# Frequency Domain Revisited – Spectral Density

Ruoqi Yu

Week 13

# Announcements

- Homework 6 will be posted later this week, and due on Dec 5.
- Project is due on Nov 23. You have to work in groups of 3-5 people. Otherwise, penalty will be applied. For instance, if you work by yourself, you will receive 20% penalty. If you work in groups of two people, you will receive 10% penalty.

# Disclaimer

Today we'll discuss several things about the frequency domain, but not too in depth. The purpose is to

1. Give you exposure to a set of tools that are available
2. Connect several things we've been talking about this semester

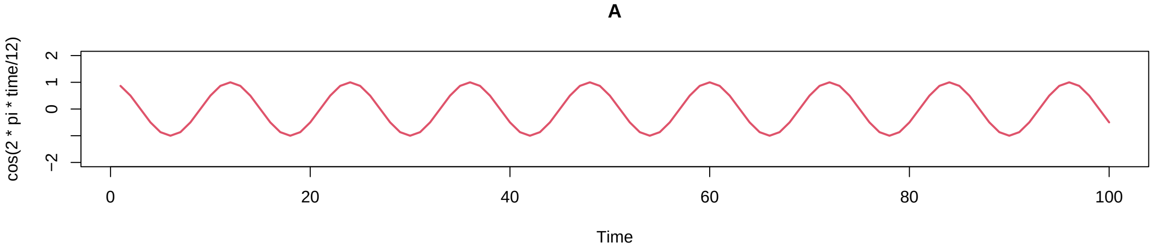
Section 1

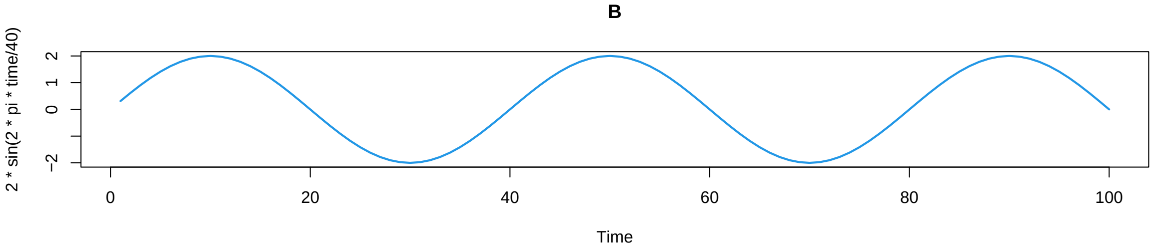
# Return to the Frequency domain

# Frequency domain

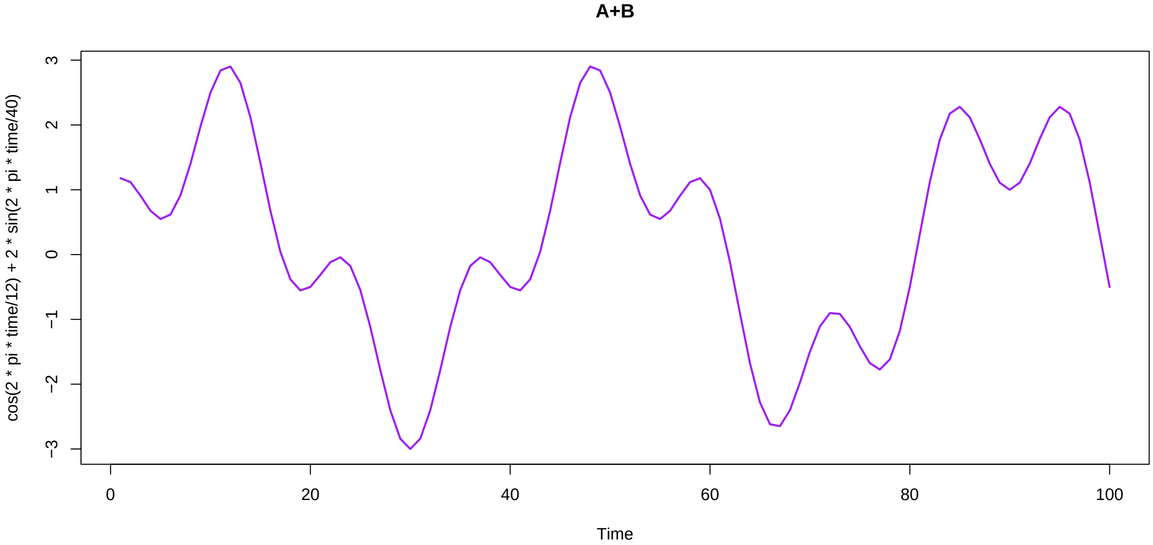
- We have largely studied the time domain approach: models for stationary processes which were directly constructed via the relationship of observations $X_t$ at different time points.
- Now, we will study a stationary process as a composition of periodic components with different frequencies.
- This is quite natural for many time series data, which are often directly driven by periodic random events.

# Example: Multiple types of seasonality



# Diagnose with time-domain only methods?



# Recall the Definition of Sinusoids

We define the set of sinusoid functions as

$$\{g(t) = R\cos(2\pi ft + \Phi) \ : \ R \in R_+, \ f \in R_+, \ \Phi \in [0, 2\pi/f)\},$$

where

- $R$ is called the *amplitude*
- $f$ is called the *frequency*
- $\Phi$ is called the *phase*
- $1/f$ is called the *period*

# Sinusoids rewritten a different way

1. With $A = R\cos(\Phi)$ and $B = -R\sin(\Phi)$ one can rewrite sinuosoids as

$$\{g(t) = A\cos(2\pi ft) + B\sin(2\pi ft) \ : \ A, B \in R, \ f \in R_+\}.$$

# Sinusoids rewritten a yet another way

2. Note that

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

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

$$\sin(2\pi f t) = \frac{\exp(2\pi i f t) - \exp(-2\pi i f t)}{2i}$$

Thus, one can rewrite sinusoids with $C = A/2 + B/(2i)$ and its complex conjugate $\overline{C} = A/2 - B/(2i)$ as

$$\{g(t) = C \exp(2\pi i f t) + \overline{C} \exp(-2\pi i f t) \ : \ C \in C, \ f \in R_+\}.$$

# Definition: Discrete Fourier Transform

For data $x_0, \ldots, x_{n-1} \in C$ the discrete Fourier transform (DFT) is given by $b_0, \ldots, b_{n-1} \in C$, where

$$b_j = \sum_{t=0}^{n-1} x_t \exp\left(-\frac{2\pi i j t}{n}\right) \text{ for } j = 0, \ldots, n-1.$$

(In R, the DFT is calculated by the function *fft()*.)

