

Lecture 12

Spectral Analysis of Time Series I

Readings: CC08 Chapter 13-14; BD16 Ch 4; SS17 Chapter 4.1-4.4

MATH 8090 Time Series Analysis
Week 12

Background

The Periodogram and
Spectral Density

Spectral Estimation

Whitney Huang
Clemson University

Background

The Periodogram and
Spectral Density

Spectral Estimation

- 1 Background
- 2 The Periodogram and Spectral Density
- 3 Spectral Estimation

- Time domain methods [Box and Jenkins, 1970]:

- Regress present on past

Example: $Y_t = \phi Y_{t-1} + Z_t$, $|\phi| < 1$, $\{Z_t\} \sim \text{WN}(0, \sigma^2)$

- Capture dynamics in terms of “velocity”, “acceleration”, etc

- Frequency domain methods [Priestley, 1981]:

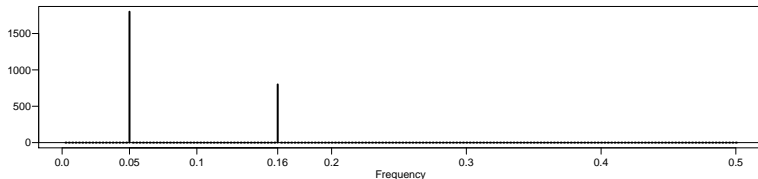
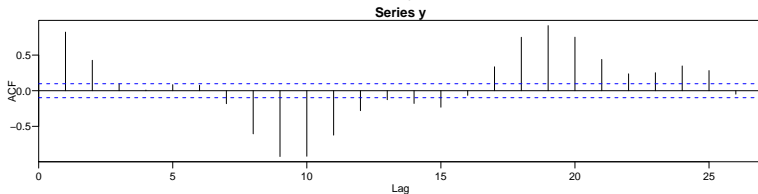
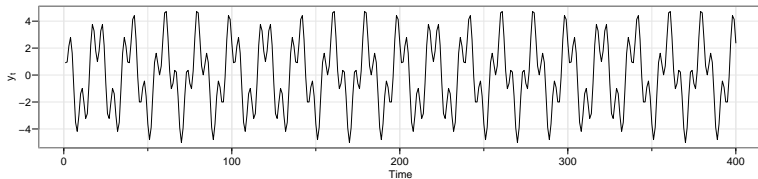
- Regress present on periodic sines and cosines

Example: $Y_t = \alpha_0 + \sum_{j=1}^p [\alpha_{1j} \cos(2\pi\omega_j t) + \alpha_{2j} \sin(2\pi\omega_j t)]$

- Capture dynamics in terms of **resonant frequencies**

Searching Hidden Periodicities

$$y_t = 3 \cos\left(2\pi\left(\frac{10}{200}\right)t\right) + 2 \cos\left(2\pi\left(\frac{32}{200}t + 0.3\right)\right)$$



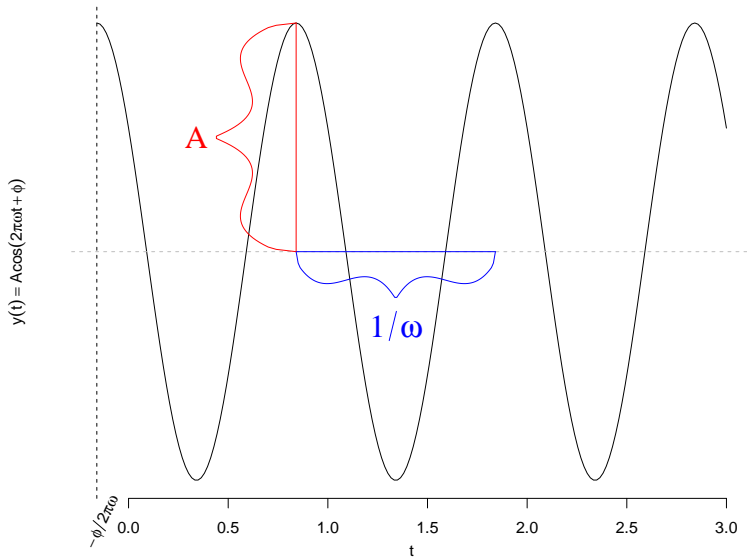
The simplest case is the **cosine wave**

$$\begin{aligned} Y_t &= A \cos(2\pi\omega t + \phi) \\ &= \alpha_1 \cos(2\pi\omega t) + \alpha_2 \sin(2\pi\omega t), \end{aligned}$$

where

- A is **amplitude**
- ω is **frequency**, in cycles per time unit
- ϕ is **phase**, determining the start point of the cosine function
- $\alpha_1 = A \cos(\phi)$, $\alpha_2 = -A \sin(\phi)$, $A = \sqrt{\alpha_1^2 + \alpha_2^2}$, $\phi = \tan^{-1} \frac{-\alpha_2}{\alpha_1}$

