STAT 153 - Introduction to Time Series
Lecture Five

Fall 2022, UC Berkeley

Aditya Guntuboyina

September 9, 2022

In the last class, we discussed the problem of fitting a single sinusoid to time series data. We derived the Bayesian Posterior for the simple model:

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

given observed data $(t_1, y_1), \dots, (t_n, y_n)$. We noted that this posterior is usually highly peaked and that it is related to the Periodogram. In this lecture, we shall first define the periodogram more formally in terms of the Discrete Fourier Transform (DFT). We shall come back to this in the next class for some additional insights into the connection between the Bayesian posterior and the periodogram.

While discussing the DFT and the periodogram, it is customary to assume that the observed time points $t_1, \dots, t_n$ are regularly spaced. We shall make this assumption and denote the times by $0, \dots, n-1$. The observed time series values will then be denoted by $y_0, y_1, \dots, y_{n-1}$. Note that the length of the observed time series is still n.

Let us start by formally introducing the Sinusoid.

1 The Sinusoid

When we say sinusoid, we refer to a function of time (t) that can be represented in the following three ways:

1. $R\cos(2\pi f t + \Phi)$. In this representation, R is called the *amplitude*, f is called the *frequency* and Φ is called the *phase*. The quantity $1/f$ is called the *period* and $2\pi f$ is termed the *angular frequency*. We shall use the notation $\omega = 2\pi f$ for the angular frequency.

2. The sinusoid can also be written as $A\cos 2\pi f t + B \sin 2\pi f t$ where $A = R\cos\Phi$ and $B = -R\sin\Phi$.

3. Yet another way of representing the sinusoid is to use complex exponentials:

$$\exp(2\pi i f t) = \cos(2\pi f t) + i \sin(2\pi f t).$$

Therefore

$$\cos(2\pi f t) = \frac{\exp(2\pi i f t) + \exp(-2\pi i f t)}{2} \text{ and } \sin(2\pi f t) = \frac{\exp(2\pi i f t) - \exp(-2\pi i f t)}{2i}.$$

Thus $A\cos 2\pi f t + B \sin 2\pi f t$ can also be written as a linear combination of $\exp(2\pi i f t)$ and $\exp(-2\pi i f t)$.

2 Discrete Sampling

We shall assume that the time variable t takes the discrete values $t = 0, 1, \dots, n-1$. It turns out that when we consider the sinusoid:

$$x_t := R\cos(2\pi f t + \phi)$$

and restrict the time t to $0, 1, \dots, n-1$, then we can write

$$x_t = R\cos(2\pi f t + \phi) = R\cos(2\pi f_0 t + \phi_0) \tag{2}$$

for some other frequency f_0 that lies between 0 and 1/2 and some ϕ_0. This frequency f_0 is said to be an **alias** of f. From now on, whenever we speak of a sinusoid of frequency f, we shall assume that the frequency $0 \le f \le 1/2$.

The reason for (2) is the following:

1. If $f < 0$, then we can write $\cos(2\pi f t + \phi) = \cos(2\pi(-f)t - \phi)$. Clearly, $-f \ge 0$.
2. If $f \ge 1$, 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 $0 \le f - [f] < 1$.
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 $0 < 1 - f \le 1/2$.

Thus the sinusoid $R\cos(2\pi f t + \phi)$ can be written as $R\cos(2\pi f_0 t + \phi_0)$ with $0 \le f_0 \le 1/2$ and a phase ϕ_0 that is possibly different from ϕ. From now on, whenever we consider the sinusoid $R\cos(2\pi f t + \phi)$ in the regularly spaced time context, we assume that $0 \le f \le 1/2$.

If $\phi = 0$, then we have $x_t = R\cos(2\pi f t)$. When $f = 0$, then $x_t = R$ and so there is no oscillation in the data at all. When $f = 1/2$, then $x_t = R\cos(\pi t) = R(-1)^t$ and so $f = 1/2$ corresponds to the maximum possible oscillation.

