MATH 7339 - Machine Learning and Statistical Learning Theory 2

Section- Introduction to Spectral Analysis

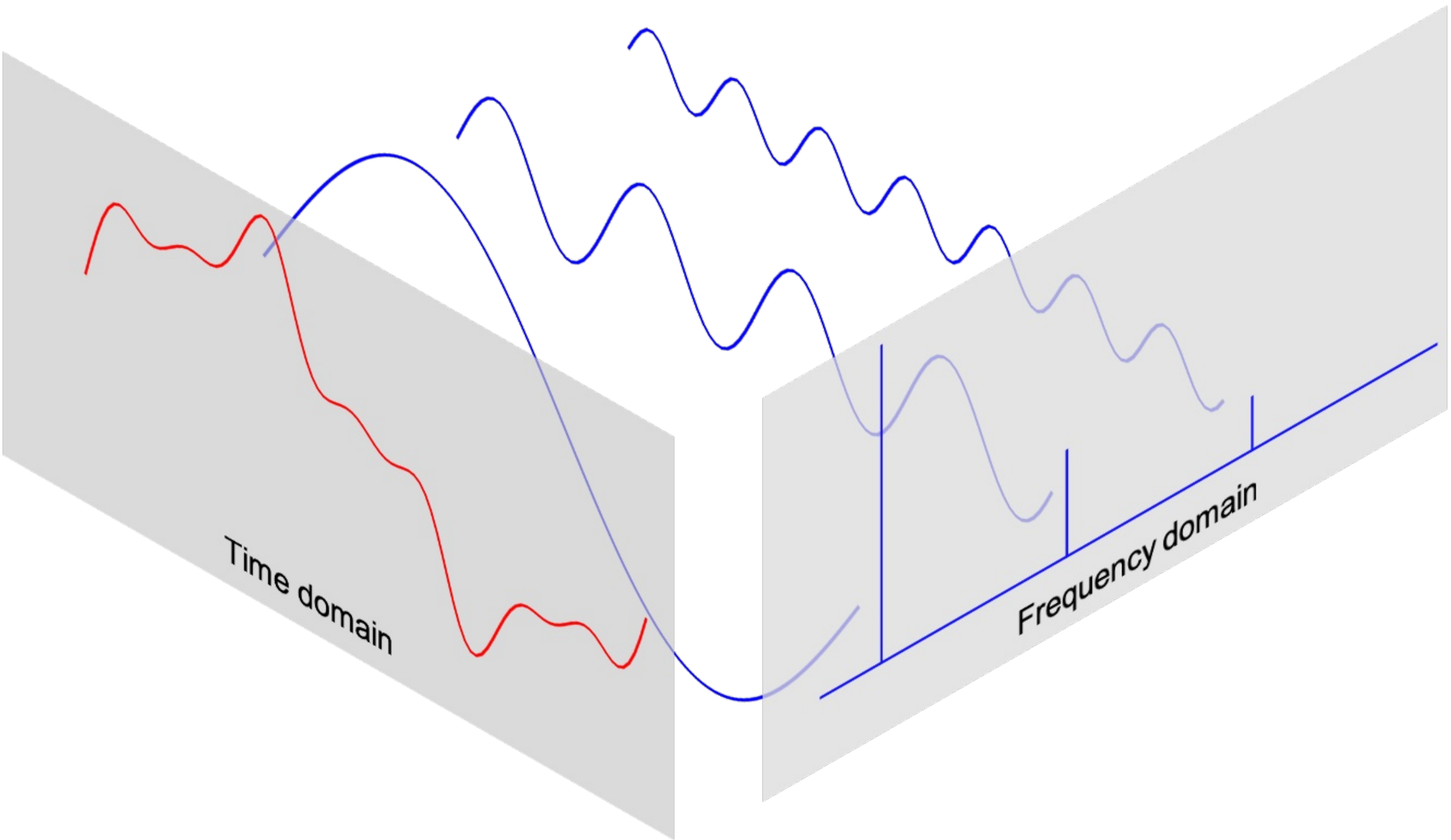
1. Motivation of Spectral Analysis
2. Properties of Spectral Analysis
3. Spectral analysis of some examples
4. Spectral density for causal ARMA processes
5. Time-invariant linear filters
6. Frequency response of a time-invariant linear filter

Motivation of Spectral Analysis

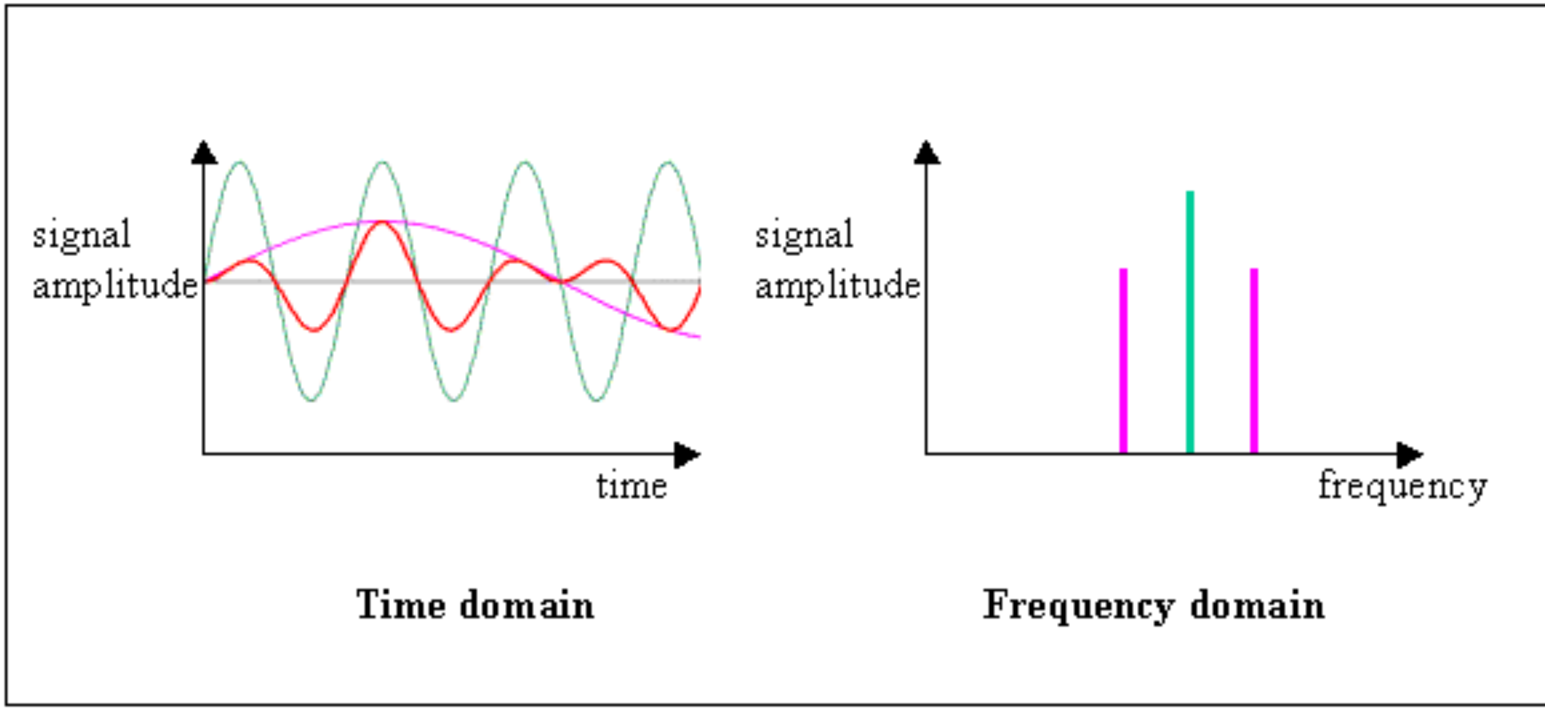
There are two primary approaches to time series.

1. One is the **time domain** approach, which we covered in our previous few sections. This approach focuses on the rules for a time series to move forward. It considers *regression of the present on the past values of the time series*. The models give an explicit formula for the current observation in terms of past observations and past white noise terms.

