Spectral Analysis

Jesús Fernández-Villaverde
University of Pennsylvania

Why Spectral Analysis?

- We want to develop a theory to obtain the business cycle properties of the data. Burns and Mitchell (1946).

- General problem of signal extraction.

- We will use some basic results in spectral (or harmonic) analysis.

- Then, we will develop the idea of a filter.

Riesz-Fisher Theorem

- First we will show that there is an intimate link between $L_2[-\pi, \pi]$ and $l_2(-\infty, \infty)$.

- Why? Because we want to represent a time series in two different ways.

- Riesz-Fischer Theorem: Let $\{c_n\}_{n=-\infty}^{\infty} \in l_2(-\infty, \infty)$. Then, there exist a function $f(\omega) \in L_2[-\pi, \pi]$ such that:

$$f(\omega) = \sum_{j=-\infty}^{\infty} c_j e^{-i\omega j}$$

- This function is called the *Fourier Transform* of the series.

Properties

- Its finite approximations converge to the infinite series in the mean square norm:

$$\lim_{n\to\infty}\int_{-\pi}^{\pi}\left|\sum_{j=-n}^{n}c_j e^{-i\omega j}-f(\omega)\right|^2 d\omega=0$$

- It satisfies the *Parseval's Relation*:

$$\int_{-\pi}^{\pi}|f(\omega)|^2\,d\omega=\sum_{n=-\infty}^{\infty}|c_n|^2$$

- Inversion formula:

$$c_k=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\omega)\,e^{i\omega k}d\omega$$

Harmonics

- The functions $\left\{e^{-i\omega j}\right\}_{j=-\infty}^{j=\infty}$ are called harmonics and constitute an orthonormal base in $L_2[-\pi,\pi]$.

- The orthonormality follows directly from:

$$\int_{-\pi}^{\pi} e^{-i\omega j} e^{i\omega k} d\omega = 0$$

  if $j \neq k$ and

$$\int_{-\pi}^{\pi} e^{-i\omega j} e^{i\omega j} d\omega = 1$$

- The fact that they constitute a base is given by the second theorem that goes in the opposite direction than Riesz-Fischer: given any function in $L_2[-\pi,\pi]$ we can find an associated sequence in $l_2(-\infty,\infty)$.

Converse Riesz-Fischer Theorem

Let $f(\omega) \in L_2[-\pi, \pi]$. Then there exist a sequence $\{c_n\}_{n=-\infty}^{\infty} \in l_2(-\infty, \infty)$ such that:

$$f(\omega) = \sum_{j=-\infty}^{\infty} c_j e^{-i\omega j}$$

where:

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) e^{i\omega k} d\omega$$

and finite approximations converge to the infinite series in the mean square norm:

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} \left| \sum_{j=-n}^{n} c_j e^{-i\omega j} - f(\omega) \right|^2 d\omega = 0$$

Remarks I

- The Riesz-Fisher theorem and its converse assure then that the Fourier transform is an *bijective* from $l_2(-\infty, \infty)$ into $L_2[-\pi, \pi]$.

- The mapping is *isometric isomorphism* since it preserves linearity and distance: for any two series $\{x_n\}_{n=-\infty}^{\infty}$, $\{y_n\}_{n=-\infty}^{\infty} \in l_2(-\infty, \infty)$ with Fourier transforms $x(\omega)$ and $y(\omega)$ we have:

$$x(\omega) + y(\omega) = \sum_{j=-\infty}^{\infty} (x_n + y_n) e^{i\omega j}$$

$$\alpha x(\omega) = \sum_{j=-\infty}^{\infty} \alpha x_n e^{i\omega j}$$

$$\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} |x(\omega) - y(\omega)|^2 d\omega\right\}^{\frac{1}{2}} = \left(\sum_{j=-\infty}^{\infty} |x_k - y_k|^2\right)^{\frac{1}{2}}$$

Remarks II

.

- There is a limit in the amount of information in a series.

- We either have concentration in the original series or concentration in the Fourier transform.