Graphical Illustration of the Cosine Wave



If

$$\begin{aligned}Y_t &= A \cos(2\pi\omega t + \phi) \\&= \alpha_1 \cos(2\pi\omega t) + \alpha_2 \sin(2\pi\omega t),\end{aligned}$$

and ϕ is random, uniformly distributed on $[-\pi, \pi)$, then:

$$\mathbb{E}(Y_t) = 0$$

$$\mathbb{E}(Y_{t+h}Y_t) = \frac{1}{2}A^2 \cos(2\pi\omega h)$$

$\Rightarrow Y_t$ is weakly stationary

Also

$$\mathbb{E}(\alpha_1) = \mathbb{E}(\alpha_2) = 0,$$

$$\mathbb{E}(\alpha_1^2) = \mathbb{E}(\alpha_2^2) = \frac{1}{2}A^2,$$

$$\text{and } \mathbb{E}(\alpha_1\alpha_2) = 0.$$

Alternatively, if the α 's have these properties, then Y_t is stationary with the same mean and autocovariances:

$$\mathbb{E}(Y_t) = 0,$$

$$\mathbb{E}(Y_{t+h}Y_t) = \frac{1}{2}A^2 \cos(2\pi\omega h).$$

Representing a Periodic Process as Multiple Sines and Cosines

More generally, if

$$Y_t = \sum_{k=1}^K [\alpha_{k,1} \cos(2\pi\omega_k t) + \alpha_{k,2} \sin(2\pi\omega_k t)],$$

where:

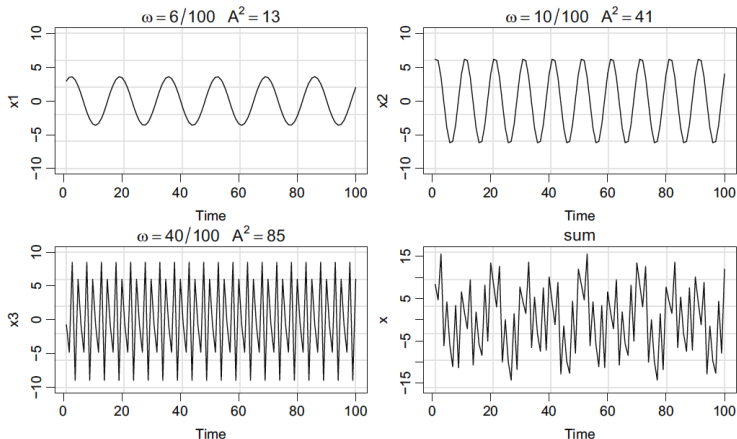
- The α 's are uncorrelated with zero mean;
- $\text{Var}(\alpha_{k,1}) = \text{Var}(\alpha_{k,2}) = \sigma_k^2$;

then Y_t is stationary with zero mean and autocovariances

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

$$\Rightarrow \gamma(0) = \text{Var}(Y_t) = \sum_{k=1}^K \sigma_k^2$$

Examples of Periodic Time Series



Source: Fig. 4.1. of Shumway and Stoffer, 2017

Folding Frequency and Aliasing

Let's consider $Y_{1,t} = \cos(2\pi(0.2)t)$ and $Y_{2,t} = \cos(2\pi(1.2)t)$

- At $t = 1$, $Y_{1,t} = \cos(0.4\pi t)$,
 $Y_{2,t} = \cos(2.4\pi t) = \cos(2\pi t + 0.4\pi t) = \cos(0.4\pi t) = Y_{1,t}$
- This is true for all integer values of t

$\Rightarrow \omega = 1.2$ is an **alias** of $\omega = 0.2$.

