

УДК 330.46; 519.86

ОСОБЛИВОСТІ ЗАСТОСУВАННЯ РЕКУРЕНТНИХ ДІАГРАМ І РЕКУРЕНТНОГО КІЛЬКІСНОГО АНАЛІЗУ ДЛЯ ДОСЛІДЖЕННЯ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ

FEATURES OF RECURRENT PLOTS AND RECURRENCE QUANTIFICATION ANALYSIS APPLICATION FOR RESEARCH OF FINANCIAL TIME SERIES

Олександр Варфоломійович ПІСКУН

к. т. н., доцент, ст. викладач кафедри

вищої математики та інформаційних технологій,

Черкаський інститут банківської справи УБС НБУ (м. Київ)

piskun@ukr.net

Oleksandr V. PISKUN

Ph.D. in Technical Sciences, Associate Professor, Senior Lecturer of Mathematics and Information Technology Department, Cherkasy Institute of Banking of the University of Banking of the National Bank of Ukraine (Kyiv)

Анотація. У статті досліджено теорію та методи рекурентного аналізу. Визначені оптимальні значення вхідних параметрів та відповідні міри для аналізу фінансових часових рядів із метою побудови інструменту моніторингу ринків.

Summary. In article the theory and methods of the recurrence analysis are considered. Optimum values of input parameters and corresponding measures for the analysis of financial time series with the purpose of tool monitoring construction of the markets are defined.

Ключові слова: *рекурентна діаграма, рекурентний кількісний аналіз, розмірність вкладень, затримка часу, фінансовий часовий ряд.*

Key words: *recurrence plot, recurrence quantification analysis, embedding dimension, time delay, financial time series.*

Постановка проблеми. Світова фінансова криза 2007-2010 років показала неспроможність більшості традиційних аналітичних методів аналізу фінансових ринків адекватно їх описувати. За своєю суттю ці підходи не були призначені для моделювання швидких змін, непередбачуваних стрибків і складних взаємодій окремих складових сучасного світового ринкового процесу. Останнім часом все більше уваги надають дослідженню фінансових часових рядів методами нелінійної

динаміки. Одними із таких методів є рекурентні діаграми та рекурентний кількісний аналіз, які не є критичними до довжини, стаціонарності часових рядів і працюють з будь-яким розподілом.

Аналіз останніх досліджень і публікацій.

Рекурентні діаграми та рекурентний кількісний аналіз вже досить широко застосовують в економіці. Наприклад, М. МакКінзі [1], А. Антоніо та К. Ворлов [2], К. Гілмор [3], Дж. Збілут [4] використовували ці методи для виявлення

природи поведінки фінансового ринку. Ф. Строззі [5, 6] застосував рекурентний кількісний аналіз для оцінки волатильності та корельованості валютних пар. А. Фабретті та М. Ауслус [7] – для визначення критичних режимів на фінансових ринках. Дж. Бастос та Дж. Кайдо [8] досліджували взаємозалежності між фондовими ринками, а також між їхньою поведінкою під час критичних подій.

Разом з тим, такого роду методи аналізу даних не можуть бути шаблонно застосовані до комплексних неприродних систем, як, наприклад, фінансові ринки.

Метою статті є визначення оптимальних значень вхідних параметрів та відповідних мір для аналізу фінансових часових рядів з метою побудови інструменту моніторингу ринків.

Обґрунтування отриманих наукових результатів. Формально, динамічна система вводиться фазовим простором, часом та законом еволюції у часі. Елементи фазового простору відображають можливі стани системи. Стан такої системи у фіксований час t може бути описаний d параметрами, які формують вектор у d -мірному фазовому просторі системи [9]:

$$\vec{x}(t) = (x_1(t), x_2(t), \dots, x_d(t))^T$$

В загальному випадку, закон еволюції у часі – це правило, що дозволяє визначити стан системи у кожний момент часу t за допомогою її попередніх станів. Для неперервних систем еволюцію у часі описують диференційними рівняннями:

$$\dot{\vec{x}}(t) = \frac{d\vec{x}(t)}{dt} = \vec{F}(\vec{x}(t))$$

$$F : \mathfrak{R}^d \rightarrow \mathfrak{R}^d.$$

Вектори $\vec{x}(t)$ визначають траєкторію у фазовому просторі.