- This limit is given by:

$$\left(\sum_{j=-\infty}^{\infty} j^2 \left|c_j\right|^2\right) \left(\int_{-\infty}^{\infty} \omega^2 \left|f(\omega)\right|^2 d\omega\right) \geq \frac{1}{4} \left\|c_j\right\|^4$$

- This inequality is called the Robertson-Schrödinger relation.

Stochastic Process

- $X \equiv \{X_t : \Omega \to \mathbb{R}^m,\ m \in \mathbb{N},\ t = 1, 2, ...\}$ is a stochastic process defined on a complete probability space $(\Omega, \Im, P)$ where:

  1. $\Omega = \mathbb{R}^{m\times\infty} \equiv \lim_{T\to\infty} \otimes_{t=0}^{T} \mathbb{R}^m$

  2. $\Im \equiv \lim_{T\to\infty} \Im^T \equiv \lim_{T\to\infty} \otimes_{t=0}^{T} \mathcal{B}\left(\mathbb{R}^m\right) \equiv \mathcal{B}\left(\mathbb{R}^{m\times\infty}\right)$ is just the Borel σ-algebra generated by the measurable finite-dimensional product cylinders.

  3. $P^T(B) \equiv P(B)\,|\Im^T \equiv P\left(Y^T \in B\right)$, $\forall B \in \Im^T$.

- Define a $T-$segment as $X^T \equiv \left(X_1', ..., X_T'\right)'$ with $X^0 = \{\emptyset\}$ and a realization of that segment as $x^T \equiv \left(x_1', ..., x_T'\right)'$.

Moments

- Moments, if they exist, can be defined in the usual way:

$$\mu_t = \mathbb{E}x_t = \int_{\mathbb{R}^m} x_t dP^T$$

$$\gamma_{tj} = \mathbb{E}x_t x_{t-j} = \int_{\mathbb{R}^{m\times(j+1)}} (x_t - \mu)\left(x_{t-j} - \mu\right) dP^T$$

- If both μ_t and $\gamma_{tj}$are independent of t for all j, X is a covariance-stationary or weakly stationary process.

- Now, let us deal with X: covariance-stationary process with zero mean (the process can be always renormalized to satisfy this requirement).

$z-$Transform

- If $\sum_{j=-\infty}^{\infty} \left|\gamma_j\right| < \infty$, define the operator $g_x : C \rightarrow C$:

$$g_x(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j$$

where C is the complex set.

- This mapping is known as the *autocovariance generating function*.

- Dividing this function by 2π and evaluating it at $e^{-i\omega}$ (where ω is a real scalar), we get the *spectrum* (or *power spectrum*) of the process x_t:

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

Spectrum

- The spectrum is the Fourier Transform of the covariances of the process, divided by 2π to normalizes the integral of $s_x(\cdot)$ to 1.

- The spectrum and the autocovariances are equivalent: there is no information in one that is not presented in other.

- That does not mean having two representations of the same information is useless: some characteristics of the series, as its serial correlation are easier to grasp with the autocovariances while others as its unobserved components (as the different fluctuations that compose the series) are much easier to understand in the spectrum.

Example

- General ARMA model: $\Phi(L)\, x_t = \theta(L)\, \varepsilon_t$.

- Since $x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$, an $MA(\infty)$ representation of x_t in the lag operator $\Psi(L)$ and variance σ^2 satisfies:

$$g_x(L) = |\Psi(L)|^2 = \Psi(L)\, \Psi\left(L^{-1}\right) \sigma^2$$

- Wold's theorem assures us that we can write any stationary ARMA as an $MA(\infty)$ and then since we can always make:

$$\theta(L) = \Phi(L)\, \Psi(L)$$

we will have

$$g_x(L) = \Psi(L)\, \Psi\left(L^{-1}\right) \sigma^2 = \frac{\theta(L)\, \theta\left(L^{-1}\right)}{\Phi(L)\, \Phi\left(L^{-1}\right)} \sigma^2$$

and then the spectrum is given by:

$$s_x(\omega) = \frac{\sigma^2}{2\pi} \frac{\theta\left(e^{-i\omega j}\right) \theta\left(e^{i\omega j}\right)}{\Phi\left(e^{-i\omega j}\right) \Phi\left(e^{i\omega j}\right)} e^{-i\omega j}$$