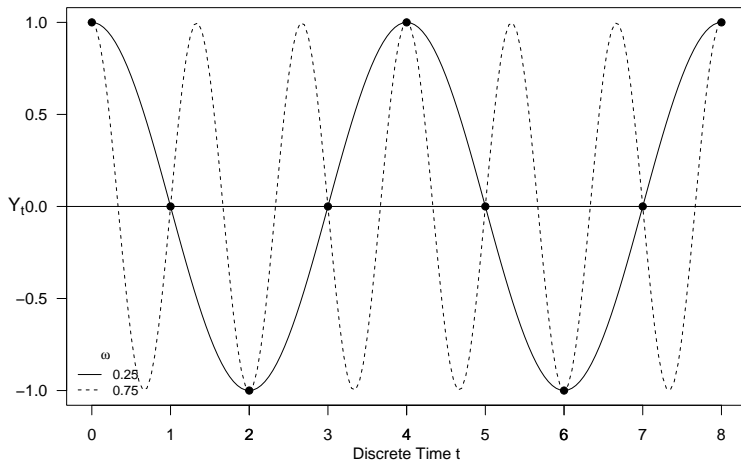
In general, all frequencies higher than $\omega = \frac{1}{2}$ have an alias
in $0 \leq \omega \leq \frac{1}{2}$

- $\omega = \frac{1}{2}$ is the **folding frequency** (aka Nyquist frequency),
because the shortest period that can be observed is $\frac{1}{\omega} = 2$.

Takeaway: It suffices to limit attention to $\omega \in [0, \frac{1}{2}]$

Illustration of Aliasing

$\omega = 0.25$ and $\omega = 0.75$ are **aliased** with one another



Background

The Periodogram and
Spectral Density

Spectral Estimation

Any time series sample y_1, y_2, \dots, y_n can be written

$$y_t = \alpha_0 + \sum_{j=1}^{(n-1)/2} [\alpha_j \cos(2\pi jt/n) + \beta_j \sin(2\pi jt/n)],$$

if n is odd; if n is even, an extra term is needed

- The (scaled) periodogram is

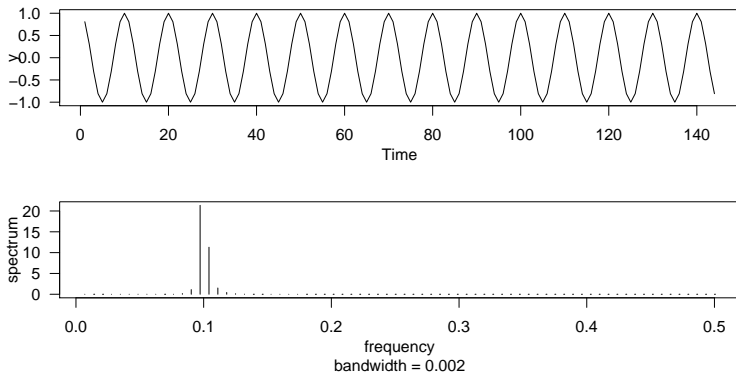
$$P(j/n) = \alpha_j^2 + \beta_j^2$$

the sample variance at each frequency component

- The R function `spectrum` can calculate and plot the periodogram

An Example: $Y_t = \cos(2\pi(0.1)t)$

```
y = cos(2 * pi * (0.1) * (1:144))  
ts.plot(y)  
spectrum(y, log = "no", main = "", type = "h")
```



Background

The Periodogram and
Spectral Density

Spectral Estimation

The Discrete Fourier Transform (DFT)

- Given data y_1, y_2, \dots, y_n , the discrete Fourier transform is

$$d(\omega_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^n y_t e^{-2\pi i \omega_j t}, \quad j = 0, 1, \dots, n-1.$$

- The frequencies $\omega_j = j/n$ are the **Fourier** or **fundamental frequencies**
- Like any other **Fourier transform**, it has an **inverse transform**:

$$y_t = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} d(\omega_j) e^{2\pi i \omega_j t}, \quad t = 1, 2, \dots, n$$

- The **periodogram** is $I(\omega_j) = |d(w_j)|^2$, $j = 0, 1, \dots, n-1$
- The scaled periodogram we used earlier is

