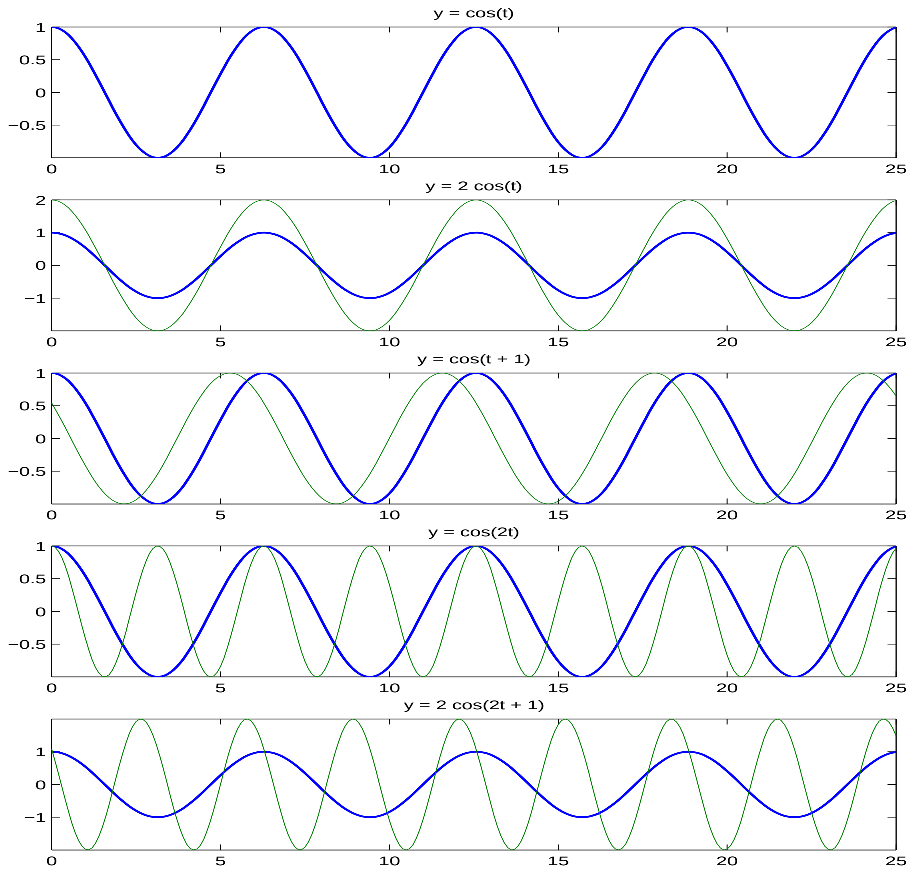
Lecture notes on Spectral Analysis

Luca Sala
IGIER, Università Bocconi
January 2007


y = cos(t)
y = 2 cos(t)
y = cos(t + 1)
y = cos(2t)
y = 2 cos(2t + 1)
0
5
10
15
20
25

Trigonometry refresh

Cosinus wave: $Y_t = A\cos(\omega t + \alpha)$

- A: amplitude

- α: phase (position of Y_t when $t = 0$)

- ω: frequency (how quickly the process obscillates w.r.t $\cos t$)

- Relation between period and frequency: $Period = 2\pi/\omega$

- $e^{it} = \cos t + i\sin t$ (limit of Taylor expansion)

- Complex number: $a + bi$

- Polar coordinates: $a + bi = Re^{i\theta} = R(\cos\theta + i\sin\theta)$
where: $R = |a + bi| = (a^2 + b^2)^{1/2}$ and $\theta = \arcsin a/R$

Fourier Theory

Any vector can be expressed as a linear combination of elements of an orthonormal basis

i.e.

$$\begin{pmatrix} 3 \\ 6 \\ 2 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 6 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Any periodic sequence of period T can be expressed as a linear combination of T elements of the orthogonal basis: $\{e^{it\lambda_j}\}$ with $\lambda_j = 2\pi j/T$, $-\pi < \lambda_j \leq \pi$

$$s_t = \sum_j g_j e^{it\lambda_j}$$

- $e^{it} = \cos t + i \sin t$: s_t is expressed as a combination of waves.

- The coefficients g_j are the coefficients of the projection of s_t on the functions $\{e^{it\lambda_j}\}$ (there is no residual as we are projecting T points on T regressors):

Example with $T = 2$ - the set of λ_j is: $[0, \pi]$

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} e^{i0} & e^{i\pi} \\ e^{2i0} & e^{2i\pi} \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \end{bmatrix} + \epsilon_t$$
$$= eg + \epsilon$$

Projecting s_t on e to recover g:

$$g = (e'e)^{-1}e's_t$$

where $e'e = I_T T$, we obtain:

$$g_j = \frac{1}{T}\sum_{r=1}^{T} e^{-ir\lambda_j} s_r \tag{1}$$

so that:

$$s_t = \sum_j \underbrace{\left[\frac{1}{T}\sum_{r=1}^{T} e^{-ir\lambda_j} s_r\right]}_{g_j} e^{it\lambda_j}$$

s_t is expressed as a sum of *rectangles* with base g_j and height $e^{it\lambda_j}$

- If T is finite, $\sum_{-\infty}^{\infty} |s_t| = \infty$

As the period $T \to \infty$, s_t can be expressed as a linear combination of an infinite number of elements of the orthogonal basis: $\{e^{it\lambda}\}$.

When $T \to \infty$, $\lambda_j - \lambda_{j-1} = 2\pi/T \to 0$.

If $\sum |s_t| < \infty$, s_t can be approximated by an integral:

$$s_t = \int_{-\pi}^{\pi} g(\lambda) e^{it\lambda} d\lambda$$

in which $g(\lambda)$ is a continuous function of λ.

The integral is called the *Fourier integral* of s_t

Fourier theory is all about finding and interpreting the coefficients g_j or the function $g(\lambda)$.

In the finite T case, the sequence is approximated using a **discrete set** of frequencies, while in the infinite T case the sequence is approximated using a **continuum of frequencies** from $-\pi$ to π.

It is possible to combine the two cases, T finite and T infinite, in one expression:

$$s_t = \int_{-\pi}^{\pi} e^{it\lambda} dG(\lambda)$$

This is called the *Fourier-Stieltjes integral* of s_t

- If T is infinite, s_t can be expressed as a linear combination over a continuous set of frequencies. If $\sum_j |s_j| < \infty$, $G(\lambda)$ will be differentiable, so that:

$$\frac{dG(\lambda)}{d\lambda} = g(\lambda)$$

and:

$$s_t = \int_{-\pi}^{\pi} e^{it\lambda} g(\lambda) d\lambda$$

In this case, $g(\lambda)$ can be obtained from (1)

$$g(\lambda) = \lim_{T\to\infty} \frac{dG(\lambda)}{\lambda_{j+1} - \lambda_j} = \lim_{T\to\infty} \frac{1/T \sum_{r=1}^{T} e^{-ir\lambda} s_r}{2\pi/T} = \frac{1}{2\pi} \sum_{r=1}^{+\infty} e^{-ir\lambda} s_r$$

This is the *Fourier transform* of the sequence s_t.

- If T is finite, as we saw above, s_t can be expressed as a linear combination over a discrete set of frequencies and $dG(\lambda)$ will be:

$$dG(\lambda) = \begin{cases} g_j & \text{if } \lambda = \lambda_j \\ 0 & \text{elsewhere} \end{cases}$$

  In this case, the function $G(\lambda)$ will have increments only at frequencies λ_j, where it increases by g_j.

Let's apply these concepts to stochastic processes.

For any $I(0)$ process x_t there exists a deterministic sequence, $\gamma(h)$, the autocovariance function.

As we just saw, we can express it as a linear combination of waves.