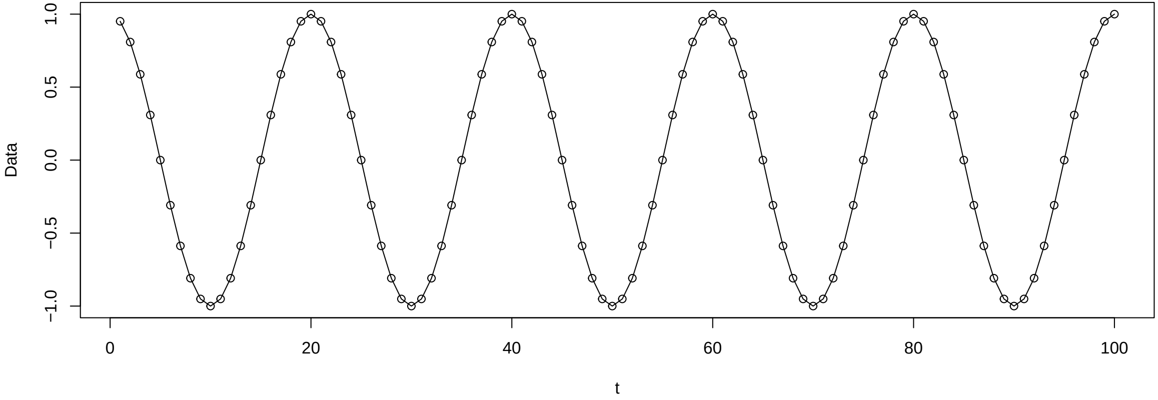
# Definition: Periodogram

For real values data $x_0, \dots, x_{n-1}$ with DFT $b_0, \dots, b_{n-1}$ the **periodogram** is defined as

$$I(j/n) = \frac{|b_j|^2}{n} \quad \text{for } j = 1, \dots, \lfloor n/2 \rfloor$$

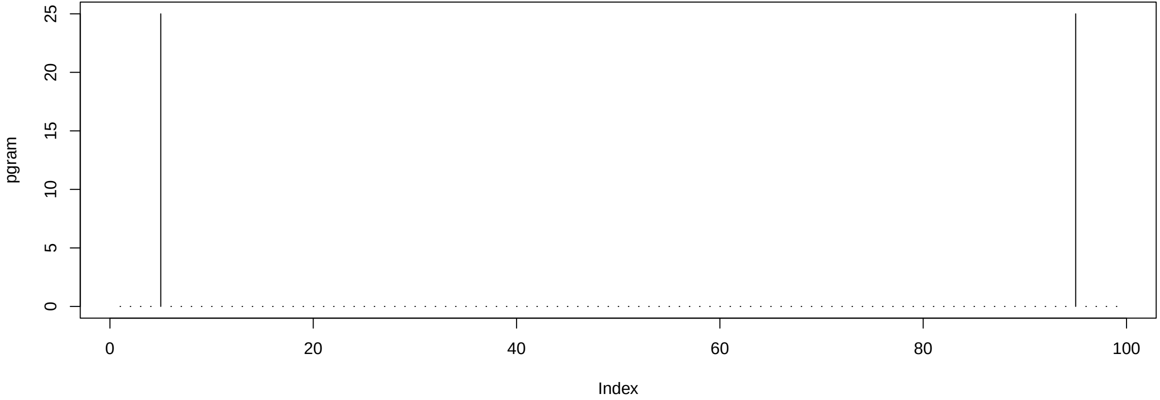
# Example Data: $cos(2\pi t * 5/100)$

```
n=100; t = 1:n; cos2 = cos(2*pi*t*(5/n))
plot(t, cos2, ylab = "Data", type = "o")
```



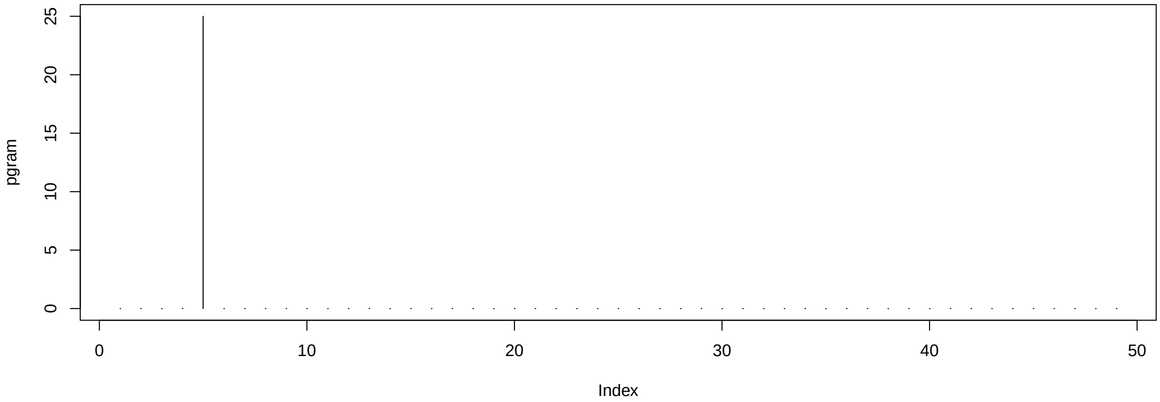
# Example: $cos(2\pi t * 5/100)$

```
pgram = abs(fft(cos2)[2:100])^2/n
plot(pgram, type = "h")
```



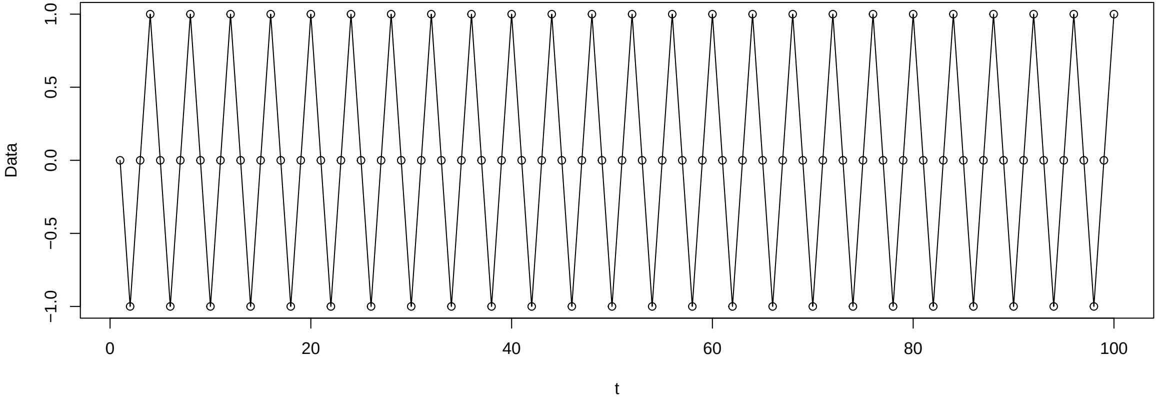
# Example Periodogram: $cos(2\pi t * 5/100)$

```
pgram = abs(fft(cos2)[2:50])^2/n
plot(pgram, type = "h")
```



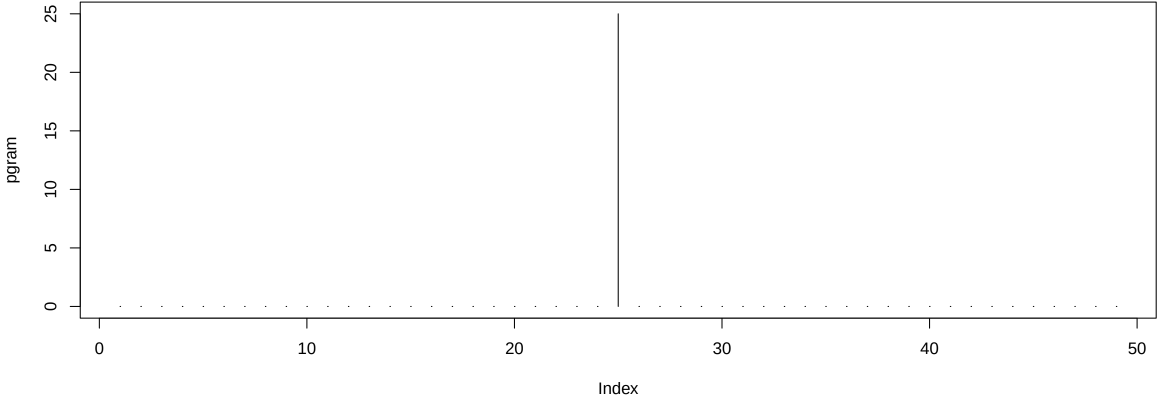
# Example Data: $cos(2\pi t * 25/100)$

```
t = 1:100; cos2 = cos(2*pi*t*(25/100))
plot(t, cos2, ylab = "Data", type = "o")
```