На практиці далеко не всі компоненти, що потрібні для побудови вектору стану, є відомими. Часто маємо дискретну оцінку лише одного компоненту. Це дає скалярний дискретний часовий ряд:

$$u_i = u(i\Delta t),$$

де $i = 1, \dots, N$; Δt – дискретна норма вибірки. У

цьому випадку необхідно провести реконструкцію фазового простору. Для цього найчастіше використовують метод часової затримки [10]:

$$\hat{\vec{x}}_i = (u_i, u_{i+\tau}, \dots, u_{i+(m-1)\tau})^T$$

де m – розмірність вкладення, τ – затримка часу. Якщо $m \geq 2d+1$, де d є кореляційною розмірністю атрактора, то за теоремою Такенса гарантується збереження топологічних структур оригінальної траєкторії [11].

Для аналізу часових рядів необхідно правильно підібрати вхідні параметри – розмірність вкладень m та затримку τ . У літературі наводять різні підходи для оцінки найменшої достатньої розмірності (наприклад, алгоритм помилкових найближчих точок FNN) та затримки (наприклад, автокореляційна функція або функція взаємної інформації) [12].

У 1987 році Дж.-Р. Екман та співавтори [13] на основі теорії рекурентності запропонували інструмент для відображення поведінки системи у m -мірному фазовому просторі, що дістав назву рекурентна діаграма (RP). M -мірна фазова траєкторія станів системи $\vec{x}(t)$ завдовжки N відображається на двовимірну квадратну двійкову матрицю розміром $N \times N$, в якій 1 (чорна точка) відповідає повторенню стану при деякому часі i в деякий інший час j , а обидві координатні осі є осями часу:

$$R_{i,j}^{m,\varepsilon_i} = \Theta(\varepsilon_i - \|\vec{x}_i - \vec{x}_j\|) \quad \vec{x} \in \mathfrak{R}^m, i, j = 1, \dots, N$$

де N – кількість досліджуваних станів x_i ; ε_i – розмір околиці точки \vec{x} у момент i ; $\|\cdot\|$ – норма та $\Theta(\cdot)$ – функція Хевісайда (рис. 1).

Аналіз рекурентних діаграм дає змогу виявити дрібномасштабні структури – текстуру, яка складається з таких елементів [9]:

- ізольовані, окремо розташовані рекурентні точки з'являються в тому разі, коли відповідні стани рідкісні або нестійкі в часі, або викликані високою турбулентністю;

- діагональні лінії $R_{i+k, j+k} = 1$ (при $k = 1, \dots, l$, де l – довжина діагональної лінії) з'являються у разі, коли сегмент траєкторії пролягає практично

