Lecture Notes 5. The Frequency Domain Approach to Time Series

These lecture notes develop a different perspective on time series from the Hilbert space approach we have studied so far. This new approach will focus on cycles in time series data.

i. spectral densities

The spectral density of a time series is defined as

\[ f_x(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_x(j) e^{-ij\omega} \quad \omega \in [-\pi, \pi] \]  

(5.1)

Since \( \cos(\omega) = \cos(-\omega) \), \( \sin(\omega) = -\sin(-\omega) \) and \( \sigma_x(j) = \sigma_x(-j) \), the spectral density can be rewritten as

\[ \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_x(j) e^{-ij\omega} = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_x(j) \cos j\omega + \frac{1}{\pi} \sum_{j=1}^{\infty} \sigma_x(j) \cos j\omega \]  

(5.2)

The spectral density is the Fourier transform of the autocovariance function. By the Riesz-Fischer theorem, this means that the spectral density function and the autocovariance function contain the same information about the stochastic process \( x_t \). One can recover the autocovariances via the formula

\[ \int_{-\pi}^{\pi} f_x(\omega) e^{ij\omega} d\omega = \sigma_x(j) \]  

(5.3)
Notice that a special case of the recovery formula is

$$\int_{-\pi}^{\pi} f_x(\omega) d\omega = \sigma_x(0). \quad (5.4)$$

In words, the integral of the spectral density equals the variance of the process. This will prove to have a deep interpretation.

We now turn to some examples of spectral density functions.

**Example 5.1. white noise process**

If $x_t = \varepsilon_t$, then the spectral density equals

$$\frac{\sigma^2}{2\pi} \quad (5.5)$$

The function is shaped as a rectangle for the interval $[-\pi, \pi]$. Each frequency produces the same value, which is the origin of the term white noise.

**Example 5.2. AR(1) Process**

If $x_t = \rho x_{t-1} + \varepsilon_t$, then the spectral density equals

$$\frac{\sigma^2}{2\pi (1 - \rho e^{-i\omega})(1 - \rho e^{i\omega})} = \frac{\sigma^2}{2\pi (1 - 2\rho c \cos\omega + \rho^2)} \quad (5.6)$$

When $\rho > 0$, then the maximum of the function is $\omega = 0$. If $\rho < 0$, then $\omega = 0$ minimum. Notice that as $\rho \to 1$, the spectral density function becomes arbitrarily large at $\omega = 0$. 

2
Example 5.3. MA(1) Process

If $x_t = \varepsilon_t + \rho \varepsilon_{t-1}$, then the spectral density equals

$$\frac{\sigma^2}{2\pi} \left\{ 1 + \rho e^{-i\omega} \right\} \left\{ 1 + \rho e^{i\omega} \right\} = \frac{\sigma^2}{2\pi} \left( 1 + 2\rho \cos \omega + \rho^2 \right)$$

(5.7)

If $\rho > 0$, then the function is maximal at $\omega = 0$ whereas if $\rho < 0$ the function is minimal. Notice that the qualitative shape of the spectral density of an MA(1) is the same as for the AR(1).

Example 5.4. AR(2) Process

If $x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \varepsilon_t$, then the spectral density equals

$$\frac{\sigma^2}{2\pi} \left\{ 1 - \rho_1 e^{-i\omega} - \rho_2 e^{-2i\omega} \right\} \left\{ 1 - \rho_1 e^{i\omega} - \rho_2 e^{2i\omega} \right\} = \frac{\sigma^2}{2\pi} \left( 1 + \rho_1^2 + \rho_2^2 + 2(\rho_1 \rho_2 - \rho_1) \cos \omega - 2\rho_2 \cos 2\omega \right)$$

(5.8)

the maximum value of this function is difficult to compute, except when $\rho_1$ and $\rho_2$ are both positive, in which case the spectral density must have a maximum at $\omega = 0$.

ii. spectral representation of a time series

This section describes a decomposition of a time series into frequency-specific components. The decomposition will illustrate a deep relationship between the linear dependence structure of a time series and its spectral density.
Theorem 5.1  Spectral representation theorem (Cramér’s theorem)