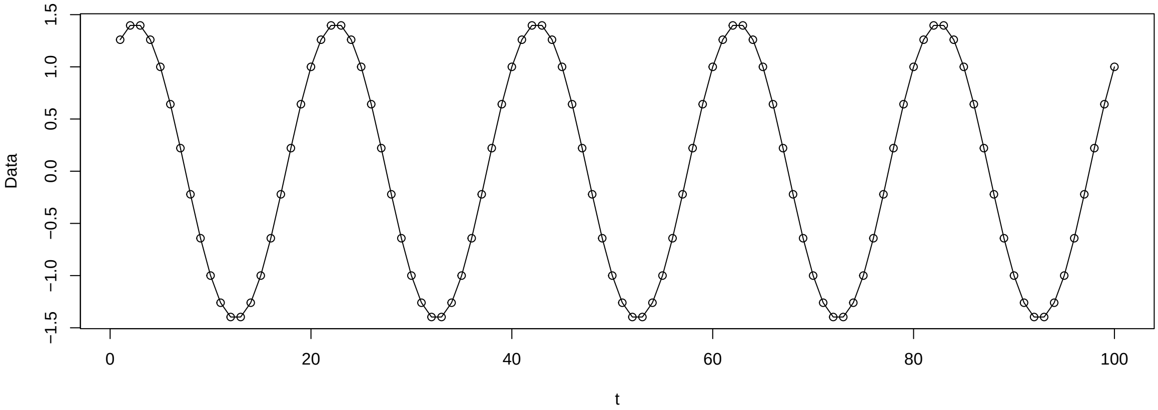
# Example Periodogram: $cos(2\pi t * 25/100)$

```
pgram = abs(fft(cos2)[2:50])^2/n
plot(pgram, type = "h")
```



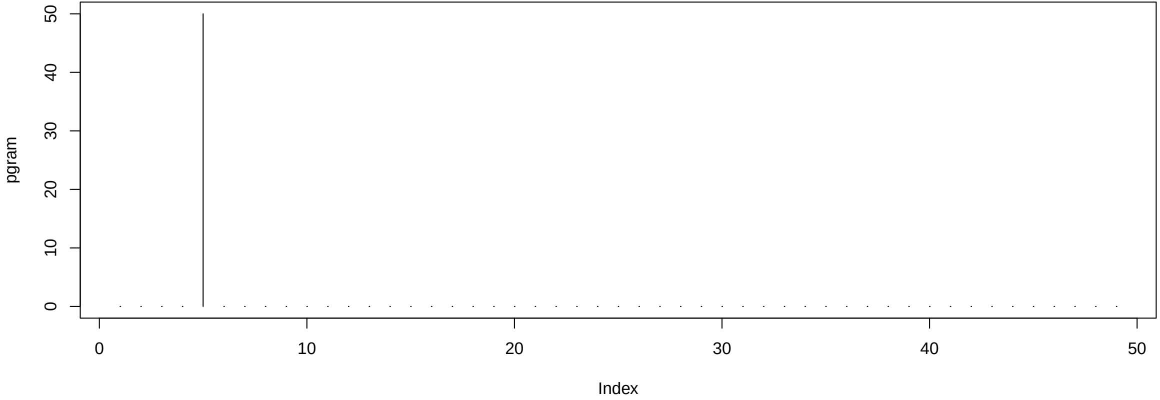
# Example Data: $cos(2\pi t * 5/100) + sin(2\pi t * 5/100)$

```
t = 1:100; cos2 = cos(2*pi*t*5/100) + sin(2*pi*t*5/100)
plot(t, cos2, ylab = "Data", type = "o")
```



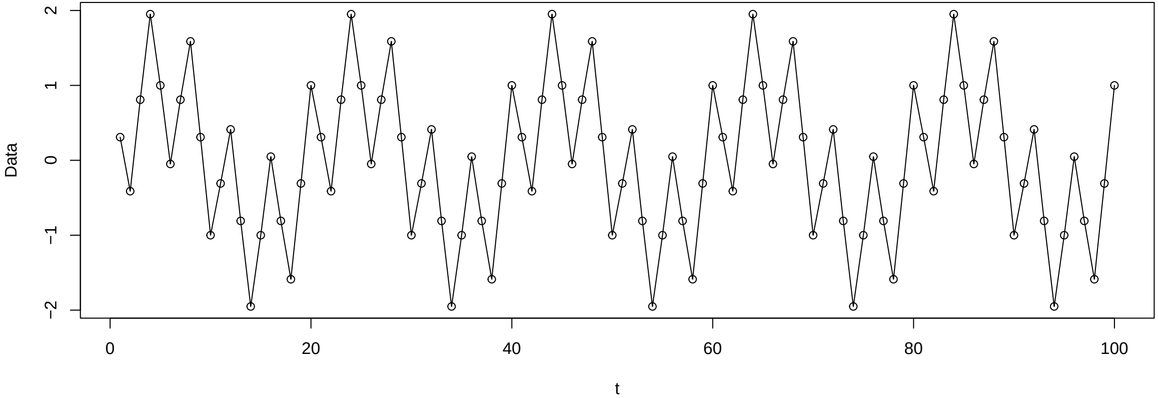
# Example Periodogram: $cos(2\pi t * 5/100) + sin(2\pi t * 5/100)$

```
pgram = abs(fft(cos2)[2:50])^2/n
plot(pgram, type = "h")
```



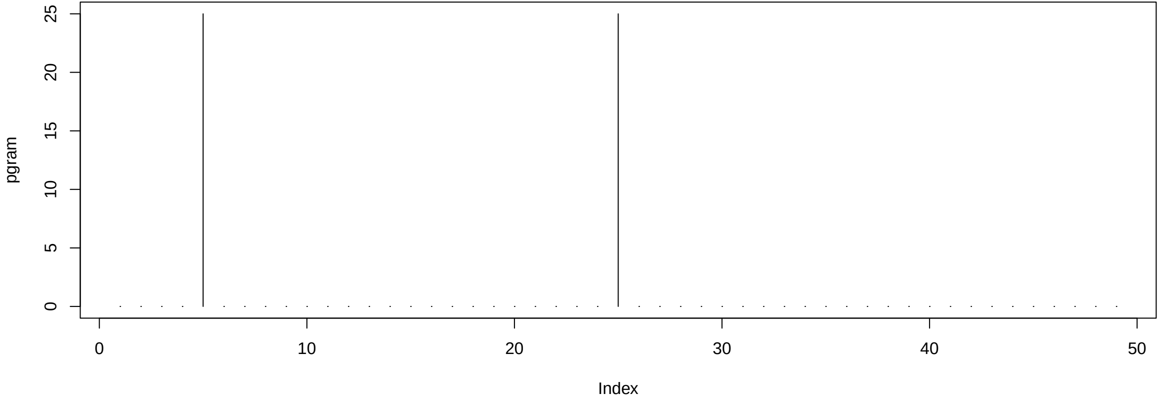
# Example Data: $cos(2\pi t * 25/100) + sin(2\pi t * 5/100)$

```r
t = 1:100; cos2 = cos(2*pi*t*(25/100)) + sin(2*pi*t*5/100)
plot(t, cos2, ylab = "Data", type = "o")
```



