Lecture 4: Frequency domain

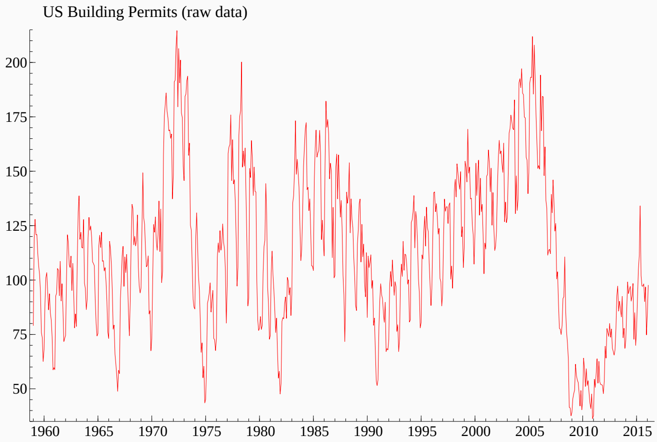
Geert Mesters (Universitat Pompeu Fabra, Barcelona GSE and VU Amsterdam)

Motivation

The study of business cycles necessarily begins with the measurement of business cycles

Baxter & King (1999)

US building permits



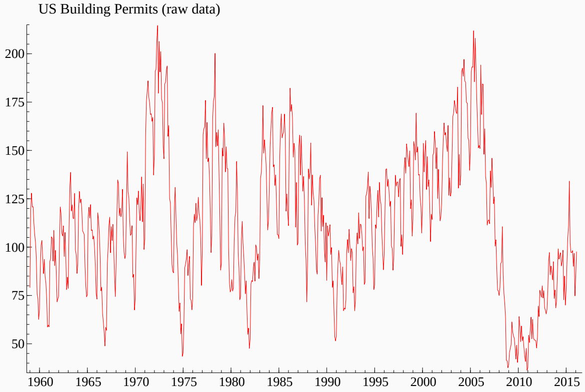
Motivation

- We are used to describing time series in terms of past shocks

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

- A more natural (literally natural) way is to think about time series in terms of sums of different regular components (e.g. cosines and sines).

US building permits



We may find a long term component, a business cycle component, a seasonal component etc... each of these can be represented by a (co)sine function

Some motivation

- Decomposing Y_t in sines and cosines amounts to studying Y_t in the frequency domain as opposed to the time domain
- Some recent uses for frequency domain methods
  - Business cycle detection (e.g. Baxter & King 1999, Beaudry, Galizia & Portier 2019)
  - Low frequency econometrics (e.g. Muller & Watson 2018)
  - HAC/HAR standard errors (e.g. Lazurus, Lewis, Stock & Watson 2017, 2019)

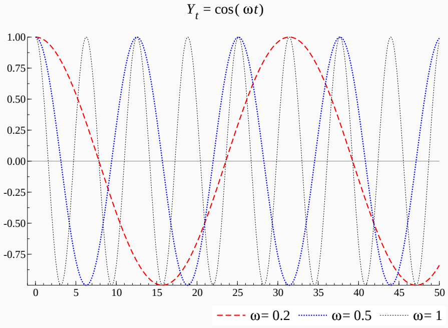
Spectral representation

A simple deterministic model

Consider the deterministic model

$$Y_t = \cos(\omega t)$$

where ω is the frequency and $2\pi/\omega$ is the period.



Maybe a bit too simplistic for an economic time series:-)

A simple stochastic model

Lets add a sine and some random weights a and b

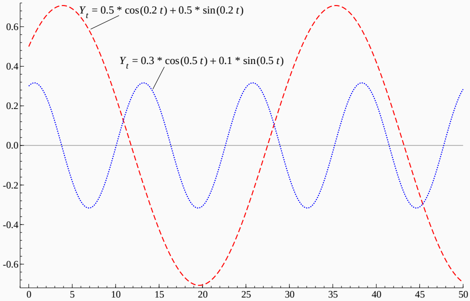
$$Y_t = a\cos(\omega t) + b\sin(\omega t)$$

where

- a and b are uncorrelated random variables $\mathrm{E}(a) = \mathrm{E}(b) = 0$ and $\mathrm{Var}(a) = \mathrm{Var}(b) = \sigma^2$
- Adding the sine and random variables allows us to change the starting point a and the amplitude $\sqrt{a^2 + b^2}$ [1]

[1] To derive the amplitude: $\frac{\partial Y_t}{\partial \omega} = -at\sin(\omega t) + bt\cos(\omega t) = 0$ which implies $\cos(\omega t) = \frac{a}{b}\sin(\omega t)$. Using $\cos(\omega t) = \sqrt{1 - \sin(\omega t)^2}$ we can solve for $\cos(\omega t) = \frac{a}{\sqrt{a^2+b^2}}$ and $\sin(\omega t) = \frac{b}{\sqrt{a^2+b^2}}$. Finally, to get the value plug these in Y_t to find the maximum value of $Y_t^{max} = \sqrt{a^2 + b^2}$.