Remark 1: If x_t is ARMA, then $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$.

$$\gamma(k) = \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda$$

$$\gamma(0) = \int_{-\pi}^{\pi} f(\lambda) d\lambda$$

Remark 2: If x_t is $I(0)$, but not ARMA, $\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty$

$$\gamma(k) = \int_{-\pi}^{\pi} e^{ik\lambda} dF(\lambda)$$

Example: periodic process

x_t is the combination of 2 waves with different frequencies with stochastic amplitudes:

$$x_t = \left[A_1 \cos(\frac{\pi}{2}t) + B_1 \sin(\frac{\pi}{2}t)\right] + \left[A_2 \cos(\frac{\pi}{4}t) + B_1 \sin(\frac{\pi}{4}t)\right]$$

A_1, A_2, B_1 and B_2 are orthogonal random variables, with:

$$V(A_1) = V(B_1) = \sigma_1^2 = 1$$
$$V(A_2) = V(B_2) = \sigma_2^2 = .6$$

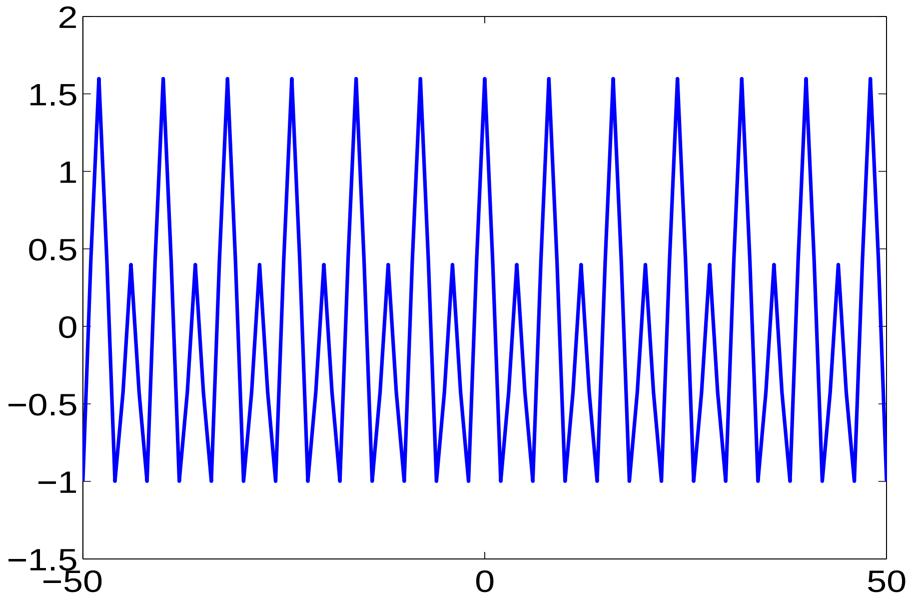
One can prove that:

$$\gamma(h) = \sigma_1^2 \cos(\frac{\pi}{2}h) + \sigma_2^2 \cos(\frac{\pi}{4}h)$$

Periodic function with period 8: $f(h) = f(h+8)$

Remark 1: As $\gamma(h)$ does not depend on t, the process is $I(0)$.

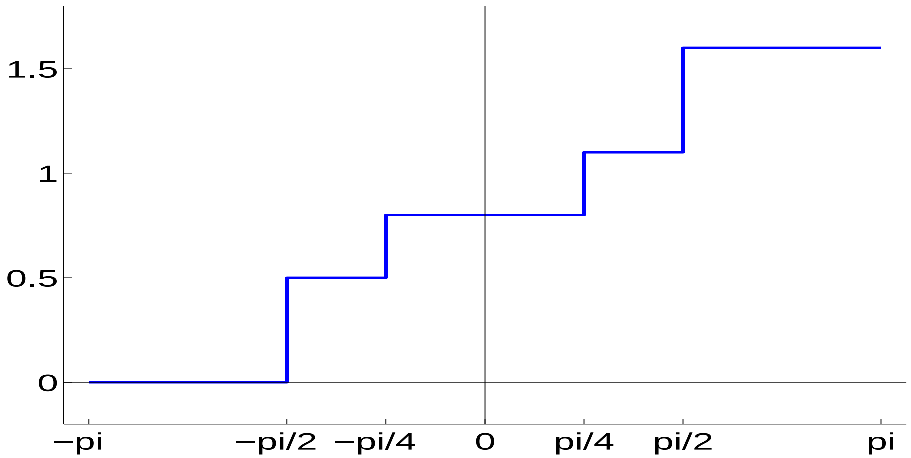
Remark 2: In this case, $\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty$



Autocovariance function of a periodic process

The function $F(\lambda)$ will take the form:

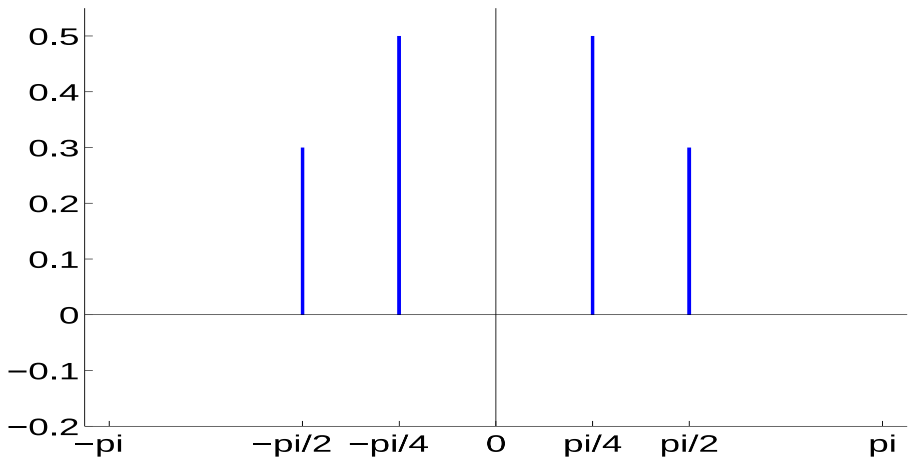
$$F(\lambda) = \begin{cases} 0 & \theta < -\frac{\pi}{2} \\ .5 & -\frac{\pi}{2} \leq \theta < -\frac{\pi}{4} \\ .8 & -\frac{\pi}{4} \leq \theta < \frac{\pi}{4} \\ 1.1 & \frac{\pi}{4} \leq \theta < \frac{\pi}{2} \\ 1.6 & \theta \geq \frac{\pi}{2} \end{cases}$$



$F(\lambda)$: **CDF of the autocov of a combination of waves**

$$dF(\lambda) = \begin{cases} .5 & \theta = -\frac{\pi}{2} \\ .3 & \theta = -\frac{\pi}{4} \\ .3 & \theta = \frac{\pi}{4} \\ .5 & \theta = \frac{\pi}{2} \\ 0 & \text{elsewhere} \end{cases}$$

The derivative $\frac{dF(\lambda)}{d\lambda}$ will be ∞ at $\lambda = \pm\frac{\pi}{2}$ and $\lambda = \pm\frac{\pi}{4}$



$dF(\lambda)$: **the "discrete probability function" of a combination of waves**

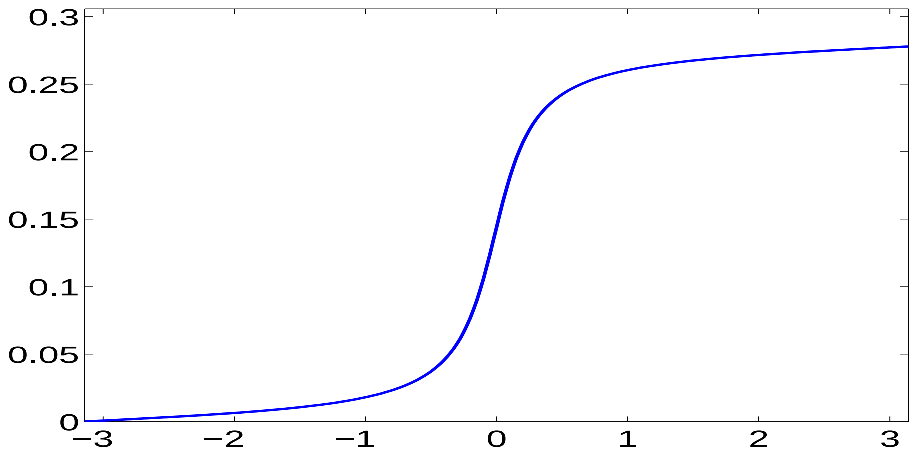
From $\gamma(h)$ and $F(\lambda)$, it is easy to verify that two properties hold:

$$\begin{aligned}\gamma(h) &= \int_{-\pi}^{\pi} e^{i\lambda h} dF(\lambda) = \sigma_1^2 \cos(\frac{\pi}{2}h) + \sigma_2^2 \cos(\frac{\pi}{4}h) \\ \gamma(0) &= \int_{-\pi}^{\pi} dF(\lambda) = \sigma_1^2 + \sigma_2^2\end{aligned}$$