Working on the Spectrum

- Since $\gamma_j = \gamma_{-j}$:

$$s_x(\omega) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j e^{-i\omega j} = \frac{1}{2\pi}\left(\gamma_0 + \sum_{j=1}^{\infty}\gamma_j\left[e^{i\omega j} + e^{-i\omega j}\right]\right)$$

$$= \frac{1}{2\pi}\left(\gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j \cos(\omega j)\right)$$

by Euler's Formula ($\frac{e^{iz}+e^{-iz}}{2} = \cos z$)

- Hence:

1. The spectrum is always real-valued.
2. It is the sum of an infinite number of cosines.
3. Since $\cos(\omega) = \cos(-\omega) = \cos(\omega + 2\pi k)$, $k = 0, \pm 1, \pm 2, ...$, the spectrum is symmetric around 0 and all the relevant information is concentrated in the interval $[0, \pi]$.

Spectral Density Function

.

- Sometimes the autocovariance generating function is replaced by the autocorrelation generating function (where every term is divide by γ_0).

- Then, we get the *spectral density function*:

$$s'_x(\omega) = \frac{1}{2\pi}\left(1 + 2\sum_{j=1}^{\infty} \rho_j \cos(\omega j)\right)$$

where $\rho_j = \gamma_j / \gamma_0$.

Properties of Periodic Functions

.

- Take the modified cosine function:

$$y_j = A \cos(\omega j - \theta)$$

- ω (measured in radians): *angular frequency* (or *harmonic* or simply the *frequency*).

- $2\pi/\omega$: *period* or whole cycle of the function.

- A: *amplitude* or range of the function.

- θ: *phase* or how much the function is translated in the horizontal axis.

Different Periods

- $\omega = 0$, the period of fluctuation is infinite, i.e. the frequency associated with a trend (stochastic or deterministic).

- $\omega = \pi$ (the Nyquist frequency), the period is 2 units of time, the minimal possible observation of a fluctuation.

- Business cycle fluctuations: usually defined as fluctuations between 6 and 32 quarters. Hence, the frequencies of interest are those comprised between $\pi/3$ and $\pi/16$.

- We can have cycles with higher frequency than π. When sampling in not continuous these higher frequencies will be imputed to the frequencies between 0 and π. This phenomenon is known as *aliasing*.

Inversion

- Using the inversion formula, we can find all the covariances from the spectrum:

$$\int_{-\pi}^{\pi} s_x(\omega)\, e^{i\omega j} d\omega = \int_{-\pi}^{\pi} s_x(\omega) \cos(\omega j)\, d\omega = \gamma_j$$

- When $j = 0$, the variance of the series is:

$$\int_{-\pi}^{\pi} s_x(\omega)\, d\omega = \gamma_0$$

- Alternative interpretation of the spectrum: integral between $[-\pi, \pi]$ of it is the variance of the series.

- In general, for some $\omega_1 \geq -\pi$ and and $\omega_2 \leq \pi$:

$$\int_{\omega_1}^{\omega_2} s_x(\omega)\, d\omega$$

is the variance associated with frequencies in the $[\omega_1, \omega_2]$

- Intuitive interpretation: decomposition of the variances of the process.

Spectral Representation Theorem

- Any stationary process can be written in its *Cramér's Representation*:

$$x_t = \int_0^{\pi} u(\omega) \cos \omega t d\omega + \int_0^{\pi} v(\omega) \sin \omega t d\omega$$

- A more general representation is given by:

$$x_t = \int_{-\pi}^{\pi} \xi(\omega) d\omega$$

where $\xi(\cdot)$ is any function that satisfies an orthogonality condition:

$$\int_{-\pi}^{\pi} \xi(\omega_1) \xi(\omega_2)' d\omega_1 d\omega_2 = 0 \text{ for } \omega_1 \neq \omega_2$$

Relating Two Time Series

- Take a zero mean, covariance stationary random process Y with realization y_t and project it against the process X, with realization x_t:

$$y_t = B(L)\, x_t + \varepsilon_t$$

where

$$B(L) = \sum_{j=-\infty}^{\infty} b_j L^j$$

and $Ex_t\varepsilon_t = 0$ for all j.