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

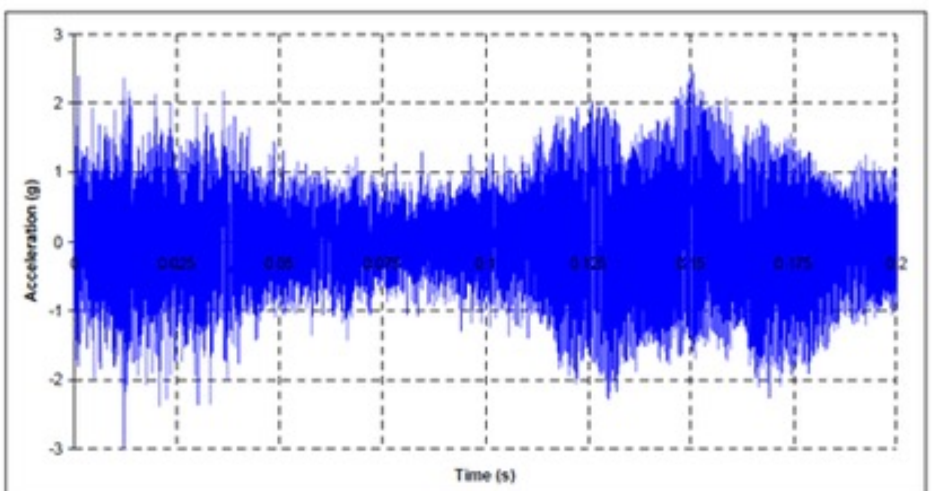


Time domain
Frequency domain

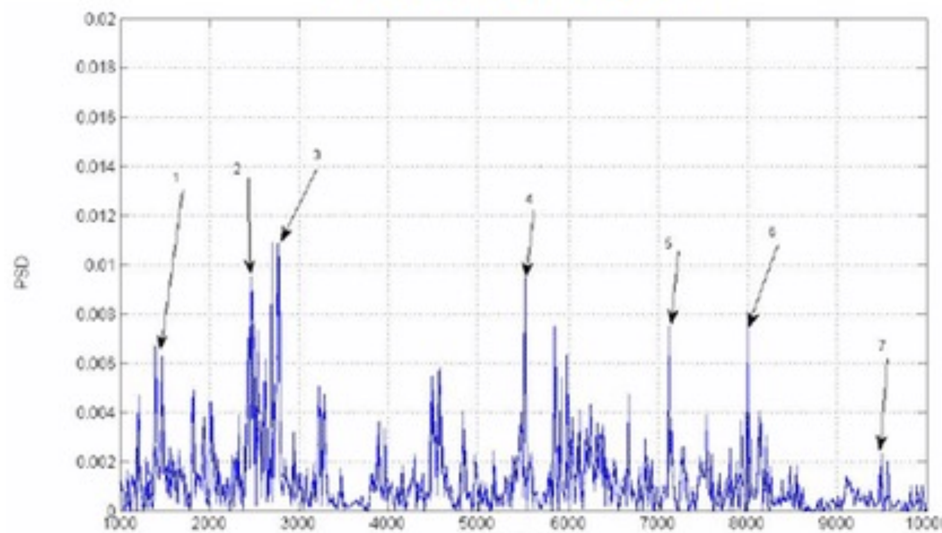


SAMPLE INPUT / OUTPUT (FROM REFERENCE PAPER)

Time Domain



Frequency Domain



Frequency Domain Approach

Idea of Spectral Analysis: decompose a stationary time series $\{X_t\}$ into a combination of sinusoids, with random (and uncorrelated) coefficients. The frequency domain approach considers regression on sinusoids.

Just as in Fourier analysis, where we decompose (deterministic) functions into combinations of sinusoids.

Frequency domain approach model the current observation as a combination of waves. *Regression of the current time on sines and cosines of various frequencies.* In Spectral Analysis,

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

Periodic functions

Consider μ_t as a **periodic function.** For example,

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

where

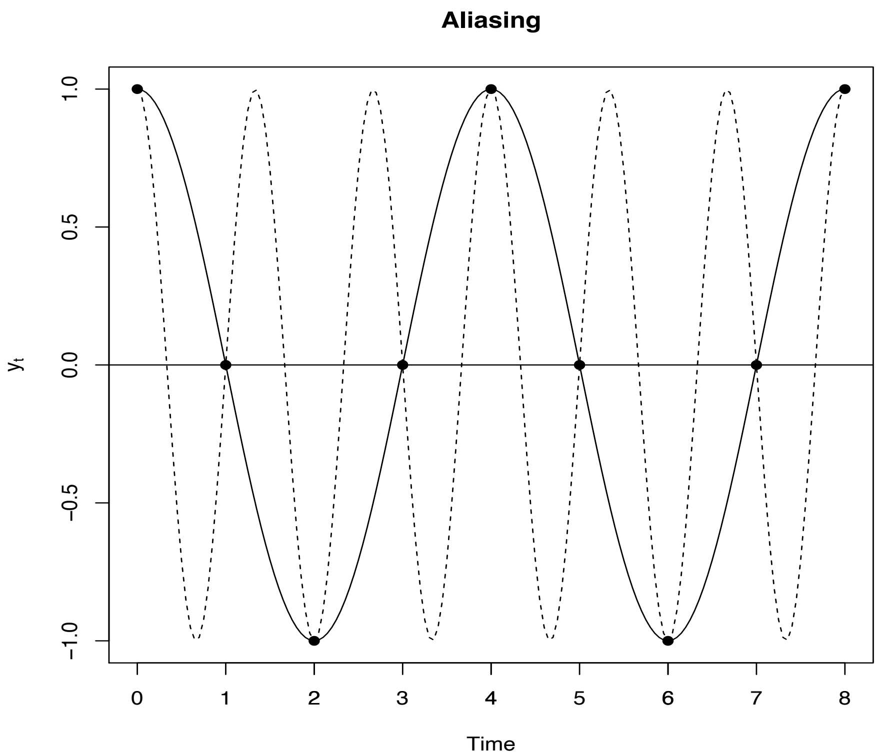
- A: **Amplitude**
- ϕ: **Phase**

- ω: **Frequency**
- $\frac{1}{\omega}$: **Period**

Period and **frequency** are inversely related.

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.



$\omega = 0.25$ (bold)

$\omega = 0.75$ (dashed)

Notice that at the discrete time points 0, 1, 2, 3,…, 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 1/4 and 3/4 are aliased with one another.

Some trigonometric identities

$$\tan\theta = \frac{\sin\theta}{\cos\theta}$$

$$\sin^2\theta + \cos^2\theta = 1$$

$$\sin(a + b) = \sin a \cos b + \cos a \sin b,$$

$$\cos(a + b) = \cos a \cos b - \sin a \sin b.$$

Having the ϕ inside the cosine function $A\cos(2\pi\omega t + \phi)$ 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

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

$$= U_1\cos(2\pi\omega t) + U_2\sin(2\pi\omega t)$$

We assume U_1 and U_2 are iid Gaussian with zero mean and fixed variance.

Periodic Time Series

Consider time series

$$X_t = A\sin(2\pi\omega t) + B\cos(\pi\omega t)$$

where A, B are uncorrelated, mean zero, variance σ^2 and $C^2 = A^2 + B^2$ and $\tan\phi = \frac{B}{A}$.

Then,

$$\mu_t = E[X_t] = 0$$

$$\gamma(h) = \sigma^2\cos(2\pi\omega h)$$

$$\gamma(h) = \gamma(t, t+h)$$

$$= \mathrm{Cov}(X_t, X_{t+h})$$

So, $\{X_t\}$ is stationary.

Multiple frequencies and amplitudes

The autocovariance of the sum of two uncorrelated time series is the sum of their autocovariances. Thus, the autocovariance of a sum of random sinusoids is a sum of sinusoids with the corresponding frequencies:

$$X_t = \sum_{j=1}^{k} A_j \sin\left(2\pi\omega_j t\right) + B_j \cos(\pi\omega_j t)$$

$$\gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(2\pi\omega h)$$

where A_j and B_j are uncorrelated, mean zero, variance σ_j^2 .

A consequence of the representation of X_t is that any stationary time series may be thought of, approximately, as the random superposition of sines and cosines oscillating at various frequencies.

What is the variance of X_t?

❑ Spectral density

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

Before we discuss that Fourier transform, we'll discuss the *Fourier transform the autocovariance function* $\gamma(h)$*,* which is the “spectral density."