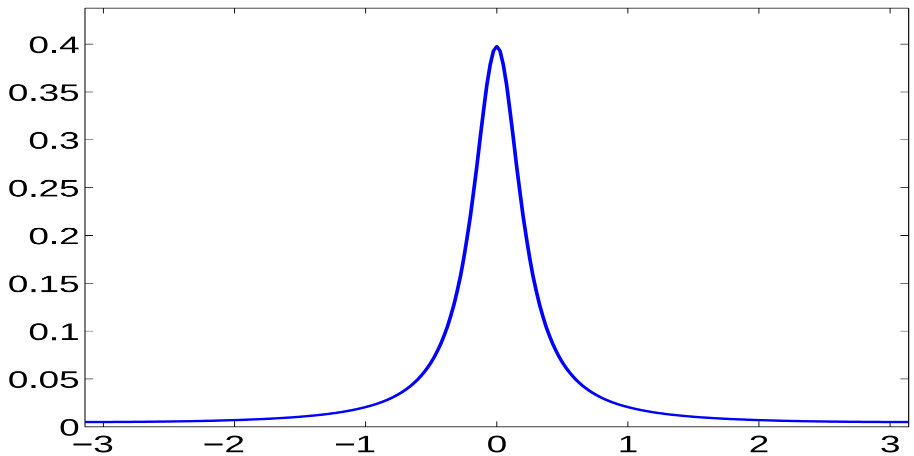
Important: the function $dF(\lambda)$ have peaks at those frequencies at which the process obscillates and the peaks are proportional to the amplitude (variance) of each wave.

This object is well-suited to identify cyclical components

In the case in which $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ a possible shape for the functions $F(\lambda)$ and $f(\lambda)$ is below:



$F(\lambda)$: **CDF of the autocov of a non-periodic process**



$f(\lambda) = \frac{dF(\lambda)}{d\lambda}$**: "density" of the autocov of a non-periodic process**

In the continuous case, the function $f(\lambda)$ will be large in correspondence of those frequencies which are important in explaining the variance of the process (more on this below).

The function $f(\lambda)$ is the *spectral density* of the process x_t with autocovariance $\gamma(h)$:

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda \tag{2}$$

Given a $f(\lambda)$ it is possible to obtain the associated $\gamma(h)$.

The Fourier Transform:

We saw above that it is also possible to go the other way around: from $\gamma(h)$ obtain $f(\lambda)$

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-ij\lambda}$$

The operation above delivers the *Fourier transform* of the sequence $\gamma(h)$: the spectral density $f(\lambda)$ is the *Fourier transform* of the sequence $\gamma(h)$.

Let us prove that if we plug the equation above for $f(\lambda)$ in the right hand side of (2), we obtain $\gamma(h)$.

$$\begin{aligned}
\gamma(h) &= \int_{-\pi}^{\pi} \underbrace{\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)e^{-ij\lambda}}_{f(\lambda)} e^{ih\lambda} d\lambda \\
&= \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)\int_{-\pi}^{\pi} e^{-ij\lambda}e^{ih\lambda}d\lambda
\end{aligned}$$

Let's evaluate the integral:

$$\text{If } h = j : \int_{-\pi}^{\pi} e^{-ij\mu}e^{ij\mu}d\mu = 2\pi$$

$$\text{If } h \neq j : \int_{-\pi}^{\pi} e^{-ij\mu}e^{ij\mu}d\mu = 0$$

Indeed, the right-hand side is $\gamma(h)$

Note that:

$$\gamma(0) = \int_{-\pi}^{\pi} f(\lambda)d\lambda$$

The spectral density decomposes the variance of the process in waves of different frequency.

The Spectral Representation Theorem

Not only the autocovariance function of any $I(0)$ process can be expressed as a weighted sum of waves, but also any $I(0)$ stochastic process can be expressed as a weighted sum of waves in which the amplitude of each wave is stochastic.

Basically, it is the limit case of our previous example:

$$x_t = \left[A_1 \cos(\frac{\pi}{2}t) + B_1 \sin(\frac{\pi}{2}t)\right] + \left[A_2 \cos(\frac{\pi}{4}t) + B_1 \sin(\frac{\pi}{4}t)\right]$$

$$x_t = \int_{-\infty}^{\infty} e^{-i\lambda t} dZ(\lambda)$$

The integral above is computed with respect to a stochastic measure: $dZ(\lambda)$ is an "orthogonal increments process".

Example: White Noise

The autocovariance function is: $\gamma(0) = \sigma^2$, $\gamma(k) = 0, k = \pm 1, \pm 2, ...$

From the definition of the spectral density:

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-ij\lambda} = \frac{\sigma^2}{2\pi}$$

All the frequencies contribute with the same weight to the variance of the process.

Filtering

Let's start from the following equation:

$$X_t = \sum_{j=-\infty}^{+\infty} A_j Y_{t-j} = A(L)Y_t$$

Y_t is the **input** process

$A(L)$ is the (linear and time-invariant) **filter**

X_t is the **output** process

The operation $X_t = A(L)Y_t$ is a convolution and can be very complicated. Its frequency domain counterpart will turn out to be very easy!!

Important result: the spectral density of the output Y_t

Assume that Y_t is $I(0)$, with spectral density $f_Y(\lambda)$. Then,

$$f_X(\lambda) = \left| \sum_{j=-\infty}^{+\infty} A_j e^{-ij\lambda} \right|^2 f_Y(\lambda) = \left| A(e^{-i\lambda}) \right|^2 f_Y(\lambda)$$

The spectral density of Y_t can be obtained by a simple multiplication.

A complicated convolution becomes a multiplication.

$$\begin{array}{rcl} A(e^{-i\lambda}) & : & \text{transfer function} \\ \left|A(e^{-i\lambda})\right| & : & \text{amplitude function} \\ \left|A(e^{-i\lambda})\right|^2 & : & \text{power function} \end{array}$$

The variance at each frequency is modified by $\left|A(e^{-i\lambda})\right|^2$

In polar form:

$$A(e^{-i\lambda}) = \underbrace{B(\lambda)}_{\text{amplitude: } \left|A(e^{-i\lambda})\right|} \; e^{i\overbrace{\phi(\lambda)}^{\text{phase shift}}}$$

Computing the spectral density for ARMA is straightforward

Let $X_t \sim ARMA(p, q)$:

$$\phi(L)X_t = \theta(L)\epsilon_t$$

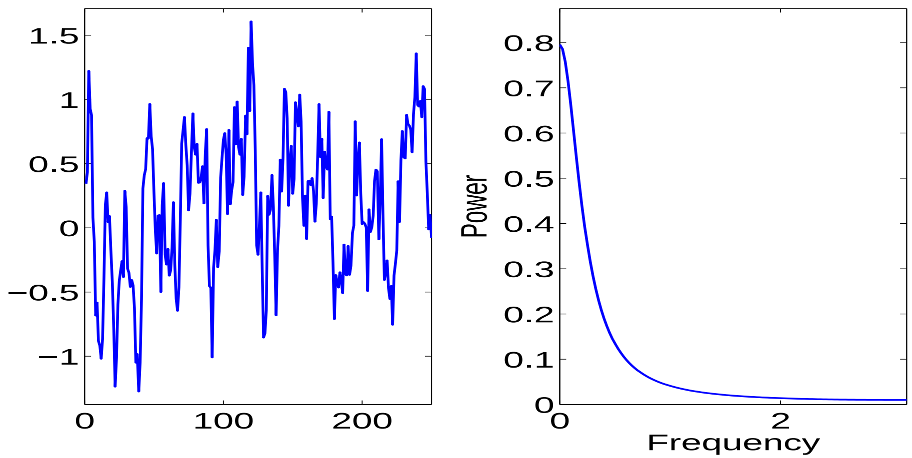
The spectral density of X_t is:

