Introduction to Time Series Analysis. Lecture 15.

Spectral Analysis

2. Spectral distribution function.
3. Wold’s decomposition.
Spectral Analysis

Idea: decompose a stationary time series \( \{X_t\} \) into a combination of sinusoids, with random (and uncorrelated) coefficients.

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

This is referred to as ‘spectral analysis’ or analysis in the ‘frequency domain,’ in contrast to the time domain approach we have considered so far.

The frequency domain approach considers regression on sinusoids; the time domain approach considers regression on past values of the time series.
A periodic time series

Consider

\[ X_t = A \sin(2\pi\nu t) + B \cos(2\pi\nu t) \]
\[ = C \sin(2\pi\nu t + \phi), \]

where \( A, B \) are uncorrelated, mean zero, variance \( \sigma^2 = 1 \), and \( C^2 = A^2 + B^2 \), \( \tan \phi = B/A \). Then

\[ \mu_t = \mathbb{E}[X_t] = 0 \]
\[ \gamma(t, t + h) = \cos(2\pi\nu h). \]

So \( \{X_t\} \) is stationary.
An aside: Some trigonometric identities

\[
\tan \theta = \frac{\sin \theta}{\cos \theta},
\]
\[
\sin^2 \theta + \cos^2 \theta = 1,
\]
\[
\sin(a + b) = \sin a \cos b + \cos a \sin b,
\]
\[
\cos(a + b) = \cos a \cos b - \sin a \sin b.
\]
A periodic time series

For $X_t = A \sin(2\pi \nu t) + B \cos(2\pi \nu t)$, with uncorrelated $A, B$ (mean 0, variance $\sigma^2$), $\gamma(h) = \sigma^2 \cos(2\pi \nu h)$.

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:

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

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

where $A_j, B_j$ are uncorrelated, mean zero, and $\text{Var}(A_j) = \text{Var}(B_j) = \sigma_j^2$. 
A periodic time series

\[ X_t = \sum_{j=1}^{k} (A_j \sin(2\pi \nu_j t) + B_j \cos(2\pi \nu_j t)), \quad \gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(2\pi \nu_j h). \]

Thus, we can represent \( \gamma(h) \) using a Fourier series. The coefficients are the variances of the sinusoidal components.

The spectral density is the continuous analog: the Fourier transform of \( \gamma \).

(The analogous spectral representation of a stationary process \( X_t \) involves a stochastic integral—a sum of discrete components at a finite number of frequencies is a special case. We won’t consider this representation in this course.)
If a time series \( \{X_t\} \) has autocovariance \( \gamma \) satisfying
\[
\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty,
\]
then we define its **spectral density** as
\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h}
\]
for \(-\infty < \nu < \infty\).
Spectral density: Some facts

1. We have \( \sum_{h=-\infty}^{\infty} |\gamma(h) e^{-2\pi i \nu h}| < \infty. \)
   This is because \( |e^{i\theta}| = |\cos \theta + i \sin \theta| = (\cos^2 \theta + \sin^2 \theta)^{1/2} = 1, \)
   and because of the absolute summability of \( \gamma. \)

2. \( f \) is periodic, with period 1.
   This is true since \( e^{-2\pi i \nu h} \) is a periodic function of \( \nu \) with period 1.
   Thus, we can restrict the domain of \( f \) to \(-1/2 \leq \nu \leq 1/2.\) (The text does this.)
3. $f$ is even (that is, $f(\nu) = f(-\nu)$).

To see this, write

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

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

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

$$= f(\nu).$$

4. $f(\nu) \geq 0$. 
5. $\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \nu h} f(\nu) d\nu$.

$$\int_{-1/2}^{1/2} e^{2\pi i \nu h} f(\nu) d\nu = \int_{-1/2}^{1/2} \sum_{j=-\infty}^{\infty} e^{-2\pi i \nu(j-h)} \gamma(j) d\nu$$

$$= \sum_{j=-\infty}^{\infty} \gamma(j) \int_{-1/2}^{1/2} e^{-2\pi i \nu(j-h)} d\nu$$

$$= \gamma(h) + \sum_{j \neq h} \frac{\gamma(j)}{2\pi i (j-h)} \left( e^{\pi i (j-h)} - e^{-\pi i (j-h)} \right)$$

$$= \gamma(h) + \sum_{j \neq h} \frac{\gamma(j) \sin(\pi (j-h))}{\pi (j-h)} = \gamma(h).$$
Example: White noise

For white noise \( \{W_t\} \), we have seen that \( \gamma(0) = \sigma_w^2 \) and \( \gamma(h) = 0 \) for \( h \neq 0 \).

Thus,

\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h}
\]

\[
= \gamma(0) = \sigma_w^2.
\]

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.
Example: AR(1)

For $X_t = \phi_1 X_{t-1} + W_t$, we have seen that $\gamma(h) = \sigma_w^2 |\phi_1|^h / (1 - \phi_1^2)$. Thus,

$$f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h} = \frac{\sigma_w^2}{1 - \phi_1^2} \sum_{h=-\infty}^{\infty} \phi_1^{|h|} e^{-2\pi i \nu h}$$

$$= \frac{\sigma_w^2}{1 - \phi_1^2} \left( 1 + \sum_{h=1}^{\infty} \phi_1^h \left( e^{-2\pi i \nu h} + e^{2\pi i \nu h} \right) \right)$$

$$= \frac{\sigma_w^2}{1 - \phi_1^2} \left( 1 + \frac{\phi_1 e^{-2\pi i \nu}}{1 - \phi_1 e^{-2\pi i \nu}} + \frac{\phi_1 e^{2\pi i \nu}}{1 - \phi_1 e^{2\pi i \nu}} \right)$$

$$= \frac{\sigma_w^2}{(1 - \phi_1^2) (1 - \phi_1 e^{-2\pi i \nu}) (1 - \phi_1 e^{2\pi i \nu})}$$

$$= \frac{\sigma_w^2}{1 - 2\phi_1 \cos(2\pi \nu) + \phi_1^2}.$$
Examples

White noise: \( \{W_t\} \), \( \gamma(0) = \sigma_w^2 \) and \( \gamma(h) = 0 \) for \( h \neq 0 \).
\[
f(\nu) = \gamma(0) = \sigma_w^2.
\]

AR(1): \( X_t = \phi_1 X_{t-1} + W_t \), \( \gamma(h) = \sigma_w^2 \phi_1^{|h|}/(1 - \phi_1^2) \).
\[
f(\nu) = \frac{\sigma_w^2}{1-2\phi_1 \cos(2\pi \nu)+\phi_1^2}.
\]
If \( \phi_1 > 0 \) (positive autocorrelation), spectrum is dominated by low frequency components—smooth in the time domain.
If \( \phi_1 < 0 \) (negative autocorrelation), spectrum is dominated by high frequency components—rough in the time domain.
Example: AR(1)

Spectral density of AR(1): $X_t = 0.9 X_{t-1} + W_t$
Example: AR(1)

Spectral density of AR(1): $X_t = -0.9 X_{t-1} + W_t$
Introduction to Time Series Analysis. Lecture 16.