We can represent $x_t = R\cos(2\pi f t + \phi)$ in terms of complex exponentials as

$$x_t = \frac{Re^{i\phi}}{2} e^{2\pi i f t} + \frac{Re^{-i\phi}}{2} e^{-2\pi i f t}.$$

We therefore have to deal with $-f$ as well here (because of the second term $e^{-2\pi i f t} = e^{2\pi i(-f)t}$) and $-f$ lies between $-1/2$ and 0. Thus when discussing sinusoids in terms of complex exponentials $e^{2\pi i f t}, t = 0, 1, \dots, n-1$, we take $f \in (-0.5, 0.5]$ (note that $f = -0.5$ leads to the same $e^{2\pi i f t}$ as $f = 0.5$ so we drop $f = -0.5$ from consideration). If one does not want to deal with negative frequencies, then we can use

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

because $\cos(2\pi(1-f)t) = \cos(2\pi t - 2\pi f t) = \cos(2\pi f t)$ (note t is an integer) and $\sin(2\pi(1-f)t) = \sin(2\pi t - 2\pi f t) = -\sin(2\pi f t)$.

Therefore, if we want to use complex exponentials $e^{2\pi i f t}$ but we do not want to deal with negative frequencies, then we can restrict f to $[0, 1)$. From here on:

1. Whenever we consider the real sinusoid $x_t = R\cos(2\pi f t + \phi)$ for $t = 0, 1, \ldots, n-1$, we restrict $f \in [0, 1/2]$. $f = 0$ corresponds to the constant R while $f = 1/2$ corresponds to $R(-1)^t$. This sinusoid oscillates according to the frequency f: smallest possible oscillation is when $f = 0$ and largest possible oscillation is for $f = 0.5$.
2. Whenever we consider the complex sinusoid $x_t = e^{2\pi i f t}$ for $t = 0, 1, \ldots, n-1$, we restrict $f \in [0, 1)$.

3 Fourier Frequencies

Consider the problem of guessing the frequency of a sinusoid from observed values sampled at n regularly spaced time points. This problem is quite easy and can be done just visually when the frequency of the sinusoid is of the form $f = j/n$ for some integer $0 \le j \le n-1$. Such frequencies are known as Fourier Frequencies. More precisely, Fourier Frequencies for sample size n is are $0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}$. The corresponding Angular Fourier Frequencies are $0, \frac{2\pi}{n}, \frac{4\pi}{n}, \ldots, \frac{2(n-1)\pi}{n}$.

Note that Fourier Frequencies change with n. For example, the frequency $f = \frac{5.5}{10} = \frac{55}{100}$ is not a Fourier Frequency when $n = 10$ but it is a Fourier Frequency when $n = 100$.

4 Sinusoidal Vectors at Fourier Frequencies

For every $0 \le j \le (n-1)$, let us define the $n \times 1$ vector

$$u^j = (1, \exp(2\pi i j/n), \exp(2\pi i 2j/n), \ldots, \exp(2\pi i (n-1)j/n))^T.$$

This vector can be interpreted as the complex sinusoid $e^{2\pi i f t}$ with Fourier frequency $f = j/n$ evaluated at the time points $t = 0, 1, \ldots, (n-1)$. It is easy to see that

1. When $j = 0$, we have $u^0 = (1, 1, \ldots, 1)$.
2. When $1 \le j \le n-1$, we have $u^j = \overline{u^{n-j}}$. Here $\bar{u}$ denotes complex conjugate of u (the complex conjugate $\bar{u}$ of a vector u is defined as the vector obtained by taking the complex conjugates of each entry of u).

The most important property of these complex valued vectors $u^0, u^1, \ldots, u^{n-1}$ is **orthogonality**. Specifically, for $0 \le j \ne k \le n-1$, we have

$$\left\langle u^j, u^k \right\rangle = 0. \tag{3}$$

Recall that the inner product between two complex valued vectors $a = (a_1, \ldots, a_n)^T$ and $b = (b_1, \ldots, b_n)^T$ is given by

$$\langle a, b \rangle = \sum_{j=1}^{n} a_j \bar{b}_j.$$

Note specially the complex conjugate of b_j above.