We can represent autovariance $\gamma(h)$ using a Fourier series. The coefficients are the variances of the sinusoidal components.

Autocovariance is in terms of lags whereas spectral density is in terms of cycles.

Recall Euler's Formula:

$$e^{ix} = \cos(x) + i\sin(x)$$

So,

$$\cos x = \frac{e^{-ix} + e^{ix}}{2}$$

$$\sin x = \frac{e^{-ix} - e^{ix}}{2}$$

Suppose $\sum_{-\infty}^{\infty} |\gamma(h)| < \infty$, then the **spectral density** is defined as

$$f(\omega) = \sum_{-\infty}^{\infty} \gamma(h) e^{-2\pi i\omega h}$$

for $-\infty < \omega < \infty$

Motivation Example

Here is an example of a spectral representation of an autocovariance function.

Let $X_t = A\ \sin\left(2\pi \frac{1}{4} t\right) + B\ \cos\left(2\pi \frac{1}{4} t\right)$

$$\gamma(h) = \sigma^2 \cos\left(2\pi \frac{1}{4} t\right) = \sigma^2 \frac{e^{-i2\pi\frac{1}{4}t} + e^{i2\pi\frac{1}{4}t}}{2} = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi\omega h}\, dF(\omega)$$

$$F(\omega) = \begin{cases} 0 & \omega < \frac{1}{4} \\ \frac{\sigma^2}{2} & -\frac{1}{4} \leq \omega < \frac{1}{4} \\ \sigma^2 & \omega \geq \frac{1}{4} \end{cases}$$

This $F(\omega)$ **always** exists for **all** stationary processes.

Proposition:

Let X_t be **stationary** with an autocovariance function $\gamma(h)$. Then there exists a **unique** monotonically increasing function $F(\omega)$, called the **spectral distribution function**, that satisfies

- $F(-\infty) = F\left(-\frac{1}{2}\right) = 0 \;\; for\ \omega \leq \frac{1}{2}$

- $F(\infty) = F\left(\frac{1}{2}\right) = \gamma(0)\ for\ \omega \geq \frac{1}{2}$

- $\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi\omega h}\, dF(\omega)$

Remark: 1. We can split F into three components: discrete, continuous, and singular.
2. If $\gamma(h)$ is absolutely summable, F is continuous: $dF(\omega) = f(\omega)d\omega$.
3. If $\gamma(h$ is a sum of sinusoids, F is discrete.

A periodic time series

$$X_t = \sum_{j=1}^{k} A_j \sin(2\pi\omega_j t) + B_j \cos(\pi\omega_j t)$$

$$= \sum_{j=1}^{k} (A_j^2 + B_j^2)^{1/2} \sin\left(2\pi\omega_j t + \tan^{-1}\frac{B_j}{A_j}\right)$$

$$\gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(2\pi\omega h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi\omega h}\, dF(\omega)$$

where the discrete distribution $F = \sigma^2(F_1 + \cdots + F_k)$

$$F_j(\omega) = \begin{cases} 0 & if\ \omega < -\omega_j \\ \frac{1}{2} & if\ -\omega_j \le \omega \le \omega_j \\ 1 & otherwise \end{cases}$$

For ARMA models, we will also have a spectral representation of the autocovariance function, but the integral will be a smoother blend without any jumps:

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi\omega h}\, dF(\omega) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi\omega h}\, f(\omega) d\omega$$

Suppose $\sum_{-\infty}^{\infty} |\gamma(h)| < \infty$, then the **spectral density** is

$$f(\omega) = \sum_{-\infty}^{\infty} \gamma(h) e^{-2\pi i\omega h}$$

for $-1/2 < \omega < 1/2$. (*f is periodic, with period 1.*)

Autocovariance is in terms of **lags** whereas spectral density is in terms of **cycles.**

Wold's decomposition

Notice that the following is is deterministic (once we've seen the past, we can predict the future without error).

$$X_t = \sum_{j=1}^{k} A_j \sin\left(2\pi\omega_j t\right) + B_j \cos(\pi\omega_j t)$$

Wold showed that every **stationary** process can be represented as

$$X_t = X_t^{(d)} + X_t^{(n)}$$

where $X_t^{(d)}$ is deterministic and $X_t^{(n)}$ is nondeterministic

(The decomposition of a spectral distribution function as $F^{(d)} + F^{(c)}$.)

Example: $X_t = A\sin(2\pi\lambda t) + \frac{\theta(B)}{\phi(B)} W_t$

Some Remarks on Spectral density

- The spectral density provides information about the relative strengths of the various frequencies for explaining the variation in the time series.
- The spectral density is also called the power spectrum.
- Remember that $\gamma(h)$ completely determines the distribution for a stationary Gaussian process. So, the spectral density also completely determines the distribution for a stationary Gaussian process.

When $h = 0$,

$$\gamma(h) = Var(X_t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\omega)d\omega$$

An interpretation is that the total integrated spectral density equals the variance of the time series. Thus the spectral density within a particular interval of frequencies can be viewed as the amount of the variance explained by those frequencies.

Proposition:

- $f(\omega) \geq 0$, because $\gamma(h)$ is non-negative definite.
- $f(\omega)$ is even, i.e. $f(\omega) = f(-\omega)$
- $f(\omega) = f(\omega + 1)$
- $f(\omega)$ is periodic, with period 1. (Since $e^{-2\pi i\omega h}$ periodic with period 1.)
- $$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi\omega h} f(\omega) d\omega$$

Example: Spectral Density of White Noise

For white noise $\{W_t\}$, we have autocovariance $\gamma(0) = \sigma^2$ and $\gamma(h) = 0$ for $h \neq 0$. Thus

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

This means all frequencies receive equal weight. This is analogous to the spectrum of white light, where all colors enter equally in white light. (Hence the term white noise.)

Example: Spectral Density of AR(1)

$$X_t = \phi X_{t-1} + W_t$$

$$\gamma(h) = \sigma^2 \frac{\phi^h}{1-\phi^2}$$

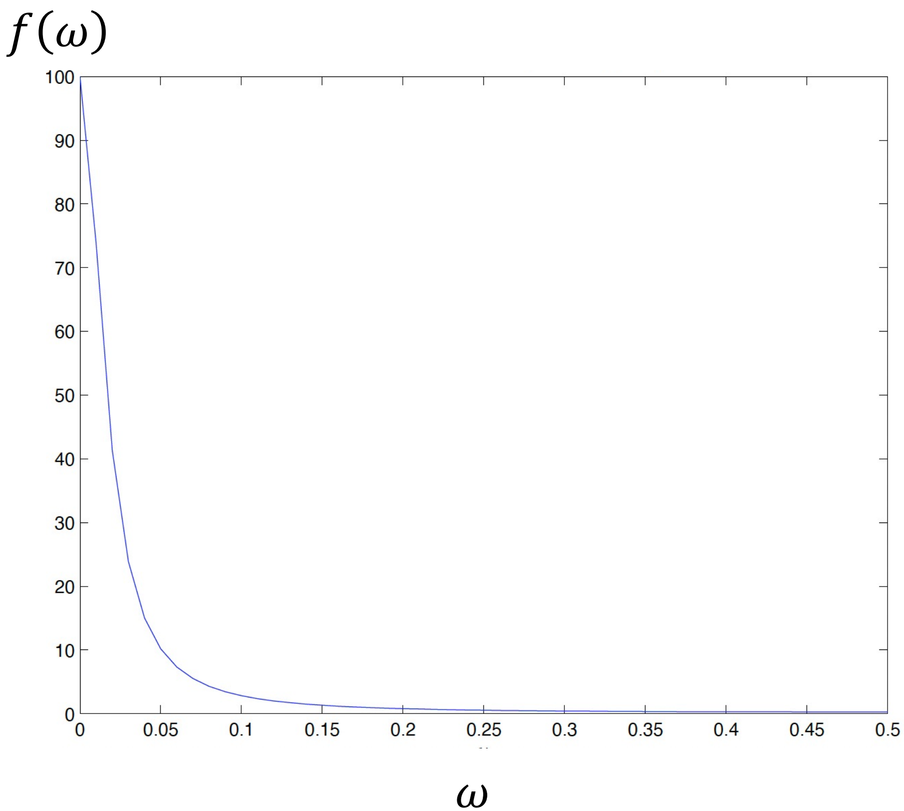
$$f(\omega) = \sum_{-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h} = \cdots$$