# Example Periodogram: $cos(2\pi t * 25/100) + sin(2\pi t * 5/100)$

```
pgram = abs(fft(cos2)[2:50])^2/n
plot(pgram, type = "h")
```



# Recall this Theorem: Connection between periodogram and $\hat{\gamma}$

For some data $x_0, \ldots, x_{n-1}$ let $\hat{\gamma}(h)$ for $h = 0, \ldots, n-1$ be its sample ACVF. Then

$$I(j/n) = \sum_{h=-(n-1)}^{n-1} \hat{\gamma}(h) \exp\left(-\frac{2\pi i j h}{n}\right) \text{ for } j = 1, \ldots, \lfloor n/2 \rfloor.$$

# Section 2

## Spectral Density

# Transition

- We've shown that every dataset can be written in terms of sinusoids.
- The magnitude of the sinusoid component with frequency $j/n$ is given by the respective periodogram $I(j/n)$.
- But this is a discrete representation and leads to leakage!
- Now we extend these definitions to the process $\{X_t\}$ itself.
- Remember that ACVF is related to the periodogram, and that leads to the following natural process-analog of the periodogram.

# Definition: Spectral Density

For a stationary process with ACVF $\gamma_X(h)$ with $\sum_{h=-\infty}^{\infty} |\gamma_X(h)| < \infty$ we define the *spectral density* as

$$f(\lambda) := \sum_{h=-\infty}^{\infty} \gamma_X(h) \exp\left(-2\pi i \lambda h\right) \text{ for } -1/2 \leq \lambda \leq 1/2.$$

# Notes on the Spectral Density

- $f$ is symmetric: $f(-\lambda) = f(\lambda)$
- $f$ is always nonnegative: $f(\lambda) \geq 0$
- Like the periodogram, the spectral density gives the strengths of sinusoids at various frequencies contributing to a stationary stochastic process.
- The spectral density and the ACVF provide equivalent information.

# Thoerem: ACVF and Spectral Density

For a stationary process with spectral density $f(\lambda)$, $-1/2 \leq \lambda \leq 1/2$, it holds for its ACVF that

$$\gamma_X(h) = \int_{-1/2}^{1/2} e^{2\pi i \lambda h} f(\lambda) d\lambda = \int_{-1/2}^{1/2} \cos(2\pi\lambda h) f(\lambda) d\lambda.$$

# Proof (page 1)

Using the definition of the spectral density we get

$$\begin{aligned}\int_{-1/2}^{1/2} e^{2\pi i\lambda h} f(\lambda) d\lambda &= \int_{-1/2}^{1/2} e^{2\pi i\lambda h} \sum_{k=-\infty}^{\infty} \gamma_X(k) e^{-2\pi i\lambda k} d\lambda \\ &= \sum_{k=-\infty}^{\infty} \gamma_X(k) \int_{-1/2}^{1/2} e^{2\pi i\lambda(h-k)} d\lambda\end{aligned}$$

note that

$$\int_{-1/2}^{1/2} e^{2\pi i\lambda(k-h)} d\lambda \neq 0 \quad \Leftrightarrow \quad k = h$$

and thus

$$\int_{-1/2}^{1/2} e^{2\pi i\lambda h} f(\lambda) d\lambda = \gamma_X(h).$$

# Proof (page 2)

For the second equality note that by symmetry $f(\lambda) = f(-\lambda)$ and our 2nd sinusoid identity, it follows that

$$\begin{aligned}
\int_{-1/2}^{1/2} e^{2\pi i\lambda h} f(\lambda)d\lambda &= \int_{-1/2}^{0} e^{2\pi i\lambda h} f(-\lambda)d\lambda + \int_{0}^{1/2} e^{2\pi i\lambda h} f(\lambda)d\lambda \\
&= \int_{0}^{1/2} e^{-2\pi i\lambda h} f(\lambda)d\lambda + \int_{0}^{1/2} e^{2\pi i\lambda h} f(\lambda)d\lambda \\
&= \int_{0}^{1/2} \left(e^{-2\pi i\lambda h} + e^{2\pi i\lambda h}\right) f(\lambda)d\lambda \\
&= \int_{0}^{1/2} 2\cos(2\pi\lambda h) f(\lambda)d\lambda \\
&= \int_{-1/2}^{1/2} \cos(2\pi\lambda h) f(\lambda)d\lambda
\end{aligned}$$

# Example: White Noise

- Suppose $\{X_t\}$ is white noise with mean zero and variance $\sigma^2$. What is the spectral density of white noise?
- Consider the MA(1) process $X_t = W_t + \theta W_{t-1}$. What is the spectral density of a MA(1) process?

# Example: White Noise

- Suppose $\{X_t\}$ is white noise with mean zero and variance $\sigma^2$. It is obvious that $\gamma(h) = 0$ for $h \neq 0$ and $\gamma(h) = \sigma^2$ for $h = 0$.
- Solve here:

$$f_X(\lambda) = \sum_{h=-\infty}^{\infty} \gamma_X(h) \exp\left(-2\pi i \lambda h\right) \text{ for } -1/2 \leq \lambda \leq 1/2.$$

- In this case, the spectral density formula simply gives

$$f_W(\lambda) = \sigma^2 \text{ for every } -1/2 \leq \lambda \leq 1/2.$$

- This means that the spectral density of white noise is flat (all frequencies are combined equally).

# Example: MA(1)

Consider the MA(1) process $X_t = W_t + \theta W_{t-1}$. The autocovariance function is given by $\gamma(0) = \sigma_W^2(1+\theta^2)$, $\gamma(\pm 1) = \theta\sigma_W^2$ and $\gamma(h)$ equals zero for every other $h$.

The spectral density then immediately gives

$$\begin{aligned} f(\lambda) &= \gamma(-1)\exp(2\pi i\lambda) + \gamma(0)\exp(0) + \gamma(1)\exp(-2\pi i\lambda) \\ &= \gamma(0) + \gamma(1)\left(\exp(2\pi i\lambda) + \exp(-2\pi i\lambda)\right) \\ &= \gamma(0) + 2\gamma(1)\cos(2\pi\lambda) \\ &= \sigma_W^2\left(1 + \theta^2 + 2\theta\cos(2\pi\lambda)\right) \text{ for } -1/2 \le \lambda \le 1/2. \end{aligned}$$

This function is increasing when $\theta < 0$ and decreasing when $\theta > 0$.

# Section 3

## Linear Time Invariant Filters

# Filters

- The general technique of linear time invariant filters: transforming one time series into another.
- Linear filters were already introduced in the context of trend estimation months ago
- We will see that they are particularly helpful within the frequency domain approaches.