Here is the proof of (3). Fix $0 \leq j \neq k \leq n-1$ and write

$$\begin{aligned}
\left\langle u^j, u^k \right\rangle &= \sum_{t=0}^{n-1} \exp\left(2\pi i \frac{j}{n} t\right) \overline{\exp\left(2\pi i \frac{k}{n} t\right)} \\
&= \sum_{t=0}^{n-1} \exp\left(2\pi i \frac{j}{n} t\right) \exp\left(-2\pi i \frac{k}{n} t\right) \\
&= \sum_{t=0}^{n-1} \exp\left(2\pi i \frac{j-k}{n} t\right) \\
&= \sum_{t=0}^{n-1} \left[\exp\left(2\pi i \frac{j-k}{n}\right)\right]^t \\
&= \frac{1 - \left(\exp\left(2\pi i \frac{j-k}{n}\right)\right)^n}{1 - \exp\left(2\pi i \frac{j-k}{n}\right)} \\
&= \frac{1 - \exp\left(2\pi i (j-k)\right)}{1 - \exp\left(2\pi i \frac{j-k}{n}\right)} \\
&= \frac{1 - \cos\left(2\pi (j-k)\right) - i \sin\left(2\pi (j-k)\right)}{1 - \exp\left(2\pi i \frac{j-k}{n}\right)} = \frac{1-1-0}{1 - \exp\left(2\pi i \frac{j-k}{n}\right)} = 0
\end{aligned}$$

This proves (3). It is also easy to see that (just take $j = k$ in the above calculation and the answer can be found in the third line)

$$\left\langle u^j, u^j \right\rangle = \|u^j\|^2 = n.$$

Therefore the n complex-valued vectors $u^0, u^1, \ldots, u^{n-1}$ are orthogonal and they all have the same squared length equal to n. This immediately implies that they form a **basis** for the space $\mathbb{C}^n$ consisting of all complex-valued vectors of length n. In other words, every complex-valued vector of length n can be written as a linear combination of $u^0, u^1, \ldots, u^n$. This observation is the foundation for the definition of the Discrete Fourier Transform.

5 The Discrete Fourier Transform (DFT)

Because $u^0, \ldots, u^{n-1}$ form a basis, we can write any $n \times 1$ vector of complex entries:

$$x = (x_0, \ldots, x_{n-1})^T$$

as a linear combination of $u^0, \ldots, u^{n-1}$. More specifically, we can write

$$x = a_0 u^0 + a_1 u^1 + \cdots + a_{n-1} u^{n-1} \tag{4}$$

Take the inner product of both sides of the above equation with u^j for a fixed j and use orthogonality so that $\left\langle u^j . u^k \right\rangle = 0$ for $k \neq j$ and the fact that $\left\langle u^j, u^j \right\rangle = n$ to obtain

$$a_j = \frac{1}{n} \left\langle x, u^j \right\rangle = \frac{1}{n} \sum_{t=0}^{n-1} x_t \exp\left(-\frac{2\pi i j t}{n}\right). \tag{5}$$

We are now ready to define the Discrete Fourier Transform (DFT). The DFT of $x_0, x_1, \ldots, x_{n-1}$ is defined by

$$b_j := \left\langle x, u^j \right\rangle = \sum_{t=0}^{n-1} x_t \exp\left(-\frac{2\pi i j t}{n}\right). \tag{6}$$

More specifically, the n (possibly) complex numbers $b_0, b_1, \ldots, b_{n-1}$ are collectively called the DFT of $x_0, \ldots, x_{n-1}$. Typically, $x_0, \ldots, x_{n-1}$ will represent observed time series data. It is important to note that even though $x_0, \ldots, x_{n-1}$ are real-valued, their DFT $b_0, \ldots, b_{n-1}$ can be complex-valued.

Here are some basic things to note about the DFT:

1. b_0 is always equal to $x_0 + \cdots + x_{n-1}$. To see this, just plug in $j = 0$ in (6).

2. For each $j = 1, \ldots, n-1$, the DFT term b_{n-j} equals the complex conjugate of b_j:
$$b_{n-j} = \bar{b}_j. \tag{7}$$
The reason for the above is
$$b_{n-j} = \sum_t x_t \exp\left(-\frac{2\pi i(n-j)t}{n}\right) = \sum_t x_t \exp\left(\frac{2\pi ijt}{n}\right) \exp(-2\pi it) = \bar{b}_j,$$
where, in the above, we used that $\exp(-2\pi it) = 1$ (because t is an integer) and that $\exp\left(\frac{2\pi ijt}{n}\right)$ is the complex conjugate of $\exp\left(-\frac{2\pi ijt}{n}\right)$. Note that, for the above argument, it is crucial that $x_0, \ldots, x_{n-1}$ are real. If some of $x_0, \ldots, x_{n-1}$ are complex, the relation (7) is no longer true.

3. The DFT of $x = (x_0, \ldots, x_{n-1})^T$ can be obtained in `R` using the command `fft(x)`. Here `fft` stands for Fast Fourier Transform which is a special efficient algorithm for computing the DFT (we will not be going over the details of the FFT algorithm).