1. Review: Spectral density
2. Examples
4. Autocovariance generating function and spectral density.
Review: Spectral density

If a time series \( \{X_t\} \) has autocovariance \( \gamma \) satisfying
\[
\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty,
\]
then we define its spectral density as
\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h}
\]
for \(-\infty < \nu < \infty\).
Review: Spectral density

1. $f(\nu)$ is real.

2. $f(\nu) \geq 0$.

3. $f$ is periodic, with period 1. So we restrict the domain of $f$ to $-1/2 \leq \nu \leq 1/2$.

4. $f$ is even (that is, $f(\nu) = f(-\nu)$).

5. $\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \nu h} f(\nu) \, d\nu$. 
Examples

White noise: \( \{W_t\} \), \( \gamma(0) = \sigma_w^2 \) and \( \gamma(h) = 0 \) for \( h \neq 0 \).
\[ f(\nu) = \gamma(0) = \sigma_w^2. \]

AR(1): \( X_t = \phi_1 X_{t-1} + W_t \), \( \gamma(h) = \sigma_w^2 \phi_1^{|h|} / (1 - \phi_1^2) \).
\[ f(\nu) = \frac{\sigma_w^2}{1 - 2\phi_1 \cos(2\pi \nu) + \phi_1^2}. \]

If \( \phi_1 > 0 \) (positive autocorrelation), spectrum is dominated by low frequency components—smooth in the time domain.
If \( \phi_1 < 0 \) (negative autocorrelation), spectrum is dominated by high frequency components—rough in the time domain.
Example: AR(1)

Spectral density of AR(1): $X_t = 0.9 X_{t-1} + W_t$
Example: AR(1)

Spectral density of AR(1): $X_t = -0.9 X_{t-1} + W_t$
Example: MA(1)

\[ X_t = W_t + \theta_1 W_{t-1}. \]

\[ \gamma(h) = \begin{cases} 
\sigma_w^2 (1 + \theta_1^2) & \text{if } h = 0, \\
\sigma_w^2 \theta_1 & \text{if } |h| = 1, \\
0 & \text{otherwise.} 
\end{cases} \]

\[ f(\nu) = \sum_{h=-1}^{1} \gamma(h) e^{-2\pi i \nu h} 
= \gamma(0) + 2 \gamma(1) \cos(2\pi \nu) 
= \sigma_w^2 \left(1 + \theta_1^2 + 2\theta_1 \cos(2\pi \nu)\right). \]
Example: MA(1)

\[ X_t = W_t + \theta_1 W_{t-1}. \]

\[ f(\nu) = \sigma_w^2 \left( 1 + \theta_1^2 + 2\theta_1 \cos(2\pi\nu) \right). \]

If \( \theta_1 > 0 \) (positive autocorrelation), spectrum is dominated by low frequency components—smooth in the time domain.

If \( \theta_1 < 0 \) (negative autocorrelation), spectrum is dominated by high frequency components—rough in the time domain.
Example: MA(1)

Spectral density of MA(1): \( X_t = W_t + 0.9 W_{t-1} \)
Example: MA(1)

Spectral density of MA(1): $X_t = W_t - 0.9 W_{t-1}$
Introduction to Time Series Analysis. Lecture 16.

1. Review: Spectral density

2. Examples


4. Autocovariance generating function and spectral density.

Recall: A periodic time series

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

\[ = \sum_{j=1}^{k} (A_j^2 + B_j^2)^{1/2} \sin(2\pi \nu_j t + \tan^{-1}(B_j/A_j)). \]

\[ E[X_t] = 0 \]

\[ \gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(2\pi \nu_j h) \]

\[ \sum_{h} |\gamma(h)| = \infty. \]
For \( X_t = A \sin(2\pi \lambda t) + B \cos(2\pi \lambda t) \), we have \( \gamma(h) = \sigma^2 \cos(2\pi \lambda h) \), and we can write

\[
\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \nu h} dF(\nu),
\]

where \( F \) is the discrete distribution

\[
F(\nu) = \begin{cases} 
0 & \text{if } \nu < -\lambda, \\
\frac{\sigma^2}{2} & \text{if } -\lambda \leq \nu < \lambda, \\
\sigma^2 & \text{otherwise}.
\end{cases}
\]
For any stationary \( \{X_t\} \) with autocovariance \( \gamma \), we can write

\[
\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \nu h} dF(\nu),
\]

where \( F \) is the spectral distribution function of \( \{X_t\} \).

We can split \( F \) into three components: discrete, continuous, and singular.

If \( \gamma \) is absolutely summable, \( F \) is continuous: \( dF(\nu) = f(\nu)d\nu \).

If \( \gamma \) is a sum of sinusoids, \( F \) is discrete.
The spectral distribution function

For $X_t = \sum_{j=1}^{k} (A_j \sin(2\pi \nu_j t) + B_j \cos(2\pi \nu_j t))$, the spectral distribution function is $F(\nu) = \sum_{j=1}^{k} \sigma_j^2 F_j(\nu)$, where

$$F_j(\nu) = \begin{cases} 
0 & \text{if } \nu < -\nu_j, \\
\frac{1}{2} & \text{if } -\nu_j \leq \nu < \nu_j, \\
1 & \text{otherwise.}
\end{cases}$$
Wold’s decomposition

Notice that \( X_t = \sum_{j=1}^{k} (A_j \sin(2\pi \nu_j t) + B_j \cos(2\pi \nu_j t)) \) is deterministic (once we’ve seen the past, we can predict the future without error).

Wold showed that every stationary process can be represented as

\[
X_t = X_t^{(d)} + X_t^{(n)},
\]

where \( X_t^{(d)} \) is purely deterministic and \( X_t^{(n)} \) is purely nondeterministic. (c.f. the decomposition of a spectral distribution function as \( F^{(d)} + F^{(c)} \).)

Example: \( X_t = A \sin(2\pi \lambda t) + \frac{\theta(B)}{\phi(B)} W_t. \)
Introduction to Time Series Analysis. Lecture 16.

