Coherence studies in time series

Julija Janeiko
Vytautas Panknas

Vilnius Gediminas Technical University

Workshop of Baltic-Nordic-Ukrainian Network
on Survey Statistics
August 21-24, 2018, Jelgava, Latvia

Table of Contents

Introduction

- Why do we need coherence?
- What does it mean in Official statistics and time series?

Cross – covariance function

Suppose we have two finite stationary times series $x_t = x_1, x_2, ..., x_n$ and $y_t = y_1, y_2, ..., y_n$. Their empirical cross – covariance function is defined as

$$\gamma_{xy} = \frac{1}{n} \sum_{t=1}^{n} ((x_{t+h} - \bar{x})(y_t - \bar{y})),$$

where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i,$$

$$t = 1, 2, \dots, n, \quad h = 0, 1, 2, \dots, n,$$

$$x_{t+n} = x_t, \quad y_{t+n} = y_t.$$

Orthogonal functions

Discrete time functions ϕ_j and ϕ_k are *orthogonal* if

$$\sum_{t \in D} \phi_k(t)\phi_j^*(t) \begin{cases} = 0, & k \neq j, \\ \neq 0, & k = j. \end{cases}$$

In our case

$$\sum_{t=1}^{n} e^{i\frac{2\pi kt}{n}} e^{-i\frac{2\pi kt}{n}} \begin{cases} = 0, & k \neq j, \\ \neq 0, & k = j. \end{cases}$$

Fourier transformation

$$\gamma_{xy}(h) = \sum_{k=-[n/2]+1}^{[n/2]} c_k e^{i\omega_k t}$$

with the frequencies $\omega_k = \frac{2\pi k}{n} \in [-\pi, \pi]$. Fourier transformation is applied to cross - covariance function. Frequencies $\omega_k = \frac{2\pi k}{n}$ are called Fourier frequencies.

Cross - spectrum

The Fourier coefficients c_k (cross-spectrum) are calculated using the formula

$$c_k = \frac{1}{n}\sum_{h=1}^{n}\gamma_{xy}(h)\mathrm{e}^{-i\omega_k t}, \quad k = -\left[\frac{n}{2}\right] + 1, ..., 0, 1, ..., \left[\frac{n}{2}\right].$$

These values are complex.

Power of the sequence and periodogram

Energy of cross - covariance function $\gamma_{xy}(h)$ is defined as

$$Energy = \sum_{h=1}^{n} \gamma_{xy}(h)$$

and the power of cross-covariance

$$Power = \frac{1}{n} \sum_{h=1}^{n} \gamma_{xy}(h) = \sum_{k=-[n/2]+1}^{[n/2]} c_k^2.$$

The coefficients $f_0 = c_0^2$, $f_{[n/2]} = |c_{[n/2]}|^2$, $f_k = |c_{-k}|^2 + |c_k|^2 = 2|c_k|^2$, $k = 0, 1, 2, ..., [n/2]$ are interpreted as the contribution of this frequency to the total power. Points (w_k, f_k) plotted as a function is called a periodogram.

Squared coherence function

A measure of strength of two processes relation at the w_k frequency is coherence coefficient defined as

$$\mathrm{coh}^2_{xy}(\omega_k) = \frac{|f_{xy}(\omega_k)|^2}{f_{xx}(\omega_k) f_{yy}(\omega_k)}. \tag{1}$$

where $f_{xx}(\omega_k)$ and $f_{yy}(\omega_k)$ are the individual spectra of the series x_t and y_t, respectively; $f_{xy}(\omega_k) = f_k$.

Points (w_k, coh^2) plotted as a function is denoted as squared coherence function.

Correlation coefficient

Let us all remember conventional Pearson's correlation coefficient determined as

$$\rho_{xy}^2 = \frac{\sigma_{yx}^2}{\sigma_x^2 \sigma_y^2}, \tag{2}$$

where σ_x^2 and σ_y^2 are variances of random variables X and Y and $\sigma_{yx} = \sigma_{xy}$ is their covariance.

Data

- Statistics Lithuania, Labor force survey data:
  - number of employed (LFE) (in thousands);
  - number of unemployed (LFU) (in thousands).
- Statistics Lithuania, Labour remuneration survey data:
  - number of employees (Emp) (in thousands);
  - resource for remuneration (RRS) (in millions of Euros).
- Labour Exchange office data:
  - number for registered unemployment (EU) (in thousands).
- Administrative data of the Social insurance institution Sodra:
  - enterprise remuneration, from which taxes are paid (RSI) (in millions of Euros)

