РОЗДІЛ І. МОДЕЛЮВАННЯ СКЛАДНИХ СИСТЕМ В УМОВАХ ГЛОБАЛЬНОЇ ФІНАНСОВОЇ КРИЗИ

MULTISCALING OF INFORMATION COMPLEXITY MEASURES

A.Kiv
Ben-Gurion University of the Negev,
V. Soloviev
Cherkasy National University named after Bogdan Khmelnitsky,
K. Solovieva
Kryvyi Rih National University

Abstract

The features of the complexity concept in social and economic systems. It is shown that the economic paradigm of complexity theory is an alternative in volatile dynamics of the global economy. Information and multiscale measures of complexity are used to analyze comparative dynamic complexity of systems in the current global financial crisis.

Key words: complex systems, algorithmic complexity, Lempel-Ziv estimation, multiscaling, multiscale entropy

Introduction. In the time of globalization life becomes more complex and unpredictable. Chaos, order, and self-organization both in nature and society emerge according to the laws of complex dynamic systems. Complex dynamic systems are already successfully investigated in natural and technical sciences, starting from atomic and molecule systems in physics and chemistry to cellular organisms and ecologic systems in biology and neural networks studied by brain theories, social and computer networks. Nowadays the peculiarities of applying complex systems theory to economic and social sciences are widely discussed.

It is worth noting that complexity problems have started to attract significant attention at the end of the XX – the beginning of the XXI century. Outstanding natural scientists, Nobel prize laureates (I. Prigogine, M. Gell-Mann, P. Anderson) as well as mathematicians A. Kolmogorov, G. Chaitin, M. Li [1-6] have created a number of fundamental works in this field. Genius

American astrophysicist Stephen Hawking called 21st century a century of complexity.

There is no single definition of complexity of the system. Having announced the start of the new monograph series («Springer: Complexity. Understanding complex systems») in 2009, «Springer» has defined complexity in the following way: «Complex systems are the ones that consist of multiple interacting agents able to produce new qualities on the level of macroscopic collective behavior, which displays itself in spontaneous yet noticeable temporal spatial or functional structures» [7]. Apart from that, the following concepts and instruments have been developed: dynamic systems, nonlinear dynamics, instabilities, catastrophes, stochastic processes, determined chaos, self-organization, turbulence theory, graphs and networks, cellular automatons, adaptive systems, genetic algorithms, computer intellect etc.

2. Complexity economics. Complexity measures. For the past 10-15 years economic science has been going through the change of the dominant paradigm and, according to a lot of authoritative economists, both theorists and practitioners [8-13], still remains in a certain «bifurcation point», which will be left when it becomes qualitatively different. As noted by Eric Beinhocker we witness the transition «from traditional to the complexity economics»[13]. In many ways it is caused by the achievements in different fields of fundamental and applied sciences, which in the past three decades have disproved one of the major postulates of the classic economic theory stating rational behavior of economic agents.

Famous Ukrainian economist A. Galchinsky considers it to be the time of «… methodological renewal of economic theory – creative mastering of conceptual postulates of functioning and development created by applied sciences (physics and mathematics)» [10]. According to Galchinsky it is the methodology of complex systems that provides the basis for detailed consideration of the problem of inter-disciplinary connections of economic theory with other social and natural sciences, particularly the problem of scientific synthesis, which gains importance in modern scientific process [11].

There are different approaches to defining the specifics of complexity. All of them are emphasize the complexity of structure, interdependence, and interaction of different components, functioning within the system. «Complexity, - write M. Zgurovsky and N. Pankratova, - is a common quality of a single multitude of different objects, which are structurally interconnected, functionally interdependent and interact with each other...» [12].

It is important to take into account that changing the emphasis from simple to complex is not simply an establishment of a certain scientific paradigm. According to I. Prigogine it is about the scientific revolution in the investigation of modern world that don't restrain the perspective by «molecules, biological or social systems».

«Complexity economics» has not become a fully distinguished scientific discipline – it is more of a field of interdisciplinary research including aspects of behavioral economics, imitational modeling, chaos theory as well as ideas, derived from physics, biology, anthropology, cognitive psychology and other natural and humanity science disciplines.

Thus, the task implies the essential reconstruction of existing methodological mechanisms rather than correcting them. The aforementioned reconstruction causes deep breaches in the research methods, disabling current principles of scientific cognition: old methodological canons don't work anymore, while the new ones don't work yet - they have not been systematically formed. Essential challenges of modern scientific process, that cause critical events in the field of scientific research, including economic analysis theory, have to be considered in the corresponding context. Complex system methodology contains wide potential abilities to solve them.