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

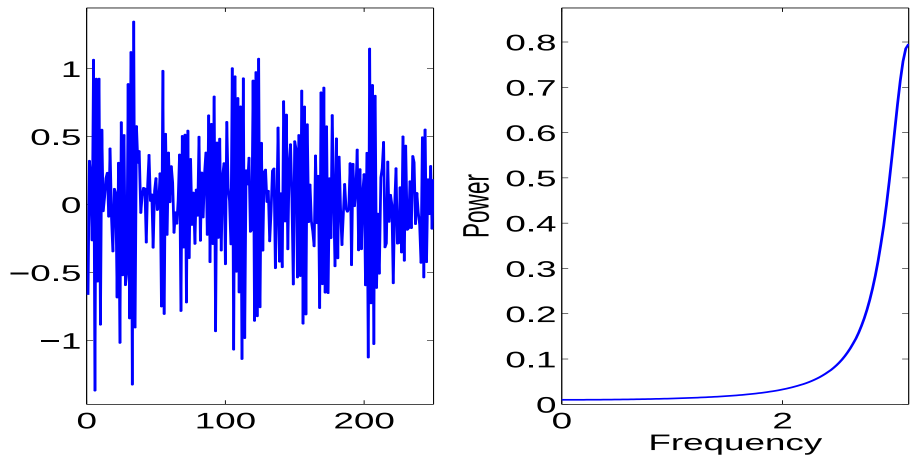
Some examples

AR(1): $X_t = \rho X_{t-1} + \eta_t$

$$f_X(\lambda) = |1 - \rho e^{-i\lambda}|^{-2} \frac{\sigma_\eta}{2\pi}$$



AR(1) with $\rho = .8$



AR(1) with $\rho = -.8$

What is the effect of a filter?

Suppose the input process is a pure cosinus wave:

$$X_t = 2\cos\lambda t = e^{i\lambda t} + e^{-i\lambda t}$$

We apply a filter $h(L)$, so that the output is: $Y_t = h(L)X_t$:

$$\begin{aligned}
Y_t = h(L)\left(e^{i\lambda t} + e^{-i\lambda t}\right) &= \\
\sum_{j=-\infty}^{+\infty} h_j \left(e^{i\lambda(t-j)} + e^{-i\lambda(t-j)}\right) &= \\
e^{i\lambda t} \underbrace{\sum_{j=-\infty}^{+\infty} h_j e^{-i\lambda j}}_{s(\lambda)e^{i\phi(\lambda)}} + e^{-i\lambda t} \underbrace{\sum_{j=-\infty}^{+\infty} h_j e^{i\lambda j}}_{s(\lambda)e^{-i\phi(\lambda)}} &= \\
s(\lambda)\left[e^{i(\lambda t+\phi(\lambda))} + e^{-i(\lambda t+\phi(\lambda))}\right] &= s(\lambda)2\cos(\lambda t + \phi(\lambda))
\end{aligned}$$

- Amplitude $s(\lambda)$

  The amplitude of the waves is modified by the function $s(\lambda)$
  $\rightarrow$ The "variance" of the wave is modified

- Phase $\phi(\lambda)$

  X_t peaks at $t = 0$, Y_t peaks at: $\lambda t + \phi(\lambda) = 0$, that is, at $t = \frac{-\phi(\lambda)}{\lambda}$ units of time.

  If $\phi(\lambda) > 0$, output leads input, if $\phi(\lambda) < 0$, output lags input.

- Phase $\phi(\lambda)$ (2)

  From

$$\sum_{j=-\infty}^{+\infty} h_j e^{-i\lambda j} = s(\lambda) e^{i\phi(\lambda)}$$

  if the filter is symmetric ($h_j = h_{-j}$), the phase shift $\phi(\lambda) = 0$.

Examples of filters - first difference: $1 - L$

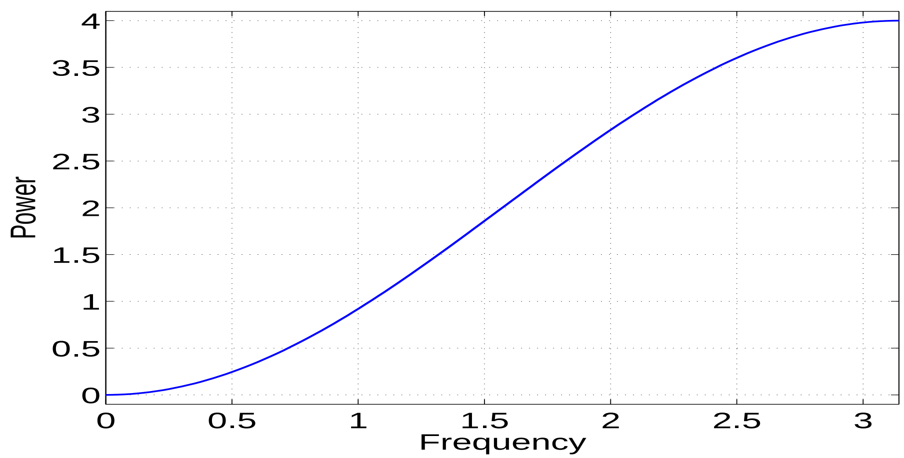
In the frequency domain:

- Transfer function: $1 - e^{-i\lambda} = 1 - \cos\lambda + i\sin\lambda$

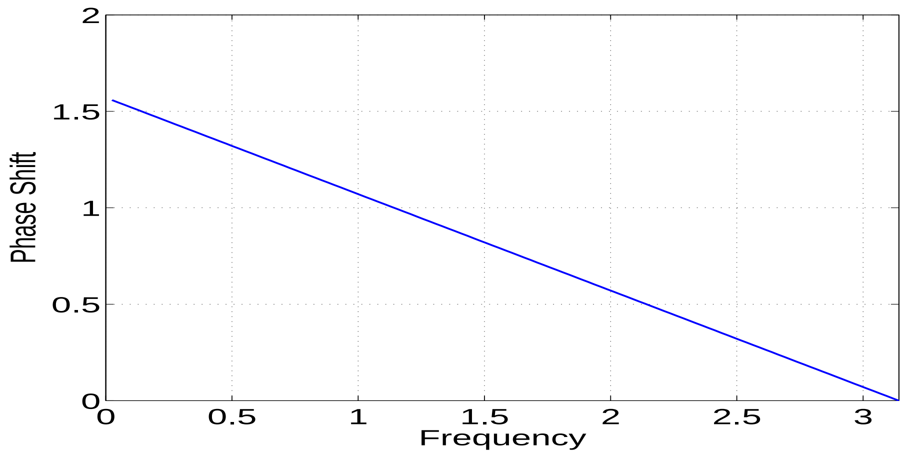
- Amplitude: $\left|1 - e^{-i\lambda}\right| = (2 - 2\cos\lambda)^{1/2}$

- Power: $\left|1 - e^{-i\lambda}\right|^2 = 2 - 2\cos\lambda$

- Phase: $\arctan\left(\frac{\sin\lambda}{1-\cos\lambda}\right) > 0$: the output is leading



Power function of $1 - L$



Phase shift of $1 - L$

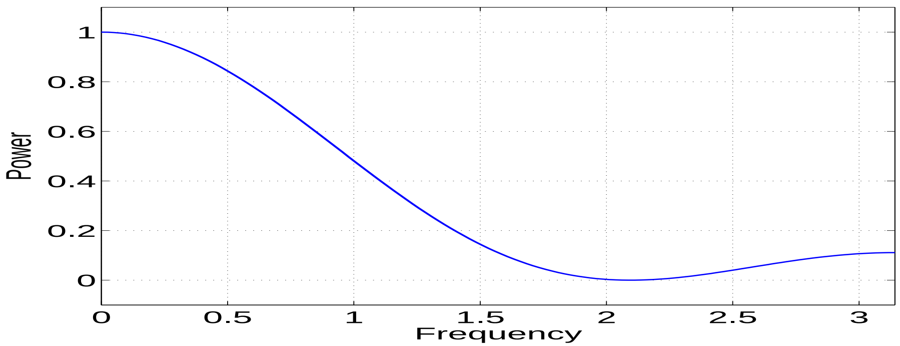
Moving average: $MA(3)$**:** $Y_t = \frac{X_{t-1}+X_t+X_{t+1}}{3} = \frac{L+1+L^{-1}}{3}X_t$

In the frequency domain:

- Transfer function: $\frac{e^{-i\lambda}+1+e^{i\lambda}}{3} = 1/3\,(1+\cos\lambda - i\sin\lambda + \cos\lambda + i\sin\lambda)$

- Power function: $\left(\frac{1+2\cos\lambda}{3}\right)^2$

- Phase shift: symmetric filter ($\phi(\lambda) = 0$)



Power function of $\frac{L+1+L^{-1}}{3}$

This filter is 0 when $\lambda = 2/3\pi$.