$$= \frac{\sigma^2}{1 - 2\phi \cos(2\pi\omega) + \phi^2}$$

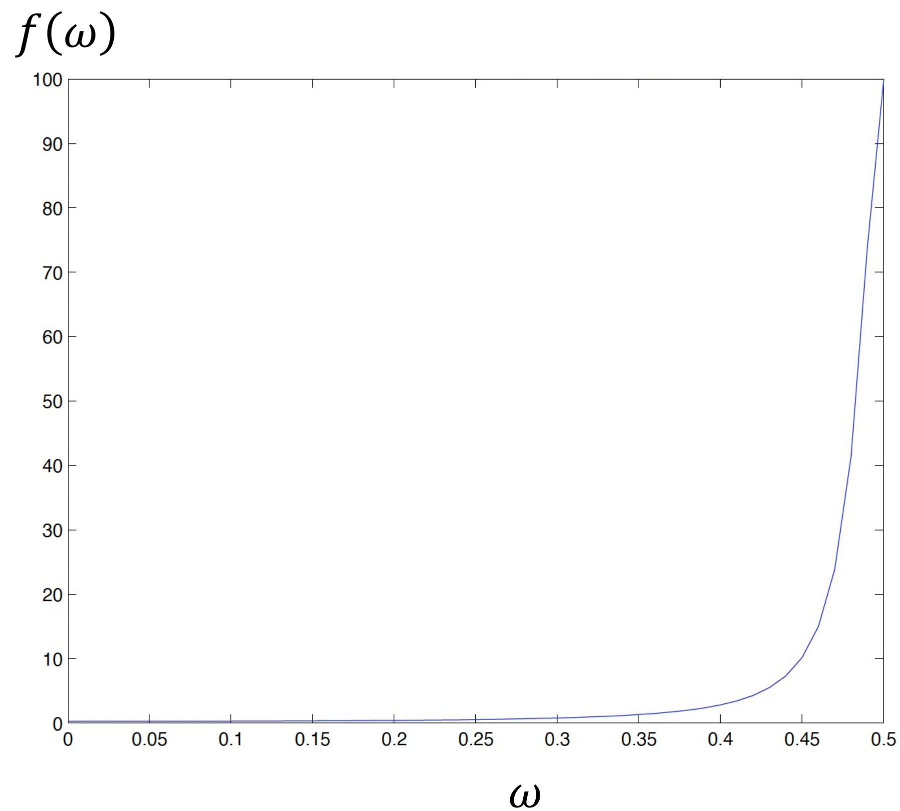
- If $\phi > 0$ (positive autocorrelation), spectrum is dominated by low frequency components—smooth in the time domain.
- If $\phi < 0$ (negative autocorrelation), spectrum is dominated by high frequency components—rough in the time domain.

Example: Spectral Density of AR(1)

$$X_t = 0.9X_{t-1} + W_t$$



$$X_t = -0.9X_{t-1} + W_t$$



Example: Spectral Density of MA(1)

$$X_t = W_t + \theta W_{t-1}$$

$$\gamma(h) = \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0 \\ \sigma^2\theta & \text{if } h = 1, or\ -1 \\ 0 & \text{others} \end{cases}$$

$$f(\omega) = \sum_{-\infty}^{\infty} \gamma(h) e^{-2\pi i\omega h} = \cdots$$

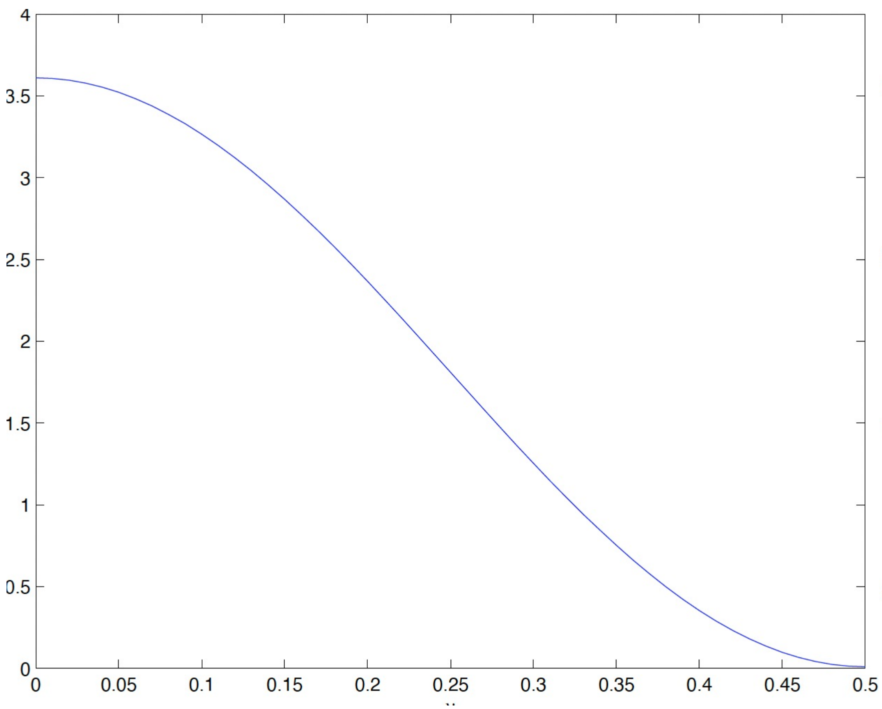
$$= \gamma(0) + 2\gamma(1)\cos(2\pi\omega)$$

$$= \sigma^2(1 + 2\theta^2 + 2\theta\cos(2\pi\omega))$$

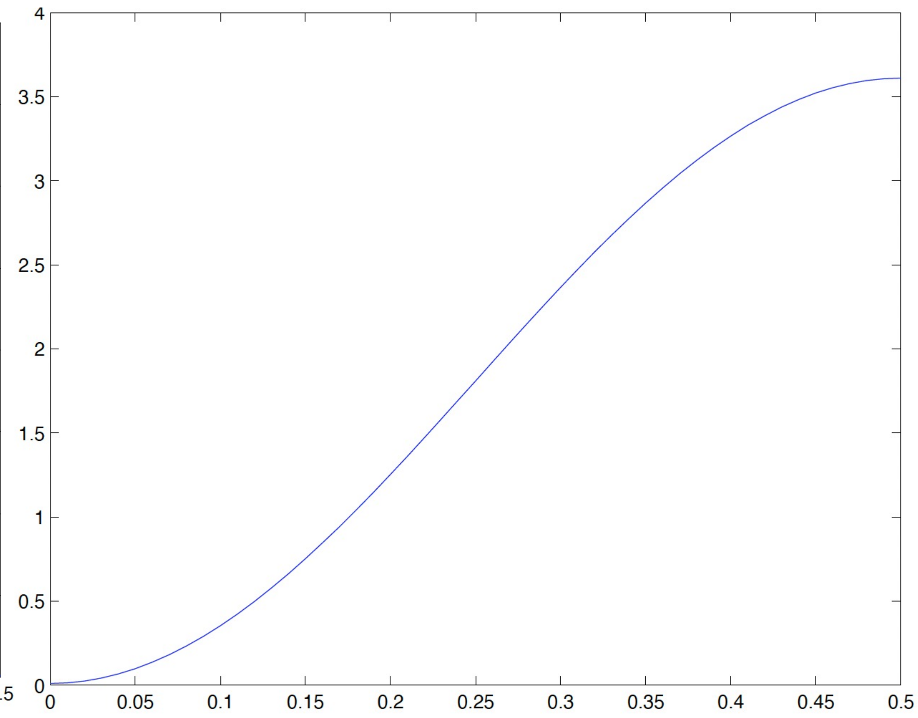
If $\theta > 0$ (positive autocorrelation), spectrum is dominated by low frequency components—smooth in the time domain.
If $\theta < 0$ (negative autocorrelation), spectrum is dominated by high frequency components—rough in the time domain.