- Adapting our previous notation of covariances to distinguish between different stationary processes we define:

$$\gamma_j^l = \mathbb{E} l_t l_{t-j}$$

- Thus:

$$
\begin{aligned}
y_t y_{t-j} &= \left(\sum_{s=-\infty}^{\infty} b_s x_{t-s}\right)\left(\sum_{r=-\infty}^{\infty} b_r x_{t-j-r}\right) \\
&\quad + \left(\sum_{s=-\infty}^{\infty} b_s x_{t-s}\right)\varepsilon_{t-j} + \left(\sum_{r=-\infty}^{\infty} b_r x_{t-j-r}\right)\varepsilon_t + \varepsilon_t \varepsilon_{t-j}
\end{aligned}
$$

- Taking expected values of both sides, the orthogonality principle implies that:

$$
\gamma_j^y = \sum_{s=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} b_s b_r \gamma_{j+r-s}^x + \gamma_j^\varepsilon
$$

- With these covariances, the computation of the spectrum of y is direct:

$$
s_y(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j^y e^{-i\omega j} = \frac{1}{2\pi}\left(\sum_{j,s,r=-\infty}^{\infty} b_r b_s \gamma_{j+r-s}^x e^{-i\omega j} + \sum_{j=-\infty}^{\infty} \gamma_j^\varepsilon e^{-i\omega j}\right)
$$

- If we define $h = j + r - s$:

$$e^{-i\omega j} = e^{-i\omega(h-r+s)} = e^{-i\omega h}e^{-i\omega s}e^{i\omega r}$$

we get:

$$\begin{aligned} s_y(\omega) &= \sum_{r=-\infty}^{\infty} b_r e^{i\omega r} \sum_{s=-\infty}^{\infty} b_s e^{-i\omega s} \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} \gamma_h^x e^{-i\omega h} + \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j^{\varepsilon} e^{-i\omega j} \\ &= B\left(e^{i\omega r}\right) B\left(e^{-i\omega s}\right) s_x(\omega) + s_{\varepsilon}(\omega) \end{aligned}$$

- Using the symmetry of the spectrum:

$$s_y(\omega) = \left|B\left(e^{-i\omega}\right)\right|^2 s_x(\omega) + s_{\varepsilon}(\omega)$$

LTI Filters

- Given a process X, a *linear time-invariant filter (LTI-filter)* is a operator from the space of sequences into itself that generates a new process Y of the form:

$$y_t = B(L)\, x_t$$

- Since the transformation is deterministic, $s_{\varepsilon}(\omega) = 0$ and we get:

$$s_y(\omega) = \left| B\left(e^{-i\omega}\right) \right|^2 s_x(\omega)$$

- $B\left(e^{-i\omega}\right)$: *frequency transform* or the *frequency response function* is the Fourier transform of the coefficients of the lag operator.

- $G(\omega) = \left| B\left(e^{-i\omega}\right) \right|$ is the *gain* (its modulus).

- $\left| B\left(e^{-i\omega}\right) \right|^2$ is the *power transfer function* (since $\left| B\left(e^{-i\omega}\right) \right|^2$ is a quadratic form, it is a real function). It indicates how much the spectrum of the series is changed at each particular frequency.

Gain and Phase

- The definition of gain implies that the filtered series has zero variance at $\omega = 0$ if and only if $B\left(e^{-i0}\right) .= \sum_{j=-\infty}^{\infty} b_j e^{-i0} = \sum_{j=-\infty}^{\infty} b_j = 0$.

- Since the gain is a complex mapping, we can write:

$$B\left(e^{-i\omega}\right) = B^a(\omega) + iB^b(\omega)$$

where $B^a(\omega)$ and B^b are both real functions.

- Then we define the *phase* of the filter

$$\phi(\omega) = \tan^{-1}\left(-\frac{B^b(\omega)}{B^a(\omega)}\right)$$

The phase measures how much the series changes its position with respect to time when the filter is applied. For a given $\phi(\omega)$, the filter shifts the series by $\phi(\omega)/\omega$ time units.