# Definition: Linear Time Invariant Filter

A linear time-invariant filter with coefficients $\{a_j\}$ for $j = \dots, -2, -1, 0, 1, 2, 3, \dots$ transforms an input time series $\{X_t\}$ into an output time series $\{Y_t\}$ via

$$Y_t = \sum_{j=-\infty}^{\infty} a_j X_{t-j}.$$

In the above definition, the coefficients $\{a_j\}$ are often assumed to satisfy $\sum_{j=-\infty}^{\infty} |a_j| < \infty$.

# Examples

- Particular types of time invariant linear filters we've already been using:
- q-step smoothing $\Rightarrow a_j = \frac{1}{2q+1}$ for $|j| \leq q$, $a_j = 0$ otherwise.
- Differencing $\Rightarrow a_0 = 1$ and $a_1 = -1$, $a_j = 0$ otherwise.
- We have seen that these two filters act very differently; one estimates trend while the other eliminates it.

# Autocovariance of Linear Time Invariant Filter

- Suppose that the input time series $\{X_t\}$ is stationary with ACVF $\gamma_X$.
- Then for the autocovariance function of $\{Y_t\}$ we observe

$$
\begin{aligned}
\gamma_Y(h) &= \text{cov}\left(Y_t, Y_{t+h}\right) \\
&= \text{cov}\left(\sum_j a_j X_{t-j}, \sum_k a_k X_{t+h-k}\right) \\
&= \sum_{j,k} a_j a_k \text{cov}(X_{t-j}, X_{t+h-k}) \\
&= \sum_{j,k} a_j a_k \gamma_X(h-k+j).
\end{aligned}
$$

- Note that the above calculation shows also that $\{Y_t\}$ is stationary.

## Spectral Density of Filters

- Let $f_X$ be the spectral density of the input $\{X_t\}$.
- Recall

$$\gamma_X(h) = \int_{-1/2}^{1/2} e^{2\pi i h\lambda} f_X(\lambda) d\lambda.$$

- Combining this with the ACVF of $\{Y_t\}$ from the last slide, we get the spectral density $f_Y$ of the output $\{Y_t\}$:

$$\begin{aligned}\gamma_Y(h) &= \sum_j \sum_k a_j a_k \int e^{2\pi i(h-k+j)\lambda} f_X(\lambda) d\lambda \\ &= \int e^{2\pi i h\lambda} f_X(\lambda) \left( \sum_j \sum_k a_j a_k e^{-2\pi i k\lambda} e^{2\pi i j\lambda} \right) d\lambda\end{aligned}$$

- We'll simplify this rearranged formula on the last line.

# Definition: Transfer Function

For a time invariant linear filter with coefficients $\{a_j\}$, we define the **transfer function**

$$A(\lambda) := \sum_j a_j e^{-2\pi i j \lambda} \text{ for } -1/2 \le \lambda \le 1/2. \tag{1}$$

# Note: Complex Numbers

- $A(\lambda)$ contains $i \Rightarrow$ complex number!
- Recall Euler's equation: $e^{ix} = \cos x + i \sin x$, and it's conjugate $e^{-ix} = \cos x - i \sin x$
- Thus

$$
\begin{aligned}
A(\lambda) &= \sum_j a_j e^{-2\pi i j \lambda} \\
&= \sum_j a_j [\cos(2\pi j \lambda) - i \sin(2\pi j \lambda)] \\
&= \left[ \sum_j a_j \cos(2\pi j \lambda) \right] - i \sum_j a_j \sin(2\pi j \lambda)
\end{aligned}
$$

- Conjugate: $\overline{A(\lambda)} = \left[\sum_j a_j \cos(2\pi j \lambda)\right] + i \sum_j a_j \sin(2\pi j \lambda) = \sum_j a_j e^{2\pi i j \lambda}$

# Using the Transfer Function

- Recall our previous equation for the ACVF of Y:

$$\gamma_Y(h) = \int e^{2\pi i h\lambda} f_X(\lambda) \left( \sum_j \sum_k a_j a_k e^{-2\pi i k\lambda} e^{2\pi i j\lambda} \right) d\lambda$$

- Applying the definition of the transfer function:

$$\gamma_Y(h) = \int e^{2\pi i\lambda h} f_X(\lambda) A(\lambda) \overline{A(\lambda)} d\lambda,$$

where, of course, $\overline{A(\lambda)}$ denotes the complex conjugate of $A(\lambda)$.

- As a result, we have

$$\gamma_Y(h) = \int e^{2\pi i\lambda h} f_X(\lambda) \left| A(\lambda) \right|^2 d\lambda.$$

- This is clearly of the form $\gamma_Y(h) = \int e^{2\pi i\lambda h} f_Y(\lambda) d\lambda$.

# Definition: Power Transfer Function

The function $\lambda \mapsto |A(\lambda)|^2$ is called the **power transfer function**.

# Use of the Power Transfer Function

- We therefore have

$$f_Y(\lambda) = f_X(\lambda)\,|A(\lambda)|^2 \text{ for } -1/2 \le \lambda \le 1/2.$$

- So what does the filter do to the spectrum? It modifies the spectrum by multiplying it with the power transfer function $|A(\lambda)|^2$.
- Depending on the value of $|A(\lambda)|^2$, some frequencies may be enhanced in the output while other frequencies will be diminished.
- Thus, the spectral density is very useful while studying the properties of a filter.
- While the autocovariance function of the output series $\gamma_Y$ depends in a complicated way on that of the input series $\gamma_X$, the dependence between the two spectral densities is very simple.

# Example: Power Transfer Function of the Differencing Filter

- Consider lag $s$ differencing $Y_t = X_t - X_{t-s}$
- Suppose $X - t$ has spectral density $f_X(\lambda)$. What is the spectral density of $Y_t$?

# Example: Power Transfer Function of the Differencing Filter

- $Y_t$ corresponds to the weights $a_0 = 1$ and $a_s = -1$ and $a_j = 0$ for all other $j$.
- Then the transfer function is given by

$$
\begin{aligned}
A(\lambda) &= \sum_j a_j e^{-2\pi i j\lambda} \\
&= a_0 e^{-2\pi i(0)\lambda} + a_s e^{-2\pi i s\lambda} \\
&= (1)e^{0} + (-1)e^{-2\pi i s\lambda} \\
&= 1 - e^{-2\pi i s\lambda} \\
&= 1 - \cos(2\pi s\lambda) + i\sin(2\pi s\lambda)
\end{aligned}
$$