1. Review: Spectral density
2. Examples
4. Autocovariance generating function and spectral density.
Suppose $X_t$ is a linear process, so it can be written
\[ X_t = \sum_{i=0}^{\infty} \psi_i W_{t-i} = \psi(B)W_t. \]
Consider the autocovariance sequence,
\[ \gamma_h = \text{Cov}(X_t, X_{t+h}) \]
\[ = E \left[ \sum_{i=0}^{\infty} \psi_i W_{t-i} \sum_{j=0}^{\infty} \psi_j W_{t+h-j} \right] \]
\[ = \sigma_w^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+h}. \]
Define the autocovariance generating function as

\[ \gamma(B) = \sum_{h=-\infty}^{\infty} \gamma_h B^h. \]

Then,

\[ \gamma(B) = \sigma_w^2 \sum_{h=-\infty}^{\infty} \sum_{i=0}^{\infty} \psi_i \psi_{i+h} B^h \]

\[ = \sigma_w^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j B^{j-i} \]

\[ = \sigma_w^2 \sum_{i=0}^{\infty} \psi_i B^{-i} \sum_{j=0}^{\infty} \psi_j B^j = \sigma_w^2 \psi(B^{-1})\psi(B). \]
Autocovariance generating function and spectral density

Notice that

\[ \gamma(B) = \sum_{h=-\infty}^{\infty} \gamma_h B^h. \]

\[ f(\nu) = \sum_{h=-\infty}^{\infty} \gamma_h e^{-2\pi i\nu h} \]

\[ = \gamma(e^{-2\pi i\nu}) \]

\[ = \sigma_w^2 \psi(e^{-2\pi i\nu}) \psi(e^{2\pi i\nu}) \]

\[ = \sigma_w^2 \left| \psi(e^{2\pi i\nu}) \right|^2. \]
For example, for an MA(q), we have $\psi(B) = \theta(B)$, so

$$f(\nu) = \sigma_w^2 \theta \left( e^{-2\pi i\nu} \right) \theta \left( e^{2\pi i\nu} \right)$$

$$= \sigma_w^2 \left| \theta \left( e^{-2\pi i\nu} \right) \right|^2 .$$

For MA(1),

$$f(\nu) = \sigma_w^2 \left| 1 + \theta_1 e^{-2\pi i\nu} \right|^2$$

$$= \sigma_w^2 \left| 1 + \theta_1 \cos(-2\pi \nu) + i\theta_1 \sin(-2\pi \nu) \right|^2$$

$$= \sigma_w^2 \left( 1 + 2\theta_1 \cos(2\pi \nu) + \theta_1^2 \right) .$$
Autocovariance generating function and spectral density

For an AR(p), we have $\psi(B) = 1/\phi(B)$, so

$$f(\nu) = \frac{\sigma_w^2}{\phi(e^{-2\pi i \nu})^2 \phi(e^{2\pi i \nu})}$$

$$= \frac{\sigma_w^2}{|\phi(e^{-2\pi i \nu})|^2}.$$

For AR(1),

$$f(\nu) = \frac{\sigma_w^2}{|1 - \phi_1 e^{-2\pi i \nu}|^2}$$

$$= \frac{\sigma_w^2}{1 - 2\phi_1 \cos(2\pi \nu) + \phi_1^2}.$$
If $X_t$ is a linear process, it can be written $X_t = \sum_{i=0}^{\infty} \psi_i W_{t-i} = \psi(B)W_t$. Then

$$f(\nu) = \sigma_w^2 |\psi(e^{-2\pi i \nu})|^2.$$ 

That is, the spectral density $f(\nu)$ of a linear process measures the modulus of the $\psi(MA(\infty))$ polynomial at the point $e^{2\pi i \nu}$ on the unit circle.
Introduction to Time Series Analysis. Lecture 17.

1. Review: Spectral distribution function, spectral density.
3. Examples.
4. Time-invariant linear filters
5. Frequency response
If a time series \{X_t\} has autocovariance \gamma satisfying
\[ \sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty, \]
then we define its **spectral density** as

\[ f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h)e^{-2\pi i \nu h} \]

for \(-\infty < \nu < \infty\). We have

\[ \gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \nu h} f(\nu) \, d\nu = \int_{-1/2}^{1/2} e^{2\pi i \nu h} \, dF(\nu), \]

where \(dF(\nu) = f(\nu) \, d\nu\).

\(f\) measures how the variance of \(X_t\) is distributed across the spectrum.
If $X_t$ is a linear process, it can be written $X_t = \sum_{i=0}^{\infty} \psi_i W_{t-i} = \psi(B) W_t$. Then

$$f(\nu) = \sigma_w^2 |\psi(e^{-2\pi i \nu})|^2.$$ 

That is, the spectral density $f(\nu)$ of a linear process measures the modulus of the $\psi(MA(\infty))$ polynomial at the point $e^{2\pi i \nu}$ on the unit circle.
Spectral density of a linear process

For an ARMA(p,q), \( \psi(B) = \theta(B)/\phi(B) \), so

\[
\begin{align*}
   f(\nu) &= \sigma_w^2 \frac{\theta(e^{-2\pi i \nu})\theta(e^{2\pi i \nu})}{\phi(e^{-2\pi i \nu})\phi(e^{2\pi i \nu})} \\
   &= \sigma_w^2 \left| \frac{\theta(e^{-2\pi i \nu})}{\phi(e^{-2\pi i \nu})} \right|^2.
\end{align*}
\]

This is known as a \textit{rational spectrum}. 
Consider the factorization of $\theta$ and $\phi$ as

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

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

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

$$
f(\nu) = \sigma_w^2 \left| \frac{\theta_q \prod_{j=1}^{q} (e^{-2\pi i \nu} - z_j)}{\phi_p \prod_{j=1}^{p} (e^{-2\pi i \nu} - p_j)} \right|^2 \left| \frac{\theta_q \prod_{j=1}^{q} |e^{-2\pi i \nu} - z_j|^2}{\phi_p \prod_{j=1}^{p} |e^{-2\pi i \nu} - p_j|^2} \right|.
$$
Rational spectra

\[ f(\nu) = \sigma_w^2 \frac{\theta_q^2 \prod_{j=1}^{q} |e^{-2\pi i \nu} - z_j|^2}{\phi_p^2 \prod_{j=1}^{p} |e^{-2\pi i \nu} - p_j|^2}. \]

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

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 $\nu$ 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 $\nu = 0.5$.

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 $\nu$ moves towards $-1$. If $\theta_1 < 0$, the zero is to the right of 1, so the spectral density is at its minimum when $\nu = 0$. 
Consider \( X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t \). Example 4.6 in the text considers this model with \( \phi_1 = 1, \phi_2 = -0.9, \) and \( \sigma_w^2 = 1 \). In this case, the poles are at \( p_1, p_2 \approx 0.5555 \pm i0.8958 \approx 1.054e^{\pm i1.01567} \approx 1.054e^{\pm 2\pi i0.16165} \).