The filter kills periodicities equal to $\frac{2\pi}{\lambda} = 2\pi\frac{3\pi}{2} = 3$

Hodrik-Prescott filter

The trend in the HP filter is defined as:

$$\underset{T_t}{min} = \sum_{t=1}^{N} \{(Y_t - T_t)^2 + \mu\,[(T_{t+1} - T_t) - (T_t - T_{t-1})]^2\}$$

The FOC gives a trend T_t:

$$(1 - L)T_t = F(L)(1 - L)Y_t$$

where:

$$F(L) = \frac{1}{1 + \mu(1 - L)^2(1 - F)^2}$$

and a cycle $C_t = B(L)Y_t$, where:

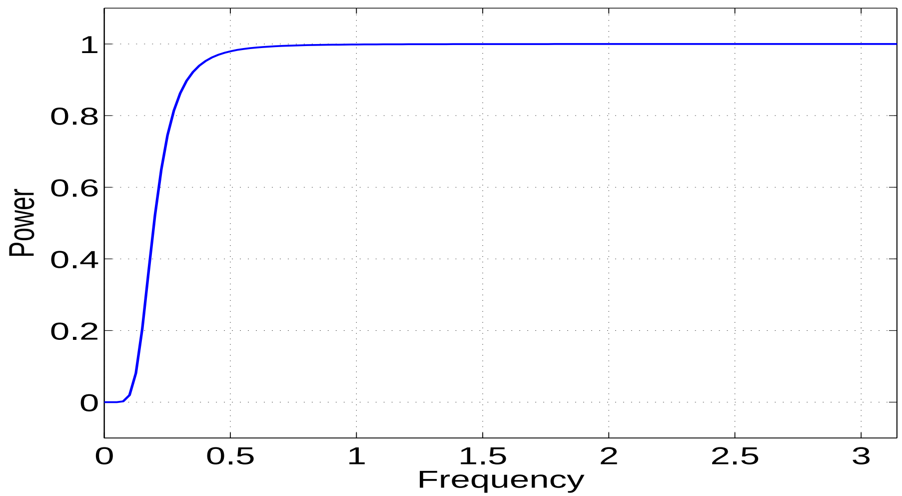
$$B(L) = \frac{\mu(1 - L)^2(1 - F)^2}{1 + \mu(1 - L)^2(1 - F)^2}$$

The transfer function for the cycle is:

$$B(e^{-i\lambda}) = \frac{4\mu(1 - \cos\lambda)^2}{1 + 4\mu(1 - \cos\lambda)^2}$$

The phase shift is zero (the filter is symmetric)

→ Problems at the beginning and at the end of the sample.



Power function of the HP filter

So far, we have seen filters starting from their time-domain form: given a sequence of coefficients $A(L)$, we studied their frequency domain features.

The usefulness of frequency domain can be appreciated going "the other way around": we want to design a filter with specific features in the frequency domain and find the coefficients to be applied to the series.

We start from a transfer function $\psi(e^{-i\lambda})$ with specific features and go back to time domain, obtaining the coefficients of the sequence $\psi(L)$.

How do we do it? We have inversion formulae:

$$\psi_k = \int_{-\pi}^{\pi} \psi(\lambda) e^{i\lambda k} d\lambda$$

The output of the filter will be:

$$y_t = \sum_{k=-\infty}^{+\infty} \psi_k x_{t-k}$$

In general, the filters $\psi(L)$ are bilateral and infinite and therefore they have to be truncated:

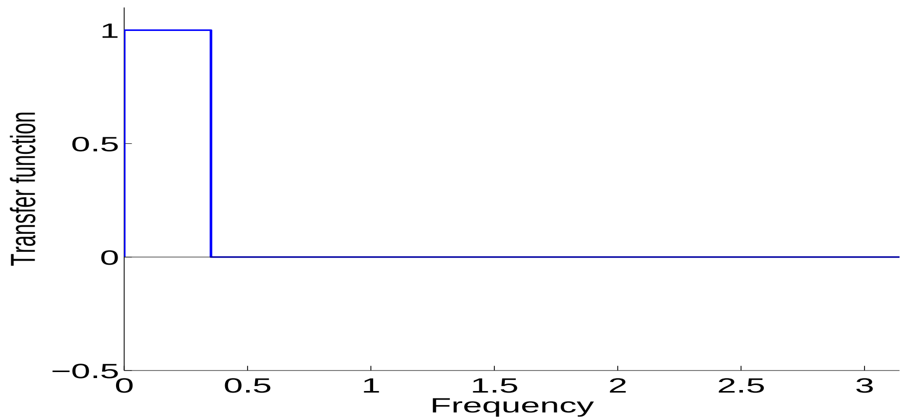
$$y_t = \sum_{k=-K}^{+K} \psi_k x_{t-k}$$

This will create approximation errors...

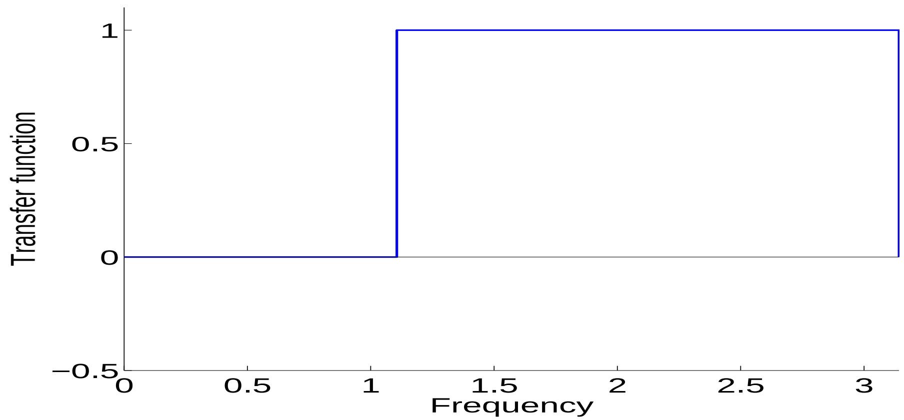
The band-pass filter

An interesting example is the band-pass filter. It has interesting implications (basically, it is what the **graphical equalizer** does in your hi-fi...) and will help us in understanding the frequency domain.

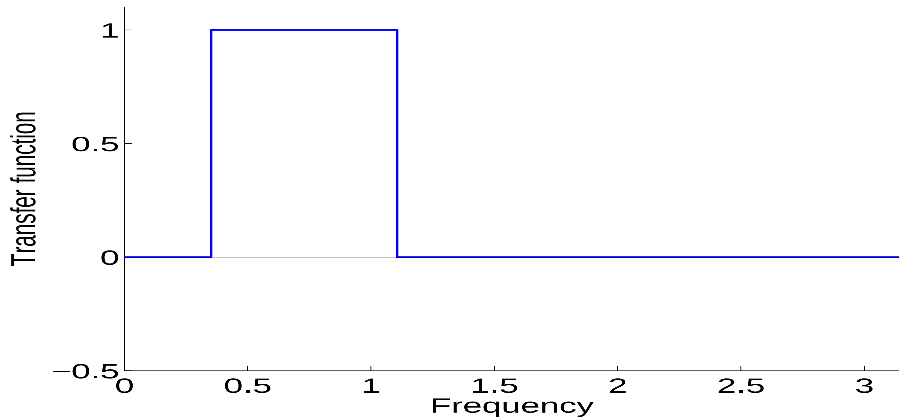
It is also interesting for economists as they are often interested in extracting particular cyclical components from the data.