Symmetric Filters I

- A filter is symmetric if $b_j = b_{-j}$:

$$B(L) = b_0 + \sum_{j=1}^{\infty} b_j \left(L^j + L^{-j}\right)$$

- Symmetric filters are important for two reasons:

1. They do not induce a phase shift since the Fourier transform of $B(L)$ will always be a real function.

2. Corners of the a $T - segment$ of the series are difficult to deal with since the lag operator can only be applied to one side.

Symmetric Filters II

If $\sum_{j=-\infty}^{\infty} b_j = 0$, symmetric filters can remove trend frequency components, either deterministic or stochastic up to second order (a quadratic deterministic trend or a double unit root):

$$\begin{aligned}
\sum_{j=-\infty}^{\infty} b_j L^j &= \sum_{j=1}^{\infty} b_j \left(L^j + L^{-j}\right) - 2 \sum_{j=-\infty}^{\infty} b_j L^j \\
&= \sum_{j=1}^{\infty} b_j \left(L^j + L^{-j} - 2\right) = - \sum_{j=1}^{\infty} b_j \left(\left(1 - L^j\right)\left(1 - L^{-j}\right)\right) \\
&= - (1 - L)\left(1 - L^{-1}\right) \sum_{j=1}^{\infty} b_j \sum_{h=-j+1}^{j-1} (k - |h|) L^h \\
&= (1 - L)\left(1 - L^{-1}\right) B'(L)
\end{aligned}$$

where we used $\left(1 - L^j\right) = (1 - L)\left(1 + L + ... + L^{j-1}\right)$ and

$$(1 + L + ...)\left(1 + L^{-1} + ...\right) = \sum_{j=1}^{\infty} b_j \sum_{h=-j+1}^{j-1} (k - |h|) L^h$$

if the sum is well defined.

Ideal Filters

- An ideal filter is an operator $B(L)$ such that the new process Y only has positive spectrum in some specified part of the domain.

- Example: a band-pass filter for the interval $\{(a, b) \cup (-b, -a)\} \in (-\pi, \pi)$ and $0 < a \leq b \leq \pi$, we need to choose $B\left(e^{-i\omega}\right)$ such that:

$$B\left(e^{-i\omega}\right) = \begin{cases} = 1, \text{ for } \omega \in (a, b) \cup (-b, -a), \\ \quad = 0 \text{ otherwise} \end{cases}$$

Since $a > 0$, this definition implies that $B(0) = 0$. Thus, a band-pass filter shuts off all the frequencies outside the region (a, b) or $(-b, -a)$ and leaves a new process that only fluctuates in the area of interest.

How Do We Build an Ideal Filter?

- Since X is a zero mean, covariance stationary process, $\sum_{n=-\infty}^{\infty} |x_n|^2 < \infty$, the Riesz-Fischer Theorem holds.

- If we make $f(\omega) = B\left(e^{-i\omega}\right)$, we can use the *inversion formula* to set the B_j:

$$b_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} B\left(e^{-i\omega}\right) e^{i\omega j} d\omega$$

- Substituting $B\left(e^{-i\omega}\right)$ by its value:

$$\begin{aligned} b_j &= \frac{1}{2\pi} \int_{-\pi}^{\pi} B\left(e^{-i\omega}\right) e^{i\omega j} d\omega = \frac{1}{2\pi} \left(\int_{-b}^{-a} e^{i\omega j} d\omega + \int_{a}^{b} e^{i\omega j} d\omega \right) \\ &= \frac{1}{2\pi} \int_{a}^{b} \left(e^{i\omega j} + e^{-i\omega j} \right) d\omega \end{aligned}$$

where the second step just follows from $\int_{-b}^{-a} e^{i\omega j} d\omega = \int_{a}^{b} e^{-i\omega j} d\omega$.

