STAT 153 & 248 - Time Series
Lecture Seven

Fall 2025, UC Berkeley

Aditya Guntuboyina

September 18, 2025

1 The Sinusoidal Model

By the sinusoidal model, we refer to:

$$y_t = \beta_0 + \beta_1 \cos(2\pi f t) + \beta_2 \sin(2\pi f t) + \epsilon_t \qquad \text{where } \epsilon_t \overset{\text{i.i.d}}{\sim} N(0, \sigma^2). \tag{1}$$

This is a simple model for time series data which show strong periodicity (like the sunspots data).

The unknown parameters in this model are $\beta_0, \beta_1, \beta_2, \sigma$ as well as the frequency parameter f. If f is assumed to be known, then clearly (1) is a multiple linear regression model:

$$y = X_f \beta + \epsilon$$

with

$$y = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} \text{ and } X_f = \begin{pmatrix} 1 & \cos(2\pi f(1)) & \sin(2\pi f(1)) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & \cos(2\pi f(n)) & \sin(2\pi f(n)) \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \text{ and } \epsilon = \begin{pmatrix} \epsilon_1 \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_n \end{pmatrix}$$

When f is unknown, this is a nonlinear regression model.

The function $t \mapsto \beta_0 + \beta_1 \cos(2\pi f t) + \beta_2 \sin(2\pi f t)$ is called a sinusoid. Before proceeding further, let us look at some basic properties and terminology related to sinusoids.

2 The Sinusoid

When we say sinusoid, we refer to the following function of time t:

$$s(t) := \beta_0 + R \cos(2\pi f t + \phi) \tag{2}$$

Here

- R is called the *amplitude*. It represents the height of the oscillation from its center line.
- f is called the *frequency*. It represents the number of oscillations in unit time. If time is measured in seconds, then the unit of f is Hertz (Hz).

- $1/f$ is called the *period*. It is the length of time to complete one full oscillation.
- ϕ is called the *phase*. Without ϕ (i.e., if $\phi = 0$), then the above sinusoid becomes $\beta_0 + R\cos(2\pi f t)$ so it starts at its maximum value at $t = 0$. Adding ϕ shifts the wave left or right in time. This captures the fact that different oscillations might 'start' at different points in their cycle.
- $2\pi f$ is called the *angular frequency*. Sometimes we use the notation $\omega = 2\pi f$ for the angular frequency. It measures the rate of change of the angle inside the cosine.

Using the formula $\cos(\alpha + \beta) = (\cos\alpha)(\cos\beta) - (\sin\alpha)(\sin\beta)$, we can represent the sinusoid (2) in the following equivalent alternative form:

$$s(t) = \beta_0 + \beta_1 \cos 2\pi f t + \beta_2 \sin 2\pi f t. \tag{3}$$

The parameters β_1, β_2 in (3) are related to R, ϕ in (2) via $\beta_1 = R\cos\phi$ and $\beta_2 = R\sin\phi$. While working with models involving sinusoids, we use the representation (3) because the parameters β_1 and β_2 appear linearly in (3).

3 Discrete sampling and restricting f to $[0, 1/2]$

Often in time series analysis, we work with equally spaced time points and assume that the time variable t takes the values $1, \dots, n$ (where n is the sample size). It turns out that if we consider the sinusoid (2) and restrict the time t to $1, \dots, n$, then we can always constrain the frequency parameter f to $[0, 1/2]$. This is a consequence of the following result.

Fact 3.1. *For every $f \in (-\infty, \infty)$ and $\phi \in (-\infty, \infty)$, there exists $f_0 \in [0, 1/2]$ and $\phi_0 \in (-\infty, \infty)$ such that*

$$s(t) = \beta_0 + R\cos(2\pi f t + \phi) = \beta_0 + R\cos(2\pi f_0 t + \phi_0) \qquad \text{for all } t = 1, \dots, n.$$

Proof. Consider the following three cases.