# Example: Power Transfer Function of the Differencing Filter

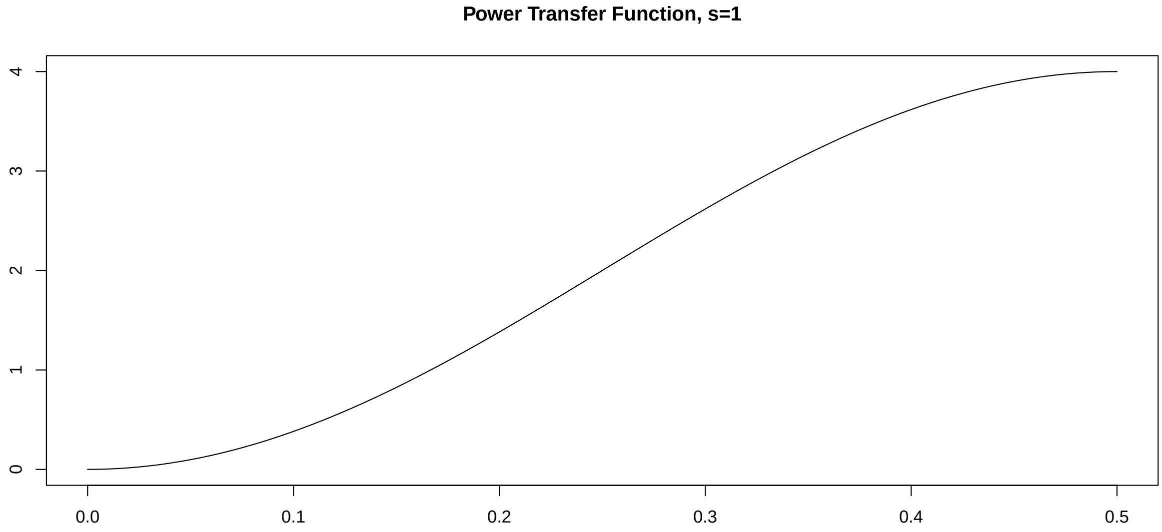
- The power transfer function:

$$
\begin{aligned}
|A(\lambda)|^2 &= \sqrt{Re(A(\lambda))^2 + Im(A(\lambda))^2}^2 \\
&= [1 - \cos(2\pi s\lambda)]^2 + \sin^2(2\pi s\lambda) \\
&= 1 - 2\cos(2\pi s\lambda) + \cos^2(2\pi s\lambda) + \sin^2(2\pi s\lambda) \\
&= 1 - 2\cos(2\pi s\lambda) + 1 \\
&= 2 - 2\cos(2\pi s\lambda)
\end{aligned}
$$

- $$f_Y(\lambda) = f_X(\lambda)(2 - 2\cos(2\pi s\lambda)) \text{ for } -1/2 \leq \lambda \leq 1/2.$$

- To understand $|A(\lambda)|^2$, we only need to consider the interval $[0, 1/2]$ because it is symmetric on $[-1/2, 1/2]$.

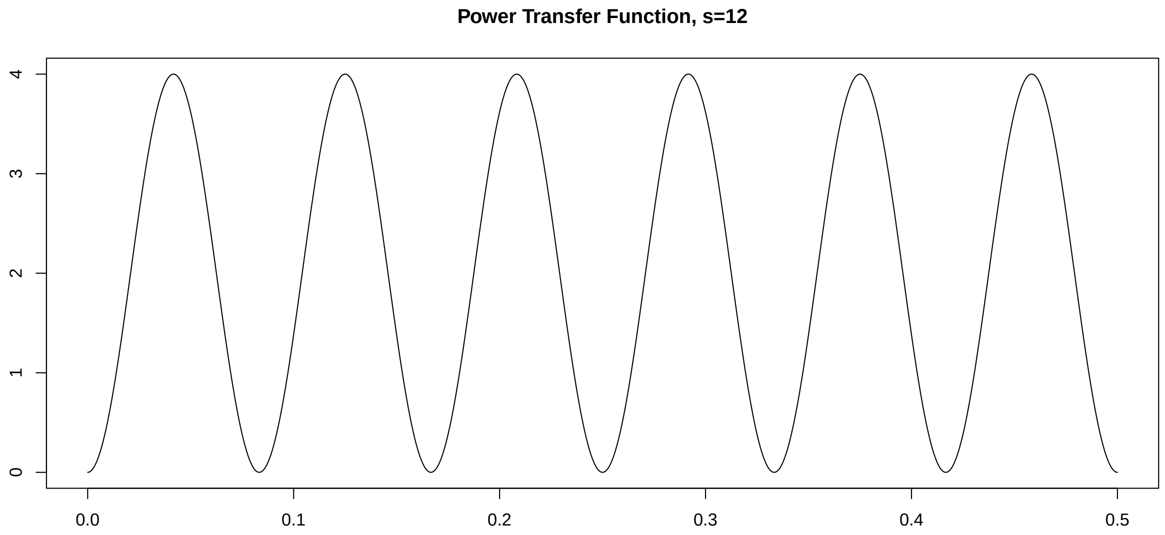
# $s = 1$



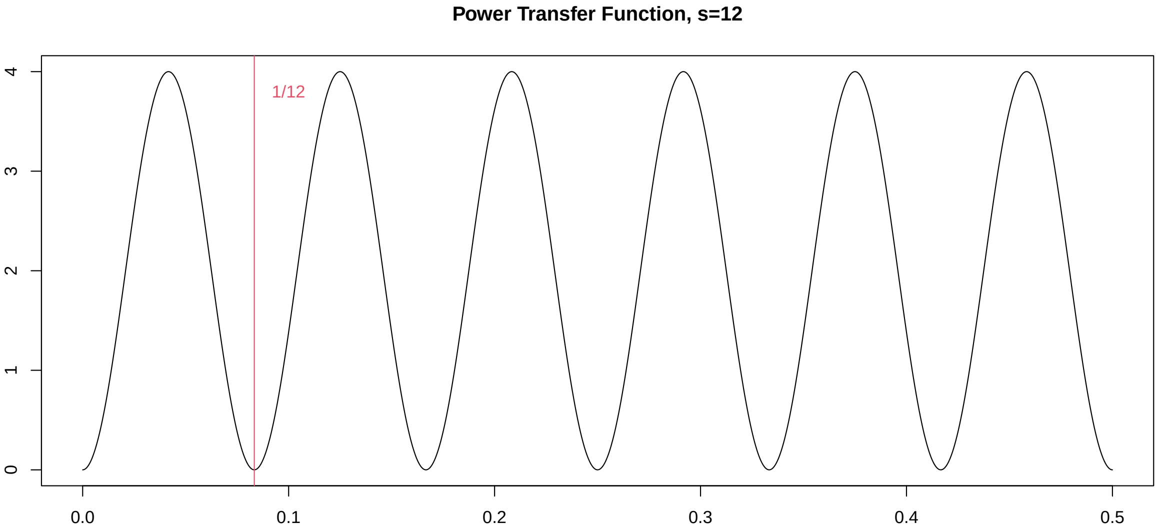
# Example: Power Transfer Function of the Differencing Filter

- When $s = 1$, the function $|A(\lambda)|^2$ is increasing on $[0, 1/2]$.
- This means that first order differencing enhances the higher frequencies in the data and diminishes the lower frequencies.
- Therefore, it will make the data "more wiggly" as it elminates low frequency elements (i.e. trend!).

# $s = 12$



# $s = 12$



# Example: Power Transfer Function of the Differencing Filter

- For higher values of $s$, the function $A(\lambda)$ goes up and down and takes the value zero for $\lambda = 0, 1/s, 2/s, \ldots$.
- In other words, it eliminates all components of period $s$.

# Example: Power Transfer Function of Smoothing Filter

- Now consider the smoothing filter which corresponds to the coefficients $a_j = 1/(2q+1)$ for $|j| \leq q$.
- What is its power transfer function?