Let \( x_t \) be a zero mean \( L^2 \) process. The \( x_t \) may be expressed as

\[
x_t = \int_{-\pi}^{\pi} e^{it\omega} dz_x(\omega)
\]  \hspace{1cm} (5.9)

where \( dz_x(\omega) \) is a complex valued random process such that

\[
i. \ E(dz_x(\omega)) = 0
\]

\[
ii. \ E(dz_x(\omega) \overline{dz_x(\omega)}) = dF_x(\omega)
\]  \hspace{1cm} (5.10)

\[
iii. \ E(dz_x(\omega) \overline{dz_x(\omega_i)}) = 0, \ \omega_i \neq \omega
\]

\[
iv. \ z_x(\omega) \text{ is unique, outside a set of measure zero.}
\]

The domain of integration is \((-\pi, \pi]\). (The use of the half-open interval is a minor technicality which I will ignore.)

What is \( z_x(\omega) \)? This is an example of a random function\(^1\). This means that for each fixed frequency \( \omega \), \( z_x(\omega) \) is a random variable. A random function is thus a collection of random variables indexed by \( \omega \). Notice that this collection is uncountably infinite. Suppose that for any fixed frequencies \( \omega_1 < \omega_2 \leq \omega_3 < \omega_4 \)

\[
E\left(\left(z_x(\omega_4) - z_x(\omega_3)\right)\overline{\left(z_x(\omega_2) - z_x(\omega_1)\right)}\right) = 0
\]  \hspace{1cm} (5.11)

\(^1\) A standard example of a random function is Brownian motion. A Brownian motion \( B(t) \) \( t \geq 0 \) is random function such that \( B(\bar{T}) \sim N(0, \bar{T}) \).
then the process is said to possess orthogonal increments.

What does this theorem tell us about the underlying structure of a time series? Recall (5.4), which stated that the variance of a process is the integral of its spectral density. Using (5.9), the variance can also be written as

$$
\sigma_x^2(0) = \sigma^2 = \mathbb{E}\left( \int_{-\pi}^{\pi} e^{it\omega} \, dz_x(\omega) \right) \overline{\int_{-\pi}^{\pi} e^{it\omega} \, dz_x(\omega)}.
$$

The use of the complex conjugate is allowed since $x_t$ is real. Since $dz_x(\omega)$ is composed of orthogonal increments, it must be the case that

$$
E\left( \int_{-\pi}^{\pi} e^{it\omega} \, dz_x(\omega) \right) \overline{\int_{-\pi}^{\pi} e^{it\omega} \, dz_x(\omega)} = 0.
$$

The expected value operator can be moved inside the integral (since the integral is a linear operator), which means

$$
E\left( \int_{-\pi}^{\pi} e^{it\omega} \, dz_x(\omega) \right) \overline{\int_{-\pi}^{\pi} e^{it\omega} \, dz_x(\omega)} = \int_{-\pi}^{\pi} E\left( dz_x(\omega) \overline{dz_x(\omega)} \right) = \int_{-\pi}^{\pi} dF_x(\omega)
$$

This indicates that the spectral density reveals how each stochastic bit $dz_x(\omega)$ of $x_t$ contributes to the overall variance of $x_t$.

If $x_t$ is real, which is of course the case for economic data, one can rewrite the spectral representation as follows. First, observe that in order for
integral in (5.9) to be real with probability 1, it is necessary for \( dz_x(\omega) = \overline{dz_x(-\omega)} \).

Why? Take a fixed \( \omega \) and \(-\omega\). In order for the sum

\[
e^{i\omega x} dz_x(\omega) + e^{-i\omega x} dz_x(-\omega)
\]

(5.15)
to be real, it is necessary that \( e^{i\omega x} dz_x(\omega) = \overline{e^{-i\omega x} dz_x(-\omega)} \). Recall that \( e^{i\omega x} = \overline{e^{-i\omega x}} \).