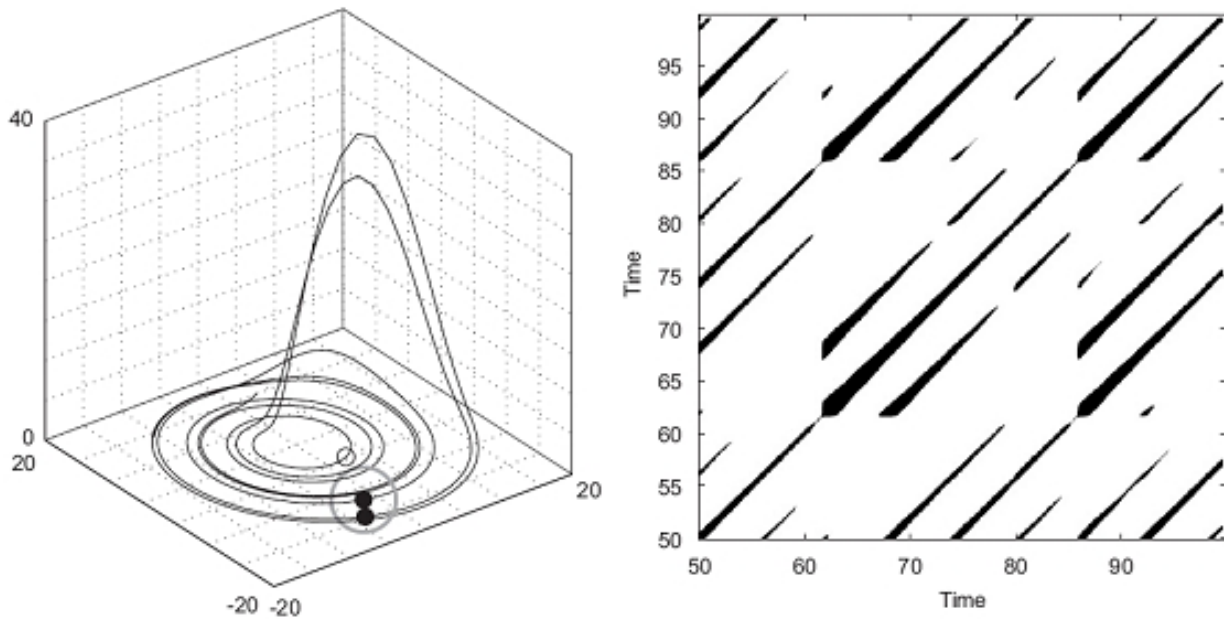


Рис. 1. Сегмент траєкторії системи Рьослера та відповідна рекурентна діаграма [9]

паралельно іншому сегменту, тобто траєкторія проходить у колі радіуса ϵ впродовж іншого сегмента. Довжина таких ліній визначається часом, протягом якого сегменти траєкторії залишаються паралельними;

- вертикальні (горизонтальні) лінії $R_{i,j+k} = 1$ (при $k = 1, \dots, u$, де u – довжина вертикальної або горизонтальної лінії) виділяють проміжки часу, протягом яких стан системи не змінюється або змінюється досить повільно.

Дж. Збілут, Ч. Вебер [14, 15], Дж. Гао [16] та Н. Марван [17] розробили низку кількісних мір, що розраховують на основі текстур у діаграмі.

Міри, що обчислюють на базі рекурентних точок:

- коефіцієнт самоподібності (RR) – оцінює густину рекурентних точок у діаграмі.

Міри, що обчислюють на основі діагональних ліній:

- детермінізм (DET) – частка рекурентних точок у діагональних лініях та усіх рекурентних точок. Оцінює рівень детермінованості системи;

- дивергенція (DIV) – зворотна величина максимальної довжини діагональної лінії. Показує рівень розходження сегментів траєкторії у фазовому просторі;

- ентропія (ENTR) – частотний розподіл

рекурентних точок у діагональних лініях. Відображає рівень складності системи;

- тренд (TREND) – лінійний коефіцієнт зниження густини рекурентних точок діагоналей. Характеризує нестационарність процесу.

Міри, що обчислюють на основі вертикальних (горизонтальних) ліній:

- час завмирання 1 (T1) – час між рекурентними точками у стовпчиках. Показує частоту рекурентності системи для кожного її стану;

- час завмирання 2 (T2) – час між початковими точками вертикальних ліній у стовпчиках. Розраховує частоту рекурентності системи до попередніх станів за винятком тангенціального руху;

- ламінарність (LAM) – частка рекурентних точок у вертикальних лініях та усіх рекурентних точок. Відображає рівень ламінарності (турбулентності) системи;

- час завмирання (TT) – частка рекурентних точок у вертикальних лініях та кількості вертикальних ліній. Визначає середню тривалість періодів ламінарної поведінки системи.

Зазначені міри сформували рекурентний кількісний аналіз (RQA).

Одним із суттєвих аспектів застосування RQA є

визначення оптимальних параметрів: розмірність m , затримка часу t та пороговий радіус ϵ [9, 12, 17-19]. Вхідні параметри підібрати відповідно до часового ряду та завдання [20]. Більш того, пропонують проводити тестування чутливості результатів при варіюванні значень параметрів.

У нашому випадку проводимо аналіз фінансових часових рядів з метою побудови інструменту моніторингу ринків. Тому необхідно визначити оптимальні параметри для такого роду досліджень. За контрольний ряд даних візьмемо значення котирувань фондового індексу S&P500 з 09.08.1999 по 11.07.2011 р. Такий період обрано не випадково. Справа у тому, що він включає в себе різні періоди функціонування ринку, зокрема і кризу.