# Example: Power Transfer Function of Smoothing Filter

- For $-1/2 \leq \lambda \leq 1/2$, the transfer function is

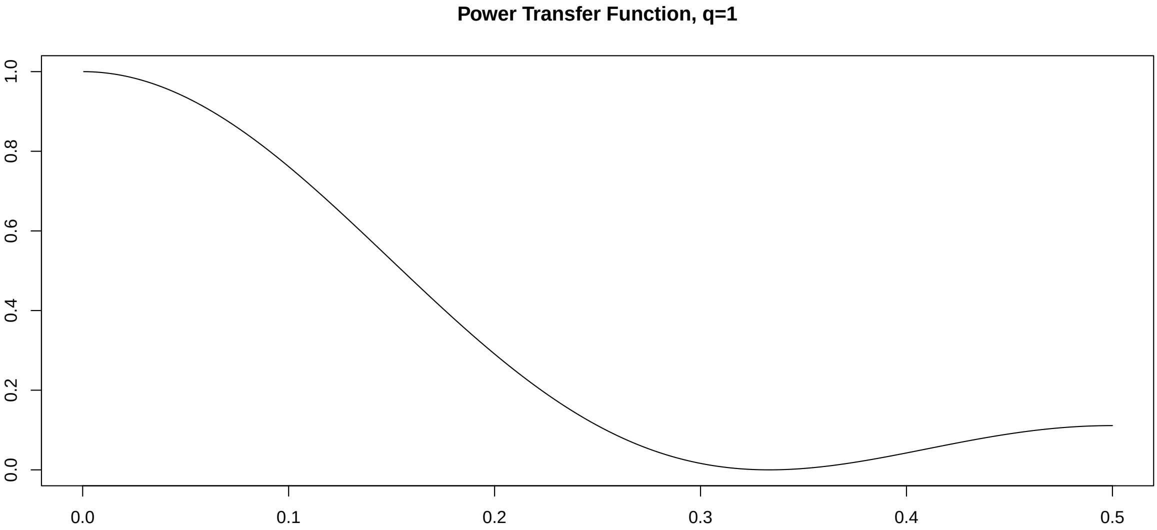
$$
\begin{aligned}
A(\lambda) &= \sum_{j=-q}^{q} \frac{1}{2q+1} e^{-2\pi i j \lambda} \\
&= \frac{\sum_{j=-1}^{-q} e^{-2\pi i j \lambda} + 1 + \sum_{j=1}^{q} e^{-2\pi i j \lambda}}{2q+1} \\
&= \frac{\sum_{j=0}^{-q} e^{-2\pi i j \lambda} - 1 + \sum_{j=0}^{q} e^{-2\pi i j \lambda}}{2q+1}
\end{aligned}
$$

- When $\lambda = 0$ it is easy to see that and $A(0) = \frac{q+1-1+q+1}{2q+1} = 1$.
- When $\lambda \neq 0$ then $\exp(2\pi i \lambda) \neq 1$ and this function can be evaluated using the geometric series formula, e.g. $\sum_{j=0}^{q} e^{-2\pi i j \lambda} = \frac{1-e^{2\pi i \lambda (q+1)}}{1-e^{2\pi i \lambda}}$.

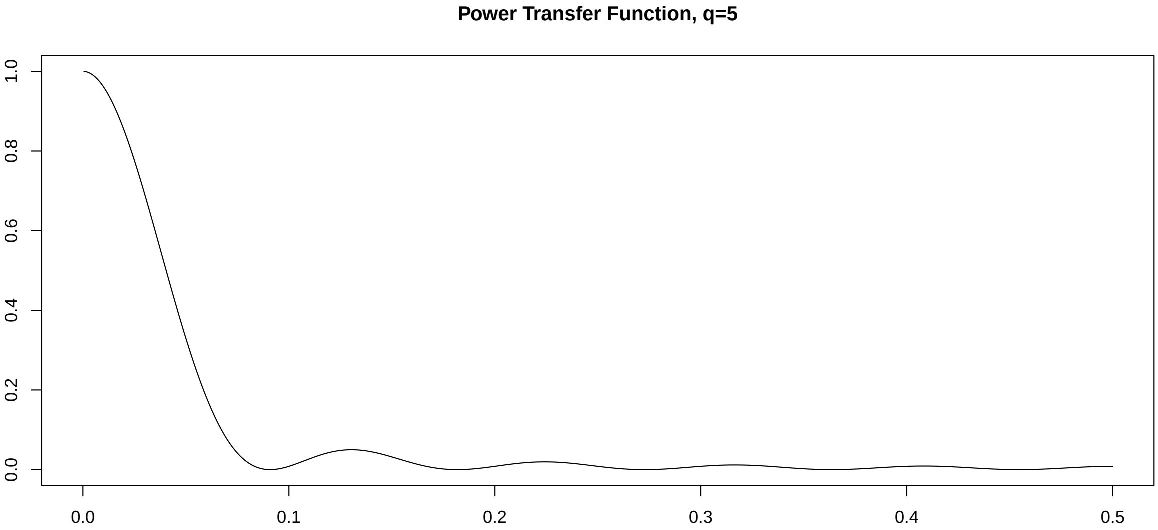
# Example: Power Transfer Function of Smoothing Filter

- Then, because
$$e^{i\theta} - 1 = \cos\theta + i\sin\theta - 1 = 2ie^{i\theta/2}\sin(\theta/2)$$
we get
$$S_q(\lambda) = \frac{\sin \pi q\lambda}{\sin \pi\lambda} e^{i\pi\lambda(q-1)}.$$
- Thus
$$S_q(\lambda) + S_q(-\lambda) = 2\frac{\sin(\pi q\lambda)}{\sin(\pi\lambda)}\cos(\pi\lambda(q-1)),$$
which implies that the transfer function is given by
$$A(\lambda) = \frac{1}{2q+1}\left(2\frac{\sin(\pi(q+1)\lambda)}{\sin(\pi\lambda)}\cos(\pi q\lambda) - 1\right).$$
- For $q$ large, it drops to zero very quickly $\Rightarrow$ the filter kills the high frequency components in the input process.

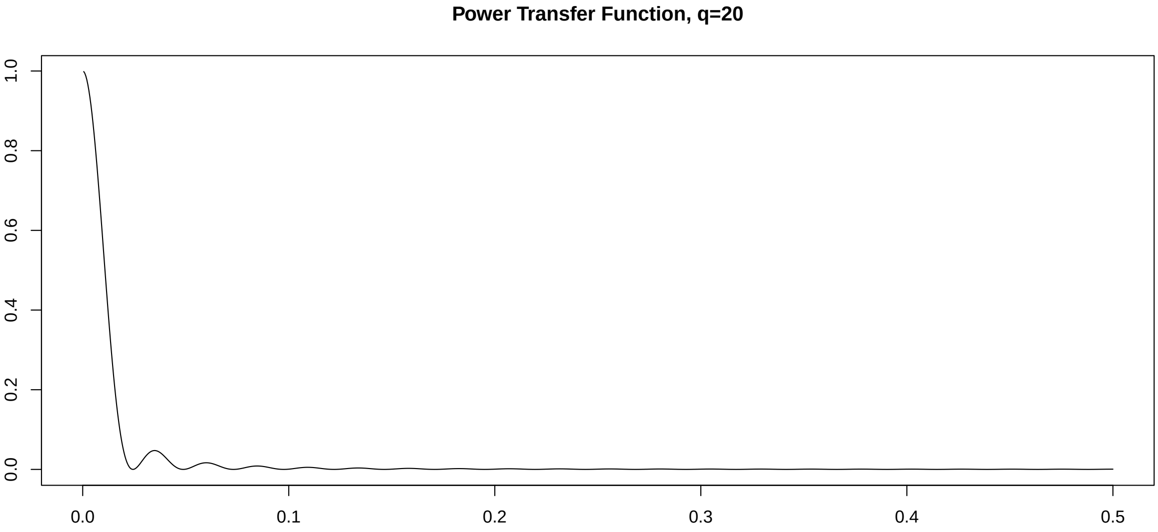
# Example: Power Transfer Function of Smoothing Filter



# Example: Power Transfer Function of Smoothing Filter



# Example: Power Transfer Function of Smoothing Filter



# Section 4

## Spectral density of ARMA process

# Spectral Density of ARMA Process

Because we know

$$f_Y(\lambda) = f_X(\lambda)\left|A(\lambda)\right|^2 \text{ for } -1/2 \leq \lambda \leq 1/2$$

we can compute the spectral density of the unique stationary solution of a causal ARMA process.