Building an Ideal Filter

- Now, using again Euler's Formula,

$$b_j = \frac{1}{2\pi}\int_a^b 2\cos\omega j d\omega = \frac{1}{\pi j}\sin.\omega j|_a^b = \frac{\sin jb - \sin ja}{\pi j} \ \forall \ j \in \mathbb{N} \setminus \{0\}$$

- Since $\sin z = -\sin(-z)$,

$$b_{-j} = \frac{\sin(-jb) - \sin(-ja)}{-\pi j} = \frac{-\sin jb + \sin ja}{-\pi j} = \frac{\sin jb - \sin ja}{\pi j} = b_j$$

- Also:

$$b_0 = \frac{1}{2\pi}\int_a^b 2\cos\omega 0 d\omega = \frac{b-a}{\pi}$$

and we have all the coefficients to write:

$$y_t = \left(\frac{b-a}{\pi} + \sum_{j=1}^{\infty}\frac{\sin jb - \sin ja}{\pi j}\left(L^j + L^{-j}\right)\right)x_t$$

- The coefficients b_j converge to zero at a sufficient rate to make the sum well defined in the reals.

Building an Low-Pass Filter

- Analogously, a low-pass filter allows only the low frequencies (between $(-b, b)$), implying a choice of $B^{\cdot}\left(e^{-i\omega}\right)$ such that:

$$B\left(e^{-i\omega}\right) = \begin{cases} = 1, & \text{for } \omega \in (-b, b)\,, \\ = 0 \text{ otherwise} \end{cases}$$

- Just set $a = 0$.

$$\begin{aligned} b_j &= b_{-j} = \frac{\sin jb}{\pi j} \; \forall \; j \in \mathbb{N} \setminus \{0\} \\ b_0 &= \frac{b}{\pi} \\ y_t &= \left(\frac{b}{\pi} + \sum_{j=1}^{\infty} \frac{\sin jb}{\pi j} \left(L^j + L^{-j} \right) \right) x_t \end{aligned}$$

Building an High-Pass Filter

- Finally a high-pass allows only the high frequencies:

$$B\left(e^{-i\omega}\right) = \begin{cases} = 1, \text{ for } \omega \in (b, \pi) \cup (-\pi, -b), \\ = 0 \text{ otherwise} \end{cases}$$

- This is just the complement of a low-pass $(-b, b)$:

$$y_t = \left(1 - \frac{b}{\pi} - \sum_{j=1}^{\infty} \frac{\sin jb}{\pi j} \left(L^j + L^{-j}\right)\right) x_t$$

Finite Sample Approximations

- With real, finite, data, it is not possible to apply any of the previous formulae since they require an infinite amount of observations.

- Finite sample approximations are this required.

- We will study two approximations:

  1. the Hodrick-Prescott filter
  2. the Baxter-King filters.

- We will be concern with the minimization of several problems:

  1. *Leakage*: the filter passes frequencies that it was designed to eliminate.
  2. *Compression*: the filter is less than one at the desired frequencies.
  3. *Exacerbation*: the filter is more than one at the desired frequencies.

HP Filter

- Hodrick-Prescott (HP) filter:
  1. It is a coarse and relatively uncontroversial procedure to represent the behavior of macroeconomic variables and their comovements.
  2. It provides a benchmark of regularities to evaluate the comparative performance of different models.

- Suppose that we have T observations of the stochastic process X, $\{x_t\}_{t=1}^{t=\infty}$.

- HP decomposes the observations into the sum of a trend component, x_t^t and a cyclical component x_t^c:

$$x_t = x_t^t + x_t^c$$

Minimization Problem

- How? Solve:

$$\min_{x_t^t} \sum_{t=1}^{T} \left(x_t - x_t^t\right)^2 + \lambda \sum_{t=2}^{T-1} \left[\left(x_{t+1}^t - x_t^t\right) - \left(x_t^t - x_{t-1}^t\right)\right]^2$$

- Intuition.

- Meaning of λ:

1. $\lambda = 0 \Rightarrow$trivial solution ($x_t^t = x_t$).

2. $\lambda = \infty \Rightarrow$linear trend.

Matrix Notation

- To compute the HP filter is easier to use matrix notation, and rewrite minimization problem as:

$$\min_{x^t} \left(x - x^t\right)' \left(x - x^t\right) + \lambda \left(Ax^t\right)' \left(Ax^t\right)$$

where $x = (x_1, ..., x_T)'$, $x^t = \left(x_1^t, ..., x_T^t\right)'$ and:

$$A = \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 1 & -2 & 1 & 0 \\ 0 & \cdots & \cdots & \cdots & 1 & -2 & 1 \end{pmatrix}_{(T-2)\times T}$$

Solution

- First order condition:

$$x^t - x + \lambda A'Ax^t = 0$$

or

$$x^t = \left(I + \lambda A'A\right)^{-1} x$$

- $(I + \lambda A'A)^{-1}$ is a sparse matrix (with density factor $(5T - 6)/T^2$). We can exploit this property.

