Applying Free Random Variables to the Analysis of Temporal Correlations in Real Complex Systems

Małgorzata Snarska, Jagiellonian University - Institute of Physics,Cracow University of Economics-Chair of Econometrics and Operations Research

Abstract

Complex systems in nature fluctuate - exhibit very rich spatio– temporal structure.
We study the dynamical properties of real complex systems such as e.g. economy and/or financial market by looking at their spectral properties eg. density of eigenvalues under umbrella of Free Random Variables Calculus - back-of-an -envelope calculation of complicated problems in Random Matrix Theory.

Classical Methodology

General idea: In the absence of information on the phenomenon - take large number of possible explanatory variables and large number of output variables, look for correlations between pairs hoping to find some signal.
Problem: Standard tools for identifying hidden spatio–temporal structures like

- Factor Component Analysis
- Principal Component Analysis

are rapidly marred by measurement noise, quantified by $r = N/T$ and caught into dimensionality curse trap. This usually leads to biased estimates and spurious correlations.
Solution: Random Matrix Theory - with reacher structures can be difficult to get the result

- Can we have a more user-friendly version of RMT, to easily incorporate dynamical parameter ?

Models considered

we will assume, that cross–correlations of N variables can be described by the two–point covariance (correlation) function,

$$\mathcal{C}_{ia,jb} \equiv \langle X_{ia} X_{jb} \rangle . \qquad (1)$$

For $X_{ia} \equiv x_{ia} - \langle x_{ia} \rangle$, which describe the fluctuations (with zero mean) of the returns around the trend, and collect them into a rectangular $N \times T$ matrix $\mathbf{X}$. The average $\langle \ldots \rangle$ is understood as taken according to some probability distribution whose functional shape is stable over time, but whose parameters may be time–dependent. With cross–covariances and auto–covariances factorized, non–random, and decoupled the temporal dependence of the distribution of variable is the same, and the structure of cross–correlations does not evolve in time

$$\mathcal{C}_{ia,jb} = C_{ij} A_{ab} \qquad (2)$$

we will consider these distinct cases:

- $C = A = I$
- $C \neq I \quad A = I$
- $C \neq A \neq I$
- Cross–correlations

malgorzata.snarska@gmail.com

Under supervision of prof. M.A.Nowak

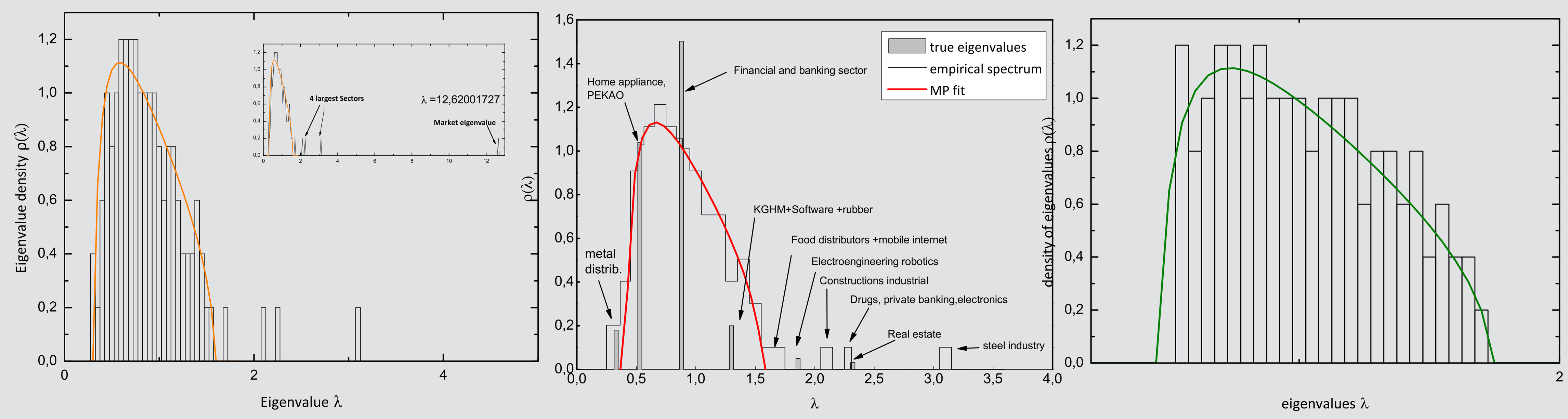
Free Random Variables

A generalization of probability theory to noncommutative random variables, such as infinite (Hermitian) random matrices $\mathbf{H}$. It relies on the concept of freeness, which is noncommutative independence.