1. If $f < 0$, then we can write $\cos(2\pi f t + \phi) = \cos(2\pi(-f)t - \phi)$. Clearly, $-f \geq 0$.
2. If $f \geq 1$, then we write (below $[f]$ is the largest integer less than or equal to f):

   $$\cos(2\pi f t + \phi) = \cos(2\pi[f]t + 2\pi(f - [f])t + \phi) = \cos(2\pi(f - [f])t + \phi),$$

   because $\cos(\cdot)$ is periodic with period 2π. Clearly $0 \leq f - [f] < 1$.
3. If $f \in [1/2, 1)$, then

   $$\cos(2\pi f t + \phi) = \cos(2\pi t - 2\pi(1 - f)t + \phi) = \cos(2\pi(1 - f)t - \phi)$$

   because $\cos(2\pi t - x) = \cos x$ for all integers t. Clearly $0 < 1 - f \leq 1/2$.

Thus the sinusoid $R\cos(2\pi f t + \phi)$ equals $R\cos(2\pi f_0 t + \phi_0)$ at all integers t for some $0 \leq f_0 \leq 1/2$ and a phase ϕ_0 that is possibly different from ϕ. □

From now on, when we discuss sinusoids $s(t) = \beta_0 + R\cos(2\pi f t + \phi)$ in the context of $t = 1, \dots, n$, we shall assume that the frequency parameter f is restricted to $[0, 1/2]$. Note also the behavior of the sinusoid for the two frequency extremes $f = 0$ and $f = 1/2$. When

$f = 0$, the sinusoid $s(t)$ is simply a constant function equal to $\beta_0 + R\cos(\phi)$. When $f = 1/2$, we have

$$s(t) = \beta_0 + R\cos(\pi t + \phi) = \beta_0 + R(\cos\phi)\cos(\pi t) = \beta_0 + R(-1)^t \cos\phi.$$

This sinusoid exhibits the maximum possible oscillation going back and forth between $\beta_0 + R\cos\phi$ and $\beta_0 - R\cos\phi$.

4 Least Squares Estimation of β, f, σ

We use exactly the same method for parameter estimation as in the change of slope model. We take a bunch of possible values of f, calculate the goodness of fit $RSS(f)$ of the resulting linear regression model with fixed f, and then use $\hat{f}$ as the minimizer of $RSS(f)$ over f. The remaining parameters (β and σ) are estimated as in usual linear regression with f fixed at $\hat{f}$. This algorithm is described below:

1. Take a grid of all possible values of f in the range $[0, 1/2]$.
2. For each frequency value f in the grid,
    a) Form the matrix X_f
    b) Do a regression of y on X_f and compute the Residual Sum of Squares $RSS(f)$
3. Take $\hat{f}$ to be the grid value which minimizes $RSS(f)$ over all the grid values.
4. Take $\hat{\beta}$ and $\hat{\sigma}$ to be the usual regression estimates (of β and σ) in the linear regression of y on $X_{\hat{f}}$.

5 Bayesian Posterior

For Bayesian inference, we need to select the prior on β, σ and f. For β and σ, we shall use, as usual:

$$\beta_0, \beta_1, \beta_2, \log\sigma \overset{\text{i.i.d}}{\sim} \text{unif}(-C, C).$$

For f, we shall use:

$$f \sim \text{unif}[0, 1/2]$$

because, as seen in Section 3, we know we can restrict f to $[0, 1/2]$.

This will let us write the joint posterior of all the parameters β, σ, f, and then we integrate out β and σ to deduce the posterior of f. This calculation is exactly the same as in the last lecture when we studied the change of slope model; the only difference being that c there is now replaced by f (also the indicator $I\{1 < c < n\}$ should be replaced by $I\{0 \le f \le 1/2\}$). The posterior for f will be given by:

$$\text{posterior}(f) \propto I\{0 \le f \le 1/2\}|X_f^T X_f|^{-1/2} \left(\frac{1}{RSS(f)}\right)^{(n-p)/2}.$$

Thus psoterior will be evaluated numerically over a grid of values of f in the range $[0, 0.5]$. The term $|X_f^T X_f|^{-1/2}$ becomes infinite when $|X_f^T X_f| = 0$ i.e., when X_f does not have full column rank. This will be the case when $f = 0$ or $f = 1/2$. We will exclude these edge cases while computing this posterior:

$$\text{posterior}(f) \propto I\{0 < f < 1/2\}|X_f^T X_f|^{-1/2} \left(\frac{1}{RSS(f)}\right)^{(n-p)/2}. \tag{4}$$

6 Efficient Computation of $RSS(f)$

$RSS(f)$ describes how well the sinusoid with frequency f fits the observed data $y_1, \dots, y_n$. It is crucial to estimation of f in the model (1), and also for calculating the Bayesian posterior (4).