Since modeling processes and using quantitative methods in economics imply measuring procedures, it is important to pay certain attention to complexity measures. I. Prigogine states that the notions of simplicity and complexity become relative in the pluralism of description languages [4], which causes multiple approaches to quantitative description of the idea of the complexity phenomenon. Therefore we proceed with the investigation of evident systems complexity using modern methods of quantitative analysis.

In this work we consider two of the most popular information complexity measures: Kolmogorov complexity and one of the entropies, particularly the sample entropy. Let us also investigate the change of the afore-mentioned measures on different time scales.

Test signals of different nature and complexity as well as stock indices of developed countries (German index DAX and the US one S&P 500 – http://finance.yahoo.com) and Ukraine (Ukraine's Stock Exchange index UX – www.ux.ua) have been chosen as research objects.

3. Information complexity measures. The most well-known and simple of information measures is the Kolmogorov complexity. The notion of Kolmogorov complexity has appeared in 1960s at the turn of algorithm, information and probability theories.

The A. Kolmogorov's idea [14], implied that to measure the amount of information in individual finite objects (not in random quantities as it was in Shannon's information theory). It turned out to be possible (although only with limited accuracy). Kolmogorov proposed to measure the amount of information in finite objects using the algorithm theory, defining object complexity as the minimal length of the program that generates this object. This definition became the basis of the algorithmic information theory as well as algorithmic probability theory: the object is considered to be random if it's complexity is close to maximal.

Therefore, according to Kolmogorov, object (e.g. a text – a symbol sequence) complexity is the length of the minimal program which outputs this text, and entropy is complexity divided by the length of this text. Unfortunately, this definition is purely theoretical. There is no precise way to define this program. However there are algorithms that actually attempt to calculate Kolmogorov complexity and entropy.

4. Evaluating Kolmogorov complexity using the Lempel-Ziv scheme. A. Lempel and J. Ziv suggested the following scheme of division words into sub-words. Indicate with x_l^r the word, consisting of letters of the word $x = a_{i1}...a_{in}$, staring from l and finishing with r, i.e. $x_l^r = a_{il}...a_{ir}$. Let us divide the word $x_1^n \in A^n$ into sub-words $\sigma_i, i = 1,...,m$ according to the following rule. Let the beginning of the word x_1^n be already divided into sub-words, i.e. be a concatenation of the sub-words $\sigma_1 \sigma_2 ... \sigma_{i-1}$ and

$x_1^n = \sigma_1 ... \sigma_{i-1} x_{li}^n$. We choose the following sub-word $\sigma_i = x_{li}^{l_{i=1}-1}$ so that the word $x_{li}^{l_{i=1}-2}$ is the longest prefix of the word x_{li}^n and is included into it as the sub-word in the word $x_{li}^{l_{i=1}-3}$, i.e. $\sigma_i = x_{l_i - d_i}^{l_{i+1}-d_i-2} a_{j_i}$, where $d_i \le l_i$. Every sub-word σ_i is defined by three numbers $(d_i, l_{i+1} - l_i j_i)$.

For example, the word a1a2a2a1a2a1a1a2a1a2a1a2 can be divided into sub-words a1, a2a2a1, a2a1a1, a2a1a2a1a2 and coded by the sequence of number triplets (1, 1, 1), (2, 1, 2), (1, 2, 1), (2, 3, 1), (4, 5, 2).

The Lempel-Ziv scheme generates the program P_{LZ}, which recreates the word from the sequence of triplets. In order to unambiguously divide binary codes of natural numbers, the first number of each triplet should be written as a binary using equally $\log l_i$ bits, the second one can be coded in the optional prefix code of natural numbers, writing down the third one requires only $\log|A|$ bits.

Let us determine the Lempel-Ziv complexity (LZC) for a time series of daily data of a stock market index. To research the LZC dynamics and compare it with other stock markets, we will calculate this complexity measure for a sub-sequence of a fixed size (window). To do this we will calculate logarithmic returns and turn them into a sequence of bits. We can give a number of differentiable states (number language). Therefore, for two different states we have 0, 1, for three – 0,1,2 etc. In case of three states, a certain b threshold is given (e.g. in points of standard deviation of normalized returns) and ret states are coded in the following manner:

$$ret = \begin{cases} 0, ret < -b \\ 1, -b \le ret \le b \\ 2, ret > b \end{cases}.$$

The algorithm executes two operations:

- adds a new bit to the existing sequence;
- copies the previously formed sequence.