Example: Spectral Density of MA(1)

$$X_t = W_t + 0.9W_{t-1}$$



$$X_t = W_t - 0.9W_{t-1}$$



❑ Spectral density for causal ARMA processes (or a linear process)

A zero-mean **causal** ARMA(p,q) process $\phi(B)X_t = \theta(B)W_t$ can be written as a linear process:

$$X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j} = \psi(B)W_t \qquad \text{where } \psi(B) = \frac{\theta(B)}{\phi(B)}$$

The autocovariance function is

$$\gamma(h) = E(X_t X_{t+h}) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$$

Define the autocovariance **generating function** as

$$\Gamma(B) := \sum_{h=-\infty}^{\infty} \gamma(h) B^h = \sum_{h=-\infty}^{\infty} \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h} B^h$$

$$= \sigma^2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \psi_j \psi_k \, B^{k-j} = \sigma^2 \sum_{j=0}^{\infty} \psi_j B^{-j} \sum_{k=0}^{\infty} \psi_k \, B^k$$

$$= \sigma^2 \psi(B^{-1}) \psi(B)$$

Lemma: If $\Gamma(B) = \sum_{h=-\infty}^{\infty} \gamma(h) B^h$, then,

$$f(\omega) = \sum_{-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h} = \Gamma\left(e^{-2\pi i \omega h}\right) = \sigma^2 \psi\left(e^{-2\pi i \omega h}\right) \psi(e^{2\pi i \omega h})$$

$$= \sigma^2 \left|\psi\left(e^{2\pi i \omega h}\right)\right|^2$$

Spectral density for a causal ARMA process

Proposition: The spectral density for a causal ARMA process can be expressed as

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

This is also called the **rational spectrum** of an ARMA(p,q).

Recall (Fundamental Theorem of Algebra) that every degree p polynomial $g(z)$ can be factorized as

$$g(z) = a(z - z_1)(z - z_2) \cdots (z - z_p)$$

where $z_1, \dots, z_p$ are complex roots.

For the MA and AR polynomials,

$$\theta(z) = \theta_q(z - z_1)(z - z_2)\cdots(z - z_q)$$

$$\phi(z) = \phi_p(z - p_1)(z - p_2)\cdots(z - p_p)$$

where $z_1, \ldots, z_q$ and $p_1, \ldots, p_p$ are called the **zeros** and **poles**.

$$f(\omega) = \sigma^2 \left|\frac{\theta\left(e^{-2\pi i\omega}\right)}{\phi\left(e^{-2\pi i\omega}\right)}\right|^2 = \sigma^2 \left|\frac{\theta_q \prod_{j=1}^{q}\left(e^{-2\pi i\omega} - z_j\right)}{\phi_p \prod_{j=1}^{p}\left(e^{-2\pi i\omega} - p_j\right)}\right|^2 = \sigma^2 \frac{\theta_q^2 \prod_{j=1}^{q}\left|e^{-2\pi i\omega} - z_j\right|^2}{\phi_p^2 \prod_{j=1}^{p}\left|e^{-2\pi i\omega} - p_j\right|^2}$$

- As ω varies from 0 to 1/2, $e^{-2\pi i\omega}$ moves clockwise around the unit circle from 1 to $e^{-\pi i} = -1$.
- And the value of $f(\omega)$ goes up as this point moves closer to (further from) the poles p_j (zeros z_j).

Remarks on Spectral Density of AR(1) and MA(1)

Recall AR(1): $\phi(z) = 1 - \phi_1 z$. The pole is at $1/\phi_1$.

- If $\phi_1 > 0$, the pole is to the right of 1, so the spectral density decreases as ω moves away from 0.
- If $\phi_1 < 0$, the pole is to the left of -1, so the spectral density is at its maximum when $\omega = 0.5$.

Recall MA(1): $\theta(z) = 1 + \theta_1 z$. The zero is at $-1/\theta_1$.

- If $\theta_1 > 0$, the zero is to the left of -1, so the spectral density decreases as moves towards -1.
- If $\theta_1 < 0$, the zero is to the right of 1, so the spectral density is at its minimum when $\omega = 0$.

Example: Spectral Density of AR(2)

Suppose we have the following AR(2) model:

$$X_t = X_{t-1} - 0.9X_{t-2} + W_t \qquad \text{with } \sigma^2 = 1$$

The roots (poles) of the AR polynomial $\phi(z) = 0.9z^2 - z + 1$ are

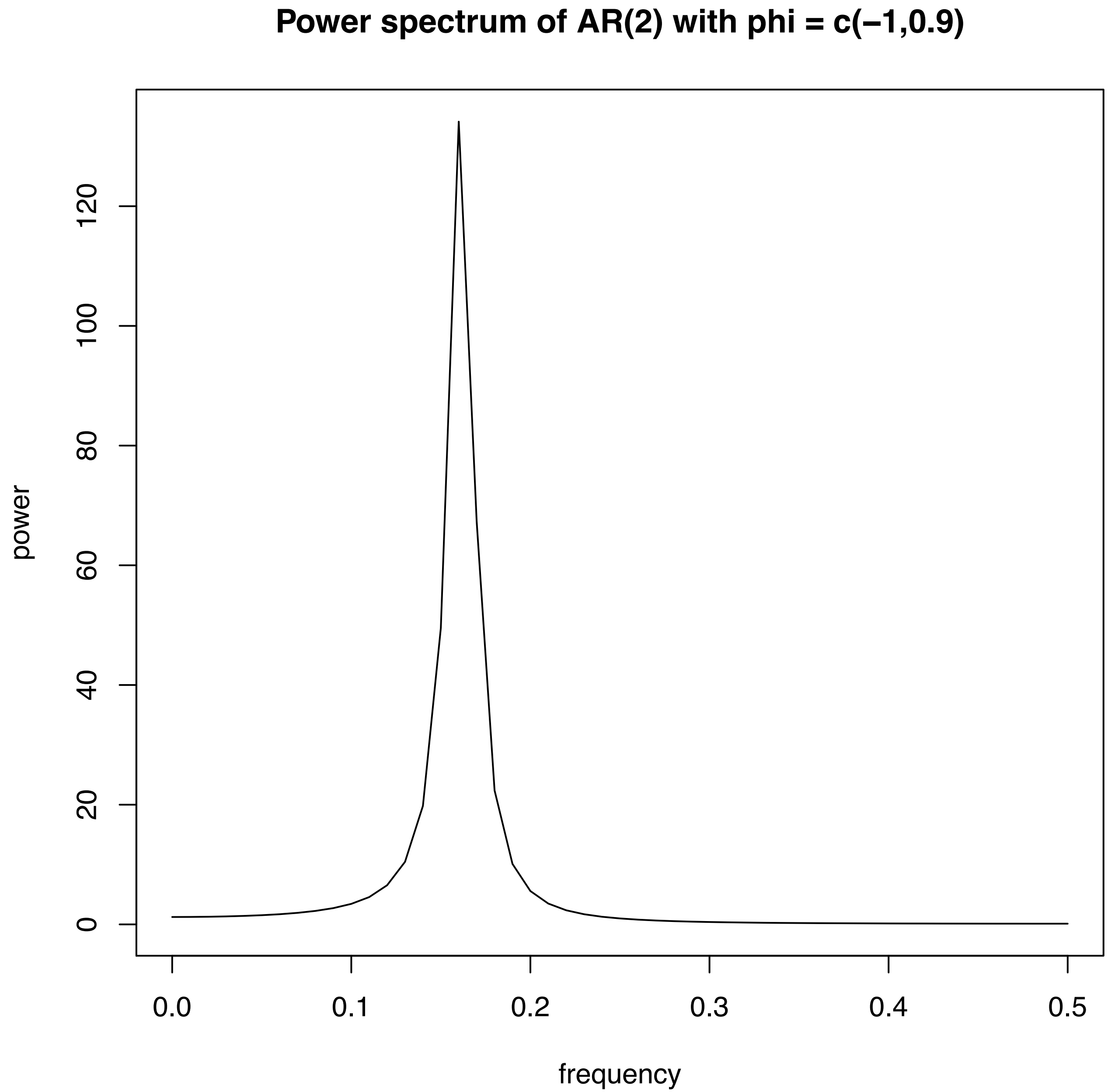
$$p1, p2 = 0.555 \pm \; i\, 0.8958.$$

Using the representation, the spectral density is

$$f(\omega) = \frac{1}{\phi_2^2 |e^{-2\pi i\omega} - p_1|^2 |e^{-2\pi i\omega} - p_2|^2}$$