A simple stochastic model: Illustration



Oke becoming more flexible!!!

A simple stochastic model: properties

The stochastic process

$$Y_t = a\cos(\omega t) + b\sin(\omega t)$$

has properties[2]

- $\mathrm{E}(Y_t) = 0$
- $\mathrm{Var}(Y_t) = \cos(\omega t)^2 \mathrm{Var}(a) + \sin(\omega t)^2 \mathrm{Var}(b) = \sigma^2$
- $\mathrm{Cov}(Y_t, Y_{t+h}) = \sigma^2 \cos(\omega h)$

[2] Recall trigonometry $\cos(x)^2 + \sin(x)^2 = 1$ and $\cos(x - y) = \cos(x)\cos(y) + \sin(x)\sin(y)$

Sums of simple stochastic models

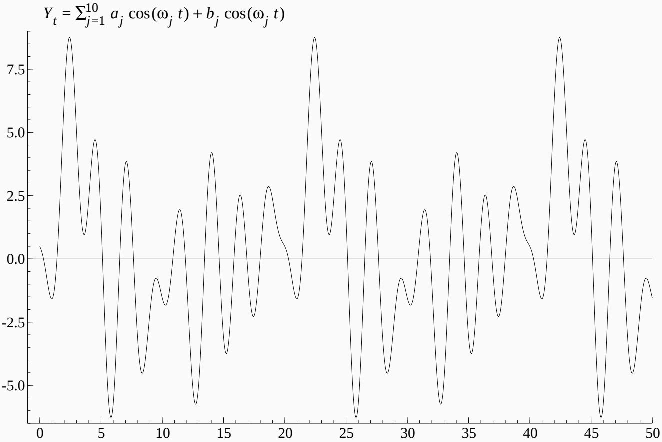
Now having only one frequency is a bit sad, so lets add a few components (i.e. add seasonal, business cycle, etc...)

$$Y_t = \sum_{j=1}^{M} a_j \cos(\omega_j t) + b_j \sin(\omega_j t)$$

where

- a_j and b_j are uncorrelated random variables
  $\mathrm{E}(a_j) = \mathrm{E}(b_j) = 0$ and $\mathrm{Var}(a_j) = \mathrm{Var}(b_j) = \sigma_j^2$
- Allowing different frequencies ω_j to have different variances σ_j^2 allows for changing the importance of different frequencies

Sums of simple stochastic models: Illustration



Now this is what we want!!!

Sums of simple stochastic models: properties

The stochastic process

$$Y_t = \sum_{j=1}^{M} a_j \cos(\omega_j t) + b_j \sin(\omega_j t)$$

has properties

- $\mathrm{E}(Y_t) = 0$
- $\mathrm{Var}(Y_t) = \sum_{j=1}^{M} \sigma_j^2$
- $\mathrm{Cov}(Y_t, Y_{t+h}) = \sum_{j=1}^{M} \sigma_j^2 \cos(\omega_j h)$

Sums of simple stochastic models: interpretation

For the stochastic process

$$Y_t = \sum_{j=1}^{M} a_j \cos(\omega_j t) + b_j \sin(\omega_j t)$$

we have

- variance is decomposed into M components σ_j^2
- each σ_j^2 corresponds to a different frequency ω_j
- covariances are decomposed into M components $\sigma_j^2 \cos(\omega_j h)$
- each $\sigma_j^2 \cos(\omega_j h)$ corresponds to a different frequency ω_j

This implies that we can attribute a certain portion of the variance to periodic random components

Sums of simple stochastic models: generalize

The stochastic model

$$Y_t = \sum_{j=1}^{M} a_j \cos(\omega_j t) + b_j \sin(\omega_j t)$$

can be generalized

- Potentially we may want a representation that allows for all frequencies in $[0, \pi]$
- In the model we take $M \to \infty$ to get spectral representation theorem