A plot of $RSS(f)$ over different frequencies f is also commonly used as an exploratory data analysis tool for identifying "which periodicities are present in the data". This tool is often used even when one is not interested in eventually fitting the simple model (1) to the observed data.

For computing $RSS(f)$, as discussed in the previous lecture, we need a grid of values for f. The most commonly used grid is given by:

$$\mathcal{F} := \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{[n/2]}{n}\right\} \tag{5}$$

where $[n/2]$ is the largest integer smaller than or equal to $n/2$.

A frequency of the form j/n where $j \in \{0, 1, 2, \dots, n-1\}$ and n is the observed data size is called a **Fourier Frequency**. So the grid (5) consists of all Fourier frequencies that are in the range $[0, 1/2]$.

The main reason for taking the grid to consist of Fourier Frequencies is that $RSS(f), f \in \mathcal{F}$ can be computed very efficiently (in time $O(n \log n)$) using a classical algorithm known as the Fast Fourier Transform (FFT). We explain the high level details behind this fact today (without going into the workings of the FFT algorithm).

6.1 Formula for $RSS(f)$ when f is a Fourier Frequency

When f is a Fourier Frequency, one can write down a more explicit formula for $RSS(f)$. For this, first note that for every f:

$$RSS(f) = \min_\beta \|y - X_f\beta\|^2 = \|y - X_f\hat{\beta}_f\|^2 \qquad \text{where } \hat{\beta}_f := (X_f^T X_f)^{-1} X_f^T y.$$

Thus

$$\begin{aligned} RSS(f) &= \|y - X_f\hat{\beta}_f\|^2 \\ &= (y - X_f\hat{\beta}_f)^T(y - X_f\hat{\beta}_f) = y^T y - \hat{\beta}_f^T X_f^T y - y^T X_f\hat{\beta}_f + \hat{\beta}_f^T X_f^T X_f \hat{\beta}_f. \end{aligned}$$

Plugging in $\hat{\beta}_f = (X_f^T X_f)^{-1} X_f^T y$ above, we get

$$\begin{aligned} RSS(f) &= y^T y - y^T X_f (X_f^T X_f)^{-1} X_f^T y - y^T X_f (X_f^T X_f)^{-1} X_f^T y \\ &\quad + y^T X_f (X_f^T X_f)^{-1} X_f^T X_f (X_f^T X_f)^{-1} X_f^T y \\ &= y^T y - y^T X_f (X_f^T X_f)^{-1} X_f^T y - y^T X_f (X_f^T X_f)^{-1} X_f^T y \\ &\quad + y^T X_f (X_f^T X_f)^{-1} X_f^T y \\ &= y^T y - y^T X_f (X_f^T X_f)^{-1} X_f^T y \end{aligned}$$

Now

$$X_f^T X_f = \begin{pmatrix} n & \sum_{t=1}^n \cos(2\pi f t) & \sum_{t=1}^n \sin(2\pi f t) \\ \sum_{t=1}^n \cos(2\pi f t) & \sum_{t=1}^n \cos^2(2\pi f t) & \sum_{t=1}^n \cos(2\pi f t)\sin(2\pi f t) \\ \sum_{t=1}^n \sin(2\pi f t) & \sum_{t=1}^n \cos(2\pi f t)\sin(2\pi f t) & \sum_{t=1}^n \sin^2(2\pi f t). \end{pmatrix}$$