4. Because of (4) and (5), we can figure out the data $x_0, \ldots, x_{n-1}$ given knowledge of the DFT coefficients $b_0, \ldots, b_{n-1}$. Specifically, we have
$$x_t = \frac{1}{n} \sum_{j=0}^{n-1} b_j \exp\left(\frac{2\pi ijt}{n}\right)$$
for $t = 0, 1, \ldots, n-1$. The above relation is sometimes known as the Inverse Fourier Transform.

Because of (7), the DFT coefficients for later indices j are determined as the complex conjugates for the DFT coefficients for earlier indices. For example, when $n = 11$, the DFT can be written as:
$$b_0, b_1, b_2, b_3, b_4, b_5, \bar{b}_5, \bar{b}_4, \bar{b}_3, \bar{b}_2, \bar{b}_1,$$
and, for $n = 12$, it is
$$b_0, b_1, b_2, b_3, b_4, b_5, b_6 = \bar{b}_6, \bar{b}_5, \bar{b}_4, \bar{b}_3, \bar{b}_2, \bar{b}_1.$$
Note that when $n = 12$, the term b_6 is necessarily real because $b_6 = \bar{b}_6$. The data and their DFT provide equivalent information. When $n = 11$, the data consists of 11 real numbers while the DFT consists of one real number (b_0) and 5 complex numbers. On the other hand, when $n = 12$, the data consists of 12 real numbers while the DFT consists of two real numbers (b_0 and b_6) and 5 complex numbers.

6 The Periodogram

The Periodogram is a way of visualizing the DFT. The DFT consists of complex numbers so it is difficult to visualize it directly. The common visualization consists of looking at the squared absolute values of the DFT. More precisely, the periodogram is defined by

$$I\left(\frac{j}{n}\right) := \frac{|b_j|^2}{n} \qquad \text{for } 0 < \frac{j}{n} \leq \frac{1}{2}.$$

One visualizes the size of the DFT terms by plotting the periodogram. Note that $j = 0$ is not plotted as b_0 is simply the sum of the data values and does not provide any information on the sinusoidal components present in the data. Further, the periodogram $I(j/n)$ is only plotted for $(j/n) \leq 0.5$ as b_j's for larger values of j are determined by earlier b_j's by complex conjugacy.

7 Basis consisting of Real Sinusoids

Even though our datasets $x_0, \dots, x_{n-1}$ consisting of real numbers, we work with complex exponentials for defining the DFT because of certain convenience in calculations. However it is possible to do the whole analysis only with real sinusoidal basis. This can be done as follows. Define

$$\mathbf{c}^j := (1, \cos(2\pi j/n), \cos(2\pi 2j/n), \dots, \cos(2\pi(n-1)j/n))^T$$

and

$$\mathbf{s}^j := (0, \sin(2\pi j/n), \sin(2\pi 2j/n), \dots, \sin(2\pi(n-1)j/n))^T.$$

These are the vectors obtained by evaulating $\cos(2\pi ft)$ and $\sin(2\pi ft)$ with Fourier Frequency $f = j/n$ at time points $t = 0, 1, \dots, (n-1)$.

Because while considering real sinusoids sampled at $t = 0, 1, \dots, n-1$, we never have to deal with frequencies outside the range $[0, 0.5]$, we shall assume now that $0 \leq j \leq \left[\frac{n}{2}\right]$ (here $\left[\frac{n}{2}\right]$ is the smallest integer less than or equal to $n/2$). Further it is easy to see that $\mathbf{s}^j$ equals the zero vector when j/n equals 0 or $1/2$. Thus when n is even, we are looking at the vectors:

$$\mathbf{c}^0, \mathbf{c}^1, \mathbf{s}^1, \dots, \mathbf{c}^{\frac{n}{2}-1}, \mathbf{s}^{\frac{n}{2}-1}, \mathbf{c}^{\frac{n}{2}}.$$

When n is odd, we are looking at

$$\mathbf{c}^0, \mathbf{c}^1, \mathbf{s}^1, \dots, \mathbf{c}^{\frac{n-1}{2}}, \mathbf{s}^{\frac{n-1}{2}}.$$

In either case, the total number of vectors equals n. It can be shown that these sets of vectors are orthogonal and hence form a basis for the n-dimensional vector space of real-valued vectors (Exercise: use our orthogonality result for u^j to prove this).