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

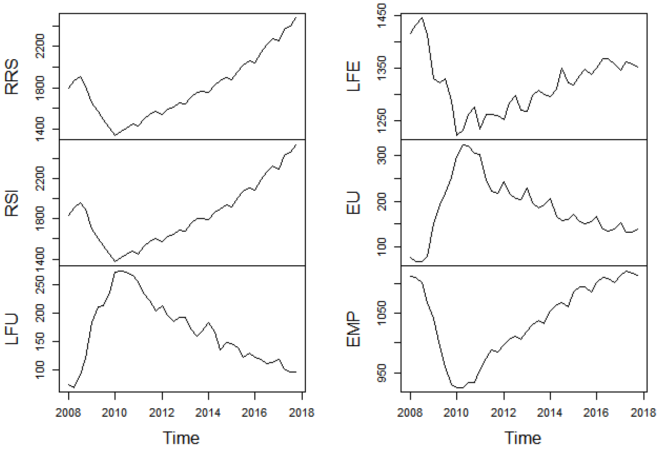


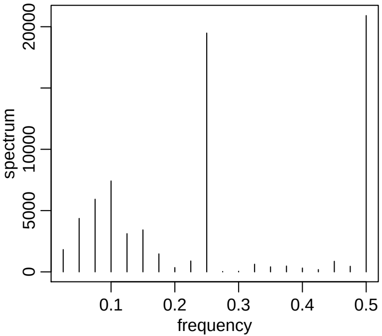
Figure 1: Graphical view of the quarterly data in 2008-2017

Correlations

Table 1: Correlation matrix of data

	RRS	RSI	LFU	LFE	EU	EMP
RRS	1.0000	0.9997	-0.8602	0.6716	-0.7274	0.9085
RSI		1.0000	-0.8625	0.6782	-0.7327	0.9082
LFU			1.0000	-0.8927	0.9512	-0.9673
LFE				1.0000	-0.9278	0.8109
EU					1.0000	-0.8962
EMP						1.0000

The differentiated time series $x_k - x_{k-1}$ and $y_k - y_{k-1}$ are used.



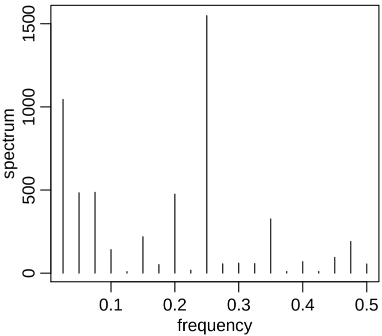


Figure 2: Periodograms of LFU (left) and RSI (right)

The frequency $\omega = 0.25 = \omega_{10}/2\pi$ with the corresponding period $T = 1/\omega = 4$ – four quarters is the most important.

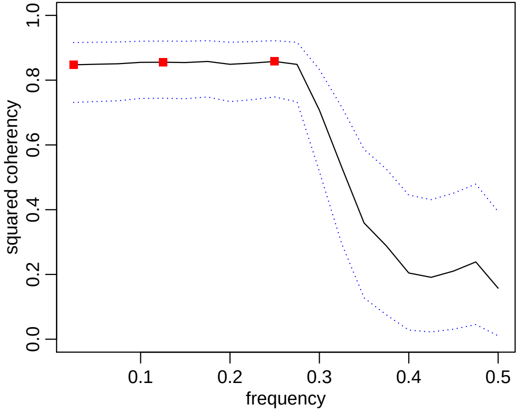


Figure 3: Coherence between LFU and RSI

Coherence coefficient values

Table 2: Coherence coefficients and squared correlation coefficient between LFU and RSI

T (years)	ω	$\mathrm{coh}^2_{xy}(\omega)$
10	0.025	0.8475
2	0.125	0.8554
1	0.25	0.8577
$\rho^2(x, y)$		0.7438

Periodograms

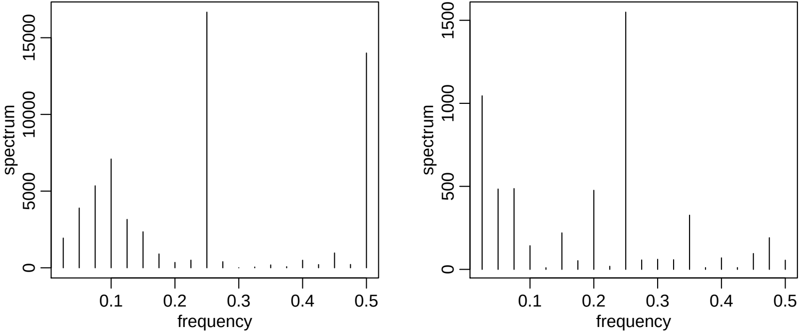


Figure 4: Periodograms of RSS (left) and RSI (right)

Coherence coefficient

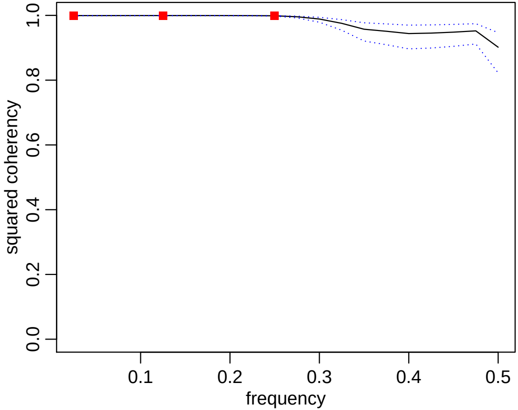


Figure 5: Coherence between RRS and RSI

Table 3: Coherence coefficients and squared correlation coefficient between RRS and RSI