Визначимо t та m . Для цього використаємо програмний пакет VRA [22]. За допомогою функції взаємної інформації визначаємо значення параметра t . На основі отриманого значення затримки часу, методом хибних найближчих точок (FNN) розраховуємо розмірність вкладень m (рис. 2).

У результаті розрахунків отримаємо такі значення вхідних параметрів: $t = 23$, $m = 20$. Для фінансового часового ряду розрахувати адекватні міри RQA з такими параметрами неможливо. Фінансові ринки є неприродними системами. Більш того, вони не тільки штучні, а й мають досить високий відсоток людського фактору. Тому

розраховані за стандартними методами вхідні параметри для них не підходять. Проаналізуємо чутливість поведінки міри LAM до різних значень вхідних параметрів. Дослідимо ряд котирувань фондового індексу S&P500 з 09.08.1999 по 11.07.2011 р. довжиною 3000 значень. За базові беремо такі значення вхідних параметрів: $t = 1$, $m = 1$, $\epsilon = 0,1$.

Розмір вікна для розрахунку мір у динаміці вибрано таким чином, щоб елімінувати стохастичну поведінку міри LAM, але й не знижувати її швидкість реагування на зміни стану ринку. Він становить 250 для рядів довжиною більше 1000 значень.

Проаналізуємо залежність поведінки міри LAM від значення розмірності вкладення m (рис. 3). Раніше було показано здатність LAM виявляти різні періоди функціонування ринку [22]. Рис. 3 є ще одним прикладом цієї властивості.

Вищі рівні розмірності m знижують значення LAM та роблять період з $x = 2010$ по $x = 3000$ важко зрозумілим для інтерпретації. При $m = 1$ маємо чітке гладке падіння LAM (з $x = 2010$ по 2460), що вказує на кризовий період; та підвищення LAM (з $x = 2461$ по 3000), що вказує на період релаксації. Інші значення m спотворюють цю динаміку ламінарності, тому оптимальним значенням розмірності вкладень $m \in 1$.

Розглянемо, як змінюється вигляд ламінарності з варіюванням порогового радіусу ϵ (рис. 4).

