Unit 17: Introduction to Spectral Analysis

Taylor Brown

Department of Statistics, University of Virginia

Fall 2020

Readings for Unit 17

Textbook chapter 4.1.

Motivation

There are two primary approaches to time series. One is the **time domain** approach, which we covered in Units 8 to 16. This approach focuses on the rules for a time series to move forward.

The other approach is the **frequency domain** approach. This approach tries to understand how differing oscillations can contribute to current observations.

1. Introduction to Spectral Analysis
2. Aliasing
3. Periodic Time Series

Time Domain Approach

Time domain approach: models which give an explicit formula for the current observation in terms of past observations and past white noise terms. “Regression of the present on the past.”

Frequency Domain Approach

Frequency domain approach: current observation as a combination of waves. “Regression of the current time on sines and cosines of various frequencies.”

Idea: decompose a stationary time series $\{x_t\}$ into a **combination** of sinusoids, with random and uncorrelated coefficients. This is also referred to **spectral analysis**.

Spectral Analysis

- Identify **dominant frequencies** within the data.
- Periodogram: **sample variance** at different of frequencies.
- Power spectrum: **population** version of the periodogram.

Spectral Analysis

Period and frequency are inversely related.

With quarterly data, there are four data points per year (cycle). The frequency (ω) 0.25 cycles per data point.

The period is

$$\frac{1}{\omega}$$

or in this case, 4.

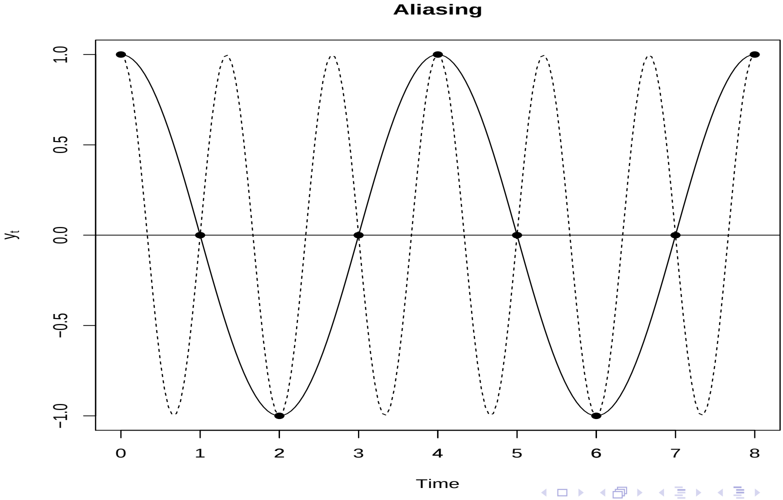
1. Introduction to Spectral Analysis
2. Aliasing
3. Periodic Time Series

Aliasing

- When $\omega = 1$, the time series makes one cycle per time unit.
- When $\omega = 0.5$, the time series makes one cycle every two time units.
- When $\omega = 0.25$, the time series makes one cycle every four time units.

Aliasing

Consider cosine curves with $\omega = \frac{1}{4}$ (bold) and $\omega = \frac{3}{4}$ (dashed).



Aliasing

Notice that at the discrete time points $0, 1, 2, 3, \cdots$ the two cosine curves have identical values. With **discrete-time** observations, we would not be able to **distinguish** between the two curves. So, the frequencies $\frac{1}{4}$ and $\frac{3}{4}$ are **aliased** with one another.

Aliasing

This is why we typically focus on frequencies between 0 and 0.5.

Higher frequencies may still be present, but they will appear as lower frequencies.

1. Introduction to Spectral Analysis

2. Aliasing

3. Periodic Time Series

Periodic Time Series

In Unit 6, we discussed periodic functions on the integers that have the following form

$$x(t) = A\cos(2\pi\omega t + \phi)$$

for $t = 0, \pm 1, \pm 2, \cdots$ where ω is the frequency, A is the amplitude, and ϕ is the phase.

Some Trigonometric Identities

$$
\begin{aligned}
\tan(\theta) &= \frac{\sin(\theta)}{\cos(\theta)}, \\
\sin^2(\theta) + \cos^2(\theta) &= 1, \\
\sin(a \pm b) &= \sin(a)\cos(b) \pm \cos(a)\sin(b), \\
\cos(a \pm b) &= \cos(a)\cos(b) \mp \sin(a)\sin(b).
\end{aligned}
$$

Periodic Time Series

Having the ϕ inside the cosine function can be problematic since if we want to do a regression, the ϕ makes this a non-linear regression. This issue is worked around using a trig identity

$$\cos(\alpha \mp \beta) = \cos(\alpha)\cos(\beta) \pm \sin(\alpha)\sin(\beta)$$

to rewrite the periodic function as

$$x_t = A\cos(\phi)\cos(2\pi\omega t) - A\sin(\phi)\sin(2\pi\omega t). \qquad (1)$$

Periodic Time Series

Now let's consider (1) differently from before. Re-write the periodic function (1) as

$$x_t = U_1 \cos(2\pi\omega t) + U_2 \sin(2\pi\omega t). \tag{2}$$

where $U_1 = A\cos(\phi)$ and $U_2 = -A\sin(\phi)$. We now assume U_1, U_2 are iid Gaussian with zero mean and fixed variance.