Thus, we have

\[
 f(\nu) = \frac{\sigma_w^2}{\phi_2^2 |e^{-2\pi i\nu} - p_1|^2 |e^{-2\pi i\nu} - p_2|^2},
\]

and this gets very peaked when \( e^{-2\pi i\nu} \) passes near \( 1.054e^{-2\pi i0.16165} \).
Example: AR(2)

Spectral density of AR(2): \( X_t = X_{t-1} - 0.9 X_{t-2} + W_t \)
Example: Seasonal ARMA

Consider $X_t = \Phi_1 X_{t-12} + W_t$.

$$\psi(B) = \frac{1}{1 - \Phi_1 B^{12}},$$

$$f(\nu) = \sigma_w^2 \frac{1}{(1 - \Phi_1 e^{-2\pi i 12\nu})(1 - \Phi_1 e^{2\pi i 12\nu})} \quad \text{for} \quad \nu \neq 0.$$

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

$$f(\nu) = \sigma_w^2 \frac{1}{1 - 2\Phi_1 \cos(24\pi \nu) + \Phi_1^2}.$$
Example: Seasonal ARMA

Spectral density of AR(1)_{12} \cdot X_t = +0.2X_{t-12} + W_t
Another view:

\[ 1 - \Phi_1 z^{12} = 0 \iff z = re^{i\theta}, \]

with \( r = |\Phi_1|^{-1/12}, \quad e^{i12\theta} = e^{-i \arg(\Phi_1)}. \)

For \( \Phi_1 > 0 \), the twelve poles are at \( |\Phi_1|^{-1/12} e^{ik \pi / 6} \) for \( k = 0, \pm 1, \ldots, \pm 5, 6 \).

So the spectral density gets peaked as \( e^{-2\pi i \nu} \) passes near \( |\Phi_1|^{-1/12} \times \{1, e^{-i \pi / 6}, e^{-i \pi / 3}, e^{-i \pi / 2}, e^{-i 2 \pi / 3}, e^{-i 5 \pi / 6}, -1\}. \)
Example: Multiplicative seasonal ARMA

Consider \((1 - \Phi_1 B^{12})(1 - \phi_1 B)X_t = W_t\).

\[
f(\nu) = \sigma_w^2 \frac{1}{(1 - 2\Phi_1 \cos(24\pi \nu) + \Phi_1^2)(1 - 2\phi_1 \cos(2\pi \nu) + \phi_1^2)}.
\]

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 \(e^{-2\pi i \nu}\) is at one of the 12th roots of 1; the AR(1) poles give a peak near \(e^{-2\pi i \nu} = 1\).
Example: Multiplicative seasonal ARMA

Spectral density of \( \frac{x_t}{(1+0.5 B)(1+0.2 B^1)} \) where \( x_t = w_t \).
Introduction to Time Series Analysis. Lecture 17.

1. Review: Spectral distribution function, spectral density.


3. Examples.

4. Time-invariant linear filters

5. Frequency response
**Time-invariant linear filters**

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

A time series \( \{Y_t\} \) is the output of a linear filter \( A = \{a_{t,j} : t, j \in \mathbb{Z}\} \) with input \( \{X_t\} \) if

\[
Y_t = \sum_{j=-\infty}^{\infty} a_{t,j} X_j.
\]

If \( a_{t,t-j} \) is independent of \( t \) (\( a_{t,t-j} = \psi_j \)), then we say that the filter is *time-invariant*.

If \( \psi_j = 0 \) for \( j < 0 \), we say the filter \( \psi \) is *causal*.

We’ll see that the name ‘filter’ arises from the frequency domain viewpoint.
Time-invariant linear filters: 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:
   \[
   \psi_j = \begin{cases}
   \frac{1}{3} & \text{if } |j| \leq 1, \\
   0 & \text{otherwise}.
   \end{cases}
   \]

3. For polynomials \( \phi(B), \theta(B) \) with roots outside the unit circle, 
   \( \psi(B) = \theta(B)/\phi(B) \) is a linear, time-invariant, causal filter.
The operation

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

is called the \textit{convolution} of \( X \) with \( \psi \).
The sequence $\psi$ is also called the *impulse response*, since the output $\{Y_t\}$ of the linear filter in response to a *unit impulse*,

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

is

$$Y_t = \psi(B)X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} = \psi_t.$$
Introduction to Time Series Analysis. Lecture 17.

1. Review: Spectral distribution function, spectral density.
3. Examples.
4. Time-invariant linear filters
5. Frequency response
Suppose that \(\{X_t\}\) has spectral density \(f_x(\nu)\) and \(\psi\) is \textit{stable}, that is, 
\[
\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.
\]
Then \(Y_t = \psi(B)X_t\) has spectral density 
\[
f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 f_x(\nu).
\]
The function \(\nu \mapsto \psi(e^{2\pi i \nu})\) (the polynomial \(\psi(z)\) evaluated on the unit circle) is known as the \textit{frequency response} or \textit{transfer function} of the linear filter.

The squared modulus, \(\nu \mapsto |\psi(e^{2\pi i \nu})|^2\) is known as the \textit{power transfer function} of the filter.
**Frequency response of a time-invariant linear filter**

For stable $\psi$, $Y_t = \psi(B)X_t$ has spectral density

$$f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 f_x(\nu).$$

We have seen that a linear process, $Y_t = \psi(B)W_t$, is a special case, since

$$f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 \sigma_w^2 = |\psi(e^{2\pi i \nu})|^2 f_w(\nu).$$

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 $\nu \mapsto |\psi(e^{2\pi i \nu})|^2$.

This is a version of the equality $\text{Var}(aX) = a^2\text{Var}(X)$, but the equality is true for the component of the variance at every frequency.

This is also the origin of the name ‘filter.’
Frequency response of a filter: Details

Why is \( f_y(\nu) = |\psi(e^{2\pi i\nu})|^2 f_x(\nu) \)? First,

\[
\gamma_y(h) = E \left[ \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} \sum_{k=-\infty}^{\infty} \psi_k X_{t+h-k} \right]
\]

\[
= \sum_{j=-\infty}^{\infty} \psi_j \sum_{k=-\infty}^{\infty} \psi_k E[X_{t+h-k}X_{t-j}]
\]

\[
= \sum_{j=-\infty}^{\infty} \psi_j \sum_{k=-\infty}^{\infty} \psi_k \gamma_x(h+j-k) = \sum_{j=-\infty}^{\infty} \psi_j \sum_{l=-\infty}^{\infty} \psi_{h+j-l} \gamma_x(l).
\]

It is easy to check that \( \sum_{j=-\infty}^{\infty} \psi_j |< \infty \) and \( \sum_{h=-\infty}^{\infty} |\gamma_x(h)| < \infty \) imply that \( \sum_{h=-\infty}^{\infty} |\gamma_y(h)| < \infty \). Thus, the spectral density of \( y \) is defined.
Frequency response of a filter: Details

\[ f_y(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h)e^{-2\pi i \nu h} \]

\[ = \sum_{h=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_j \sum_{l=-\infty}^{\infty} \psi_{h+j-l} \gamma_x(l)e^{-2\pi i \nu h} \]

\[ = \sum_{j=-\infty}^{\infty} \psi_j e^{2\pi i \nu j} \sum_{l=-\infty}^{\infty} \gamma_x(l)e^{-2\pi i \nu l} \sum_{h=-\infty}^{\infty} \psi_{h+j-l} e^{-2\pi i \nu (h+j-l)} \]

\[ = \psi(e^{2\pi i \nu j}) f_x(\nu) \sum_{h=-\infty}^{\infty} \psi_h e^{-2\pi i \nu h} \]

\[ = |\psi(e^{2\pi i \nu j})|^2 f_x(\nu). \]
For a linear process $Y_t = \psi(B)W_t$, $f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 \sigma_w^2$.