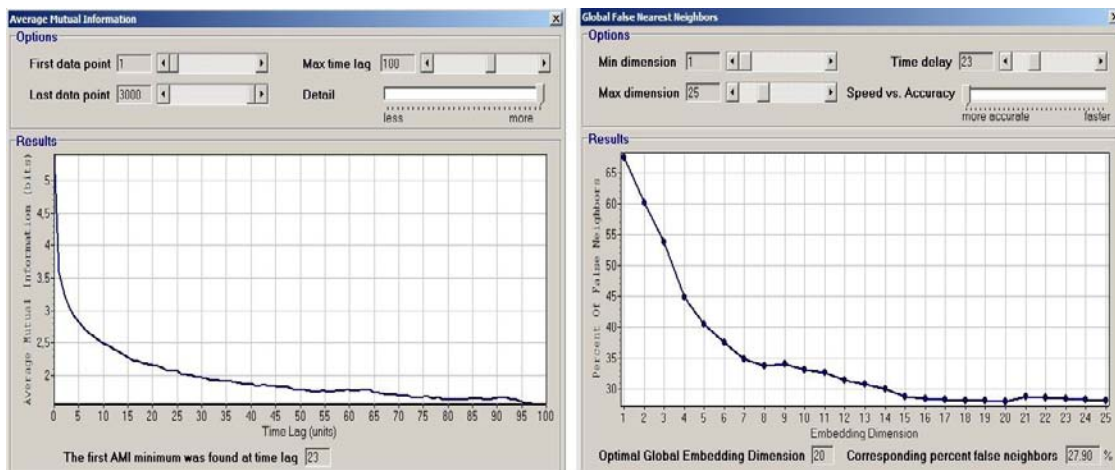


Рис. 2. Визначення оптимальних значень затримки часу t та розмірності вкладень m

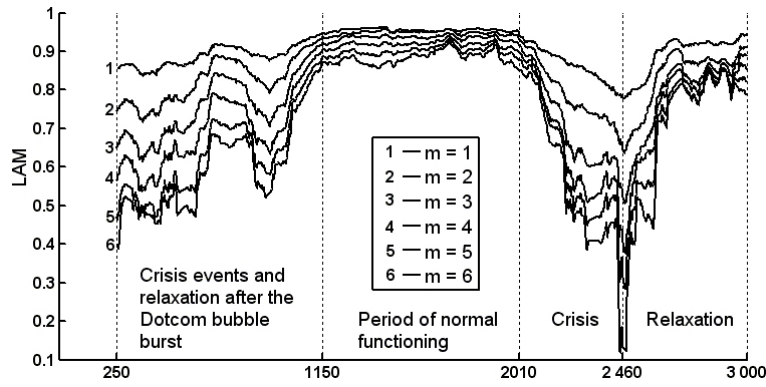


Рис. 3. Міри LAM, розраховані при різних значеннях розмірності вкладень

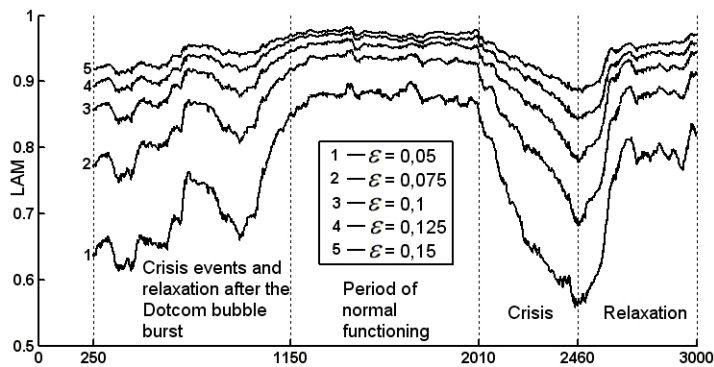


Рис. 4. Міри LAM, розраховані при різних значеннях порогового радіусу

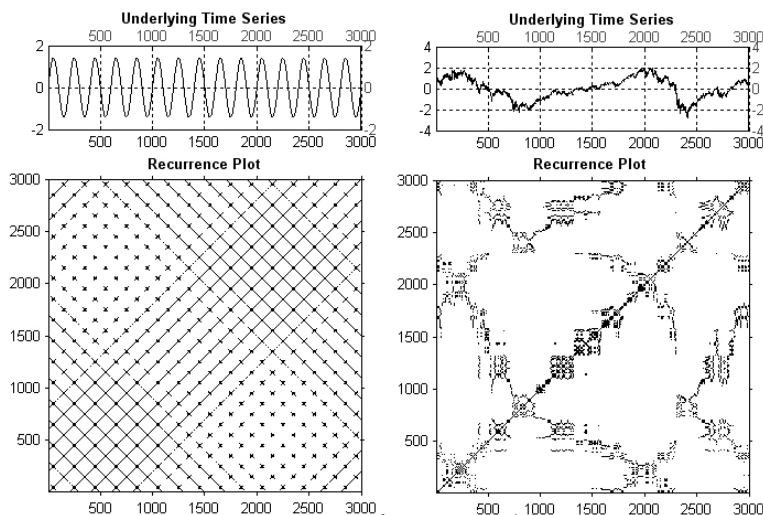


Рис. 5. Рекурентні діаграми функції $\sin(x)$ та індексу S&P500 з 09.08.1999 по 11.07.2011 р. відповідно