Since \( z_x(\omega) \) is random, \( dz_x(\omega) = \overline{dz_x(-\omega)} \) is needed to hold with probability 1 in order to ensure the integral is real. Similarly, one can show that the restriction of real data implies that \( E(du_x(\omega)dv_x(\omega)) = 0 \)

Rewriting \( dz_x(\omega) = du_x(\omega) - idv_x(\omega) \) the requirement that \( dz_x(\omega) = \overline{dz_x(-\omega)} \) requires

\[
\begin{align*}
du_x(\omega) &= du_x(-\omega) \\
dv_x(\omega) &= -dv_x(-\omega)
\end{align*}
\]

(5.16)

Recall that by the spectral representation theorem \( E(dz_x(\omega)dz_x(-\omega)) = 0 \). This requirement can be rewritten

\[
E(du_x(\omega)du_x(-\omega)) + iE(du_x(\omega)dv_x(-\omega)) - \\
- iE(dv_x(\omega)du_x(-\omega)) + E(dv_x(\omega)dv_x(-\omega))
\]

(5.17)

= 0

Using (5.16), in order for (5.17) to hold at all frequencies, it must be the case that

\[
E(du_x(\omega)du_x(\omega)) - E(dv_x(\omega)dv_x(\omega)) = 0,
\]

which implies that

\[
E(du_x(\omega)^2) = E(dv_x(\omega)^2).
\]

We now rewrite the spectral representation of \( x_i \) as
\[ x_t = \int_{-\pi}^{\pi} e^{i\omega t} dz_x(\omega) = \int_{-\pi}^{\pi} \left( \cos \omega t + i \sin \omega t \right) \left( du_x(\omega) - idv_x(\omega) \right) = \int_{-\pi}^{\pi} \cos \omega t du_x(\omega) - i \int_{-\pi}^{\pi} \cos \omega t dv_x(\omega) + i \int_{-\pi}^{\pi} \sin \omega t du_x(\omega) + \int_{-\pi}^{\pi} \sin \omega t dv_x(\omega) \]  

(5.18)

The first and fourth integrals in the third line of (5.18) are even functions, meaning \( g(-y) = g(y) \), whereas the second and third are odd functions, meaning that \( g(-y) = -g(y) \). Therefore

\[ x_t = \int_{-\pi}^{\pi} \left( \cos \omega t du_x(\omega) + \sin \omega t dv_x(\omega) \right) = \int_{0}^{\pi} \left( \cos \omega t U_x(\omega) + \sin \omega t V_x(\omega) \right) \]  

(5.19)

where

\[ dU_x(\omega) = du_x(\omega) + du_x(-\omega) = 2du_x(\omega) \]  

(5.20)

and

\[ dV_x(\omega) = dv_x(\omega) - dv_x(-\omega) = 2dv_x(\omega). \]  

(5.21)

**Example 5.5. Deterministic Seasonal**

If \( x_t = e_t + \cos \omega t \), then the spectral density will not exist at \( \bar{\omega} \). The reason for this is that the one frequency will contribute a non-eligible amount of variance at \( \bar{\omega} \). In this case, \( dF_s \) will still exist; in this case \( \bar{\omega} \) represents a jump point.

One way to think about the spectral density function is to define it as
\[ f_x(\omega) = \alpha \frac{\sigma_x^2}{2\pi} + \beta \delta(\omega - \omega) \] (5.22)

where \( \delta(\cdot) \) is the so-called Dirac or delta function. This is a function with the properties i. \( \delta(0) = \infty; 0 \) otherwise and ii. \( \int_{-\infty}^{\infty} \delta(\omega - \omega)g(\omega)d\omega = g(0) \). This is called a generalized function. The Dirac function allows one to work with jumps in the spectral distribution function.

In general, we can write the spectral density function as

\[ f_x(\omega) = \alpha f(\omega) + \sum_k \beta_k \delta(\omega - \omega_k) \] (5.23)

where \( f(\omega) \) is continuous.

**iii. filters**

Often, one works with a time series that is a transformation of another, i.e.

\[ y_t = \beta(L)x_t \] (5.24)

In this case, \( \beta(L) \) is known as a filter. The frequency domain allows for a number of insights into the effects of filters.