Spectral representation theorem

Theorem

If $\{Y_t, t \in \mathbb{Z}\}$ *is a zero mean covariance stationary process it has a representation*

$$Y_t = \int_0^\pi [a(\omega)\cos(\omega t) + b(\omega)\sin(\omega t)]\mathrm{d}\omega$$

where the random processes $a(\cdot)$ *and* $b(\cdot)$ *are mean zero, uncorrelated and for any ordered sequence of frequencies* $0 < \omega_1 < \omega_2 < \omega_3 < \omega_4 < \pi$,

$$\mathrm{E}\left[\left(\int_{\omega_1}^{\omega_2} a(\omega)\mathrm{d}\omega\right)\left(\int_{\omega_3}^{\omega_4} a(\omega)\mathrm{d}\omega\right)\right] = 0$$

and

$$\mathrm{E}\left[\left(\int_{\omega_1}^{\omega_2} b(\omega)\mathrm{d}\omega\right)\left(\int_{\omega_3}^{\omega_4} b(\omega)\mathrm{d}\omega\right)\right] = 0$$

Some comments

- It looks bad, but the interpretation is exactly the same as for the model with a discrete number of frequencies :
  Any stationary time series may be thought of, approximately, as the random superposition of sines and cosines oscillating at various frequencies
- We do not discuss the proof (it's quite involved), see Brockwell and Davis (1991) Section 4.8
- The spectral representation theorem is the frequency domain equivalent of the Wold decomposition

Describing the autocovariance function

- For the simple process we could decompose the autocovariance function into contributions from different periodic components. Can we still do this?
- The answer is yes.

Spectral density or Population spectrum

Theorem

The Spectral density or Population spectrum of the mean zero stationary time series process $\{Y_t, t \in \mathbb{Z}\}$ that has autocovariances $\gamma_Y(h)$ is implicitly defined as

$$\gamma_Y(h) = \int_{-\pi}^{\pi} e^{i\omega h} s_Y(\omega) \mathrm{d}\omega$$

which can be inverted to yield

$$s_Y(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_Y(j) e^{-i\omega j}$$

See Brockwell and Davis (1991) Chapter 4 for a derivation.

Interpretation spectral density

- Since $e^{i\omega h} = \cos(\omega h) + i \sin(\omega h)$ it follows that

$$\gamma_Y(h) = \int_{-\pi}^{\pi} (\cos(\omega h) + i \sin(\omega h))\, s_Y(\omega) \mathrm{d}\omega$$

  such that $s_Y(\omega)$ can be viewed at the function that gives weight to (co)sines with different frequencies

- Equivalently, can think about $s_Y(\omega)$ as a density function, in the same way we think about probability density functions

Interpretation spectral density (continued)

- $s_Y(\omega)$ has the same interpretation as σ_j in the simple model
- $s_Y(\omega)$ measures the variance that is due to frequency ω
- Example: for $h = 0$: $\gamma_Y(0) = \int_{-\pi}^{\pi} s_Y(\omega) \mathrm{d}\omega$
- The spectral density and the autocovariances contain the same "information"

Spectral density at 0

A very interesting frequency is $\omega = 0$. Here we have

$$s_Y(0) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_Y(j)$$

but hey, this we recognize from lecture 1 where we found for the sample mean of a stationary time series

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y_t \xrightarrow{d} N\left(0, \sum_{h=-\infty}^{\infty} \gamma_Y(h)\right)$$

So for stationary processes the long run variance of the sample mean is equivalent to (2π times) the spectral density at frequency zero

Spectral density at 0

Implies

- Estimating the long run variance is the same as estimating the spectral density at frequency zero
- Clarifies the difficulty of the problem
- This motivates thinking about HAC standard errors from the Frequency domain perspective; see Lazurus, Lewis, Stock & Watson (2017, 2019)

Examples spectral densities

Example spectral density

White noise

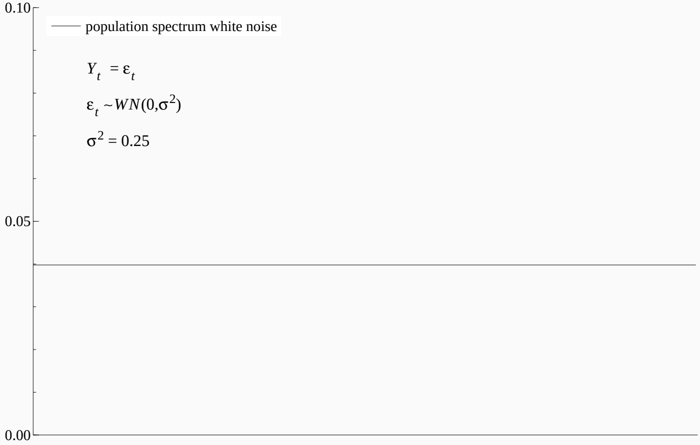
For the white noise process

$$Y_t = \epsilon_t \qquad \epsilon_t \sim WN(0, \sigma^2)$$

the spectral density function becomes

$$\begin{aligned} s_Y(\omega) &= \tfrac{1}{2\pi} \textstyle\sum_{j=-\infty}^{\infty} \gamma_Y(j) e^{-i\omega j} \\ &= \tfrac{1}{2\pi} \gamma_Y(0) = \tfrac{1}{2\pi} \sigma^2 \end{aligned}$$