The peaks of the spectral density for this process occurs when

$e^{-2\pi i\omega}$ is near $1.054e^{-2\pi i\, 0.16165}$


Power spectrum of AR(2) with phi = c(−1,0.9)
power
0
20
40
60
80
100
120
0.0
0.1
0.2
0.3
0.4
0.5
frequency

Example: Seasonal ARMA

Suppose we have the following Seasonal AR model:

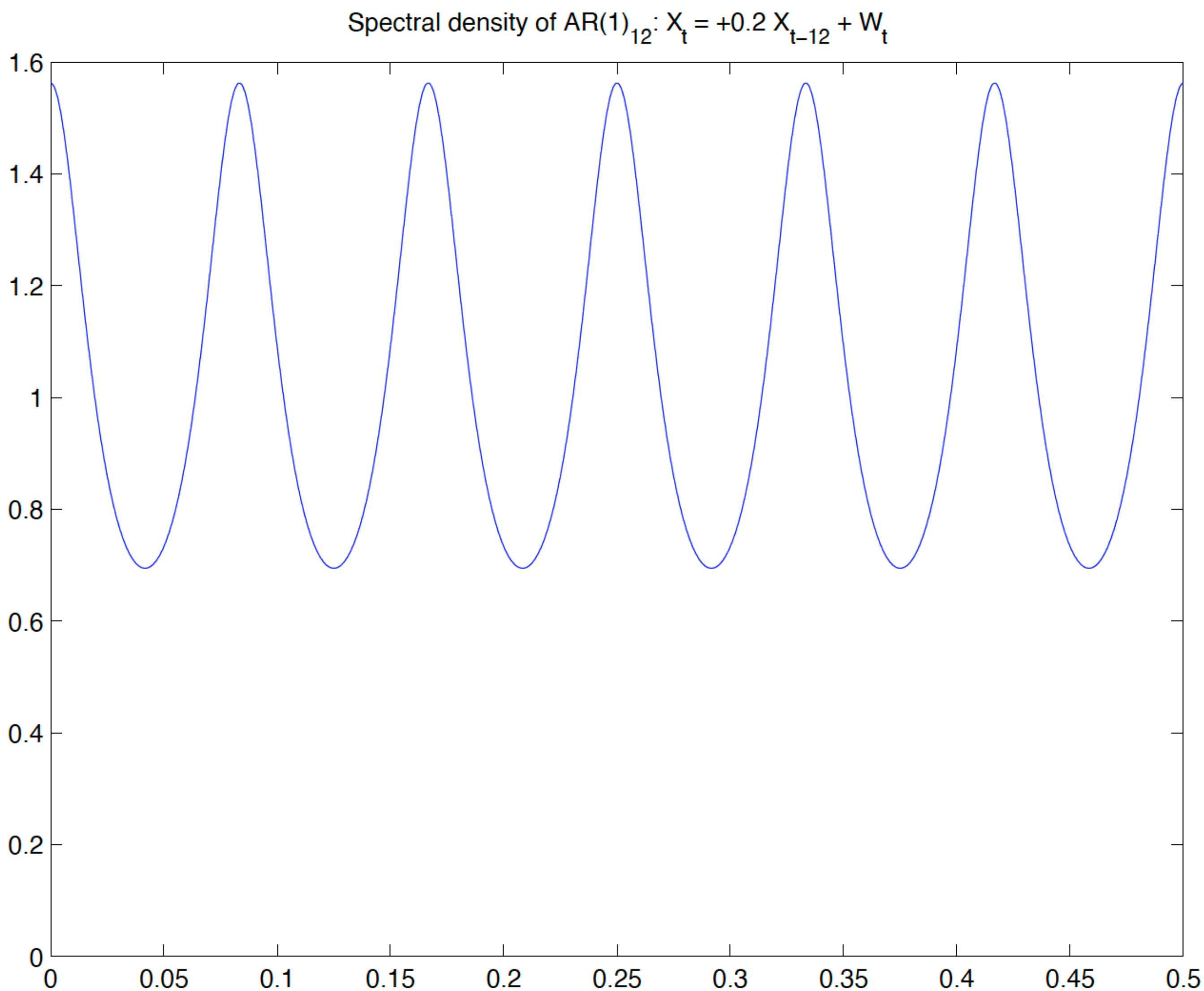
$$X_t = \phi_1 X_{t-12} + W_t$$

$$\psi(B) = \frac{1}{1 - \phi_1 B^{12}}$$

$$f(\omega) = \sigma^2 \frac{1}{(1 - \phi_1 e^{-2\pi i 12\omega})(1 - \phi_1 e^{2\pi i 12\omega})}$$

$$= \sigma^2 \frac{1}{1 - 2\phi_{12}\cos(24\pi\omega) + \phi_{12}^2}$$

Notice that $f(\omega)$ is periodic with period $1/12$


Spectral density of AR(1)12: Xt = +0.2 Xt−12 + Wt
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5

Example: Multiplicative seasonal ARMA

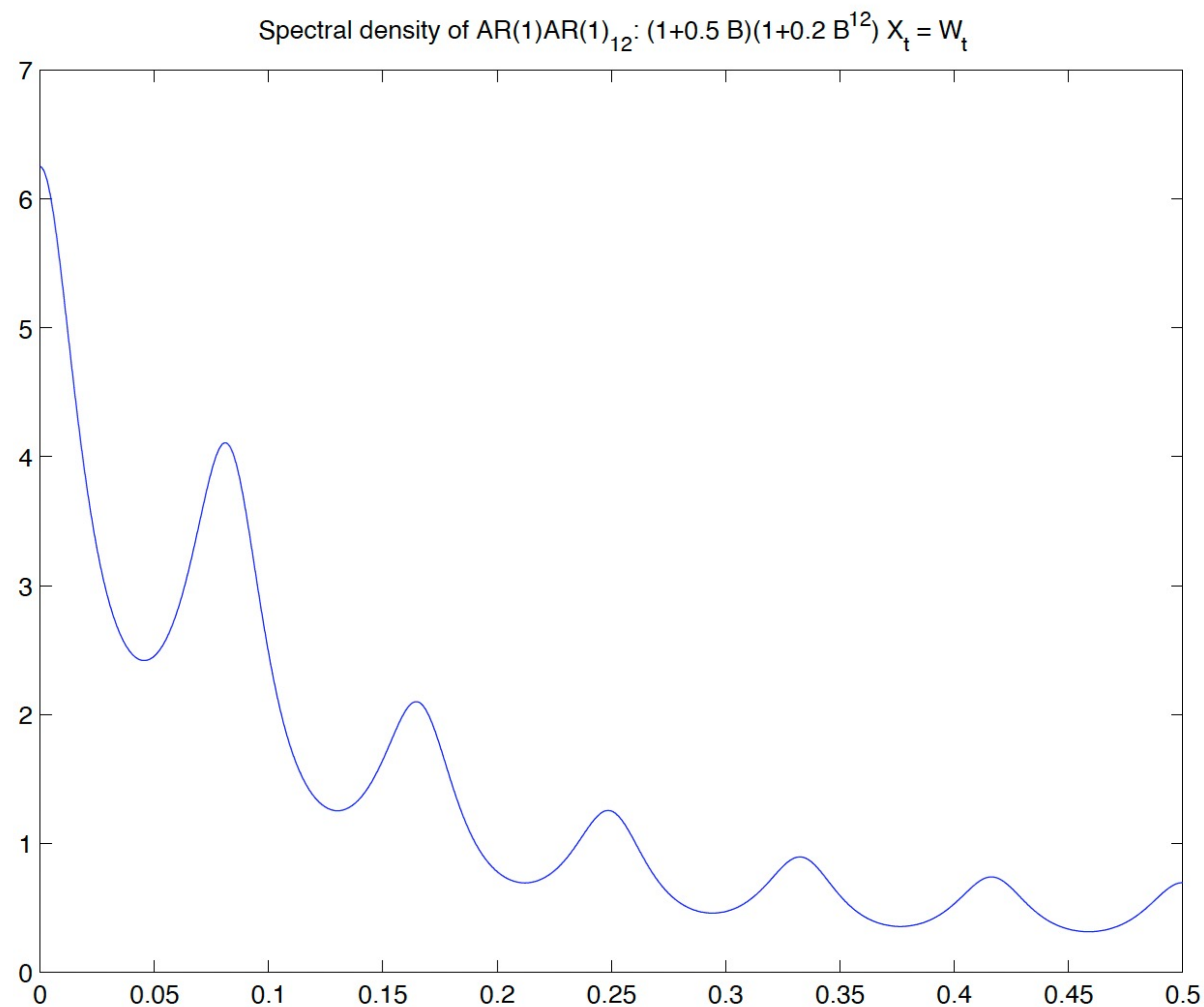
Suppose we have the following Seasonal AR model:

$$(1-\phi_{12}B^{12})(1-\phi_1 B)X_t = W_t$$

$$f(\omega) = \sigma^2 \frac{1}{(1-2\phi_{12}\cos(24\pi\omega)+\phi_{12}^2)(1-2\phi_1\cos(2\pi\omega)+\phi_1^2)}$$

This is a scaled product of the AR(1) spectrum and the (periodic) $AR(1)_{12}$ spectrum.