$$P(\omega_j) = (4/n)I(\omega_j)$$

- In terms of sample autocovariances: $I(0) = n\bar{y}^2$, and for $j \neq 0$,

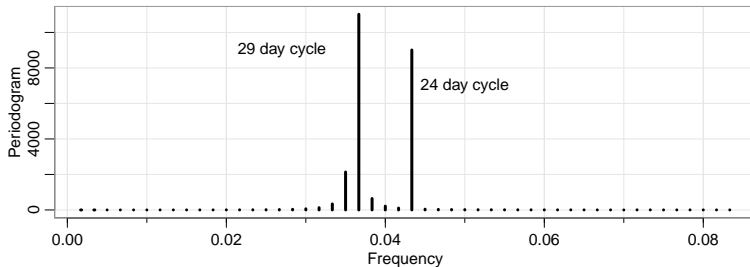
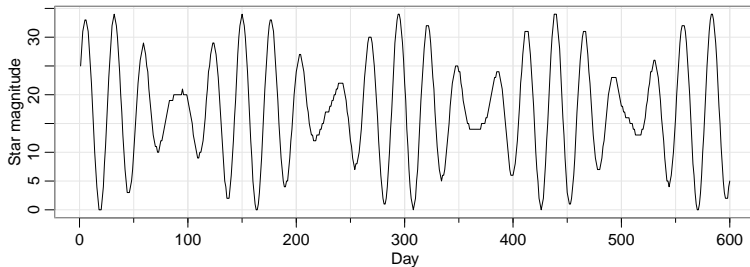
$$\begin{aligned} I(\omega_j) &= \sum_{h=-(n-1)}^{n-1} \hat{\gamma}(h) e^{-2\pi i \omega_j h} \\ &= \hat{\gamma}(0) + 2 \sum_{h=1}^{n-1} \hat{\gamma}(h) \cos(2\pi \omega_j h). \end{aligned}$$

Star Magnitude Example [Example 4.3, Shumway & Stoffer, 2017]

Background

The Periodogram and
Spectral Density

Spectral Estimation



- The periodogram shows which frequencies are strong in a finite sample $\{y_1, y_2, \dots, y_n\}$
- What about a population model for Y_t , such as a stationary time series?
- The **spectral density** plays the corresponding role

The Mathematics of the Spectrum

Every weakly stationary time series Y_t with autocovariances $\gamma(h)$ has a non-decreasing spectrum or spectral distribution function $F(\omega)$ for which

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i \omega h} dF(\omega) = 2 \int_0^{\frac{1}{2}} \cos(2\pi \omega h) dF(\omega).$$

If $F(\omega)$ is *absolutely continuous*, it has a spectral density function $f(\omega) = F'(\omega)$, and

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i \omega h} f(\omega) d\omega = 2 \int_0^{\frac{1}{2}} \cos(2\pi \omega h) f(\omega) d\omega$$

The autocovariance and the spectral distribution function contain the same information

Under various conditions on $\gamma(h)$, such as

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$$

$f(\omega)$ can be written as the sum

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega_j h} = \gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos(2\pi \omega_j h)$$

Properties of the spectral density:

- $f(\omega) \geq 0$;
- $f(-\omega) = f(\omega)$;
- $\int_{-\frac{1}{2}}^{\frac{1}{2}} f(\omega) d\omega = \gamma(0) < \infty$

Example: White Noise

For white noise $\{Z_t\}$, we have seen that $\gamma(0) = \sigma_Z^2$ and $\gamma(h) = 0$ for $h \neq 0$. Thus,

$$\begin{aligned} f(\omega) &= \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h} \\ &= \gamma(0) = \sigma_Z^2 \end{aligned}$$

That is, the spectral density is constant across all frequencies: each frequency in the spectrum contributes equally to the variance.

This is the origin of the name *white noise*: it is like white light, which is a uniform mixture of all frequencies in the visible spectrum

Background

The Periodogram and
Spectral Density

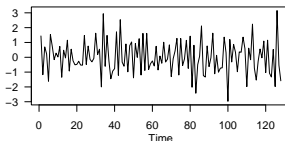
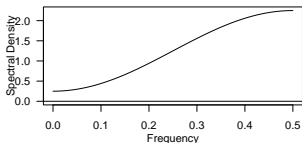
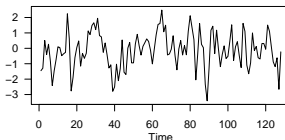
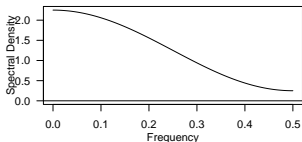
Spectral Estimation

Examples: MA(1)

An MA(1) process $Y_t = \theta Z_{t-1} + Z_t$ is a simple filtering of white noise. Therefore, we have the (power) transfer function of the MA filter is:

$$\begin{aligned}|1 + \theta e^{-2\pi i \omega}|^2 &= (1 + \theta e^{-2\pi i \omega})(1 + \theta e^{2\pi i \omega}) \\&= 1 + \theta^2 + \theta(e^{2\pi i \omega} + e^{-2\pi i \omega}) \\&= 1 + \theta^2 + 2\theta \cos(2\pi \omega).\end{aligned}$$

Thus, we have: $f(\omega) = [1 + \theta^2 + 2\theta \cos(2\pi \omega)] \sigma_Z^2$