Example spectral density



Spectral density for linear processes

Theorem

Filters in the frequency domain If $\{X_t\}$ is any zero mean possibly complex valued stochastic process with spectral density $s_X(\omega)$ and $\{Y_t\}$ is the process

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} \qquad \sum_{j=-\infty}^{\infty} |\psi_j| < \infty$$

then $\{Y_t\}$ is stationary with spectral density

$$s_Y(\omega) = |\psi(e^{-i\omega})|^2 s_X(\omega)$$

where $\psi(e^{-i\omega}) = \sum_{j=-\infty}^{\infty} \psi_j e^{-i\omega j}$

Proof: see Brockwell and Davis (1991) Theorem 4.4.1

Some terminology

The previous proposition is the frequency domain equivalent of filtering in the time domain.

Some terminology

- $\psi(L) = \sum_{j=-\infty}^{\infty} \psi_j L^j$: linear time-invariant filter
- $\psi(e^{i\omega})$ is the transfer function of the filter (as it transfers power from one frequency to another)
- $|\psi(e^{i\omega})|^2$ is the power transfer function

Example spectral density

MA(1)

For the MA(1)

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1} \qquad \epsilon_t \sim WN(0, \sigma^2)$$

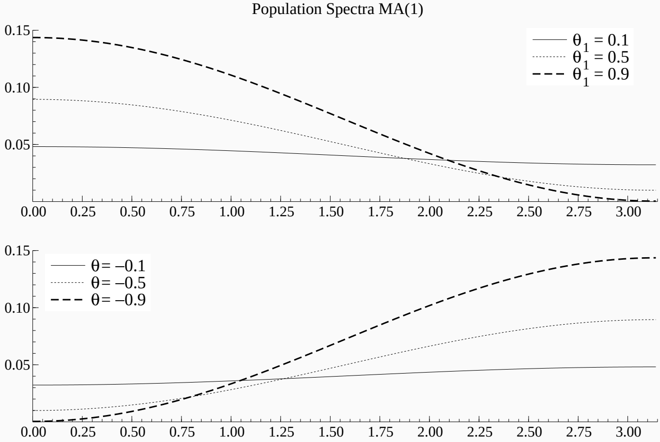
the spectral density function becomes (from the definition)

$$\begin{aligned} s_Y(\omega) &= \tfrac{1}{2\pi} \textstyle\sum_{j=-\infty}^{\infty} \gamma_Y(j) e^{-i\omega j} \\ &= \tfrac{1}{2\pi} \left[\sigma^2(1+\theta^2) + \sigma^2\theta e^{i\omega} + \sigma^2\theta e^{-i\omega}\right] \\ &= \tfrac{1}{2\pi}\sigma^2(1+\theta_1^2 + 2\theta_1 \cos(\omega)) \end{aligned}$$

as

$$\cos(\omega) = \frac{e^{i\omega} + e^{-i\omega}}{2}$$

Example spectral density



Examples spectral density

AR(1)

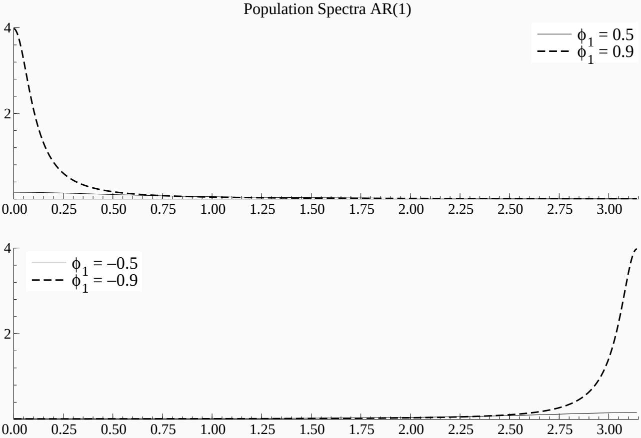
For the AR(1)

$$Y_t = \phi_1 Y_{t-1} + \epsilon_t \qquad \epsilon_t \sim WN(0, \sigma^2)$$

the spectral density function becomes (applying the power transfer function to the white noise model)

$$\begin{aligned}
s_Y(\omega) &= \left|\phi(e^{-i\omega})\right|^{-2} s_\epsilon(\omega) \\
&= \frac{1}{2\pi}\frac{\sigma^2}{(1-\phi_1 e^{i\omega})(1-\phi_1 e^{-i\omega})} \\
&= \frac{1}{2\pi}\frac{\sigma^2}{(1-\phi_1 e^{i\omega}-\phi_1 e^{-i\omega}+\phi_1^2)} \\
&= \frac{1}{2\pi}\frac{\sigma^2}{(1+\phi_1^2-2\phi_1\cos(\omega))}
\end{aligned}$$