Classical Probability	Noncommutative probability (FRV)
x - random variable, $p(x)$	H - random matrix, $P(H)$
pdf	spectral density $\varrho(\lambda)d\lambda$
characteristic function $g_x(z) \equiv \langle e^{izx} \rangle$	Green's function $G_H(z) = \frac{1}{N}\left\langle \mathrm{Tr}\frac{1}{z\cdot\mathbf{1}-H}\right\rangle$ or M - transform $M(z) = zG_H(z) - 1$
independence	freeness
Addition of independent r.v.: The logarithm of the characteristic function, $r_x(z) \equiv \log g_x(z)$, is additive, $r_{x_1+x_2}(z) = r_{x_1}(z) + r_{x_2}(z)$	Addition of f.r.v. The Blue's function $G_H(B_H(z)) = B_H(G_H(z)) = z$, is additive, $B_{H_1+H_2}(z) = B_{H_1}(z) + B_{H_2}(z) - \frac{1}{z}$
Multiplication of independent r.v.: Reduced to the addition problem via the exponential map, owing to $e^{x_1}e^{x_2} = e^{x_1+x_2}$	Multiplication of free r.v.: The N - transform, $M_H(N_H(z)) = N_H(M_H(z)) = z$, is multiplicative $N_{H_1H_2}(z) = \frac{z}{1+z}N_{H_1}(z)N_{H_2}(z)$

Equal – time correlations

- data: 100 time series of returns observed during 990 consecutive days
- correlation structure is related to the non–synchronous character of financial transactions
- the evolution of "true" eigenvalues is governed by non–stationary random variables



a) $C = A = 1$ b) $C \neq 1 \quad A = 1$ c) Randomized

VARMA(1, 1)

We introduce weak spatio–temporal correlation structure by VARMA(q_1, q_2) model

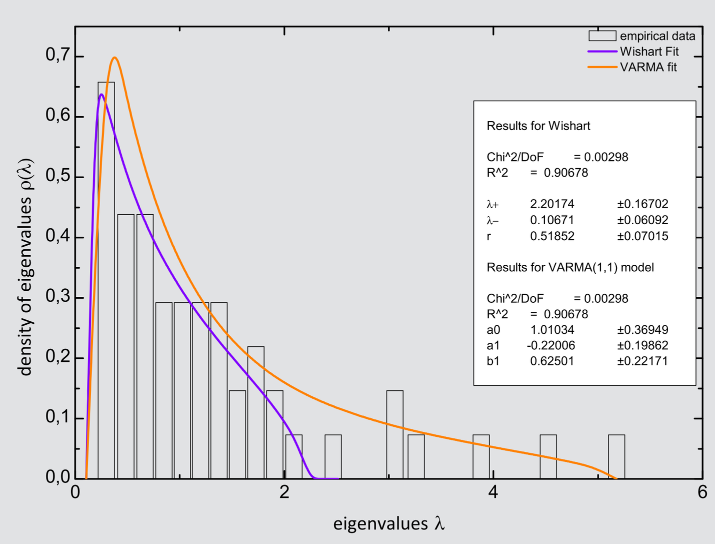
$$Y_{ia} - \sum_{\beta=1}^{q_1} b_\beta Y_{i,a-\beta} = \sum_{\alpha=0}^{q_2} a_\alpha \epsilon_{i,a-\alpha}. \qquad (3)$$