Periodic Time Series

Generalize (2) to include **multiple** frequencies and amplitudes with

$$x_t = \sum_{k=1}^{q} U_{k1} \cos(2\pi\omega_k t) + U_{k2} \sin(2\pi\omega_k t) \tag{3}$$

where the U_{k1} and the U_{k2} are independent and $N(0, \sigma_k^2)$ and the ω_k are distinct frequencies.

Periodic Time Series

A consequence of the representation given by (3) is that any **stationary** time series may be thought of, approximately, as the random superposition of sines and cosines oscillating at various frequencies.

Periodic Time Series

Let's derive the moments of (3).

$$
\begin{aligned}
E[x_t] &= E[\sum_{k=1}^{q} U_{k1}\cos(2\pi\omega_k t) + U_{k2}\sin(2\pi\omega_k t)] \\
&= \sum_{k=1}^{q} E[U_{k1}\cos(2\pi\omega_k t) + U_{k2}\sin(2\pi\omega_k t)] \\
&= \sum_{k=1}^{q} E[U_{k1}\cos(2\pi\omega_k t)] + E[U_{k2}\sin(2\pi\omega_k t)] \\
&= \sum_{k=1}^{q} E[U_{k1}]\cos(2\pi\omega_k t) + E[U_{k2}]\sin(2\pi\omega_k t) \\
&= 0
\end{aligned}
$$

Periodic Time Series

$$
\begin{aligned}
\gamma(h) &= \operatorname{Cov}\left(x_{t+h}, x_t\right) \\
&= \\
&= \\
&= \\
&= \\
&= \sum_{k=1}^{q} \sigma_k^2 \cos\left(2\pi\omega_k h\right)
\end{aligned}
$$

Periodic Time Series

$$\gamma(h) = \sum_{k=1}^{q} \sigma_k^2 \cos(2\pi\omega_k h)$$

When we plug in $h = 0$, we get

$$\text{Var}(x_t) = \sum_{k=1}^{q} \sigma_k^2$$

Periodic Time Series

Question: What is the implication of these derivations?

$$x_t = \sum_{k=1}^{q} U_{k1}\cos(2\pi\omega_k t) + U_{k2}\sin(2\pi\omega_k t) \tag{4}$$

where $U_{k1}^2 + U_{k2}^2 = A_k^2$. Also
$\mathrm{Var}[U_{k1} + U_{k2}] = E[U_{k1}^2] + E[U_{k2}^2] = 2\sigma_k^2$.
Think of it as a regression model:

$$x_t = \sum_{k=1}^{q} \beta_{k1}\cos(2\pi\omega_k t) + \beta_{k2}\sin(2\pi\omega_k t) \tag{5}$$

$\hat{\beta}_{k1}^2 + \hat{\beta}_{k2}^2$ is the scaled periodogram, it helps estimate σ_k^2 which is the population parameter corresponding to a realized A_k^2.

Periodic Time Series

Question: What is the implication of these derivations?

$$x_t = \sum_{k=1}^{q} U_{k1}\cos(2\pi\omega_k t) + U_{k2}\sin(2\pi\omega_k t) \tag{4}$$

where $U_{k1}^2 + U_{k2}^2 = A_k^2$. Also
$\mathrm{Var}[U_{k1} + U_{k2}] = E[U_{k1}^2] + E[U_{k2}^2] = 2\sigma_k^2$.
Think of it as a regression model:

$$x_t = \sum_{k=1}^{q} \beta_{k1}\cos(2\pi\omega_k t) + \beta_{k2}\sin(2\pi\omega_k t) \tag{5}$$

$\hat{\beta}_{k1}^2 + \hat{\beta}_{k2}^2$ is the scaled periodogram, it helps estimate σ_k^2 which is the population parameter corresponding to a realized A_k^2.

Periodic Time Series

Question: What is the implication of these derivations?

$$x_t = \sum_{k=1}^{q} U_{k1}\cos(2\pi\omega_k t) + U_{k2}\sin(2\pi\omega_k t) \qquad (4)$$

where $U_{k1}^2 + U_{k2}^2 = A_k^2$. Also
$\mathrm{Var}[U_{k1} + U_{k2}] = E[U_{k1}^2] + E[U_{k2}^2] = 2\sigma_k^2$.
Think of it as a regression model:

$$x_t = \sum_{k=1}^{q} \beta_{k1}\cos(2\pi\omega_k t) + \beta_{k2}\sin(2\pi\omega_k t) \qquad (5)$$

$\hat{\beta}_{k1}^2 + \hat{\beta}_{k2}^2$ is the scaled periodogram, it helps estimate σ_k^2 which is the population parameter corresponding to a realized A_k^2.

Next...

Computing the Fourier transform of the data is faster than fitting a linear regression.

Before we discuss that Fourier transform, we'll discuss what happens when you Fourier transform the autocovariance function. This is the "spectral density."