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

$$f_y(\nu) = \sigma_w^2 \left| \frac{\theta(e^{-2\pi i \nu})}{\phi(e^{-2\pi i \nu})} \right|^2 = \sigma_w^2 \frac{\theta^2 \prod_{j=1}^{q} |e^{-2\pi i \nu} - z_j|^2}{\phi^2 \prod_{j=1}^{p} |e^{-2\pi i \nu} - p_j|^2},$$

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

**Frequency response: Examples**
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
\[ \psi(e^{-2\pi i \nu}) = D_k(2\pi \nu) = \frac{1}{2k + 1} \sum_{j=-k}^{k} e^{-2\pi i j \nu} \]
\[ = \begin{cases} 
1 & \text{if } \nu = 0, \\
\frac{\sin(2\pi (k+1/2) \nu)}{(2k+1) \sin(\pi \nu)} & \text{otherwise}.
\end{cases} \]
Example: Moving average

Transfer function of moving average (k=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

\[ \psi(e^{-2\pi i \nu}) = 1 - e^{-2\pi i \nu}, \]

so

\[ |\psi(e^{-2\pi i \nu})|^2 = 2(1 - \cos(2\pi \nu)). \]
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.
Introduction to Time Series Analysis. Lecture 18.

2. Frequency response of linear filters.
3. Spectral estimation
4. Sample autocovariance
5. Discrete Fourier transform and the periodogram
**Review: Spectral density**

If a time series \( \{X_t\} \) has autocovariance \( \gamma \) satisfying
\[
\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty,
\]
then we define its **spectral density** as
\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h}
\]
for \(-\infty < \nu < \infty\). We have
\[
\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \nu h} f(\nu) \, d\nu.
\]
Review: Rational spectra

For a linear time series with $MA(\infty)$ polynomial $\psi$,

$$f(\nu) = \sigma_w^2 |\psi(e^{2\pi i \nu})|^2.$$

If it is an ARMA(p,q), we have

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

$$= \sigma_w^2 \frac{\theta_q^2 \prod_{j=1}^q |e^{-2\pi i \nu} - z_j|^2}{\phi_p^2 \prod_{j=1}^p |e^{-2\pi i \nu} - p_j|^2},$$

where $z_1, \ldots, z_q$ are the zeros (roots of $\theta(z)$) and $p_1, \ldots, p_p$ are the poles (roots of $\phi(z)$).
Review: Time-invariant linear filters

A filter is an operator; given a time series \( \{ X_t \} \), it maps to a time series \( \{ Y_t \} \). A linear filter satisfies

\[
Y_t = \sum_{j=-\infty}^{\infty} a_{t,j} X_j.
\]

time-invariant: \( a_{t,t-j} = \psi_j \):

\[
Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}.
\]

causal: \( j < 0 \) implies \( \psi_j = 0 \).

\[
Y_t = \sum_{j=0}^{\infty} \psi_j X_{t-j}.
\]
The operation

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

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

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

is

$$Y_t = \psi(B)X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} = \psi_t.$$
Introduction to Time Series Analysis. Lecture 18.

2. Frequency response of linear filters.
3. Spectral estimation
4. Sample autocovariance
5. Discrete Fourier transform and the periodogram
Suppose that \( \{ X_t \} \) has spectral density \( f_x(\nu) \) and \( \psi \) is stable, that is, \( \sum_{j=-\infty}^{\infty} |\psi_j| < \infty \). Then \( Y_t = \psi(B)X_t \) has spectral density
\[
f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 f_x(\nu).
\]
The function \( \nu \mapsto \psi(e^{2\pi i \nu}) \) (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, \( \nu \mapsto |\psi(e^{2\pi i \nu})|^2 \) is known as the power transfer function of the filter.
For stable $\psi$, $Y_t = \psi(B)X_t$ has spectral density

$$f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 f_x(\nu).$$

We have seen that a linear process, $Y_t = \psi(B)W_t$, is a special case, since $f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 \sigma_w^2 = |\psi(e^{2\pi i \nu})|^2 f_w(\nu)$.

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 $\nu \mapsto |\psi(e^{2\pi i \nu})|^2$.

This is a version of the equality $\text{Var}(aX) = a^2 \text{Var}(X)$, but the equality is true for the component of the variance at every frequency.

This is also the origin of the name ‘filter.’
Why is $f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 f_x(\nu)$? First,

$$
\gamma_y(h) = E \left[ \sum_{j=\infty}^{\infty} \psi_j X_{t-j} \sum_{k=\infty}^{\infty} \psi_k X_{t+h-k} \right]
= \sum_{j=\infty}^{\infty} \psi_j \sum_{k=\infty}^{\infty} \psi_k E[X_{t+h-k} X_{t-j}]
= \sum_{j=\infty}^{\infty} \psi_j \sum_{k=\infty}^{\infty} \psi_k \gamma_x(h + j - k) = \sum_{j=\infty}^{\infty} \psi_j \sum_{l=\infty}^{\infty} \psi_{h+j-l} \gamma_x(l).
$$

It is easy to check that $\sum_{j=\infty}^{\infty} |\psi_j| < \infty$ and $\sum_{h=\infty}^{\infty} |\gamma_x(h)| < \infty$ imply that $\sum_{h=\infty}^{\infty} |\gamma_y(h)| < \infty$. Thus, the spectral density of $y$ is defined.
Frequency response of a filter: Details

\[ f_y(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h} \]

\[ = \sum_{h=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_j \sum_{l=-\infty}^{\infty} \psi_{h+j-l} \gamma_x(l) e^{-2\pi i \nu h} \]

\[ = \sum_{j=-\infty}^{\infty} \psi_j e^{2\pi i \nu j} \sum_{l=-\infty}^{\infty} \gamma_x(l) e^{-2\pi i \nu l} \sum_{h=-\infty}^{\infty} \psi_{h+j-l} e^{-2\pi i \nu(h+j-l)} \]

\[ = \psi(e^{2\pi i \nu j}) f_x(\nu) \sum_{h=-\infty}^{\infty} \psi_h e^{-2\pi i \nu h} \]

\[ = |\psi(e^{2\pi i \nu j})|^2 f_x(\nu). \]
**Frequency response: Examples**

For a linear process $Y_t = \psi(B)W_t$, $f_y(\nu) = |\psi(e^{2\pi i \nu})|^2 \sigma_w^2$.