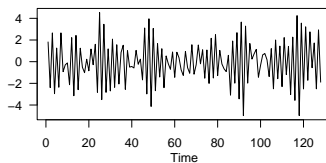
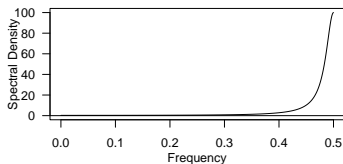
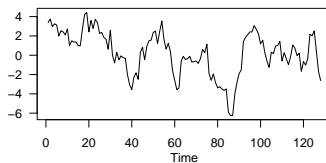
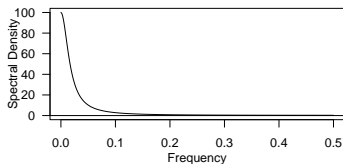


Example: AR(1)

For an AR(1) $Y_t = \phi Y_{t-1} + Z_t$, we have

$$[1 + \phi^2 - 2\phi \cos(2\pi\omega)] f(\omega) = \sigma_Z^2$$

Thus, we have: $f(\omega) = \frac{\sigma_Z^2}{1 + \phi^2 - 2\phi \cos(2\pi\omega)}$



Background

The Periodogram and
Spectral Density

Spectral Estimation

- **ARMA**: using results about linear filtering, we shall show that the spectral density of the ARMA(p, q) process

$$\phi(B)Y_t = \theta(B)Z_t$$

is

$$f(\omega) = \sigma_Z^2 \frac{|\theta(e^{-2\pi i\omega})|^2}{|\phi(e^{-2\pi i\omega})|^2}$$

- Note that this gives the characteristic polynomials $\phi(\cdot)$ and $\theta(\cdot)$ a concrete meaning: *they determine how strongly the series tends to fluctuate at each frequency*

Estimating Spectral Density Using Periodogram

If n is large

$$\begin{aligned}\mathbb{E}[I(\omega_j)] &\approx \sum_{h=-(n-1)}^{n-1} \gamma(h) e^{-2\pi i \omega_j h} \\ &\approx \sum_{-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega_j h} = \gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos(2\pi \omega_j h) \\ &= f(\omega_j) \text{ 😊}.\end{aligned}$$

- Heuristically, the spectral density is the approximate expected value of the periodogram
- Conversely, the periodogram can be used as an estimator of the spectral density
- But the periodogram values have only two degrees of freedom each, which makes it a poor estimate 😞

Recall: the discrete Fourier transform

$$d(\omega_j) = n^{-\frac{1}{2}} \sum_{t=1}^n y_t e^{-2\pi i \omega_j t}, \quad j = 0, 1, \dots, n-1,$$

and the periodogram

$$I(\omega_j) = |d(\omega_j)|^2, \quad j = 0, 1, \dots, n-1,$$

where ω_j is one of the Fourier frequencies

$$\omega_j = \frac{j}{n}.$$

Periodogram is the squared modulus of the DFT

For $j = 0, 1, \dots, n-1$

$$\begin{aligned}d(\omega_j) &= n^{-\frac{1}{2}} \sum_{t=1}^n y_t e^{-2\pi i \omega_j t} \\&= n^{-\frac{1}{2}} \sum_{t=1}^n y_t \cos(2\pi \omega_j t) - i \times n^{-\frac{1}{2}} \sum_{t=1}^n y_t \sin(2\pi \omega_j t) \\&= d_{\cos}(\omega_j) - i \times d_{\sin}(\omega_j).\end{aligned}$$

- $d_{\cos}(\omega_j)$ and $d_{\sin}(\omega_j)$ are the **cosine transform** and **sine transform**, respectively, of y_1, y_2, \dots, y_n
- The periodogram is $I(\omega_j) = d_{\cos}(\omega_j)^2 + d_{\sin}(\omega_j)^2$

Sampling Properties of the Periodogram

For convenience, suppose that n is odd: $n = 2m + 1$

- **White noise:** orthogonality properties of sines and cosines mean that

$d_{\cos}(\omega_1), d_{\sin}(\omega_1), d_{\cos}(\omega_2), d_{\sin}(\omega_2), \dots, d_{\cos}(\omega_m), d_{\sin}(\omega_m)$
have zero mean, variance $\frac{\sigma_Z^2}{2}$, and uncorrelated

- **Gaussian white noise:**

$d_{\cos}(\omega_1), d_{\sin}(\omega_1), d_{\cos}(\omega_2), d_{\sin}(\omega_2), \dots, d_{\cos}(\omega_m), d_{\sin}(\omega_m)$
are i.i.d. $N(0, \frac{\sigma_Z^2}{2})$

- So for Gaussian white noise

$$I(\omega_j) \sim \frac{\sigma_Z^2}{2} \times \chi_2^2$$

The periodogram is not a consistent estimator of the spectral density (why?)