Examples spectral density



Examples spectral density

ARMA(p,q)

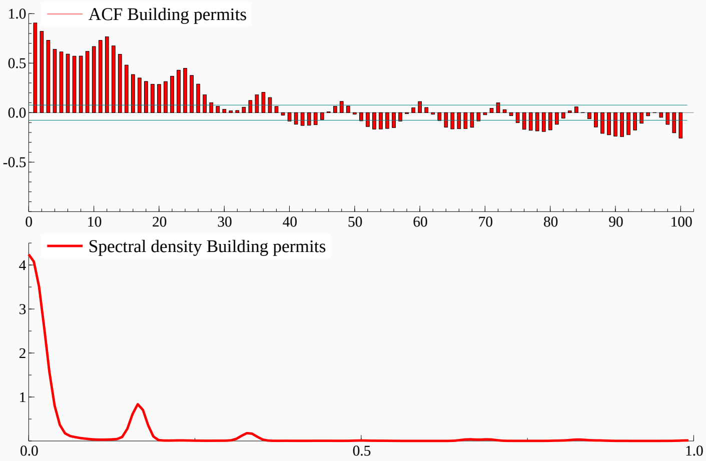
For the general ARMA(p,q) process

$$Y_t - \phi_1 Y_{t-1} - \ldots - \phi_p Y_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

the spectral density function becomes

$$s_Y(\omega) = \sigma^2 \frac{(1 + \theta_1 e^{i\omega} + \ldots + \theta_q e^{iq\omega})(1 + \theta_1 e^{-i\omega} + \ldots + \theta_q e^{-iq\omega})}{(1 - \phi_1 e^{i\omega} - \ldots - \phi_p e^{ip\omega})(1 + \phi_1 e^{-i\omega} - \ldots - \phi_p e^{-ip\omega})}$$

ACF and Spectral density building permits



Discrete Fourier transform

Finite Sample Representation

To get a description of the finite sample $Y_1, \ldots, Y_T$ in the frequency domain, we show that any fixed number of observations can be written as a weighted average of cosines and sines function

Finite Sample Representation

To see this let

$$\boldsymbol{Y}_T = \begin{bmatrix} Y_1 \\ \vdots \\ Y_T \end{bmatrix} \qquad \boldsymbol{Y}_T \in \mathbb{C}^T$$

where $\mathbb{C}^T$ is the set of T-dimensional complex numbers.

We define the Fourier frequencies

$$\omega_k = \frac{2\pi k}{T} \qquad k = \lfloor -(T-1)/2 \rfloor, \ldots, \lfloor T/2 \rfloor$$

where $\lfloor a \rfloor$ denotes the largest integer smaller then or equal to a [3] and notice that $\omega_k \in (-\pi\,, \pi]$.

[3] Just a convenient way of avoiding to distinguish between T even and T odd.

Finite Sample Representation

We introduce the T-dimensional vectors

$$\boldsymbol{e}_k = \frac{1}{\sqrt{T}} \begin{bmatrix} e^{i\omega_k 1} \\ e^{i\omega_k 2} \\ \vdots \\ e^{i\omega_k T} \end{bmatrix} \qquad k = \lfloor -(T-1)/2 \rfloor, \ldots, \lfloor T/2 \rfloor$$

which satisfy [4]

$$\boldsymbol{e}_k^* \boldsymbol{e}_j = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

and this implies that $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_T\}$ form a *basis* for $\mathbb{C}^T$.

[4] check this for yourself, follows from cosine and sine properties

Finite Sample Representation

Given that $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_T\}$ form a *basis* for $\mathbb{C}^T$ we can express any $\boldsymbol{Y}_T = (Y_1, \ldots, Y_T)' \in C^T$ as

$$\boldsymbol{Y}_T = \sum_{k=\lfloor -(T-1)/2 \rfloor}^{\lfloor T/2 \rfloor} a_k \boldsymbol{e}_k$$

The coefficients a_k can be found by pre-multiplying both sides with $\boldsymbol{e}_k^*$, i.e.

$$a_k = \boldsymbol{e}_k^* \boldsymbol{Y}_T = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y_t e^{-i\omega_k t}$$

The sequence $\{a_k\}$ is called the **discrete Fourier transform** of the numbers $Y_1, \ldots, Y_T$.