Algorithmic complexity is the number of such operations, required to create a given sequence. In other words, Lempel-Ziv complexity measures the number of different sub-lines and the speed of them repeating during the original time series.

For a random series of the n length algorithmic complexity is calculated using the formula $LZC_r = n/\log(n)$. In that case relative algorithmic complexity is found as a relation of the received complexity to the random one: $LZC = LZC / LZC_r$.

On fig. 1a Lempel-Ziv algorithmic complexity is calculated for test signals: periodic function $\sin x$, white (wnoise) and flicker (fnoise) noises and a complex biological signal –ECG fragment (ECG).

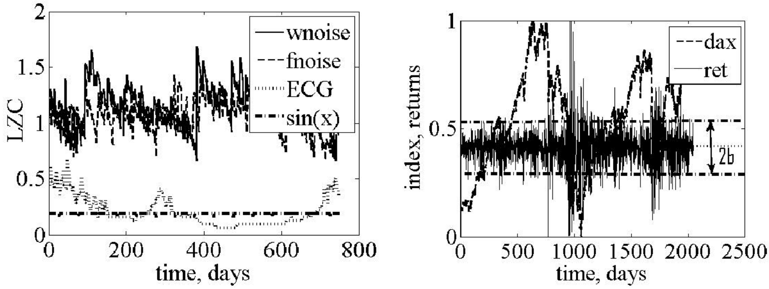


Fig. 1. a) Dynamics of algorithmic complexity for binary coded elements of different signal fragments of 1000 points, window size 250 points; b) relative dynamics of daily data of German stock index (dax) and normalized returns (ret), calculated to define the b parameter

In case of test signals and three states an evidently unnatural result is obtained: flicker and white noises depict the same level of complexity, while the complexity of an ECG-signal is almost the same as the one peculiar to a simple periodical signal. Obviously, periodic signal is the simplest one, the next being white noise and after that flicker noise. Biologic signal is also complex and contains correlations [8,9].

It is worth noting that Kolmogorov complexity rather measures the level of chaos in the system, leaving the «inner structures» of the object without any attention. Therefore, algorithmic complexity is unable to describe the complexity of real signals. Complex signals depict peculiar complexity on different spatial and temporal scales, that is have scale invariable characteristics [8], produce power distribution laws [9], which are responsible for realization of the less expected events (crises, shocks etc.).

In order to overcome such difficulties multiscale and fractal methods which will be considered in the following chapter.

5. Multiscale entropy. In the practical realization of entropy calculations for noisy time series analysis an Approximate Entropy (ApEn) or Sample Entropy (SampEn) evaluation algorithm was used. Since the detailed descriptions of both algorithms can be found in [15], we will provide short notes as to the terms of their execution. Since SampEn is more accurate the following calculations will be conducted for it or on its basis.

The inputs for SampEn include a time series and two parameters, *m* and *r*. *m* characterizes the embedding dimension, while *r* is a threshold criterion which allows us to consider two arbitrary vectors as the same («filtering factor»). We consider the subsequences of time series elements S_N, consisting of *m* numbers, taken starting from number *i*, and called vectors $p_m(i)$.

For P_m of all vectors of *m* length it is possible to calculate:

$$C_{im}(r) = \frac{n_{im}(r)}{N - m + 1},$$

where $n_{im}(r)$ is the number of vectors in P_m, similar to vector $p_m(i)$ (taking into account the chosen similarity criterion *r*). $C_{im}(r)$ is a part of vectors of *m* length, similar to a vector of the same length with elements starting from number *i*. For this time series values $C_{im}(r)$ are calculated for every vector in P_m, as well as the mean value $C_m(r)$, which shows the distribution of similar vectors of *m* length in S_N. Approximate Entropy for time series S_N using vectors of *m* length and similarity criterion *r* are defined using the formula:

$$\text{SampEn}\,(S_N, m, r) = \ln\left(\frac{C_m(r)}{C_{m+1}(r)}\right),$$

that is, as a natural logarithm of vector of *m* length repeats relation to vectors of *m*+1 repeating.

Let us highlight that SampEn is functionally dependable on one step of differentiation showing the level of uncertainty for the next count predicted with the process history. In other words, this kind of entropy describes the level of information loss on every next step. This is the reason why such parameters can be used to analyze events that are multiscale in their essence.