Now suppose that $f \in (0, 0.5)$ and suppose that f is a Fourier frequency i.e., it is of the form $f = j/n$ for some integer j. Then it turns out that

$$\begin{aligned}
&\sum_{t=1}^{n} \cos(2\pi f t) = 0 \qquad \sum_{t=1}^{n} \sin(2\pi f t) = 0 \\
&\sum_{t=1}^{n} \cos^2(2\pi f t) = \frac{n}{2} \qquad \sum_{t=1}^{n} \sin^2(2\pi f t) = \frac{n}{2} \\
&\sum_{t=1}^{n} \cos(2\pi f t)\sin(2\pi f t) = 0
\end{aligned} \tag{6}$$

As a result, for such f,

$$X_f^T X_f = \begin{pmatrix} n & 0 & 0 \\ 0 & n/2 & 0 \\ 0 & 0 & n/2 \end{pmatrix} \quad \text{so that} \quad (X_f^T X_f)^{-1} = \begin{pmatrix} 1/n & 0 & 0 \\ 0 & 2/n & 0 \\ 0 & 0 & 2/n \end{pmatrix}$$

This gives

$$\begin{aligned}
&y^T X_f (X_f^T X_f)^{-1} X_f^T y \\
&= y^T \begin{pmatrix} 1 & \cos(2\pi f(1)) & \sin(2\pi f(1)) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & \cos(2\pi f(n)) & \sin(2\pi f(n)) \end{pmatrix} \begin{pmatrix} 1/n & 0 & 0 \\ 0 & 2/n & 0 \\ 0 & 0 & 2/n \end{pmatrix} \begin{pmatrix} 1 & \cdot & \cdot & \cdot & 1 \\ \cos(2\pi f(1)) & \cdot & \cdot & \cdot & \cos(2\pi f(n)) \\ \sin(2\pi f(1)) & \cdot & \cdot & \cdot & \sin(2\pi f(n)) \end{pmatrix} y \\
&= \begin{pmatrix} \sum_t y_t & \sum_t y_t \cos(2\pi f t) & \sum_t y_t \sin(2\pi f t) \end{pmatrix} \begin{pmatrix} 1/n & 0 & 0 \\ 0 & 2/n & 0 \\ 0 & 0 & 2/n \end{pmatrix} \begin{pmatrix} \sum_t y_t \\ \sum_t y_t \cos(2\pi f t) \\ \sum_t y_t \sin(2\pi f t) \end{pmatrix} \\
&= \frac{1}{n}\left(\sum_t y_t\right)^2 + \frac{2}{n}\left(\sum_t y_t \cos(2\pi f t)\right)^2 + \frac{2}{n}\left(\sum_t y_t \sin(2\pi f t)\right)^2 \\
&= n\bar{y}^2 + \frac{2}{n}\left(\sum_t y_t \cos(2\pi f t)\right)^2 + \frac{2}{n}\left(\sum_t y_t \sin(2\pi f t)\right)^2.
\end{aligned}$$

Therefore for Fourier frequencies in the range $(0, 0.5)$, we get

$$RSS(f) = y^T y - n\bar{y}^2 - \frac{2}{n}\left(\sum_t y_t \cos(2\pi f t)\right)^2 - \frac{2}{n}\left(\sum_t y_t \sin(2\pi f t)\right)^2$$

or equivalently

$$RSS(f) = \sum_t (y_t - \bar{y})^2 - \frac{2}{n}\left(\sum_t y_t \cos(2\pi f t)\right)^2 - \frac{2}{n}\left(\sum_t y_t \sin(2\pi f t)\right)^2 \tag{7}$$

When $f = 0$ and $f = 1/2$, the above formula needs to be slightly modified. When $f = 0$, the sinusoidal model (1) simply becomes:

$$y_t = \beta_0 + \beta_1 + \epsilon_t \qquad \text{with } \epsilon_t \overset{\text{i.i.d}}{\sim} N(0, \sigma^2).$$