General case:

$d_{\cos}(\omega_1), d_{\sin}(\omega_1), d_{\cos}(\omega_2), d_{\sin}(\omega_2), \dots, d_{\cos}(\omega_m), d_{\sin}(\omega_m)$,
have zero mean and are approximately uncorrelated, and

$$\text{Var}[d_{\cos}(\omega_j)] \approx \text{Var}[d_{\sin}(\omega_j)] \approx \frac{1}{2}f(\omega_j),$$

where $f(\omega_j)$ is the spectral density function

If Y_t is Gaussian,

$$\frac{I(\omega_j)}{\frac{1}{2}f(\omega_j)} = \frac{d_{\cos}(\omega_j)^2 + d_{\sin}(\omega_j)^2}{\frac{1}{2}f(\omega_j)} \approx \chi_2^2,$$

and $I(\omega_1), I(\omega_2), \dots, I(\omega_m)$ are approximately independent

The periodogram is not a consistent estimator!

Recall:

$$\frac{I(\omega_j)}{\frac{1}{2}f(\omega_j)} = \frac{d_{\cos}(\omega_j)^2 + d_{\sin}(\omega_j)^2}{\frac{1}{2}f(\omega_j)} \approx \chi_2^2,$$

and $I(\omega_1), I(\omega_2), \dots, I(\omega_m)$ are approximately independent

Problem: $I(\omega_j)$ is an approximately unbiased estimator of $f(\omega_j)$ but with too few degrees of freedom ($\text{df} = 2$) to be useful. Specifically, $I(\omega) \sim \frac{1}{2}f(\omega)\chi_2^2$, which implies

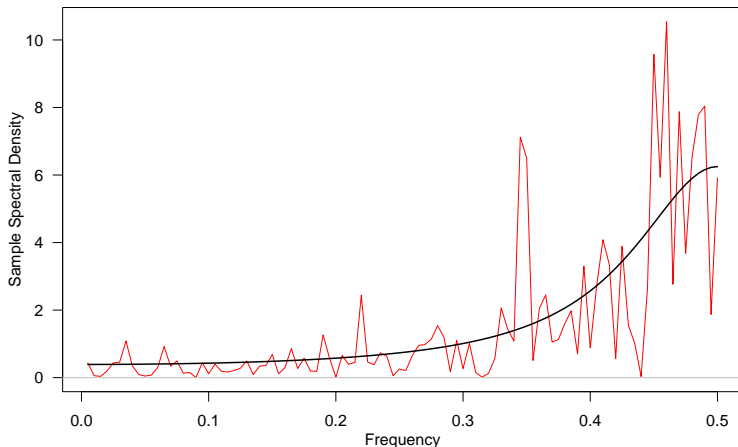
$$\mathbb{E}[I(\omega)] \approx f(\omega)$$

and

$$\text{Var}[I(\omega)] \approx f^2(\omega)$$

Consequently, $\text{Var}[I(\omega)] \stackrel{n \rightarrow \infty}{\neq} 0$ and thus the periodogram is not a consistent estimator of the spectral density

Smoothing the Periodogram

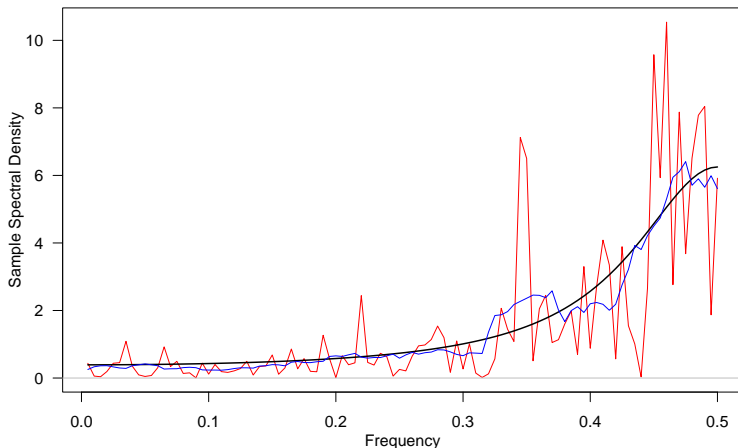


Main idea: “average” the values of the periodogram over “small” intervals of frequencies to reduce the estimation variability