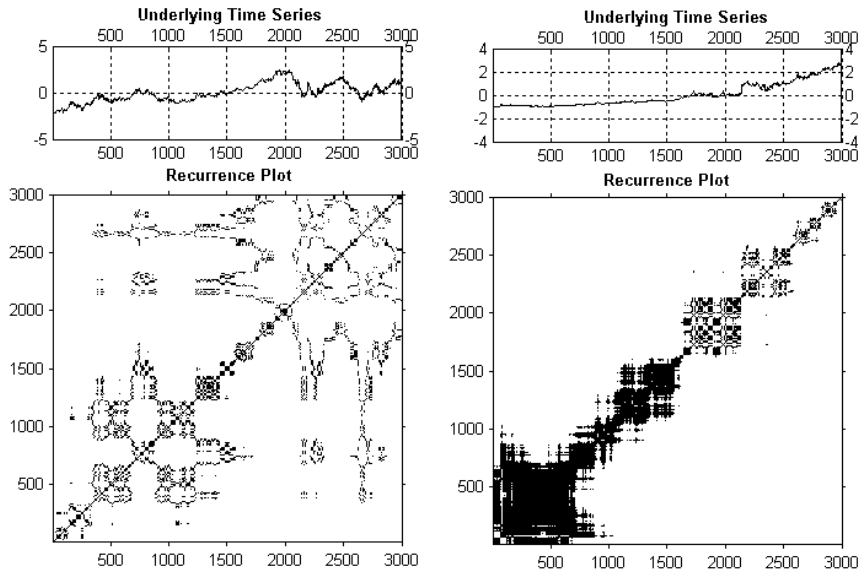


Рис. 6. Рекурентні діаграми валютної пари eur/usd з 19.12.2002 по 11.07.2011 р. та золота з 10.08.1999 по 29.07.2011 р. відповідно