Properties I

- To study the properties of the HP filter is however more convenient to stick with the original notation.

- We write the first-order condition of the minimization problem as:

$$\begin{aligned} 0 &= -2\left(x_t - x_t^t\right) + 2\lambda\left[\left(x_t^t - x_{t-1}^t\right) - \left(x_{t-1}^t - x_{t-2}^t\right)\right] - \\ & 4\lambda\left[\left(x_{t+1}^t - x_t^t\right) - \left(x_t^t - x_{t-1}^t\right)\right] + 2\lambda\left[\left(x_{t+2}^t - x_{t+1}^t\right) - \left(x_{t+1}^t - x_t^t\right)\right] \end{aligned}$$

- Now, grouping coefficients and using the lag operator L:

$$\begin{aligned} x_t &= x_t^t + \left(\lambda\left[1 - 2L + L^2\right] - 2\lambda\left[L^{-1} - 2 + L\right] + \lambda\left[L^{-2} - 2L^{-1} + 1\right]\right) x_t^t \\ &= \left[\lambda L^{-2} - 4\lambda L^{-1} + (6\lambda + 1) - 4\lambda L + \lambda L^2\right] \\ &= \left[1 + \lambda(1 - L)^2\left(1 - L^{-1}\right)^2\right] = F(L)\, x_t^t \end{aligned}$$

Properties II

- Define the operator $C(L)$ as:

$$C(L) = (F(L) - 1)\left(F(L)^{-1}\right) = \frac{\lambda (1-L)^2 \left(1-L^{-1}\right)^2}{1+\lambda (1-L)^2 \left(1-L^{-1}\right)^2}$$

- Now, if we let, for convenience, $T = \infty$, and since

$$\begin{aligned} x_t^t &= B(L)\, x_t \\ x_t^c &= (1 - B(L))\, x_t \end{aligned}$$

we can see that

$$B(L) = F(L)^{-1}$$
$$1 - B(L) = 1 - F(L)^{-1} = (F(L) - 1)\left(F(L)^{-1}\right) = C(L)$$

i.e., the cyclical component x_t^c is equal to:

$$x_t^c = \frac{\lambda (1-L)^2 \left(1-L^{-1}\right)^2}{1+\lambda (1-L)^2 \left(1-L^{-1}\right)^2} x_t$$

a stationary process if x_t is integrated up to fourth order.

Properties III

- Remember that:

$$s_{x^c}(\omega) = \left|1 - B\left(e^{-i\omega}\right)\right|^2 s_x(\omega) = \left|C\left(e^{-i\omega}\right)\right|^2 s_x(\omega)$$

- We can evaluate at $e^{-i\omega}$ and taking the module, the gain of the cyclical component is:

$$\begin{aligned} G^c(\omega) &= \left|\frac{\lambda\left(1-e^{-i\omega}\right)^2\left(1-e^{i\omega}\right)^2}{1+\lambda\left(1-e^{-i\omega}\right)^2\left(1-e^{i\omega}\right)^2}\right| \\ &= \frac{4\lambda\left(1-\cos(\omega)\right)^2}{1+4\lambda\left(1-\cos(\omega)\right)^2} \end{aligned}$$

  where we use the identity $\left(1-e^{-i\omega}\right)\left(1-e^{i\omega}\right) = 2\left(1-\cos(\omega)\right)$.

- This function gives zero weight to zero frequencies and close to unity on high frequency, with increases in λ moving the curve to the left.

- Finally note that since the gain is real, $\phi(\omega) = 0$ and the series is not translated in time.

Butterworth Filters

- Filters with gain:

$$G(\omega) = \left[1 + \left(\frac{\sin\left(\frac{\omega}{2}\right)}{\sin\left(\frac{\omega_0}{2}\right)}\right)^{2d}\right]^{-1}$$

  that depends on two parameters, d, a positive integer, and ω_0, that sets the frequency for which the gain is 0.5.