The MA representation of a process may be thought of as describing a given process created by applying a filter to white noise. The following theorem describes how this filtering applies to the spectral representation of the process.

**Theorem 5.2. Construction of spectral representation from white noise spectral representation.**

Suppose that \( \varepsilon_t \) is a white noise process such that
\[ \varepsilon_t = \int_{-\pi}^{\pi} e^{j\omega t} dz_x(\omega) \quad (5.25) \]

Let \( x_t = \beta(L)\varepsilon_t \). Then

\[ x_t = \int_{-\pi}^{\pi} e^{j\omega t} \beta(e^{-j\omega}) dz_x(\omega) \quad (5.26) \]

which means that

\[ dz_x(\omega) = \beta(e^{-j\omega}) dz_x(\omega) \quad (5.27) \]

in the spectral representation of \( x_t \).

Pf.

\[ \sum_{j=-\infty}^{\infty} \beta_j \varepsilon_{t-j} = \sum_{j=-\infty}^{\infty} \beta_j \int_{-\pi}^{\pi} e^{j(t-j)\omega} dz_x(\omega) \]

\[ = \int_{-\pi}^{\pi} e^{j\omega t} \sum_{j=-\infty}^{\infty} \beta_j e^{-j\omega} dz_x(\omega) \quad (5.28) \]

\[ = \int_{-\pi}^{\pi} e^{j\omega t} \beta(e^{-j\omega}) dz_x(\omega) \]

\[ = \int_{-\pi}^{\pi} e^{j\omega t} dz_x(\omega) \]

where

\[ z_x(\lambda) = \int_{-\pi}^{\pi} \beta(e^{-j\omega}) dz_x(\omega) \quad (5.29) \]

It is straightforward to verify that \( dz_x(\lambda) \) possesses all the necessary requirements for the spectral representation. In particular,
\[ i. \ E(dz_x(\omega)) = \beta(e^{-i\omega})E(dz_\varepsilon(\omega)) = 0. \]

\[ ii. \ E(dz_x(\omega)^\dagger dz_x(\omega)) = \beta(e^{-i\omega})\beta(e^{i\omega})E(dz_\varepsilon(\omega)dz_\varepsilon(\omega)) = \beta(e^{-i\omega})\beta(e^{i\omega})\sigma_\varepsilon(0)2\pi \]

\[ iii. \ E(dz_x(\omega)^\dagger dz_x(\omega)) = \beta(e^{-i\omega})\beta(e^{i\omega})E(dz_\varepsilon(\omega)^\dagger dz_\varepsilon(\omega)) = 0 \text{ if } \omega_i \neq \omega_j. \]

(Uniqueness is also preserved.)

The relationship between the properties of \( dz_x(\omega) \) and \( dz_\varepsilon(\omega) \) can be understood as follows. Let \( \beta(e^{-i\omega}) = \gamma(\omega)e^{i\phi(\omega)} \). Observe that one can always do this by the standard properties of complex numbers. The two components of this polar representation of the Fourier transform of the filter have distinct effects on the stochastic process. The effect of \( \gamma(\omega) \) may be seen from the calculation

\[ E(dz_x(\omega)^\dagger dz_x(\omega)) = \gamma(\omega)E(dz_\varepsilon(\omega)^\dagger dz_\varepsilon(\omega)) = \gamma(\omega)\sigma_\varepsilon^22\pi \]

The first feature of any filter is that changes the length of \( dz_x(\omega) \) by \( \gamma(\omega) \). This is referred to as the gain of the filter at \( \omega \) and illustrates how the filter alters the relative variance contributions of different frequencies.

Second, consider the effect of \( e^{i\phi(\omega)} \) on \( e^{it\omega} \). From the form of the spectral representation, the effect is to alter the complex exponential in the sense that \( e^{it\omega}e^{i\phi(\omega)} = e^{i(t\omega + \phi(\omega))} \). This shifts the sine and cosine functions which comprise the complex exponential by \( \phi(\omega) \). This is called the phase shift.