Finite Sample Representation

Now observe that the tth element of

$$\boldsymbol{Y}_T = \sum_{k=\lfloor -(T-1)/2 \rfloor}^{\lfloor T/2 \rfloor} a_k \boldsymbol{e}_k$$

can be written as

$$\begin{aligned} Y_t &= \sum_{k=\lfloor -(T-1)/2 \rfloor}^{\lfloor T/2 \rfloor} a_k e^{i\omega_k t} \\ &= \sum_{k=\lfloor -(T-1)/2 \rfloor}^{\lfloor T/2 \rfloor} a_k [\cos(\omega_k t) + i \sin(\omega_k t)] \end{aligned}$$

which shows that we can represent any finite set of numbers in terms of cosine and sine waves.[5]

[5] Proposition 6.2 in Hamilton establishes the same.
Notice the special properties of the Fourier frequencies!!!

Estimating the spectral density

Estimation

Next, we discuss estimation of the spectral density $s_Y(\omega)$. We discuss

- Non-parametric estimation
- Parametric estimation

Definition :

The **periodogram** for $\{Y_1, \ldots, Y_T\}$ is the function

$$I_T(\omega) = \frac{1}{T}\left|\sum_{t=1}^{T} Y_t e^{-i\omega t}\right|^2$$

Remark: Note that when ω_k is a Fourier frequency then $I_T(\omega) = |a_k|^2$ (check this). The value of the periodogram at such frequency ω_k is thus the contribution to this sum of squares from the "frequency ω_k"

Periodogram as the sample equivalent

The next theorem shows that the periodogram is the sample equivalent of $2\pi s_Y(\omega)$. Recall that

$$2\pi s_Y(\omega) = \sum_{j=-\infty}^{\infty} \gamma_Y(j) e^{-i\omega j}$$

Theorem

For any real numbers $\{Y_1, \ldots, Y_T\}$ and ω_k being one of the non-zero Fourier frequencies $2\pi k/T \in (-\pi\ , \pi]$ then

$$I_T(\omega_k) = \sum_{j=-T+1}^{T-1} \hat{\gamma}_Y(j) e^{-i\omega_k j}$$

Proof of proposition:

We have that $\sum_{t=1}^{T} e^{i\omega_k t} = 0$ if $\omega_k \neq 0$.[6] This implies that we can subtract the sample mean $\bar{Y}_T$ from Y_t in the definition of the periodogram. We get

$$\begin{aligned} I_T(\omega_k) &= \frac{1}{T}\left|\sum_{t=1}^{T} Y_t e^{-i\omega_k t}\right|^2 \\ &= \frac{1}{T}\left|\sum_{t=1}^{T}(Y_t - \bar{Y}_T) e^{-i\omega_k t}\right|^2 \\ &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}(Y_t - \bar{Y}_T)(Y_s - \bar{Y}_T) e^{-i\omega_k (t-s)} \\ &= \sum_{j=-T+1}^{T-1} \hat{\gamma}_Y(j) e^{-i\omega_k j} \end{aligned}$$

[6] More appropriately here is to say we have $\sum_{t=1}^{T} e^{-i\omega_k t} = 0$ if $\omega_k \neq 0$. This can be proved by taking $z = e^{-i\omega_k}$, $\sum_{t=1}^{T} z^t = \frac{1-z^T}{1-z}$ which is equal to zero as $z^T = e^{-i\omega_k T} = e^{-i2\pi k} = 1$

Periodogram

The periodogram $I_T(\omega_k)$ captures the portion of the variance that can be attributed to cycles of different frequencies ω_k

Unfortunately, using the periodogram for estimating the spectral density has some serious limitations that are illustrated on the next slides.

Unbiased but large confidence bounds

First the good news, consider $Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, for large T and $\omega_k \neq 0$ we can show that

$$\frac{I_T(\omega)/\pi}{s_Y(\omega)} \approx \chi^2(2)$$

and since a $\chi^2(2)$ variable has mean 2 we find that

$$\mathrm{E}\left(\frac{I_T(\omega)/\pi}{s_Y(\omega)}\right) \approx 2$$

implying that

$$\mathrm{E}\left(I_T(\omega)/2\pi\right) \approx s_Y(\omega)$$

So the sample periodogram offers an approximately unbiased estimate of the population spectrum.