To overcome such difficulties it is suggested to use Multiscale Entropy Analysis (MSE), where SampEn is used as a measure of entropy on different scales of initial time series decomposition.

MSE includes two consequently executed procedures: 1) coarse graining of the initial time series – averaging the data on non-intersecting segments, which size (averaging size) was increased by one at each transition to the next scale; 2) calculating SampEn on every scale.

The coarse graining process implies averaging consequent series counts within non-intersecting windows, which size τ – increases at each transition to the next scale (fig. 2). Every element of the «grained» time series $y_j^{(\tau)}$ corresponds to the equation:

$$y_j^{\tau} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \, , \; 1 \le j \le N/\tau \, ,$$

where τ characterizes the scale factor.

The length of every «granulated» series depends on the size of the window and equals N/τ. If the scale equals 1 «granulated» series is simply identical to the initial one. For each obtained «granulated» time series SampEn was calculated as the scale function.

As you can see on scale 1 SampEn is the highest for white noise. But starting from scales over 7-8, the ECG signal of a healthy person becomes the most complex. The stock market signal complexity is comparable to the complexity of a biological signal. For three states (fig. 3b) multiscale character of the entropy is evident in case of stock markets. In case of a shuffled series (sp sh) entropy quickly decreases, which is the evidence of the complexity loss.

On fig. 3 we can see the results of the MSE-analysis of different signals.

Let us introduce window complexity measures that are determined as a sum of LZC or MSEC values on all scales for each window (MSE Complexity)

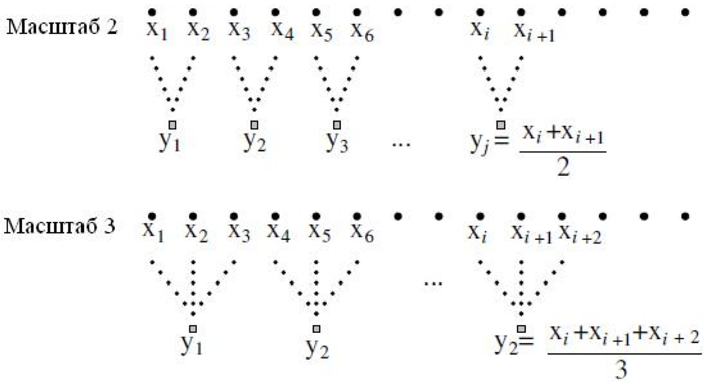
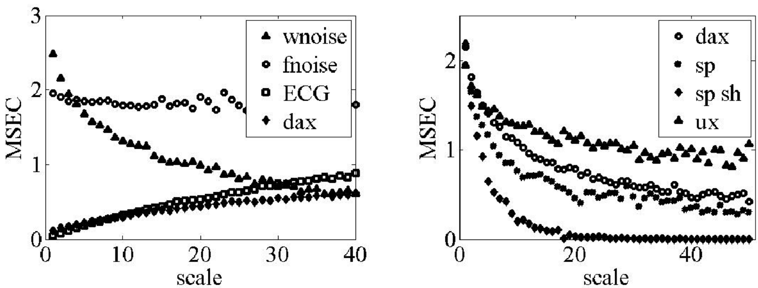


Fig.2. Schematic illustration of the coarse-graining process of initial time series for scales 2 and 3



a) b)

Fig. 3. a) Dependence of multiscale entropy on scale for binary coded signals; b) multiscale entropy dynamics for German (dax), US (sp) and Ukrainian (ux) stock indices

On figure 4 typical dependencies of multiscale complexity measures are displayed using DAX stock index data of 1992-2012.

As it is seen from fig. 3a, Lempel-Ziv complexity measure is characterized by rather low values during non-crisis periods. During pre-crisis periods, when the returns increase (volatility clusterization), LZC quickly rises as well. Multiscale MSEC measure, on the contrary, rises before crises and decreases during the crisis (fig. 3b). Apart from that, the introduced complexity measures: depict universal temporal behavior for different stock markets.