# Theorem: Spectral Density of ARMA Process

Let $\{X_t\}$ be a stationary causal ARMA process $\phi(B)X_t = \theta(B)W_t$ with $\phi$ and $\theta$ having no common roots.

Then, for the definition of spectral density $f_X$ of $\{X_t\}$ that uses the ACVF:

$$f(\lambda) := \sum_{h=-\infty}^{\infty} \gamma_X(h) \exp\left(-2\pi i \lambda h\right) \text{ for } -1/2 \leq \lambda \leq 1/2.$$

it holds that

$$f_X(\lambda) = \sigma_W^2 \frac{|\theta(e^{-2\pi i j \lambda})|^2}{|\phi(e^{-2\pi i j \lambda})|^2} \text{ for } -1/2 \leq \lambda \leq 1/2$$

# Proof

- Let $U_t = \phi(B)X_t = \theta(B)W_t$.
- First, write the spectral density of $U_t = \phi(B)X_t$ in terms of that of $\{X_t\}$: specifically $U_t$ can be viewed as the output of a filter applied to $X_t$.
- The filter is given by $a_0 = 1$ and $a_j = -\phi_j$ for $1 \le j \le p$ and $a_j = 0$ for all other $j$.
- Let $A_\phi(\lambda)$ denote the transfer function of this filter.
- Then we have

$$f_U(\lambda) = \left|A_\phi(\lambda)\right|^2 f_X(\lambda).$$

# Proof (page 2)

- Similarly, using the fact that $U_t = \theta(B)W_t$ and that the spectral density of white noise is constant, $f_W(\lambda) = \sigma_W^2$, we write

$$f_U(\lambda) = |A_\theta(\lambda)|^2 f_W(\lambda) = \sigma_W^2 |A_\theta(\lambda)|^2$$

  where $A_\theta(\lambda)$ is the transfer function of the filter with coefficients $a_0 = 1$ and $a_j = \theta_j$ for $1 \leq j \leq q$ and $a_j = 0$ for all other $j$.

- Equating the two $f_U(\lambda)$,

$$f_X(\lambda) = \frac{|A_\theta(\lambda)|^2}{|A_\phi(\lambda)|^2}\sigma_W^2 \text{ for } -1/2 \leq \lambda \leq 1/2.$$

# Proof (page 3)

- Now
$$\begin{aligned} A_\phi(\lambda) &= 1 - \phi_1 e^{-2\pi i(1)\lambda} - \phi_2 e^{-2\pi i(2)\lambda} - \cdots - \phi_p e^{-2\pi i(p)\lambda} \\ &= \phi(e^{-2\pi i\lambda}) \end{aligned}$$
- Note that the denominator $A_\phi(\lambda)$ is non-zero for all $\lambda$ because of stationarity.
- Similarly $A_\theta(\lambda) = \theta(e^{-2\pi i\lambda})$, which completes the proof:
$$f_X(\lambda) = \sigma_W^2 \frac{|\theta(e^{-2\pi i\lambda})|^2}{|\phi(e^{-2\pi i\lambda})|^2} \text{ for } -1/2 \leq \lambda \leq 1/2$$

# Example

- MA(1)
- AR(1)
- AR(2)

# Example: MA(1)

For the MA(1) process: $X_t = W_t + \theta W_{t-1}$, we have $\phi(z) = 1$ and $\theta(z) = 1 + \theta z$. Therefore

$$\begin{aligned} f_X(\lambda) &= \sigma_W^2 \left|1 + \theta e^{-2\pi i\lambda}\right|^2 \\ &= \sigma_W^2 \left|1 + \theta \cos 2\pi\lambda - i\theta \sin 2\pi\lambda\right|^2 \\ &= \sigma_W^2 \left[(1 + \theta \cos 2\pi\lambda)^2 + \theta^2 \sin^2 2\pi\lambda\right] \\ &= \sigma_W^2 \left[1 + \theta^2 + 2\theta \cos 2\pi\lambda\right] \text{ for } -1/2 \leq \lambda \leq 1/2. \end{aligned}$$

Check that for $\theta = -1$, the quantity $1 + \theta^2 + 2\theta \cos(2\pi\lambda)$ equals the power transfer function of the first differencing filter.

# Example: AR(1)

For AR(1): $X_t - \phi X_{t-1} = W_t$, we have $\phi(z) = 1 - \phi z$ and $\theta(z) = 1$. Thus

$$f_X(\lambda) = \sigma_W^2 \frac{1}{|1 - \phi e^{2\pi i \lambda}|^2} = \frac{\sigma_W^2}{1 + \phi^2 - 2\phi \cos 2\pi\lambda}$$

for $-1/2 \leq \lambda \leq 1/2$.

# Example: AR(2)

For the AR(2) model: $X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = W_t$, we have $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$ and $\theta(z) = 1$. Here it can be shown that

$$f_X(\lambda) = \frac{\sigma_W^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1-\phi_2)\cos 2\pi\lambda - 2\phi_2 \cos 4\pi\lambda}$$

for $-1/2 \le \lambda \le 1/2$.

# Section 5

# Parametric Spectral Density Estimation

# Parametric Spectral Density Estimation

- Want to estimate the spectral density of a stationary process?
- One approach: consider a parametric ARMA model $\phi(B)X_t = \theta(B)W_t$
- Estimate its parameters $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$
- Plug in these estimates into the ARMA spectral density equation.
- For convenience, usually a parametric spectral estimator is obtained by fitting an AR(p) model, where the order $p$ is determined by model selection such as AIC or BIC.
- The following theorem shows that any spectral density can be approximated arbitrary close by the spectrum of an AR process, see Property 4.7 in TSA4e.

# Theorem: AR Spectal Approximation

Let $g(\lambda)$ be the spectral density of a stationary process. Then, given $\epsilon > 0$, there is a time series with the representation

$$\phi(B)X_t = W_t,$$

for some finite order $p$ polynomial $\phi$ and some white noise $W_t$ with variance $\sigma^2$, such that

$$|f_X(\lambda) - g(\lambda)| < \epsilon \quad \text{for all } \lambda \in [-1/2, 1/2].$$

Moreover, the roots of $\phi$ outside the unit circle.

# Notes

- Unfortunately, this Theorem does not tell us how large $p$, it might be very large in some cases.
- In R, we can use the function *spec.ar* to fit the best model via AIC and plot the resulting spectrum.
- In the following, we will not discuss properties of these estimates further, but rather will consider a different class of estimates for the spectral density of a stationary process, which does not rely on some specific parametric model assumptions.
- For further reading on parametric density estimation see TSA4e Chapter 4.5.