(\omega(t+1)) + \alpha_2\sin(\omega(t+1))\right) - \left(\alpha_1\cos(\omega t) + \alpha_2\sin(\omega t)\right) \\ &= \alpha_1\left(2\cos\omega\cos(\omega(t+1)) - \cos(\omega t)\right) + \alpha_2\left(2\cos\omega\sin(\omega(t+1)) - \sin(\omega t)\right) \\ &= \alpha_1\left(\cos(\omega t)\left(2\cos^2\omega - 1\right) - \sin(\omega t)2\sin\omega\cos\omega\right) \\ &+ \alpha_2\left(\sin(\omega t)\left(2\cos^2\omega - 1\right) + \cos(\omega t)2\sin\omega\cos\omega\right) \\ &= \alpha_1\left(\cos(\omega t)\cos(2\omega) - \sin(\omega t)\sin(2\omega)\right) + \alpha_2\left(\sin(\omega t)\cos(2\omega) + \cos(\omega t)\sin(2\omega)\right) \\ &= \alpha_1\cos(\omega(t+2)) + \alpha_2\sin(\omega(t+2)) = h_{t+2}.\end{aligned}$$

Using this for $t = 0, 1, 2, \dots$ proves that (2) is the unique solution to (4).

To summarize, an alternative way of describing a sinusoid of frequency ω is via the difference equation (4) which is equivalent to

$$s_{t+2} = (2\cos\omega)s_{t+1} - s_t + 2(1 - \cos\omega)\mu.$$

Based on this equation, Yule proposed the model:

$$Y_{t+2} = \theta Y_{t+1} - Y_t + c + Z_{t+2} \tag{5}$$

with two parameters θ and c (and the additional noise parameter σ in $Z_{t+2} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$). Note that this is also a single sinusoid plus noise model but now the noise is in a different place. To better understand the difference between (5) and the earlier model:

$$Y_t = \mu + \alpha_1\cos(\omega t) + \alpha_2\sin(\omega t) + \epsilon_t, \tag{6}$$

consider the following physical situation where sinusoids naturally arise (see e.g., page 2 of the Fourier Analysis book by Stein and Shakarchi). Consider a mass m that is attached to a horizontal spring, which itself is attached to fixed wall, and assume that the system lies on a frictionless surface. Suppose that μ is the location of the center of the mass when the spring is neither compressed or stretched. When the spring is compressed or stretched and released, the mass undergoes simple harmonic motion.

Let $y(t)$ denote the displacement of the mass at time t. Hooke's law says that the force exerted by the spring on the mass is given by $F = -\kappa\,(y(t) - \mu)$ where $\kappa > 0$ is the spring constant. By Newton's law (note that the acceleration is given by $y''(t)$), we have

$$-\kappa\,(y(t) - \mu) = my''(t)$$

This is same as

$$y''(t) = -\omega^2\,(y(t) - \mu) \qquad \text{where } \omega := \sqrt{\frac{k}{m}}$$

whose general solution is the sinusoid $y(t) = \mu + \alpha_1 \cos(\omega t) + \alpha_2 \sin(\omega t)$. In the context of this physical situation, the two different noisy sinusoid models ((6) and (5)) can be understood as follows. We are taking measurements of the displacement Y_t at various times t.

Model (6): Here our measurements are noisy and every measurement is corrupted by an unknown noise which we are terming ϵ_t and modeling as $N(0, \sigma^2)$.

Model (5): Here there is no measurement error and our measurement mechanism is perfect. However the actual oscillation of the mass is not perfectly sinusoidal and is affected by noise. For example, imagine, as Yule put it, that some kids are randomly throwing stones at the mass (sometimes from the left and sometimes from the right) while it is oscillating.

It is very interesting to note that observations generated from Model (5) are much smoother compared to observations generated from Model (6). Yule used this to argue that (5) is a better model for the sunspots data compared to (6).

2 The Autoregressive Model

Yule (1927) also fit models to the sunspots dataset that are more complicated compared to (5) and introduced the Autoregressive Model (of order 2) in this process. The AR(2) model is given by

$$Y_{t+2} = \phi_1 Y_{t+1} + \phi_2 Y_t + c + Z_{t+2}$$

Note that (5) can be seen as a simpler version of the above model where the ϕ_2 parameter is set to the value -1. More commonly, the model is written (by lowering the indices by 2) as:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t \qquad \text{with } Z_t \overset{\text{i.i.d}}{\sim} N(0, \sigma^2). \tag{7}$$

(7) can also be written as a usual regression model as:

$$Y_t = c + \phi_1 X_{t1} + \phi_2 X_{t2} + Z_t \qquad \text{with } Z_t \overset{\text{i.i.d}}{\sim} N(0, \sigma^2),$$

where the two covariates X_1 and X_2 are given by $X_{t1} = Y_{t-1}$ and $X_{t2} = Y_{t-2}$. Unlike in usual regression, here, the covariates are not given by separate variables but just the response variable lagged by one and two time units respectively. Because the regressions are just lagged versions of the response variable, this regression is called AUTOregression (auto referring to the fact that the response is regressed on lagged versions of itself).

As will be discussed next week, usual linear model fitting techniques can be employed to fit the model (7) to data. These will lead to parameter estimates $\hat{c}, \hat{\phi}_1, \hat{\phi}_2, \hat{\sigma}$. Using the fitted model, future values can be predicted by recursing the equation:

$$Y_t = \hat{c} + \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2} \qquad \text{for } t = T+1, T+2, \ldots$$

with Y_T and Y_{T-1} set to the observed values y_T and y_{T-1} respectively. For the sunspots data, these predictions seem to follow a *damped* sinusoid. Indeed, fitting the AR(2) model to the sunspots data for the time period $1700 - 1969$ led to the model:

$$Y_{t+2} = 24.11 + 1.38 Y_{t+1} - 0.69 Y_t + Z_{t+2}$$

which gives the prediction equation:

$$Y_t = 24.11 + 1.38 Y_{t-1} - 0.69 Y_{t-2}$$

for the future values of sunspot numbers from 1970 onwards. This equation can also be written as

$$Y_t - 77.77 = 1.38 \left(Y_{t-1} - 77.77\right) - 0.69 \left(Y_{t-2} - 77.77\right)$$

Thus the predictions for $U_t := Y_t - 77.77$ are given by recursing the equation:

$$U_t = 1.38 U_{t-1} - 0.69 U_{t-2} \tag{8}$$

for $t = T+1, T+2, \ldots$ (note that U_{T-1} and U_T are observed from the data). We shall see later that the general solution of (8) is of the form:

$$c_1 \left(1.2\right)^{-t} \cos\left(0.59t + c_2\right)$$

for two constants c_1 and c_2. The above is clearly a damped sinusoid (the sinusoid $\cos(0.59t + c_2)$ is damped by the factor $(1.2)^{-t}$).

3 Recommended Reading for Today

1. The paper Yule [1] is available freely online.
2. A very nice account of Yule's influential 1927 paper is Chapter 6 of the 2011 book "The Foundations of Modern Time Series Analysis" by T. C. Mills. (available for free from the library website).

References

[1] Yule, G. U. (1927). On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character 226*(636-646), 267–298.