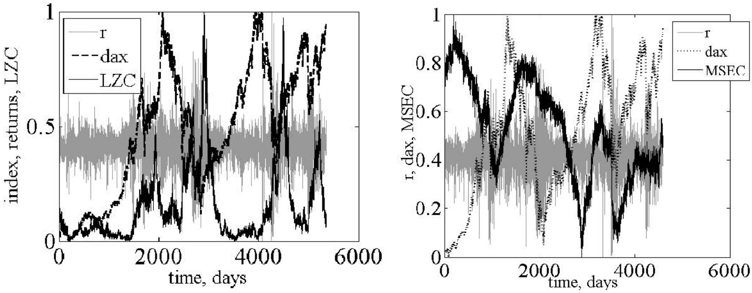


Fig. 4. Multiscale complexity measures in time: a) Lempel-Ziv; b) MSE. Window size equaled 500 days, step 1, maximal scale 20. The initial series (dax) and its returns (r) are also displayed

6. Conclusions and research perspectives. Thus, two new multiscale measures of economic complexity have been introduced and applied within the new paradigm – multiscale Lempel-Ziv measures and sample entropies. It is displayed that during crisis afore-mentioned measures change in a certain manner and are, therefore, possible to use as forerunning indicators of critical events.

Subsequent investigations will be directed at the formalization of complexity measures of network structures, which are the most common kind of a structural organization of complex socio-economic systems.

Literature:

1. Anderson P.W. More Is Different / P.W.Anderson // Science. - 1972. -V.177, No 4017.- P.393-396.
2. Gell-Mann M. What Is Complexity / M.Gell-Mann // Complexity. - 1995. -V.1, No 1.- P.16-18.
3. Prigojin I. Kost escho ne broshena [Electronniy resurs] Sayt S.P.Kurduymova «Sinergetika» Rejim dostupa: http://spkurdyumov.narod.ru/pprigoj.htm
4. Nikolis G., Prigojin I. Poznanie slojnogo. Vvedenie / G.Nokolis - M.: LKI, 2008.- 354 s.
5. Prigojin I. Konec opredelennosti. Vremya, haos i novie zakony prirody. / I Prigojin. – Ijevsk: NIC «Reguliarnaya I haoticheskaya dinamika», 2000.- 208 s.
6. Prigojin I. Ot suchestvuuchego k voznikayuchemu: Vremia

I slojnost v fizicheskih naukah / I.Prigojin – Perevod s angl. Seriya "Sinergetika: ot proshlogo k buduchemu". Izd.3 URSS-2006. - 296 s.
7. Arshinov V.I. Sinergetika konvergiruet so slojnostyu / V.Arshinov // Voprosy filosofii, 2011, №4.-S.73-83.
8. Miguel M.S., Johnson J.H., Kertesz J. et.al. Challenges in Complex Systems Science [Electronniy resurs] – Rejim dostupa:: arXiv:1204.4928v1 [nlin.AO] 22 Apr 2012.
9. Newman M.E.J. Complex Systems: A Survey. [Electronniy resurs] – Rejim dostupa: arXiv:1112.1440v1 [cond-mat.stat-mech] 6 Dec 2011.
10. Galchinskiy A. Globalni transformacii. Konceptualni alternativi. Metodologichni aspekti / Anatoliy Galchinskiy – K.: Lybid, 2006.- 312 s.
11. Galchinskiy A. Metodologiya skladnih sistem / A.Galchinskiy // Ekonomika Ukrainy - 2007. - №8. –S.4-18.
12. Zgurovskiy M.Z. Sistemniy analiz. Problemy, metodologiya, prilogeniya / M.Z. Zgurovskiy, N.D. Pankratova**.** – Kiev: Naukova dumka, 2005. – 744 s.
13. Beinhocker E.D. The Origin of Wealth. Evolution, complexity, and the radical remaking of economics // Harvard Business School Press, Boston, Massachusetts.- 2006.- 527 p.
14. Uspenskiy V.A., Vereschagin N.K., Sen A. / V.A. Uspenskiy // Kolmogorovskaya slojnost I algoritmicheskaya sluchaynost. - M.:MCNMO, 2010. - 556 s.: il.
15. Pincus S.M. Approximate entropy as a measure of system complexity. / S.M.Pincus // Proc Natl Acad Sci USA. -1991.-V88.-P.2297-2301.
16. Costa M., Goldberger A.L., Peng C.-K. Multiscale entropy analysis of biological signals / M.Costa // Phys Rev E. – 2005.-V.71.- P.021906.
17. Derbencev V.D., Serdyuk O.A., Solovyov V.M., Sharapov O.D. Sinergetichni ta ekonofizichni metody doslidjennya dynamichnyh ta strukturnyh harakterystyk ekonomichnyh sistem – Cherkassy: Brama-Ukraina, 2010. – 300 s.