The $AR(1)_{12}$ poles give peaks when $e^{-2\pi i\omega}$ is at one of the 12th roots of 1; the AR(1) poles give a peak near $e^{-2\pi i\omega} = 1$.



Spectral density of AR(1)AR(1)12: (1+0.5 B)(1+0.2 B12) Xt = Wt
7
6
5
4
3
2
1
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5

❑ Time-invariant linear filters

Definition: A time series $\{Y_t\}$ is the output of a **linear filter** $A = \{a_{t,j}\}$ with input $\{X_t\}$ if

$$Y_t = \sum_{j=-\infty}^{\infty} a_{t,j} X_{t-j}$$

1. If $a_{t,j}$ is independent of t, $(i.e., a_{t,j} = \psi_j)$, then we say that the filter is **time-invariant**.
2. If $\psi_j = 0$ for $j < 0$, we say the filter ψ_j is **causal**.

A **filter** is an operator; given a time series $\{X_t\}$, it maps to a time series $\{Y_t\}$.

For example, we can think of a linear process

$$X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$$

as the output of a causal linear filter with a white noise input.

Examples:

1. $Y_t = X_{-t}$ is linear, but not time-invariant.

2. $Y_t = \frac{1}{3}(X_{t-1} + X_t + X_{t+1})$ is linear, time-invariant, but not causal:

$$\psi_j = \begin{cases} \frac{1}{3} & if\ |j| \le 1 \\ 0 & otherwise \end{cases}$$

3. For polynomials $\phi(B)\ and\ \theta(B)$ with roots outside the unit circle, $\psi(B) = \frac{\theta(B)}{\phi(B)}$ is a linear, time-invariant, causal filter.

4. The linear operation

$$\sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$$

is called the **convolution** of X with ψ

The sequence ψ is also called the **impulse response**, since the output $\{Y_t\}$ of the linear filter in response to a **unit impulse**,

$$X_t = \begin{cases} 1 & if\ t = 0 \\ 0 & otherwise \end{cases}$$

is

$$Y_t = \psi(B)X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} = \psi_t$$

❑ Frequency response of a time-invariant linear filter

Consider a time-invariant linear filter

$$Y_t = \psi(B)X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$$

Suppose $\{X_t\}$ has spectral density $f_x(\omega)$ and ψ is **stable**, that is, $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$.
Then $Y_t = \psi(B)X_t$ has **spectral density**

$$f_y(\omega) = \left|\psi\left(e^{2\pi i\omega}\right)\right|^2 f_x(\omega)$$

The function $\omega \to \psi\left(e^{2\pi i\omega}\right)$ (the polynomial $\psi(z)$ evaluated on the unit circle) is known as the **frequency response** or **transfer function** of the linear filter.

The squared modulus, $\omega \to \left|\psi\left(e^{2\pi i\omega}\right)\right|^2$ is known as the **power transfer function** of the filter.

When we pass a time series $\{X_t\}$ through a linear filter, the spectral density is multiplied, frequency-by-frequency, by the squared modulus of the frequency response $\omega \to \left|\psi\left(e^{2\pi i\omega}\right)\right|^2$

This is a version of the equality $Var(aX) = a2Var(X)$, but the equality is true for the component of the variance at every frequency.

This is also the origin of the name ‘filter.’

For example, a linear process is a special case of linear filter

$$Y_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$$

$$f_y(\omega) = \left|\psi\left(e^{2\pi i\omega}\right)\right|^2 f_w(\omega)$$

For an ARMA model, $\psi(B) = \frac{\theta(B)}{\phi(B)}$, the model $\{Y_t\}$ has rational spectrum,

$$f_y(\omega) = \sigma^2 \left|\frac{\theta\left(e^{-2\pi i\omega}\right)}{\phi\left(e^{-2\pi i\omega}\right)}\right|^2 = \sigma^2 \frac{\theta_q^2 \prod_{j=1}^{q} \left|e^{-2\pi i\omega} - z_j\right|^2}{\phi_p^2 \prod_{j=1}^{p} \left|e^{-2\pi i\omega} - p_j\right|^2}$$

where p_j and z_j are the poles and zeros of the rational function $z \to \theta(z)/\phi(z)$