Unbiased but large confidence bounds

This is off course good news, but a $\chi^2(2)$ variable has large confidence bounds.

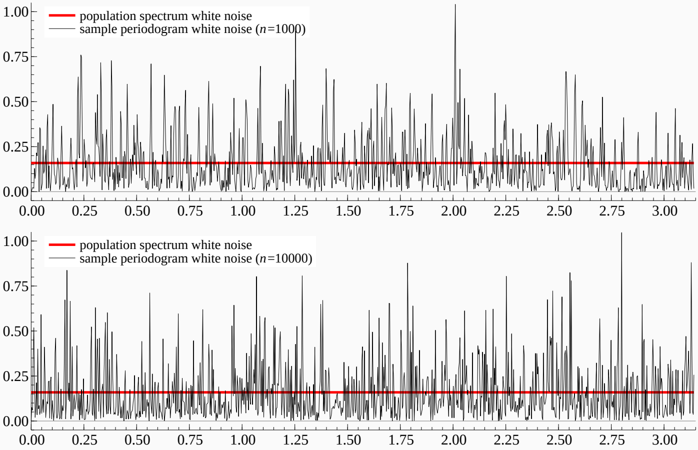
- 95% of the times it will fall between 0.05 and 7.4
- This is a large large confidence region

Inconsistent

More fundamentally, as $T \to \infty$ it does **not** hold that $I_T(\omega)/2\pi \xrightarrow{p} s_Y(\omega)$. Intuitively this can be seen by noting that

- $I_T(\omega)/2\pi = \frac{1}{2\pi}\sum_{j=-T+1}^{T-1} \hat{\gamma}_Y(j)e^{-i\omega j}$ depends on $T-1$ estimates $\hat{\gamma}_Y(j)$ so there are infinite "parameters" in the limit
- These parameters cannot all be consistently estimated!!!

Example inconsistency white noise



Non-parametric estimate population spectrum

We can solve this problem using non-parametric kernel estimators. Assume that $s_Y(\omega)$ is close to $s_Y(\lambda)$ for ω close to λ. In this way we can use the information from multiple frequencies to get consistent estimates. Consider $\omega_k = 2\pi k/n$ and we estimate the spectral density by

$$\tilde{s}_Y(\omega_k) = \frac{1}{2\pi} \sum_{m=-h}^{h} K(\omega_{k+m}, \omega_k) I_T(\omega_{k+m})$$

where h is a bandwidth parameter that determines how many different frequencies to use and $K(\omega_{k+m}, \omega_k)$ is the kernel that determines how much weight to give to each frequency. The kernel weights sum to unity

$$\sum_{m=-h}^{h} K(\omega_{k+m}, \omega_k) = 1$$

A simple distance based example for the kernel is

$$K(\omega_{k+m}, \omega_k) = \frac{h+1-|m|}{(h+1)^2}$$

for which we get

$$\tilde{s}_Y(\omega_k) = \frac{1}{2\pi} \sum_{m=-h}^{h} \frac{h+1-|m|}{(h+1)^2} I_T(\omega_{k+m})$$

which for $h = 2$ becomes

$$\begin{aligned} \tilde{s}_Y(\omega_k) \quad = \quad & \tfrac{1}{2\pi} \left(\tfrac{1}{9} I_T(\omega_{k-2}) + \tfrac{2}{9} I_T(\omega_{k-1}) + \tfrac{3}{9} I_T(\omega_k) \right. \\ & \left. + \tfrac{2}{9} I_T(\omega_{k+1}) + \tfrac{1}{9} I_T(\omega_{k+2}) \right) \end{aligned}$$

Non-parametric estimate population spectrum

Alternatively we can impose the smoothing by directly weighting the sample autocovariances. For example if

$$I_T(\omega_k)/2\pi = \frac{1}{2\pi}\left[\hat{\gamma}_Y(0) + 2\sum_{j=1}^{T-1}\hat{\gamma}_Y(j)\cos(\omega_k j)\right]$$

we can take

$$\tilde{s}_Y(\omega_k) = \frac{1}{2\pi}\left[\hat{\gamma}_Y(0) + 2\sum_{j=1}^{T-1}k_j^*\hat{\gamma}_Y(j)\cos(\omega_k j)\right]$$

were k_j^* are weights that smooth out the autocovariances, e.g. the popular Bartlett kernel weights are given by

$$k_j^* = \begin{cases} 1 - \frac{j}{q+1} & \text{for } j = 1, 2, \ldots, q, \\ 0 & \text{for } j > q \end{cases}$$

Non-parametric estimate population spectrum

An important question is how to choose bandwidth parameters h and q?