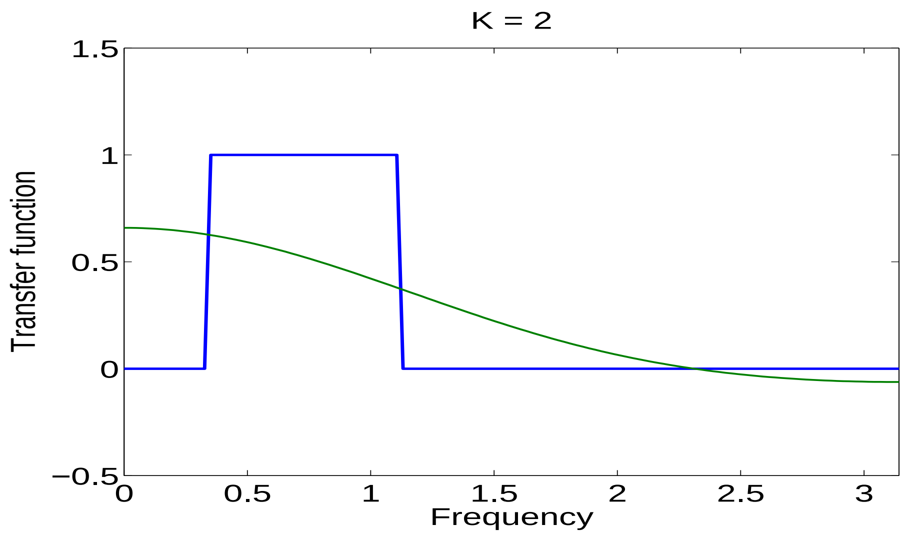
Theoretical low pass filter



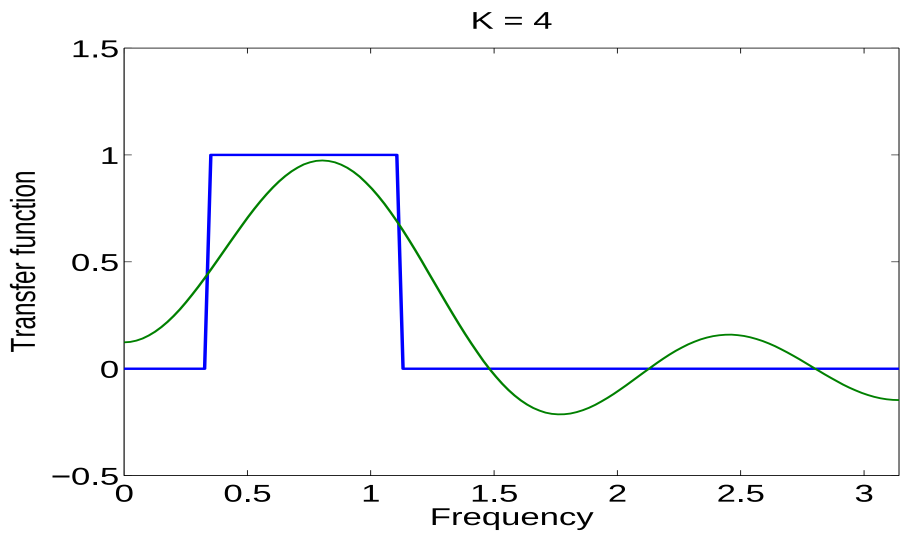
Theoretical high pass filter



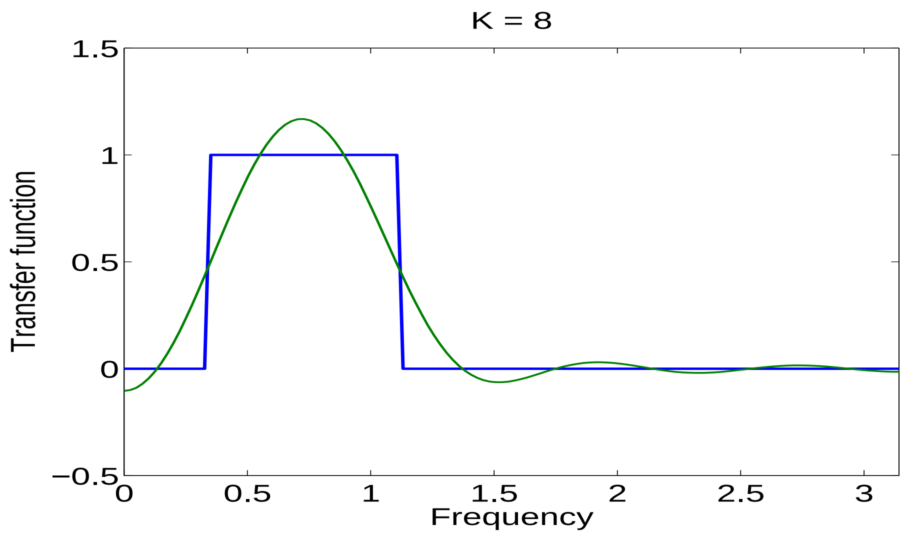
Theoretical band pass filter



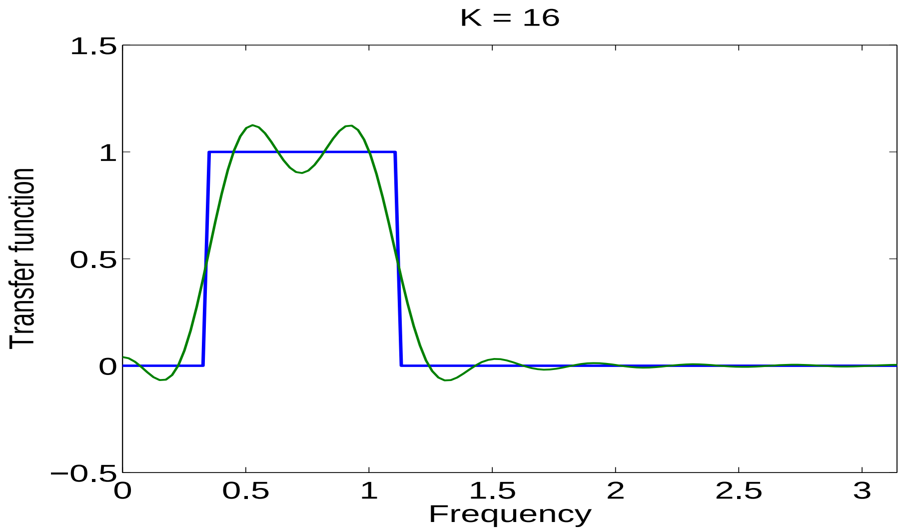
Approximated band pass filter



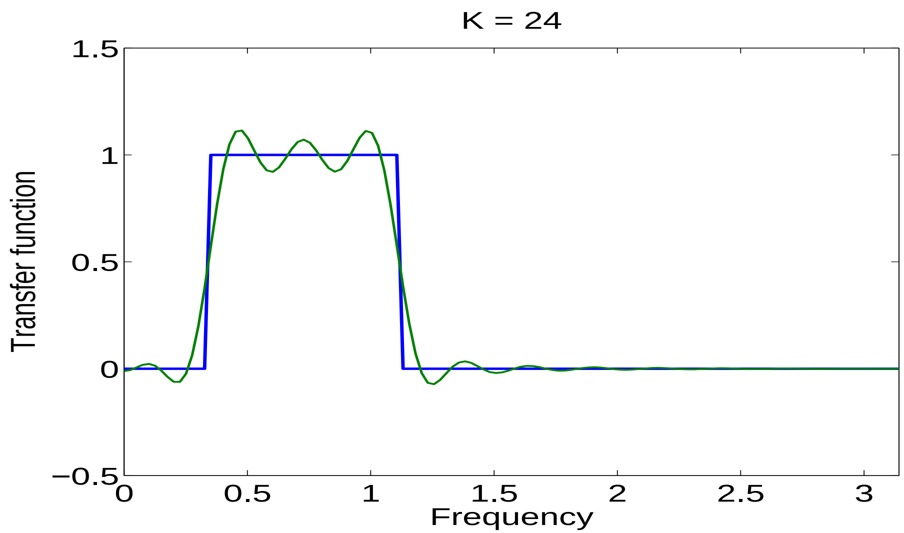
Approximated band pass filter



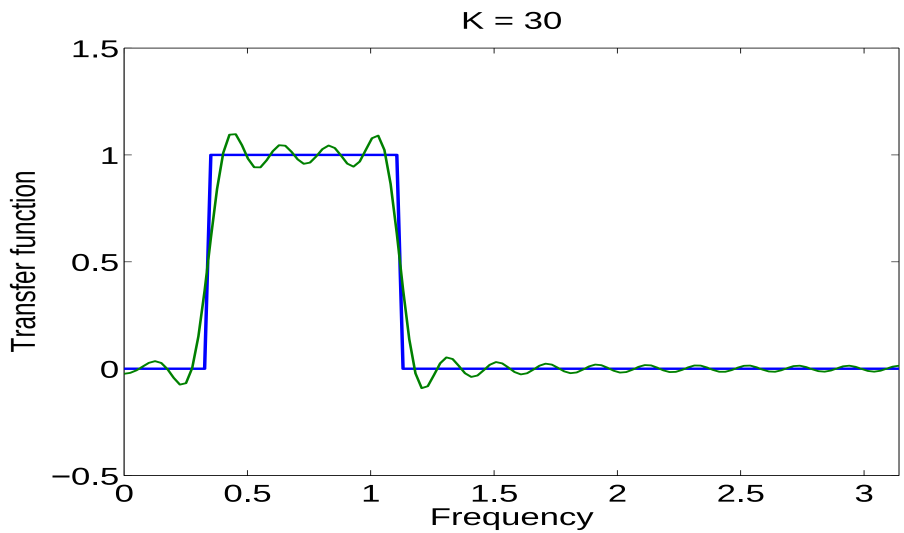
Approximated band pass filter



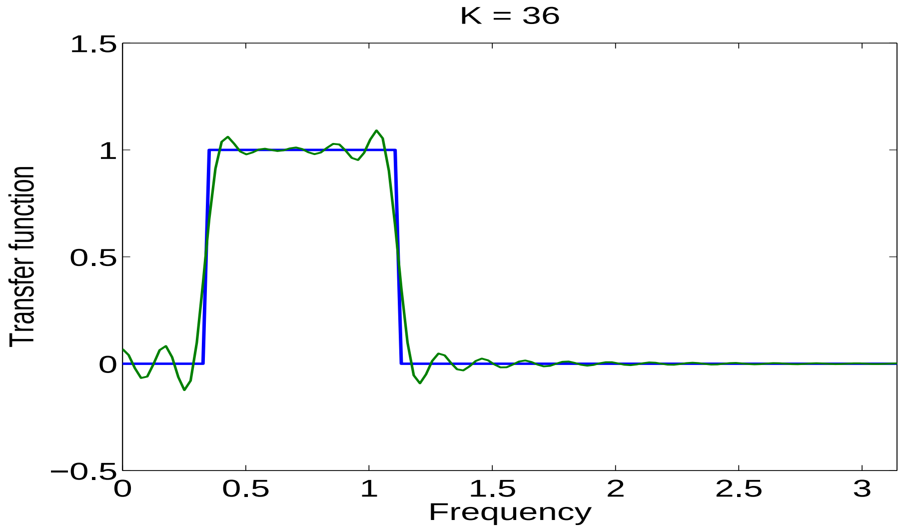
Approximated band pass filter