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

$$f_y(\nu) = \sigma_w^2 \left| \frac{\theta(e^{-2\pi i \nu})}{\phi(e^{-2\pi i \nu})} \right|^2$$

$$= \sigma_w^2 \theta_q^2 \prod_{j=1}^{q} |e^{-2\pi i \nu} - z_j|^2 \frac{\phi_p^2}{\prod_{j=1}^{p} |e^{-2\pi i \nu} - p_j|^2},$$

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

\[ \psi(e^{-2\pi i \nu}) = D_k(2\pi \nu) = \frac{1}{2k + 1} \sum_{j=-k}^{k} e^{-2\pi i j \nu} \]

\[ = \begin{cases} 1 & \text{if } \nu = 0, \\ \frac{\sin(2\pi (k+1/2)\nu)}{(2k+1)\sin(\pi \nu)} & \text{otherwise}. \end{cases} \]
Example: Moving average

Transfer function of moving average (k=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

\[ \psi(e^{-2\pi i \nu}) = 1 - e^{-2\pi i \nu}, \]

so

\[ |\psi(e^{-2\pi i \nu})|^2 = 2(1 - \cos(2\pi \nu)). \]
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.
Introduction to Time Series Analysis. Lecture 18.

2. Frequency response of linear filters.
3. Spectral estimation
4. Sample autocovariance
5. Discrete Fourier transform and the periodogram
Estimating the Spectrum: Outline

- We have seen that the spectral density gives an alternative view of stationary time series.
- Given a realization \( x_1, \ldots, x_n \) of a time series, how can we estimate the spectral density?
- One approach: replace \( \gamma(\cdot) \) in the definition

\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h)e^{-2\pi i \nu h},
\]

with the sample autocovariance \( \hat{\gamma}(\cdot) \).
- Another approach, called the periodogram: compute \( I(\nu) \), the squared modulus of the discrete Fourier transform (at frequencies \( \nu = k/n \)).
Estimating the spectrum: Outline

- These two approaches are identical at the Fourier frequencies $\nu = k/n$.
- The asymptotic expectation of the periodogram $I(\nu)$ is $f(\nu)$. We can derive some asymptotic properties, and hence do hypothesis testing.
- Unfortunately, the asymptotic variance of $I(\nu)$ is constant. It is not a consistent estimator of $f(\nu)$.
- We can reduce the variance by smoothing the periodogram—averaging over adjacent frequencies. If we average over a narrower range as $n \rightarrow \infty$, we can obtain a consistent estimator of the spectral density.
Estimating the spectrum: Sample autocovariance

Idea: use the sample autocovariance \( \hat{\gamma}(\cdot) \), defined by

\[
\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad \text{for } -n < h < n,
\]

as an estimate of the autocovariance \( \gamma(\cdot) \), and then use a sample version of

\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h)e^{-2\pi i \nu h},
\]

That is, for \(-1/2 \leq \nu \leq 1/2\), estimate \( f(\nu) \) with

\[
\hat{f}(\nu) = \sum_{h=-n+1}^{n-1} \hat{\gamma}(h)e^{-2\pi i \nu h}.
\]
Another approach to estimating the spectrum is called the periodogram. It was proposed in 1897 by Arthur Schuster (at Owens College, which later became part of the University of Manchester), who used it to investigate periodicity in the occurrence of earthquakes, and in sunspot activity.


To define the periodogram, we need to introduce the *discrete Fourier transform* of a finite sequence $x_1, \ldots, x_n$. 

**Estimating the spectrum: Periodogram**
Introduction to Time Series Analysis. Lecture 18.

2. Frequency response of linear filters.
3. Spectral estimation
4. Sample autocovariance
5. Discrete Fourier transform and the periodogram
For a sequence \((x_1, \ldots, x_n)\), define the \textit{discrete Fourier transform (DFT)} as \((X(\nu_0), X(\nu_1), \ldots, X(\nu_{n-1}))\), where

\[
X(\nu_k) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t e^{-2\pi i \nu_k t},
\]

and \(\nu_k = k/n\) (for \(k = 0, 1, \ldots, n - 1\)) are called the \textit{Fourier frequencies}. (Think of \(\{\nu_k : k = 0, \ldots, n - 1\}\) as the discrete version of the frequency range \(\nu \in [0, 1]\).)

First, let’s show that we can view the DFT as a representation of \(x\) in a different basis, the \textit{Fourier basis}. 

\begin{center}
\textbf{Discrete Fourier transform}
\end{center}
Consider the space $\mathbb{C}^n$ of vectors of $n$ complex numbers, with inner product $\langle a, b \rangle = a^* b$, where $a^*$ is the complex conjugate transpose of the vector $a \in \mathbb{C}^n$.

Suppose that a set $\{\phi_j : j = 0, 1, \ldots, n - 1\}$ of $n$ vectors in $\mathbb{C}^n$ are orthonormal:

$$\langle \phi_j, \phi_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

Then these $\{\phi_j\}$ span the vector space $\mathbb{C}^n$, and so for any vector $x$, we can write $x$ in terms of this new orthonormal basis,

$$x = \sum_{j=0}^{n-1} \langle \phi_j, x \rangle \phi_j.$$

Consider the following set of $n$ vectors in $\mathbb{C}^n$:

$$\left\{ e_j = \frac{1}{\sqrt{n}} (e^{2\pi i \nu_j}, e^{2\pi i 2\nu_j}, \ldots, e^{2\pi i n\nu_j})^t : j = 0, \ldots, n - 1 \right\} .$$

It is easy to check that these vectors are orthonormal:

$$\langle e_j, e_k \rangle = \frac{1}{n} \sum_{t=1}^{n} e^{2\pi it(\nu_k - \nu_j)} = \frac{1}{n} \sum_{t=1}^{n} \left( e^{2\pi i (k-j)/n} \right)^t$$

$$= \begin{cases} 1 & \text{if } j = k, \\ \frac{1}{n} e^{2\pi i (k-j)/n} \frac{1 - (e^{2\pi i (k-j)/n})^n}{1 - e^{2\pi i (k-j)/n}} & \text{otherwise} \end{cases}$$
where we have used the fact that $S_n = \sum_{t=1}^{n} \alpha^t$ satisfies
\[ \alpha S_n = S_n + \alpha^{n+1} - \alpha \]
and so $S_n = \alpha(1 - \alpha^n)/(1 - \alpha)$ for $\alpha \neq 1$.

So we can represent the real vector $x = (x_1, \ldots, x_n)' \in \mathbb{C}^n$ in terms of this orthonormal basis,
\[
x = \sum_{j=0}^{n-1} \langle e_j, x \rangle e_j = \sum_{j=0}^{n-1} X(\nu_j) e_j.
\]

That is, the vector of discrete Fourier transform coefficients
$(X(\nu_0), \ldots, X(\nu_{n-1}))$ is the representation of $x$ in the Fourier basis.
An alternative way to represent the DFT is by separately considering the real and imaginary parts,

$$X(\nu_j) = \langle e_j, x \rangle = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e^{-2\pi it\nu_j} x_t$$

$$= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \cos(2\pi t\nu_j) x_t - i \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sin(2\pi t\nu_j) x_t$$

$$= X_c(\nu_j) - iX_s(\nu_j),$$

where this defines the sine and cosine transforms, $X_s$ and $X_c$, of $x$. 
Introduction to Time Series Analysis. Lecture 19.