There is nothing in this argument that does not immediately generalize to any filter. Hence for \( y_t = \beta(L)x_t \)
\[ y_t = \int \pi \omega \beta(e^{-i\omega})dz_t(\omega) \] 

(5.32)

and one can make the same sorts of arguments about the effects of the filter in terms of the gain and the phase shift.

**Example 5.5. differencing**

If \( y_t = (1-L)x_t \), then \( f_y(\omega) = (2-2\cos \omega)f_x(\omega) \)

Notice that for a differenced series, \( f_y(0) = 0 \). This makes intuitive sense. Differencing eliminates the part of the process common to all observations.

**Example 5.6. averaging**

Suppose that we define \( \beta(L) \) such that

\[
\beta_j = \frac{1}{T}, \quad j = 0...T-1, \quad 0 \text{ otherwise} \tag{5.33}
\]

This means that the filter averages the \( x_t \) process. In this case,

\[
\beta(e^{-i\omega}) = T^{-1} \sum_{j=0}^{T-1} e^{-ij\omega}. \tag{5.34}
\]

Note that \( \beta(e^{-i\omega}) = 1 \) if \( \omega = 0 \).

Further,

\[
\beta(e^{-i\omega}) = T^{-1} \frac{1-e^{-iT\omega}}{1-e^{-i\omega}} \quad \omega \neq 0
\]

\[
\beta(e^{-i\omega}) = 1 \quad \omega = 0.
\]
\[ \beta(e^{-i\omega}) \beta(e^{i\omega}) = T^{-2} \left( \frac{2 - 2\cos T \omega}{2 - 2\cos \omega} \right) = T^{-2} \frac{\sin^2 \left( \frac{T \omega}{2} \right)}{\sin^2 \left( \frac{\omega}{2} \right)} \] (5.35)

This is a function which converges to the indicator function \( I_{\omega=0} \). The spectral representation of an averaged series will therefore converge to \( d_x(0) \). This makes sense; when we average all the elements of \( x_t \), the part common to all elements is what remains; this is what the zero frequency element \( d_x(0) \) captures.

The spectral density function of the averaged series is

\[ f_x(\omega) \cdot T^{-2} \cdot \frac{\sin^2 \left( \frac{T \omega}{2} \right)}{\sin^2 \left( \frac{\omega}{2} \right)} \] (5.36)

a function whose integral over \([-\pi, \pi]\) will converge to zero. This implies that a form of the law of large numbers holds for weakly stationary time series, so long as the spectral density is bounded at 0.

**Example 5.7. Band pass filter.**

Take a time series \( x_t \) and suppose we wish to create a series \( y_t \) that removes the part of the Cramér representation that corresponds to those frequencies above some specified value \( \overline{\omega} \), so that

\[
\begin{align*}
    f_y(\omega) &= f_x(\omega) \text{ if } |\omega| \leq \overline{\omega} \\
    f_y(\omega) &= 0 \text{ if } |\omega| > \overline{\omega}.
\end{align*}
\] (5.37)
This would imply that there exists a filter $\beta(L)$ such that

$$β(e^{-iω})β(e^{iω}) = 1 \text{ if } |ω| \leq \bar{ω}, \text{ 0 otherwise.} \quad (5.38)$$

What polynomial $\beta(\cdot)$ has this property? We construct it as follows. Let

$$β(e^{-iω}) = γ(ω)$$

which requires that $β(e^{-iω})$ is symmetric and 2-sided. Hence, in order to recover $β_j$, we can use the Fourier inversion formula,

$$β_j = \frac{1}{2π} \int_{-\infty}^{\infty} e^{iωj} dω = \frac{1}{\pi} \int_{0}^{\infty} \cos(ωj) dω = \frac{\sin(\bar{ω}j)}{\pi j} \quad (5.39)$$

This filter was proposed by Robert Engle to allow for band spectrum regression. The idea was to allow one to analyze regressions based on the “long run” parts of various time series. Notice that the filter does not preserve various temporal relationships in the data, so that filtered data will not obey restrictions that economic theory (for example) places on the original series. This criticism also applies to the so-called Hodrick-Prescott filter.