Example: Moving average

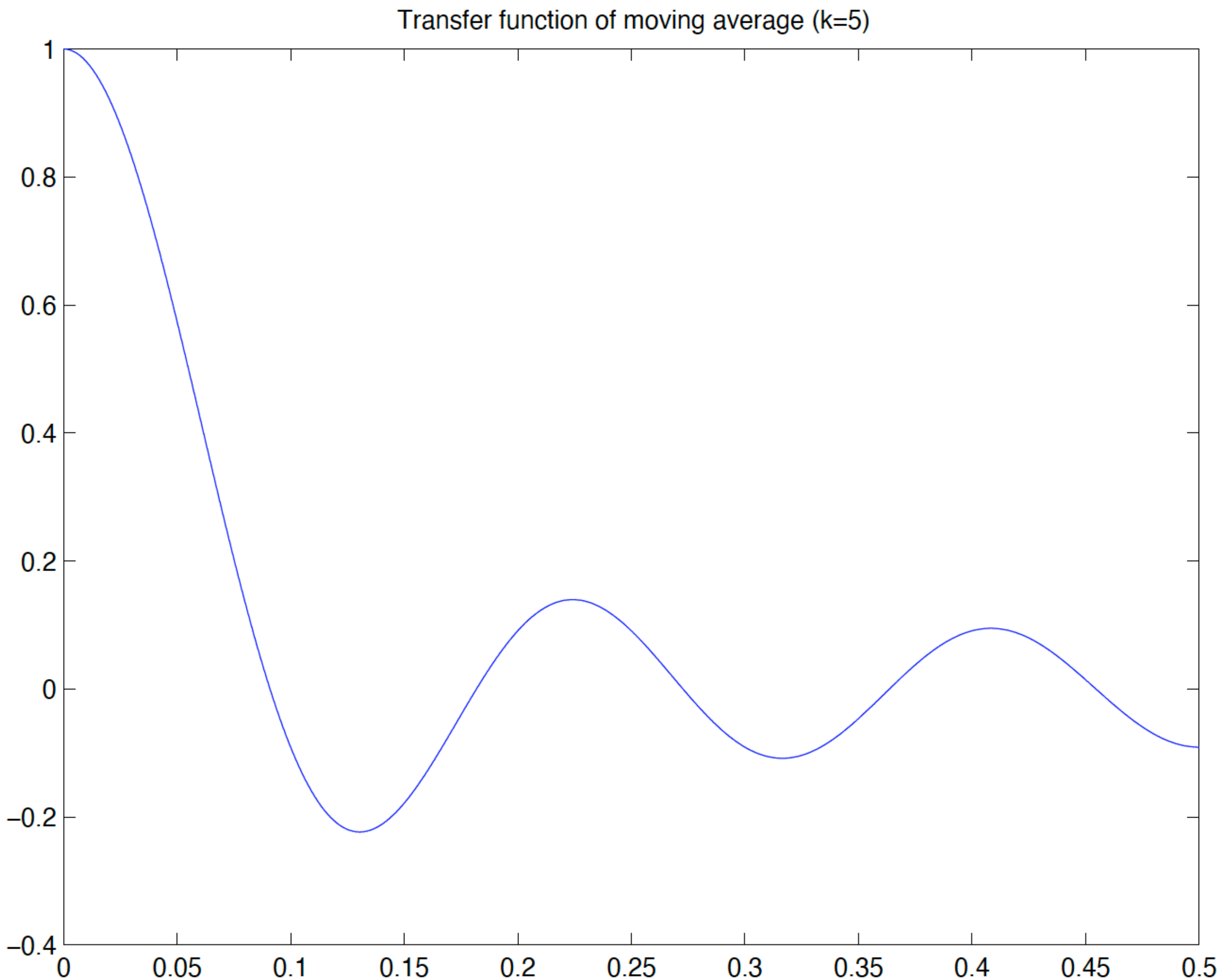
Consider the moving average

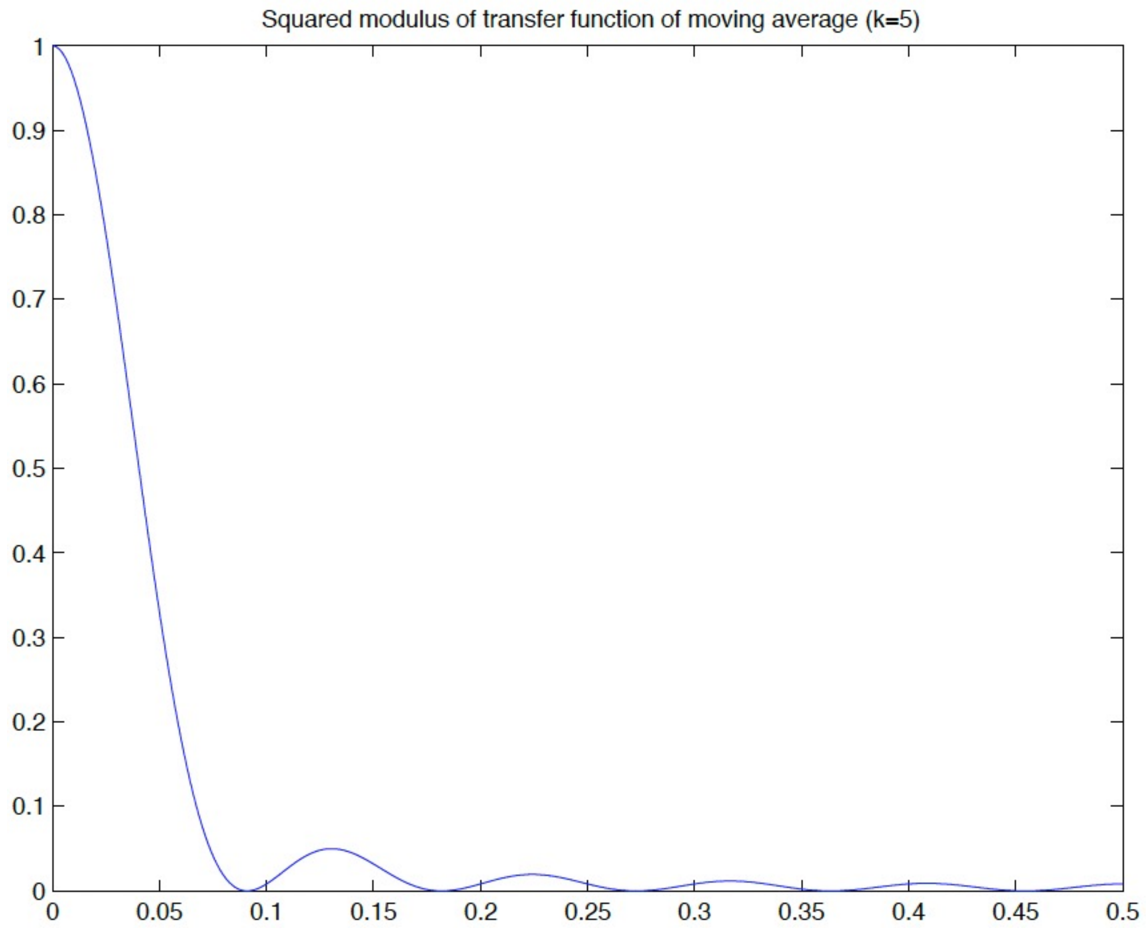
$$Y_t = \frac{1}{2k+1} \sum_{j=-k}^{k} X_{t-j}$$

This is a time invariant linear filter (but it is not causal). Its transfer function is the Dirichlet kernel

$$\psi\left(e^{-2\pi i\omega}\right) = D_k(2\pi\omega) = \frac{1}{2k+1} \sum_{j=-k}^{k} e^{-2\pi i j\omega}$$

$$= \begin{cases} 0 & \text{If } \omega = 0 \\ \dfrac{\sin\left(2\pi\left(k+\frac{1}{2}\right)\omega\right)}{(2k+1)\sin(\pi\omega)} & \text{otherwise} \end{cases}$$


Transfer function of moving average (k=5)
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5



This is a low-pass filter: It preserves low frequencies and diminishes high frequencies. It is often used to estimate a monotonic trend component of a series.

Example: Differencing

Consider the first difference

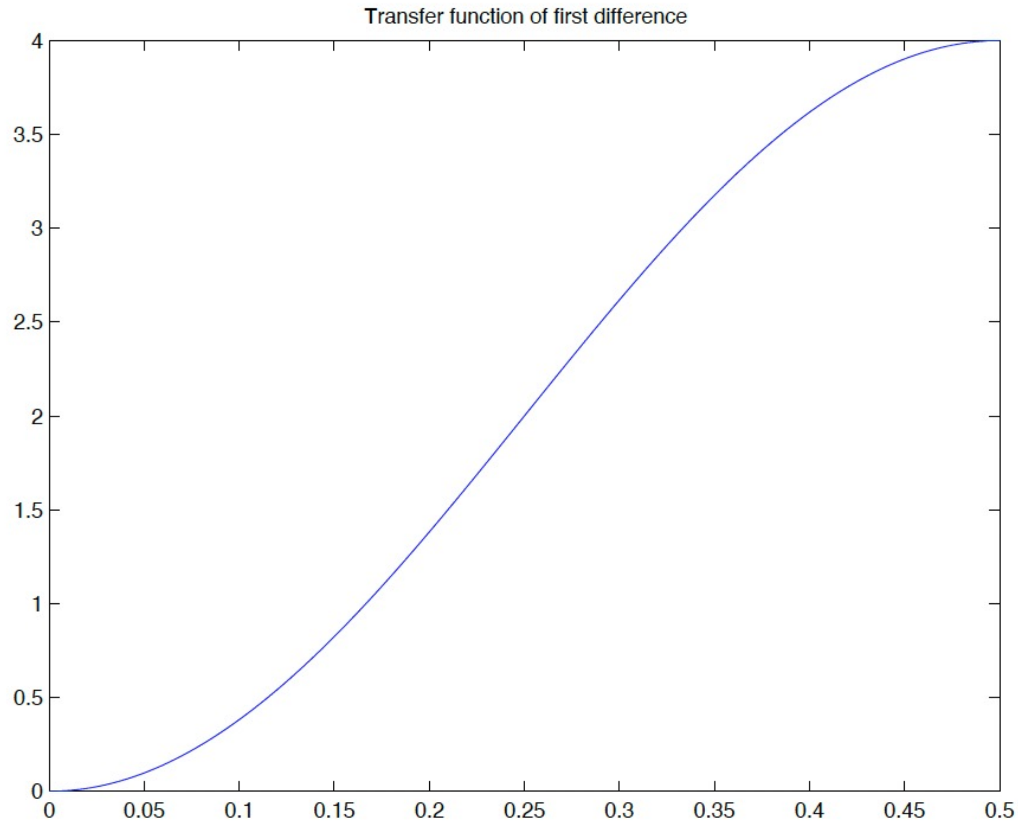
$$Y_t = (1 - B)X_t$$

This is a time invariant, causal, linear filter. Its transfer function is

$$\psi\left(e^{-2\pi i\omega}\right) = 1 - e^{-2\pi i\omega}$$

So

$$\left|\psi\left(e^{-2\pi i\omega}\right)\right|^2 = 2(1 - \cos(2\pi\omega))$$



This is a high-pass filter: It preserves high frequencies and diminishes low frequencies. It is often used to eliminate a trend component of a series.

- "Time Series Analysis and Its Applications", 4th ed. 2017, by Shumway and Stoffer.

Sections 4.1, 4.2, 4.3