T (years)	ω	$\mathrm{coh}^2_{xy}(\omega)$
10	0.025	0.9995
2	0.125	0.9995
1	0.25	0.9988
$\rho^2(x, y)$		0.9993

Periodograms

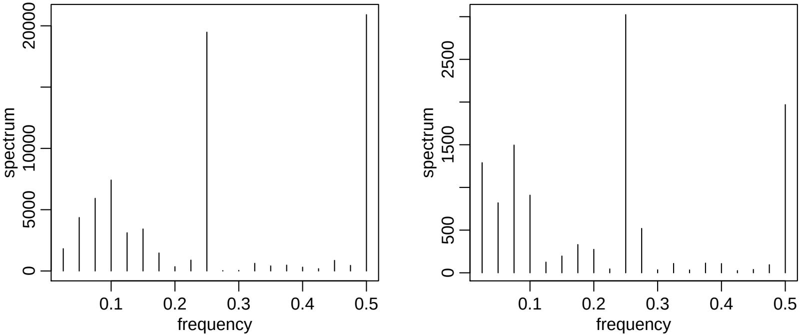


Figure 6: Periodograms of LFU (left) and EU (right)

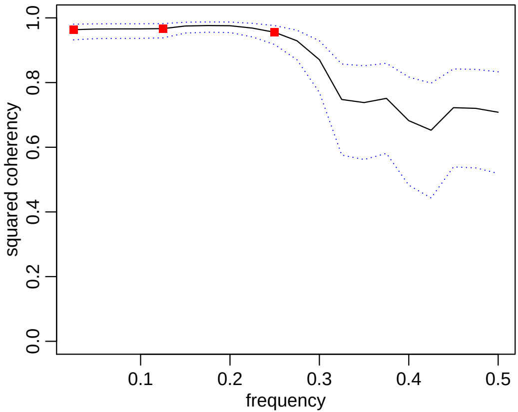


Figure 7: Coherence between LFU and EU

Table 4: Coherence coefficients and squared correlation coefficient between LFU and EU

T (years)	ω	$\mathrm{coh}^2_{xy}(\omega)$
10	0.025	0.9634
2	0.125	0.9668
1	0.25	0.9554
$\rho^2(x,y)$		0.9048

Periodograms

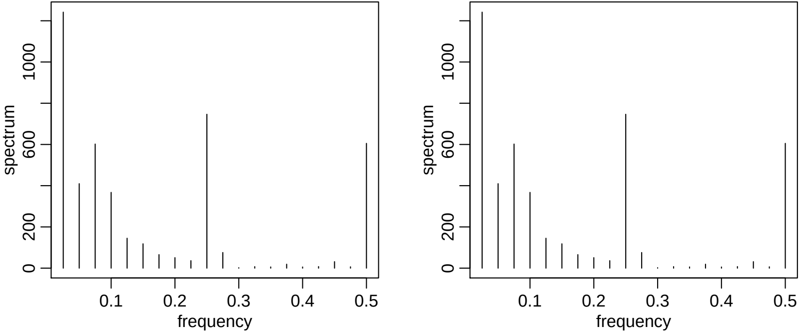


Figure 8: Periodograms of LFE (left) and EMP (right)

Coherence coefficient

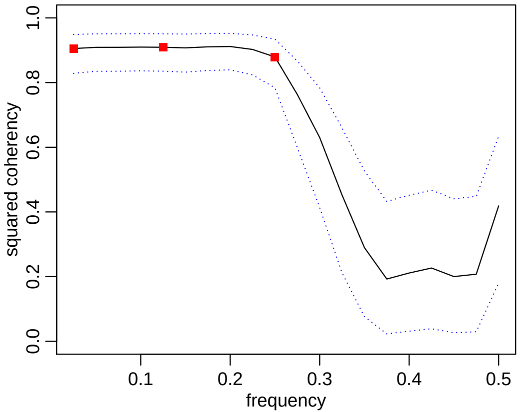


Figure 9: Coherence between LFE and EMP

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

Table 5: Coherence coefficients and squared correlation coefficient between LFE and EMP

T (years)	ω	$\mathrm{coh}_{xy}^2(\omega)$
10	0.025	0.9052
2	0.125	0.9091
1	0.25	0.8791
$\rho^2(x, y)$		0.6575

Conclusions

1. Open question – Pearsons correlation coefficient or coherence?
2. Great tool to measure the strength of relationship between two processes.
3. Difference of populations
4. Coherence coefficient values are higher than squared correlation.

References

Shumway R. H. and Stoffer D. S. (2006), *Time Series Analysis and Its Applications*, Springer Texts in Statistics, DOI = 10.1007/978-3-319-52452-8.

Stoffer, D. (2017), *astsa: Applied Statistical Time Series Analysis*, R package version 1.8, `https://CRAN.R-project.org/package=astsa`.

Wei W. W. S. (2006), *Time Series Analysis: Univariate and Multivariate Methods*, Pearson Education, `https://books.google.lt/books?id=aYOQAQAAIAAJ`.

Thank you for your attention!