for $q_1 = q_2 = 1$, using the FRV multiplication formula

$$z = rMN_A(rM)N_C(M), \qquad (4)$$

we end up with $6-th$ order polynomial equation for $M \equiv M_c(z)$

$$r^4 a_0^2 a_1^2 (a_0^2 - a_1^2)^2 M^6 + \ldots = 0 \qquad (5)$$



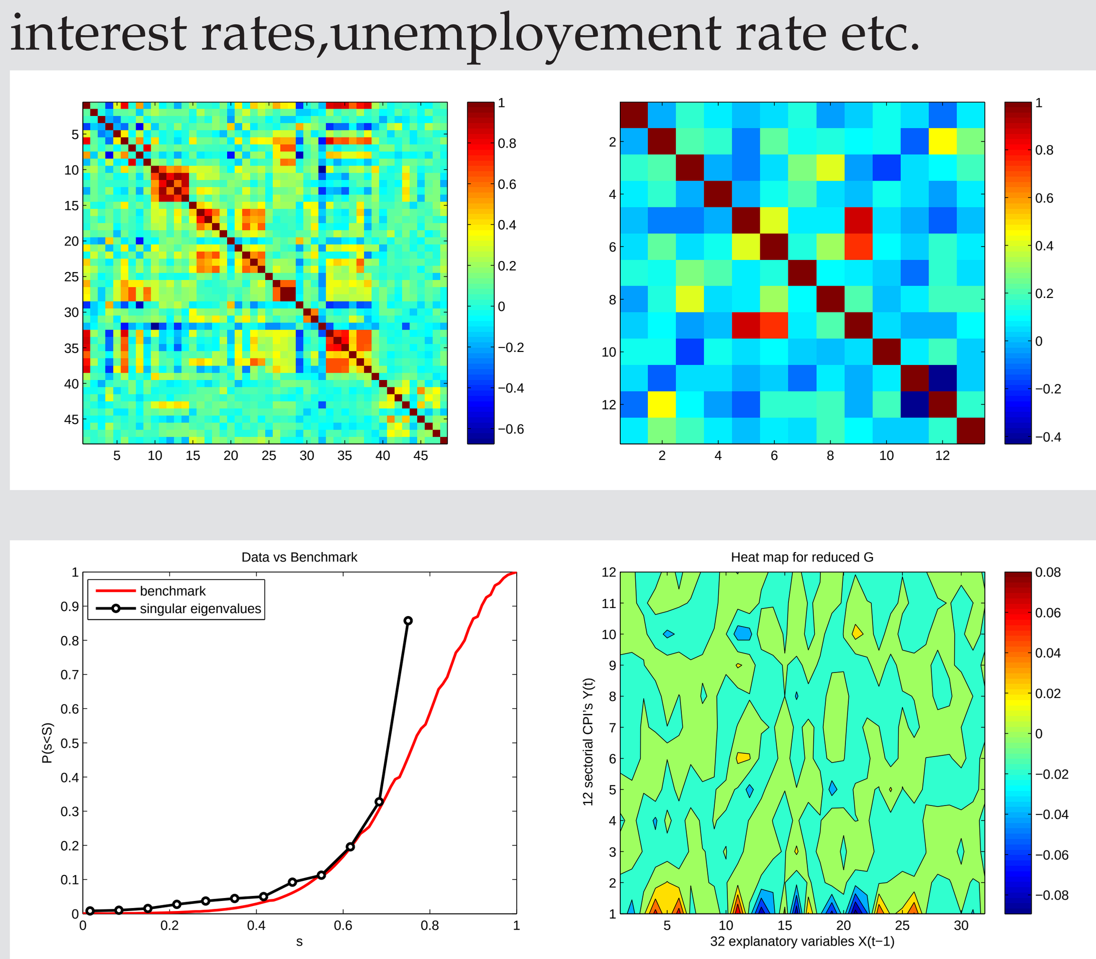
Empirical spectrum for macroeconomic data is similar to the one from financial markets, with largest eigenvalues separated. Simple VARMA(1, 1) temporal structure fits well the data.

Cross-Correlations

We are interested in the correlations between two matrices of nonequal size. We remove internal correlations inside X and Y and consider the SVD of a matrix $M \times N$

$$G = \hat{Y}\hat{X}^T \qquad (6)$$

of $\hat{Y}$ - $M = 15$ sectorial CPI's and $\hat{X}$ - $M = 37$ different macroeconomic indicators like GDP, interest rates,unemployement rate etc.



Inflation is described by relatively few common factors, like eg. foreign exchange reserves or the confidence level indicator.