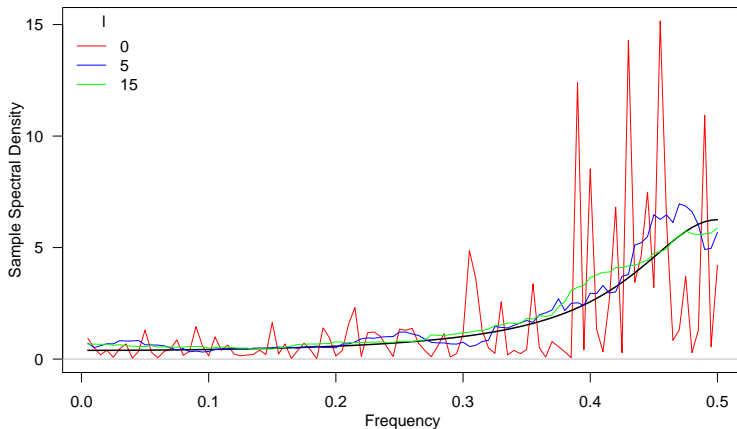
Averaged Periodogram [Daniell Spectral Window]

Use the band $[\omega_{j-l}, \omega_{j+l}]$ containing $L = 2l + 1$ Fourier frequencies:

$$\bar{f}(\omega_j) = \frac{1}{L} \sum_{k=-l}^l I(\omega_{j+k})$$



Tuning Parameter: l



A large l can effectively reduce the estimation variability but can also introduce bias

Bias and Variance

Let's assume the true spectral density does not change much locally, then a Taylor expansion produces

$$\begin{aligned}\mathbb{E}[\bar{f}(\omega)] &\approx \sum_{k=-l}^l W_l(k) f\left(\omega + \frac{k}{n}\right) \\ &\approx \sum_{k=-l}^l W_l(k) \left[f(\omega) + \frac{k}{n} f'(\omega) + \frac{1}{2} \left(\frac{k}{n}\right)^2 f''(\omega) \right] \\ &\approx f(\omega) + \frac{1}{n^2} \frac{f''(\omega)}{2} \sum_{k=-l}^l k^2 W_l(k)\end{aligned}$$

$$\text{Bias} \approx \frac{1}{n^2} \frac{f''(\omega)}{2} \sum_{k=-l}^l k^2 W_l(k)$$

$$\text{Variance} \approx f^2(\omega) \sum_{k=-l}^l W_l^2(k)$$

Example: for Daniell rectangular spectral window, we have

$$\text{bias} = \frac{2}{n^2(2l+1)} \left(\frac{l^3}{3} + \frac{l^2}{2} + \frac{l}{6} \right) \text{ and variance } \frac{1}{2l+1}$$

The distribution of $\frac{\nu \bar{f}(\omega)}{f(\omega)}$ can be approximated by $\chi_{df=\nu}^2$, where

$$\nu = \frac{2}{\sum_{k=-l}^l W_l^2(k)}$$

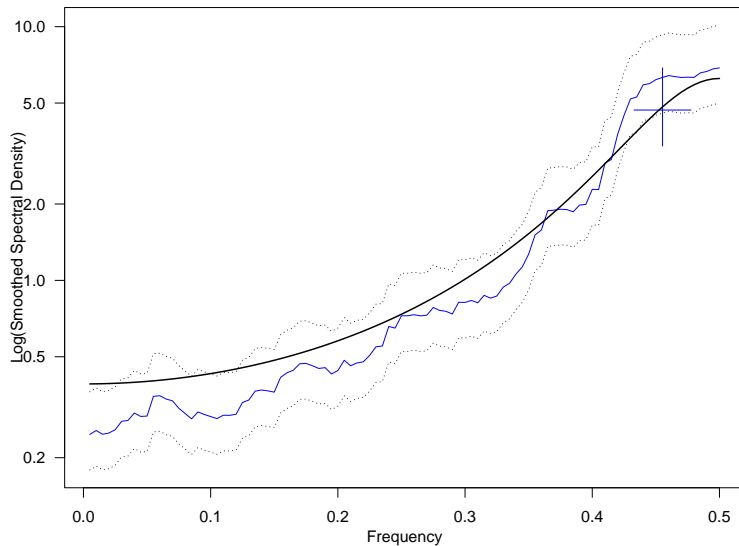
$\Rightarrow 100(1 - \alpha)\%$ CI for $f(\omega)$

$$\frac{\nu \bar{f}(\omega)}{\chi_{df=\nu, 1-\frac{\alpha}{2}}^2} < f(\omega) < \frac{\nu \bar{f}(\omega)}{\chi_{df=\nu, \frac{\alpha}{2}}^2}$$