There is only one effective parameter coefficient parameter $(\beta_0 + \beta_1)$ here which will be estimated by $\bar{y}$ so that RSS becomes

$$RSS(0) = \sum_t (y_t - \bar{y})^2. \tag{8}$$

When $f = 1/2$ and n is even (when n is odd, $1/2$ cannot be a Fourier frequency so it will not be considered), the model (1) becomes

$$y_t = \beta_0 + \beta_1 \cos(\pi t) + \epsilon_t = \beta_0 + \beta_1 (-1)^t + \epsilon_t.$$

For this model, it is easy to check that

$$X^T X = \begin{pmatrix} n & 0 \\ 0 & n \end{pmatrix}$$

so that

$$RSS(1/2) = \sum_t (y_t - \bar{y})^2 - \frac{1}{n}\left(\sum_t y_t (-1)^t\right)^2 \tag{9}$$

Let us not worry too much about the edge cases $f = 0$ and $f = 1/2$, and focus on the formula (7). Note again that this formula holds whenever $f \in (0, 0.5)$ and f is a Fourier frequency (i.e., nf is an integer).

6.2 Proof of the identities in (6)

Note that $0 < f < 1/2$ and that nf is an integer.

$$\begin{aligned}
&\sum_{t=1}^{n} \cos(2\pi f t) \\
&= \frac{1}{2}\sum_{t=1}^{n} e^{2\pi i f t} + \frac{1}{2}\sum_{t=1}^{n} e^{-2\pi i f t} \\
&= \frac{1}{2}\frac{e^{2\pi i f}}{e^{2\pi i f} - 1}\left(e^{2\pi i n f} - 1\right) + \frac{1}{2}\frac{e^{-2\pi i f}}{e^{-2\pi i f} - 1}\left(e^{-2\pi i n f} - 1\right) \\
&= \frac{1}{2}\frac{e^{2\pi i f}}{e^{2\pi i f} - 1}\left(\cos(2\pi n f) - 1 + i \sin(2\pi n f)\right) + \frac{1}{2}\frac{e^{-2\pi i f}}{e^{-2\pi i f} - 1}\left(\cos(2\pi n f) - 1 - i \sin(2\pi n f)\right) \\
&= 0
\end{aligned}$$

because $\cos(2\pi n f) = \cos(2\pi(\text{integer})) = 1$ and $\sin(2\pi n f) = \sin(2\pi(\text{integer})) = 0$.

Similarly

$$\begin{aligned}
&\sum_{t=1}^{n} \sin(2\pi f t) \\
&= \frac{1}{2i}\sum_{t=1}^{n} e^{2\pi i f t} - \frac{1}{2i}\sum_{t=1}^{n} e^{-2\pi i f t} \\
&= \frac{1}{2i}\frac{e^{2\pi i f}}{e^{2\pi i f} - 1}\left(e^{2\pi i n f} - 1\right) - \frac{1}{2i}\frac{e^{-2\pi i f}}{e^{-2\pi i f} - 1}\left(e^{-2\pi i n f} - 1\right) \\
&= \frac{1}{2i}\frac{e^{2\pi i f}}{e^{2\pi i f} - 1}\left(\cos(2\pi n f) - 1 + i \sin(2\pi n f)\right) - \frac{1}{2i}\frac{e^{-2\pi i f}}{e^{-2\pi i f} - 1}\left(\cos(2\pi n f) - 1 - i \sin(2\pi n f)\right) \\
&= 0
\end{aligned}$$

Next note

$$\sum_{t=1}^{n} \cos^2(2\pi f t) = \sum_{t=1}^{n} \frac{1 + \cos(4\pi f t)}{2} = \frac{n}{2} + \frac{1}{2}\sum_{t=1}^{n} \cos(4\pi f t)$$

and

$$\sum_{t=1}^{n} \sin^2(2\pi ft) = \sum_{t=1}^{n} \frac{1-\cos(4\pi ft)}{2} = \frac{n}{2} - \frac{1}{2}\sum_{t=1}^{n}\cos(4\pi ft).$$