- Higher values of d make the slope of the band more vertical while smaller values of ω_0 narrow the width of the filter band.

Butterworth Filters and the HP Filter

- Using the fact that $1 - \cos(\omega) = 2\sin^2\left(\frac{\omega}{2}\right)$, the HP gain is:

$$G^c(\omega) = 1 - \frac{1}{1 + 16\lambda \sin^4\left(\frac{\omega}{2}\right)}$$

- Since $B(L) = 1 - C(L)$, the gain of the trend component is just:

$$G^{x_t^t}(\omega) = 1 - G^c(\omega) = \frac{1}{1 + 16\lambda \sin^4\left(\frac{\omega}{2}\right)}$$

a particular case of a Butterworth filter of the sine version with $d = 2$ and $\omega_0 = 2\arcsin\left(\frac{1}{2\lambda^{0.25}}\right)$.

Ideal Filter Revisited

- Recall that in our discussion about the ideal filter $y_t = \sum_{j=-\infty}^{\infty} b_j L^j x_t$ we derived the formulae:

$$
\begin{aligned}
b_j &= b_{-j} = \frac{\sin jb - \sin ja}{\pi j} \; \forall \; j \in \mathbb{N} \\
b_0 &= \frac{b}{\pi}
\end{aligned}
$$

  for the frequencies $a = \frac{2\pi}{p_u}$ and $b = \frac{2\pi}{p_l}$ where $2 \leq p_l < p_u < \infty$ are the lengths of the fluctuations of interest.

- Although this expressions cannot we evaluate for all integers, they suggest the feasibility of a more direct approach to band-pass filtering using some form of these coefficients.

Baxter and King (1999)

- Take the first k coefficients of the above expression.

.

- Then, since we want a gain that implies zero variance at $\omega = 0$, we normalize each coefficient as:

$$\widetilde{b_j} = b_j - \frac{b_0 + 2\sum_{j=1}^{k} b_j}{2k+1}$$

to get that $\sum_{j=-k}^{k} \widetilde{b_j} = 0$ as originally proposed by Craddock (1957).

- This strategy is motivated by the remarkable result that the solution to the problem of minimizing

$$Q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| B\left(e^{-i\omega}\right) - B_k\left(e^{-i\omega}\right) \right|^2 d\omega$$

where $B(\cdot)$ is the ideal band-pass filter and $B_k(\cdot)$ is its k-approximation, is simply to take the first k coefficients b_j from the inversion formula and to make all the higher order coefficients zero.

- To see that, write the minimization problem s.t. $\sum_{j=-k}^{k} \widetilde{b_j} = 0$ as:

$$\mathcal{L} = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \boldsymbol{\Psi}\left(\omega, \widetilde{b}\right) \boldsymbol{\Psi}\left(\omega, \widetilde{b}\right)' d\omega + \lambda \sum_{j=-k}^{k} \widetilde{b_j}$$

where

$$\boldsymbol{\Psi}\left(\omega, \widetilde{b}\right) = \left[B\left(e^{-i\omega}\right) - \sum_{j=-\infty}^{\infty} \widetilde{b}_j e^{-i\omega j}\right]$$

- First order conditions:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\omega j}\boldsymbol{\Psi}\left(\omega, \widetilde{b}\right)' + \boldsymbol{\Psi}\left(\omega, \widetilde{b}\right)' e^{i\omega j} d\omega = -\lambda$$

$$0 = \sum_{j=-k}^{k} \widetilde{b_j}$$

for $j = 1, ..., k$.

- Since the harmonics $e^{-i\omega j}$ are orthonormal

$$
\begin{aligned}
0 &= 2\left(b_j - \widetilde{b_j}\right) + \lambda \\
\lambda &= 2 - \frac{\sum_{j=-k}^{k} b_j}{2k+1}
\end{aligned}
$$

for $j = 1, ..., k$, that provides the desired result.

- Notice that, since the ideal filter is a step function, this approach suffers from the same *Gibbs phenomenon* that arises in Fourier analysis (Koopmans, 1974).