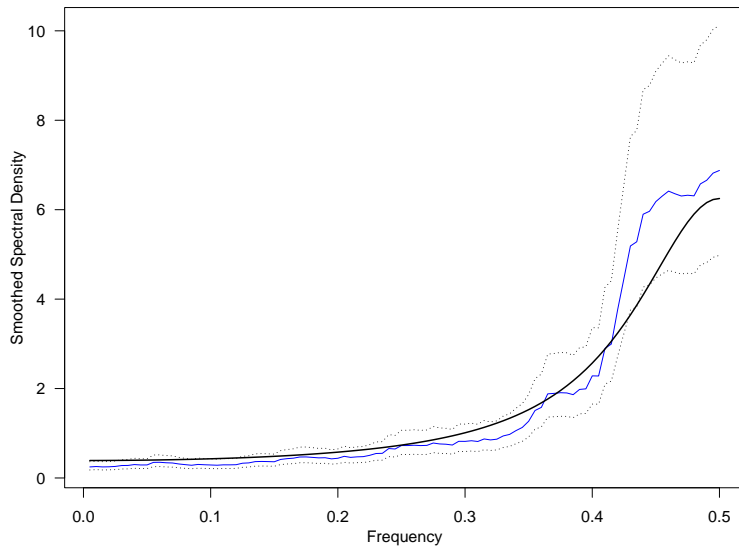
Taking logs we obtain an interval for the logged spectrum:

$$\log[\bar{f}(\omega)] + \log\left[\frac{\nu}{\chi_{\nu, 1-\frac{\alpha}{2}}^2}\right] < \log[f(\omega)] < \log[\bar{f}(\omega)] + \log\left[\frac{\nu}{\chi_{\nu, \frac{\alpha}{2}}^2}\right]$$

Pointwise Confidence Intervals for $f(\omega)$: Log-Scale



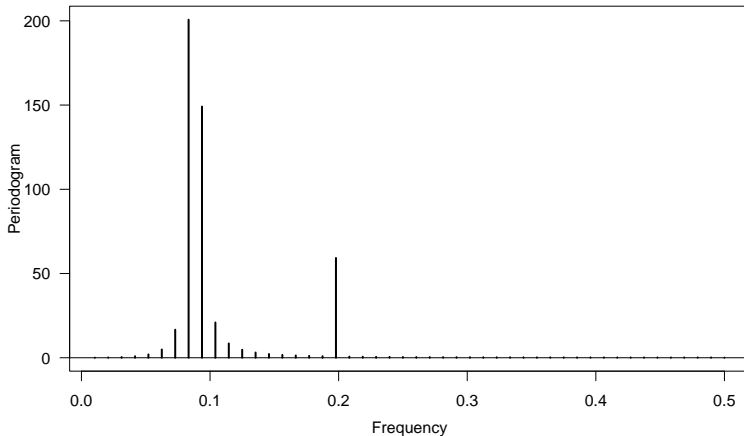
Pointwise Confidence Intervals for $f(\omega)$: Original Scale



Spectral Leakage

Much of the previous discussion has assumed that the frequencies of interest are the Fourier frequencies, i.e., $\omega_j = \frac{j}{n}$. What happens if that is not the case?

Example: $Y_t = 3 \cos(2\pi(0.088)t) + \sin(2\pi(\frac{19}{96})t)$, $t = 1, \dots, 96$

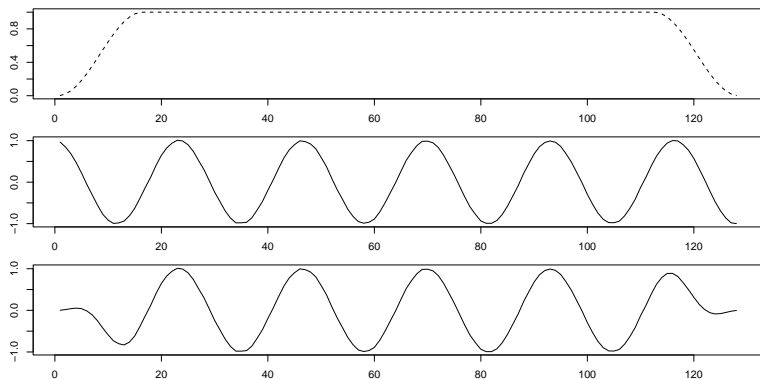


Tapering is one method used to mitigate the issue of **spectral leakage**, where power at non-Fourier frequencies leaks into nearby Fourier frequencies

Background

The Periodogram and
Spectral Density

Spectral Estimation



Tapering (Cont'd)

Background

The Periodogram and
Spectral Density

Spectral Estimation