- The periodogram is unbiased but has large variance
- Smoothing introduced bias but reduces variance
- A practical guide is to choose the bandwidth to minimize the mean squared error in the estimate, see Hurvich (1985)

Illustration GDP

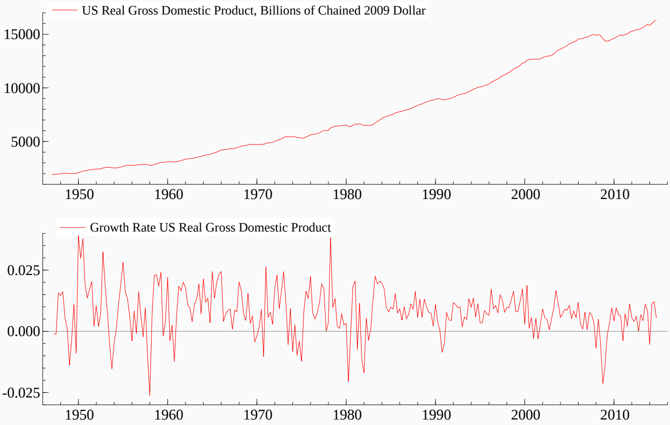


Illustration GDP

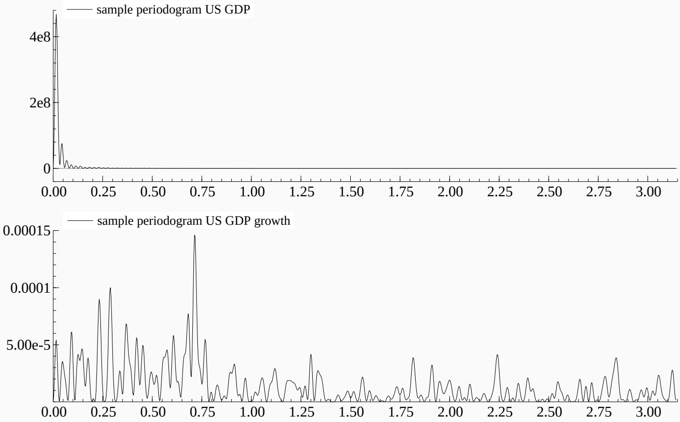
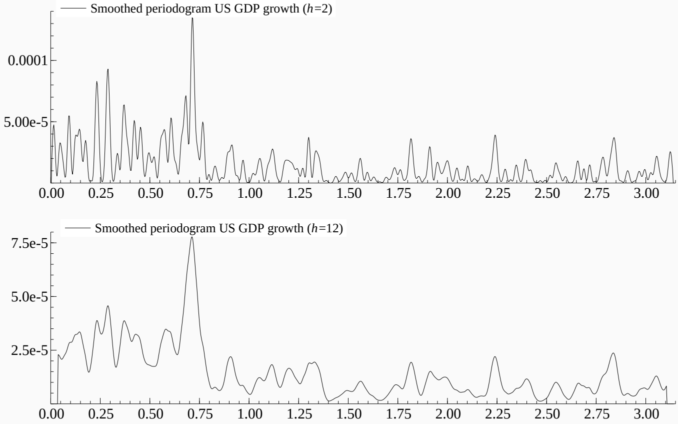


Illustration GDP



Parametric estimate population spectrum

Alternatively, an easy way to consistently estimate the spectral density is to assume that the data are generated by a particular ARMA(p,q) model

$$Y_t - \phi_1 Y_{t-1} - \ldots - \phi_p Y_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

estimate the parameters by maximum likelihood to get $\hat{\boldsymbol{\theta}} = (\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q)'$. and plug these into

$$s_Y(\omega) = \sigma^2 \frac{(1 + \theta_1 e^{i\omega} + \ldots + \theta_q e^{iq\omega})(1 + \theta_1 e^{-i\omega} + \ldots + \theta_q e^{-iq\omega})}{(1 - \phi_1 e^{i\omega} - \ldots - \phi_p e^{ip\omega})(1 - \phi_1 e^{-i\omega} - \ldots - \phi_p e^{-ip\omega})}$$

This yields a consistent estimate for $s_Y(\omega)$, *but* it requires knowing the correct ARMA(p,q) model.

Overview

Thus, we can estimate the spectral density by

- Non-parametric smoothing of the periodogram
- using the parametric functional form of the ARMA model spectral density

The estimated spectrum informs us which frequencies are important for capturing the variation in $\{Y_t\}$.

References & Material

Reference & Material

- References:
  R.H. Shumway & D.S. Stoffer, "Time Series Analysis and Its Applications with R examples", Chapter 4