The quantity $\sum_t \cos(4\pi ft)$ which appears in both the above terms turns out to be zero because

$$\begin{aligned}
&\sum_{t=1}^{n}\cos(4\pi ft) \\
&= \frac{1}{2}\sum_{t=1}^{n} e^{4\pi ift} + \frac{1}{2}\sum_{t=1}^{n} e^{-4\pi ift} \\
&= \frac{1}{2}\frac{e^{4\pi if}}{e^{4\pi if}-1}\left(e^{r\pi inf}-1\right) + \frac{1}{2}\frac{e^{-4\pi if}}{e^{-4\pi if}-1}\left(e^{-4\pi inf}-1\right) \\
&= \frac{1}{2}\frac{e^{4\pi if}}{e^{4\pi if}-1}\left(\cos(4\pi nf)-1+i\sin(4\pi nf)\right) + \frac{1}{2}\frac{e^{-4\pi if}}{e^{-4\pi if}-1}\left(\cos(4\pi nf)-1-i\sin(4\pi nf)\right) \\
&= 0
\end{aligned}$$

because $\cos(4\pi nf) = \cos(4\pi(\text{integer})) = 1$ and $\sin(4\pi nf) = \sin(4\pi(\text{integer})) = 0$.

Finally

$$\begin{aligned}
&\sum_{t=1}^{n}\cos(2\pi ft)\sin(2\pi ft) \\
&= \sum_{t=1}^{n}\sin(4\pi ft) \\
&= \frac{1}{2i}\sum_{t=1}^{n} e^{4\pi ift} - \frac{1}{2i}\sum_{t=1}^{n} e^{-4\pi ift} \\
&= \frac{1}{2i}\frac{e^{4\pi if}}{e^{4\pi if}-1}\left(e^{r\pi inf}-1\right) - \frac{1}{2i}\frac{e^{-4\pi if}}{e^{-4\pi if}-1}\left(e^{-4\pi inf}-1\right) \\
&= \frac{1}{2i}\frac{e^{4\pi if}}{e^{4\pi if}-1}\left(\cos(4\pi nf)-1+i\sin(4\pi nf)\right) - \frac{1}{2i}\frac{e^{-4\pi if}}{e^{-4\pi if}-1}\left(\cos(4\pi nf)-1-i\sin(4\pi nf)\right) \\
&= 0
\end{aligned}$$

The intermediate calculations above include the terms $e^{2\pi if}-1$ and $e^{4\pi if}-1$ in the denominators. We need to make sure that these terms are not zero (otherwise the above proofs would not be valid).

$$e^{2\pi if} - 1 = \cos(2\pi f) - 1 + i\sin(2\pi f)$$

which cannot be zero because $\cos(2\pi f) < 1$ for $f \in (0, 0.5)$, and

$$e^{4\pi if} - 1 = \cos(4\pi f) - 1 + i\sin(4\pi f)$$

which also cannot be zero because $\cos(4\pi f) < 1$ for $f \in (0, 0.5)$.

6.3 The Periodogram

Given a time series dataset $y_1, \dots, y_n$, its periodogram is the function $I(f), 0 < f < 1/2$, defined as follows:

$$I(f) := \frac{1}{n}\left(\sum_{t=1}^{n} y_t\cos(2\pi ft)\right)^2 + \frac{1}{n}\left(\sum_{t=1}^{n} y_t\sin(2\pi ft)\right)^2 \qquad \text{for } f \in (0, 0.5)$$

From the formula (7), we have

$$RSS(f) = \sum_t (y_t - \bar{y})^2 - 2I(f) \qquad \text{when } f \in (0, 0.5) \text{ is a Fourier Frequency.}$$

The periodogram $I(f)$ can be written in the following alternative way:

$$I(f) = \frac{1}{n} \left| \sum_{t=1}^{n} y_t e^{-2\pi i f t} \right|^2$$

where $|\cdot|$ denotes complex modulus. As we shall discuss in detail next lecture,

$$\sum_{t=1}^{n} y_t \exp(-2\pi i f t)$$

(when f is a Fourier frequency) is closely related to the Discrete Fourier Transform (DFT) of $y_1, \ldots, y_n$. The DFT can be efficiently computed using the Fast Fourier Transform (FFT) algorithm. This gives a way of computing $I(f)$ and $RSS(f)$ at Fourier frequencies efficiently.