2. The periodogram and sample autocovariance.
3. Asymptotics of the periodogram.
Estimating the Spectrum: Outline

• We have seen that the spectral density gives an alternative view of stationary time series.

• Given a realization \( x_1, \ldots, x_n \) of a time series, how can we estimate the spectral density?

• One approach: replace \( \gamma(\cdot) \) in the definition

\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h},
\]

with the sample autocovariance \( \hat{\gamma}(\cdot) \).

• Another approach, called the periodogram: compute \( I(\nu) \), the squared modulus of the discrete Fourier transform (at frequencies \( \nu = k/n \)).
Estimating the spectrum: Outline

- These two approaches are identical at the Fourier frequencies \( \nu = k/n \).
- The asymptotic expectation of the periodogram \( I(\nu) \) is \( f(\nu) \). We can derive some asymptotic properties, and hence do hypothesis testing.
- Unfortunately, the asymptotic variance of \( I(\nu) \) is constant. It is not a consistent estimator of \( f(\nu) \).
Review: Spectral density estimation

If a time series \( \{X_t\} \) has autocovariance \( \gamma \) satisfying
\[
\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty,
\]
then we define its spectral density as
\[
f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h)e^{-2\pi i \nu h}
\]
for \(-\infty < \nu < \infty\).
Review: Sample autocovariance

Idea: use the sample autocovariance $\hat{\gamma}(\cdot)$, defined by

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad \text{for } -n < h < n,$$

as an estimate of the autocovariance $\gamma(\cdot)$, and then use

$$\hat{f}(\nu) = \sum_{h=-n+1}^{n-1} \hat{\gamma}(h)e^{-2\pi i \nu h}$$

for $-1/2 \leq \nu \leq 1/2$. 
For a sequence \((x_1, \ldots, x_n)\), define the **discrete Fourier transform (DFT)** as 
\((X(\nu_0), X(\nu_1), \ldots, X(\nu_{n-1}))\), where

\[
X(\nu_k) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t e^{-2\pi i \nu_k t},
\]

and \(\nu_k = k/n\) (for \(k = 0, 1, \ldots, n - 1\)) are called the **Fourier frequencies**.

(Think of \(\{\nu_k : k = 0, \ldots, n - 1\}\) as the discrete version of the frequency range \(\nu \in [0, 1]\).)

First, let’s show that we can view the DFT as a representation of \(x\) in a different basis, the **Fourier basis**.
Consider the space \( \mathbb{C}^n \) of vectors of \( n \) complex numbers, with inner product \( \langle a, b \rangle = a^* b \), where \( a^* \) is the complex conjugate transpose of the vector \( a \in \mathbb{C}^n \).

Suppose that a set \( \{ \phi_j : j = 0, 1, \ldots, n - 1 \} \) of \( n \) vectors in \( \mathbb{C}^n \) are orthonormal:

\[
\langle \phi_j, \phi_k \rangle = \begin{cases} 
1 & \text{if } j = k, \\
0 & \text{otherwise.}
\end{cases}
\]

Then these \( \{ \phi_j \} \) span the vector space \( \mathbb{C}^n \), and so for any vector \( x \), we can write \( x \) in terms of this new orthonormal basis,

\[
x = \sum_{j=0}^{n-1} \langle \phi_j, x \rangle \phi_j.
\]
Consider the following set of $n$ vectors in $\mathbb{C}^n$:

$$\left\{ e_j = \frac{1}{\sqrt{n}} \left( e^{2\pi i \nu_j}, e^{2\pi i 2\nu_j}, \ldots, e^{2\pi i n\nu_j} \right)' : j = 0, \ldots, n - 1 \right\}. $$

It is easy to check that these vectors are orthonormal:

$$\langle e_j, e_k \rangle = \frac{1}{n} \sum_{t=1}^{n} e^{2\pi it(\nu_k - \nu_j)} = \frac{1}{n} \sum_{t=1}^{n} \left( e^{2\pi i(k-j)/n} \right)^t$$

$$= \begin{cases} 
1 & \text{if } j = k, \\
\frac{1}{n} e^{2\pi i(k-j)/n} \frac{1-(e^{2\pi i(k-j)/n})^n}{1-e^{2\pi i(k-j)/n}} & \text{otherwise}
\end{cases}$$

$$= \begin{cases} 
1 & \text{if } j = k, \\
0 & \text{otherwise,}
\end{cases}$$
where we have used the fact that $S_n = \sum_{t=1}^{n} \alpha^t$ satisfies
$\alpha S_n = S_n + \alpha^{n+1} - \alpha$ and so $S_n = \alpha(1 - \alpha^n)/(1 - \alpha)$ for $\alpha \neq 1$.

So we can represent the real vector $x = (x_1, \ldots, x_n)' \in \mathbb{C}^n$ in terms of this orthonormal basis,

$$x = \sum_{j=0}^{n-1} \langle e_j, x \rangle e_j = \sum_{j=0}^{n-1} X(\nu_j)e_j.$$ 

That is, the vector of discrete Fourier transform coefficients
$(X(\nu_0), \ldots, X(\nu_{n-1}))$ is the representation of $x$ in the Fourier basis.
An alternative way to represent the DFT is by separately considering the real and imaginary parts,

\[ X(\nu_j) = \langle e_j, x \rangle = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e^{-2\pi i t \nu_j} x_t \]

\[ = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \cos(2\pi t \nu_j) x_t - i \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sin(2\pi t \nu_j) x_t \]

\[ = X_c(\nu_j) - iX_s(\nu_j), \]

where this defines the sine and cosine transforms, \( X_s \) and \( X_c \), of \( x \).
The periodogram is defined as

\[
I(\nu_j) = |X(\nu_j)|^2
\]

\[
= \frac{1}{n} \left| \sum_{t=1}^{n} e^{-2\pi it\nu_j} x_t \right|^2 
= X_c^2(\nu_j) + X_s^2(\nu_j).
\]

\[
X_c(\nu_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \cos(2\pi t\nu_j) x_t, 
\]

\[
X_s(\nu_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sin(2\pi t\nu_j) x_t.
\]
Periodogram

Since $I(\nu_j) = |X(\nu_j)|^2$ for one of the Fourier frequencies $\nu_j = j/n$ (for $j = 0, 1, \ldots, n - 1$), the orthonormality of the $e_j$ implies that we can write

$$x^* x = \left( \sum_{j=0}^{n-1} X(\nu_j) e_j \right)^* \left( \sum_{j=0}^{n-1} X(\nu_j) e_j \right)$$

$$= \sum_{j=0}^{n-1} |X(\nu_j)|^2 = \sum_{j=0}^{n-1} I(\nu_j).$$

For $\bar{x} = 0$, we can write this as

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_{t=1}^{n} x_t^2 = \frac{1}{n} \sum_{j=0}^{n-1} I(\nu_j).$$
This is the discrete analog of the identity

\[ \sigma_x^2 = \gamma_x(0) = \int_{-1/2}^{1/2} f_x(\nu) d\nu. \]

(Think of \( I(\nu_j) \) as the discrete version of \( f(\nu) \) at the frequency \( \nu_j = j/n \), and think of \( (1/n) \sum_{\nu_j} \) as the discrete version of \( \int_{\nu} \cdot d\nu \).)
Estimating the spectrum: Periodogram

Why is the periodogram at a Fourier frequency (that is, $\nu = \nu_j$) the same as computing $f(\nu)$ from the sample autocovariance?