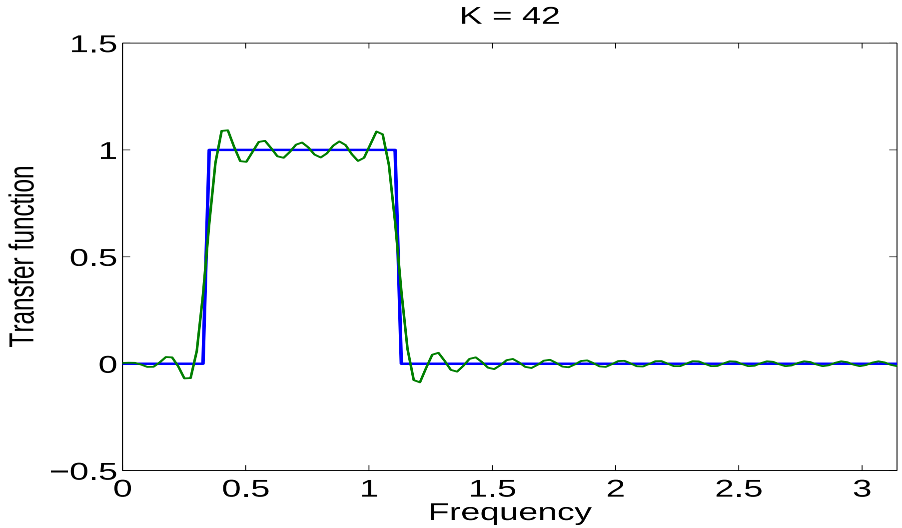
Approximated band pass filter



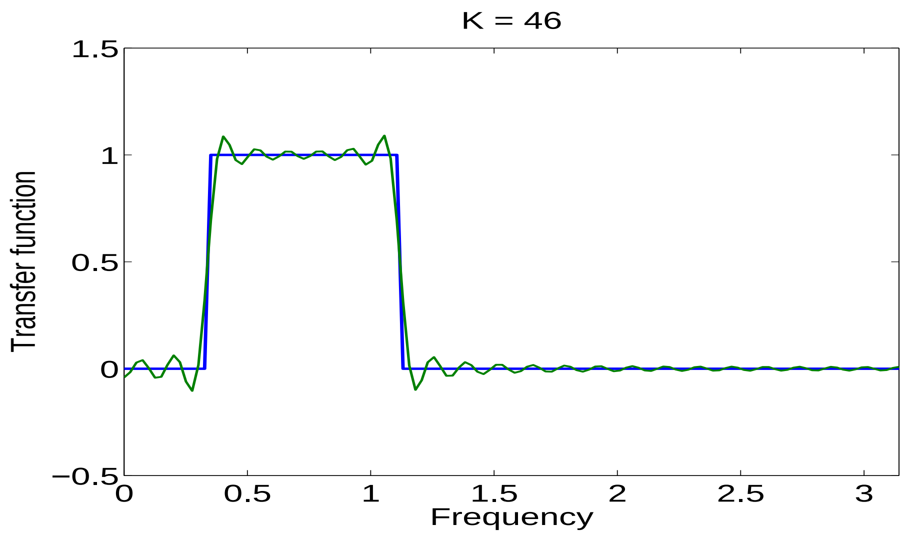
Approximated band pass filter



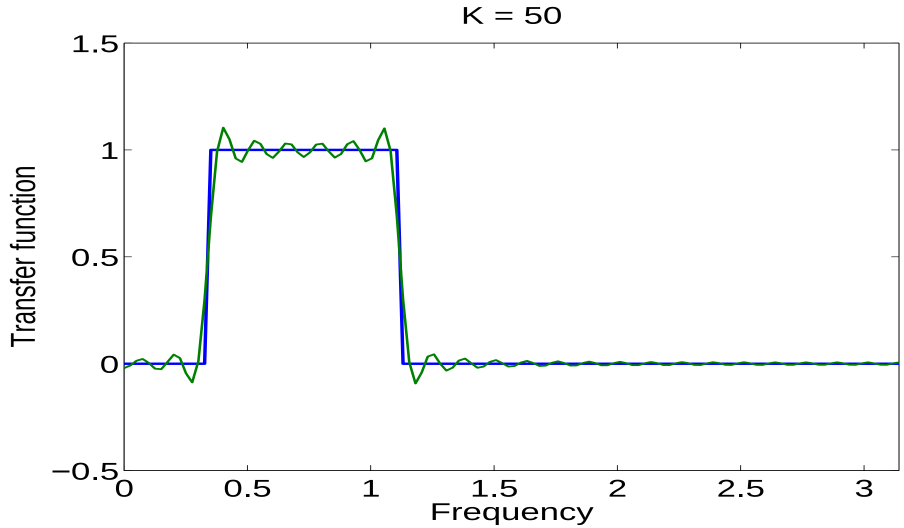
Approximated band pass filter



Approximated band pass filter



Approximated band pass filter



Approximated band pass filter

Multivariate Analysis

If X_t is bivariate, then the autocovariance matrix is:

$$\Gamma(h) = \begin{bmatrix} \gamma_{11}(h) & \gamma_{12}(h) \\ \gamma_{21}(h) & \gamma_{22}(h) \end{bmatrix}$$