Almost the same—they are not the same at $\nu_0 = 0$ when $\bar{x} \neq 0$.

But if either $\bar{x} = 0$, or we consider a Fourier frequency $\nu_j$ with $j \in \{1, \ldots, n-1\}$, \ldots
Estimating the spectrum: Periodogram

\[ I(\nu_j) = \frac{1}{n} \left| \sum_{t=1}^{n} e^{-2\pi it\nu_j} x_t \right|^2 = \frac{1}{n} \left| \sum_{t=1}^{n} e^{-2\pi it\nu_j} (x_t - \bar{x}) \right|^2 \]

\[ = \frac{1}{n} \left( \sum_{t=1}^{n} e^{-2\pi it\nu_j} (x_t - \bar{x}) \right) \left( \sum_{t=1}^{n} e^{2\pi it\nu_j} (x_t - \bar{x}) \right) \]

\[ = \frac{1}{n} \sum_{s,t} e^{-2\pi i(s-t)\nu_j} (x_s - \bar{x})(x_t - \bar{x}) = \sum_{h=-n+1}^{n-1} \hat{\gamma}(h) e^{-2\pi ih\nu_j}, \]

where the fact that \( \nu_j \neq 0 \) implies \( \sum_{t=1}^{n} e^{-2\pi it\nu_j} = 0 \) (we showed this when we were verifying the orthonormality of the Fourier basis) has allowed us to subtract the sample mean in that case.
Asymptotic properties of the periodogram

We want to understand the asymptotic behavior of the periodogram $I(\nu)$ at a particular frequency $\nu$, as $n$ increases. We’ll see that its expectation converges to $f(\nu)$.

We’ll start with a simple example: Suppose that $X_1, \ldots, X_n$ are i.i.d. $N(0, \sigma^2)$ (Gaussian white noise). From the definitions,

$$X_c(\nu_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \cos(2\pi t \nu_j) x_t, \quad X_s(\nu_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sin(2\pi t \nu_j) x_t,$$

we have that $X_c(\nu_j)$ and $X_s(\nu_j)$ are normal, with

$$\mathbb{E}X_c(\nu_j) = \mathbb{E}X_s(\nu_j) = 0.$$
Asymptotic properties of the periodogram

Also,

\[
\text{Var}(X_c(\nu_j)) = \frac{\sigma^2}{n} \sum_{t=1}^{n} \cos^2(2\pi t \nu_j)
\]

\[
= \frac{\sigma^2}{2n} \sum_{t=1}^{n} (\cos(4\pi t \nu_j) + 1) = \frac{\sigma^2}{2}.
\]

Similarly, \(\text{Var}(X_s(\nu_j)) = \sigma^2 / 2\).
Asymptotic properties of the periodogram

Also,

\[
\begin{align*}
\text{Cov}(X_c(\nu_j), X_s(\nu_j)) &= \frac{\sigma^2}{n} \sum_{t=1}^{n} \cos(2\pi t \nu_j) \sin(2\pi t \nu_j) \\
&= \frac{\sigma^2}{2n} \sum_{t=1}^{n} \sin(4\pi t \nu_j) = 0,
\end{align*}
\]

\[
\begin{align*}
\text{Cov}(X_c(\nu_j), X_c(\nu_k)) &= 0 \\
\text{Cov}(X_s(\nu_j), X_s(\nu_k)) &= 0 \\
\text{Cov}(X_s(\nu_j), X_s(\nu_k)) &= 0 \\
\text{Cov}(X_c(\nu_j), X_s(\nu_k)) &= 0.
\end{align*}
\]

for any \( j \neq k \).
Asymptotic properties of the periodogram

That is, if \( X_1, \ldots, X_n \) are i.i.d. \( N(0, \sigma^2) \) (Gaussian white noise; \( f(\nu) = \sigma^2 \)), then the \( X_c(\nu_j) \) and \( X_s(\nu_j) \) are all i.i.d. \( N(0, \sigma^2/2) \). Thus,

\[
\frac{2}{\sigma^2} I(\nu_j) = \frac{2}{\sigma^2} \left( X_c^2(\nu_j) + X_s^2(\nu_j) \right) \sim \chi^2_2.
\]

So for the case of Gaussian white noise, the periodogram has a chi-squared distribution that depends on the variance \( \sigma^2 \) (which, in this case, is the spectral density).
Asymptotic properties of the periodogram

Under more general conditions (e.g., normal \( \{X_t\} \), or linear process \( \{X_t\} \) with rapidly decaying ACF), the \( X_c(\nu_j) \), \( X_s(\nu_j) \) are all asymptotically independent and \( N(0, f(\nu_j)/2) \).

Consider a frequency \( \nu \). For a given value of \( n \), let \( \hat{\nu}^{(n)} \) be the closest Fourier frequency (that is, \( \hat{\nu}^{(n)} = j/n \) for a value of \( j \) that minimizes \(|\nu - j/n|\)). As \( n \) increases, \( \hat{\nu}^{(n)} \to \nu \), and (under the same conditions that ensure the asymptotic normality and independence of the sine/cosine transforms), \( f(\hat{\nu}^{(n)}) \to f(\nu) \).

In that case, we have

\[
\frac{2}{f(\nu)} I(\hat{\nu}^{(n)}) = \frac{2}{f(\nu)} \left( X_c^2(\hat{\nu}^{(n)}) + X_s^2(\hat{\nu}^{(n)}) \right) \xrightarrow{d} \chi^2_2.
\]
Asymptotic properties of the periodogram

Thus,

\[
E[I(\hat{\nu}(n))] = \frac{f(\nu)}{2} E \left( \frac{2}{f(\nu)} \left( X_c^2(\hat{\nu}(n)) + X_s^2(\hat{\nu}(n)) \right) \right)
\]

\[
\rightarrow \frac{f(\nu)}{2} E(Z_1^2 + Z_2^2) = f(\nu),
\]

where \( Z_1, Z_2 \) are independent \( N(0, 1) \). Thus, the periodogram is asymptotically unbiased.