The *spectral density matrix*, will have as element $(1, 2)$, the *cross-spectrum*:

$$f_{12}(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} e^{-i\lambda j} \gamma_{12}(j)$$

The *spectral density matrix* will be:

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} e^{-i\lambda j} \Gamma(j)$$

As in the univariate case, the inversion formula:

$$\int_{-\pi}^{\pi} f_{12}(\nu) e^{ih\nu} d\nu = \gamma_{12}(h)$$

Notice moreover, that:

$$\int_{-\pi}^{\pi} f_{12}(\nu) d\nu = \gamma_{12}(0)$$

- $f_{ii}(\lambda)$ is real and symmetric about $\lambda = 0$
- $f_{ij}(\lambda)$ is complex and non-symmetric about $\lambda = 0$ (as in general, $\gamma_{ij}(h) \neq \gamma_{ij}(-h)$)
- $f_{12}(\lambda) = \underbrace{c_{12}(\lambda)}_{\text{cospectrum}} - i \underbrace{q(\lambda)}_{\text{quadrature}}$
- In polar form: $f_{12}(\lambda) = \underbrace{\alpha_{12}(\lambda)}_{\text{amplitude}} e^{i\phi_{12}(\lambda)}$

- Squared coherency:

$$|H_{12}|^2 = \left| \frac{f_{12}(\lambda)}{[f_{11}(\lambda) f_{22}(\lambda)]^{1/2}} \right|^2$$

• The squared coherence is a measure of linear association defined frequency-by-frequency between the stochastic processes $X_{1,t}$ and $X_{2,t}$

• The squared coherency is the R^2 of the regression of the stochastic process $X_{1,t}$ on the stochastic process $X_{2,t}$.

Suppose that we want to find the filter $\Psi(L)$ that minimizes the mean squared distance between $X_{1,t}$ and $X_{2,t}$:

$$E\left|X_{2,t} - \sum_{j=-\infty}^{\infty} \Psi_j X_{1,t-j}\right|^2$$

This is nothing but a projection of X_{2t} on X_{1t} (a projection of an infinite dimensional vector on another infinite dimensional vector).

The coefficients of the projection can be obtained with the OLS formula ($Var^{-1} \cdot Cov$):

$$\Psi(e^{-i\lambda}) = \frac{f_{X_2 X_1(\lambda)}}{f_{X_1(\lambda)}}$$

We can obtain the spectral density of the projection with the formula for the filtering (X_{1t} is the input and the projection is the output):

$$\underbrace{f_{X_2|X_1}(\lambda)}_{\text{output}} = \underbrace{\left|\frac{f_{X_2X_1}(\lambda)}{f_{X_1}(\lambda)}\right|^2}_{\text{filter}} \underbrace{f_{X_1}(\lambda)}_{\text{input}}$$

From the equation above and the definition of the squared coherency, one can immediately see that:

$$|H_{21}|^2 = \left|\frac{f_{X_2X_1}(\lambda)}{[f_{X_2}(\lambda)f_{X_1}(\lambda)]^{1/2}}\right|^2 = \frac{f_{X_2|X_1}(\lambda)}{f_{X_2}(\lambda)}$$

$|H_{21}|^2$ defines the fraction of the variance of the process X_{2t} explained by the projection on X_{1t}.

It is clear that $0 \leq |H_{12}|^2 \leq 1$

Frequency domain and cointegration

Suppose the vector Y_t is composed by $I(1)$ processes, but that there exist cointegration relations.

The Wold representation: $\Delta Y_t = \Psi(L)\epsilon_t$

In the frequency domain: $f_{\Delta Y} = \left|\Psi(e^{-i\lambda})\right|^2 \frac{\sigma^2}{2\pi}$

We know from the theory of cointegration that $\Psi(1)$ has reduced rank. Here, $\Psi(e^{-i\lambda}) = \Psi(1)$ when $\lambda = 0$.

Therefore the spectral density of the first differences of a cointegrated vector has reduced rank at zero frequency (see Phillips and Ouliaris, 1987)

Decomposing X_t into orthogonal processes: an important property of the frequency domain

1. We obtained the spectral density by projecting on a set of orthonormal functions

2. OLS in the case in which the regressors are orthogonal...

Suppose: $\quad Y_{1,t} = B_1(L)X_t \quad$ and $\quad Y_{2,t} = B_2(L)X_t$

The cross-spectrum between Y_1 and Y_2 is:

$$f_{Y_1,Y_2}(\lambda) = \left|B_1(e^{-i\lambda})\right|^2 \left|B_2(e^{-i\lambda})\right|^2 f_X(\lambda)$$

If:

$$B_1(e^{-i\lambda}) = \begin{cases} 1 & \text{if } \lambda \in [a, b] \\ 0 & \text{elsewhere} \end{cases}$$

and:

$$B_2(e^{-i\lambda}) = \begin{cases} 1 & \text{if } \lambda \in [c, d] \\ 0 & \text{elsewhere} \end{cases}$$

$$\text{with} \quad [a, b] \cap [c, d] = \emptyset$$

$$\left|B_1(e^{-i\lambda})\right|^2 \left|B_2(e^{-i\lambda})\right|^2 = 0$$

$$\Downarrow$$

$$f_{Y_1,Y_2}(\lambda) = 0$$

The processes $Y_{1,t}$ and $Y_{2,t}$ are **orthogonal**

$\longrightarrow$ The frequency domain decorrelates: any processes is decomposed in a set of processes orthogonal one to the other.

Estimation

Suppose the data are in the $(N \times 1)$ vector $x = [x_1, x_2, ...x_N]'$

If we assume that the x are values of a periodic function of period N, we know that each x_i can be seen as linear combination of harmonics.

The Discrete Fourier Transfom

The DFT is a sequence $\{a_j, j \in F_N\}$, where:

$$F_N = \left\{ j \in Z, -\pi < \omega_j \equiv \frac{2\pi j}{N} \leq \pi \right\}$$

and:

$$a_j = \langle x, e_j \rangle = N^{-1/2} \sum_{t=1}^{N} x_t e^{-it\omega_j}$$

with $e_j = [e^{i\omega_j} e^{i2\omega_j} ... e^{iN\omega_j}] N^{-1/2}$

The Periodogram

$$I(\omega_j) \equiv |a_j|^2 = |\langle x, e_j \rangle|^2 = N^{-1} \left| \sum_{t=1}^{N} x_t e^{-it\omega_j} \right|^2$$

Note that:

$$\|x\|^2 = \sum_{j \in F_N} I(\omega_j)$$

$I(\omega_j)$ decomposes the variance in N orthogonal components.

From:

$$I(\omega_j) = N^{-1} \left| \sum_{t=1}^{N} x_n e^{-it\omega_j} \right|^2$$

it is possible to show that:

$$I(\omega_j) = \sum_{|k| \leq N} \hat{\gamma}(k) e^{-ik\omega_j}$$

where:

$$\hat{\gamma}(k) \equiv N^{-1} \sum_{t=1}^{N-k} x_{t+k} x_t'$$

is the sample autocovariance function

If we recall the definition of the spectral density:

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-i\lambda j}$$

A natural estimator seems to be:

$$\hat{f}(\omega_j) = \frac{I(\omega_j)}{2\pi}$$

$\hat{f}(\omega_j) = \frac{I(\omega_j)}{2\pi}$ is asymptotically unbiased, but it is not consistent.

Its variance does not go to zero, as $N \to \infty$

We can however prove that estimates at different frequencies are asymptotically **orthogonal**

The property of orthogonality is very important: we can use a sort of *law of large numbers* argument: averaging different ordinates in a neighbourhood (kernel estimator). It works!!

"Smoothing the periodogram"

$$\hat{f}(\omega_j) = \frac{1}{2\pi} \sum_{|k| \leq N} W_N(k) I_N(\omega_{j+k})$$

One example, $W_N(k) = (2m+1)^{-1}$

"Smoothing the autocovariances"

$$\hat{f}(\omega_j) = \frac{1}{2\pi} \sum_{|k| \leq N} W_N^{\gamma}(k) \hat{\gamma}(k) e^{-ik\omega_j}$$

Parametric estimator

Estimate an ARMA:

$$\hat{\phi}(L)X_t = \hat{\theta}(L)\epsilon_t$$

and then compute:

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