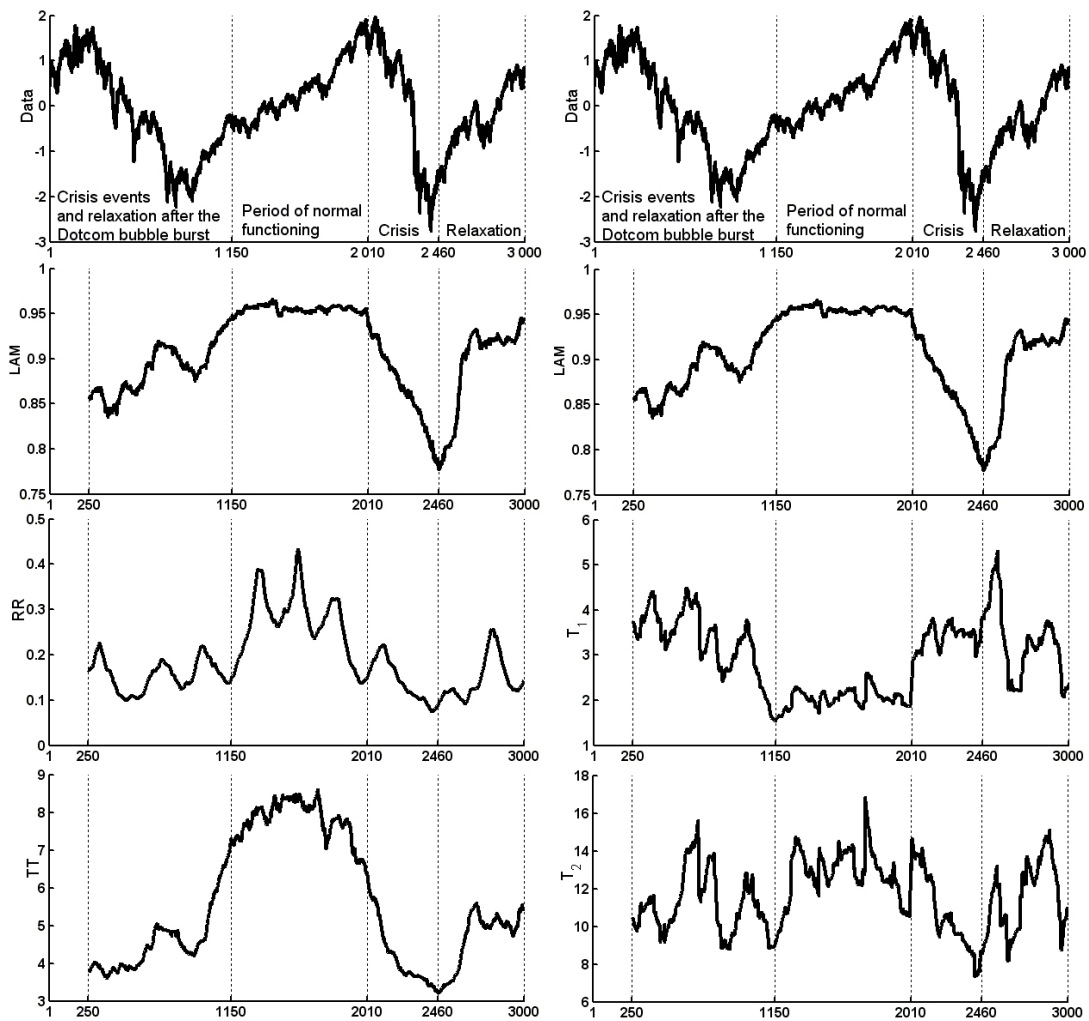


Рис. 7. Котировання фондового індексу S&P500 з 09.08.1999 по 11.07.2011 р. та відповідні міри LAM, TT, T1 та T2

Різні значення порогового радіусу суттєво не змінюють вигляд міри LAM. Більші радіуси (за 0,1) підвищують значення LAM та зменшують падіння міри під час кризового періоду ($x = 2010 - 2460$). Менші радіуси (за 0,1), навпаки, знижують значення LAM та збільшують падіння міри. При зміні порогового радіусу результати не покращуються, тому вибираємо його рівним 0,1.

Інший аспект коректного застосування RQA – це вибір відповідних для дослідження мір [21]. Проаналізуємо рекурентні діаграми різних процесів (рис.5, 6).

Для аналізу траєкторії синуса треба використовувати міри, що розраховують на основі діагональних ліній, оскільки вертикальні лінії у діаграмі не представлені. У випадку фінансових та товарних ринків потрібно брати міри, що розраховуються на базі вертикальних (горизонтальних) ліній, оскільки економічні часові ряди формують лише кластери рекурентних точок, а діагональні лінії відсутні.

Таким чином, для аналізу ринків можуть бути використані міри RR, LAM, TT, T1 та T2.

Розглянемо поведінку цих мір у різні періоди функціонування ринку (рис. 7).

Візуальний аналіз міри LAM дозволяє розпізнавати різні періоди функціонування ринку. Коли ринок нормально функціонує, траєкторія міри має горизонтальний тренд у деякому діапазоні. Падіння LAM свідчить про кризові явища, а підвищення – про релаксацію та відновлення.

Інші міри не мають такої чіткої динаміки, ідентифікувати точки переходу між періодами досить складно. Тому з усіх мір рекурентного кількісного аналізу LAM є найбільш придатною для дослідження стану фінансового ринку.

Висновки. У цій роботі розглянуто теорію та методи рекурентного аналізу. За допомогою загальноприйнятих алгоритмів та емпіричного аналізу встановлено, що для роботи з фінансовими часовими рядами для завдань, коли важливо відстежувати стан системи, треба застосовувати такі вхідні параметри: m та τ дорівнюють 1, $\epsilon - 0,1$. Найбільш інформативною мірою є міра ламінарність.

Список використаних джерел

1. McKenzie M. D. Chaotic behavior in national stock market indices / M. D. McKenzie // *Global Finance J.* – 2001. – №12. – P. 35–52.
2. Antoniou A. Recurrence Plot and financial time series analysis / A. Antoniou, C. E. Vorlow // *Neural Network World.* – 2000. – № 10. – P. 131–145.
3. Gilmore C. G. An examination of nonlinear dependence in exchange rates, using recent methods from chaos theory / C. G. Gilmore // *Global Finance Journal.* – 2001. – V. 12. – № 1. – P. 139–151.
4. Zbilut J. P. Use of Recurrence Quantification Analysis in Economic Time Series / J. P. Zbilut // *New Economic Windows: Economics – Complex Windows*, Eds.: M. Salzano, A. Kirman. – Milan: Springer, 2005. – P. 91 – 104.
5. Strozzi F. Application of non linear time series analysis techniques to high frequency currency exchange data / F. Strozzi, J. M. Zaldivar, J. P. Zbilut

// *Physica A.* – 2002. – № 312. – P.520–538.

6. Application of non-linear time series analysis techniques to the nordic spot electricity market data / F. Strozzi, E. Gutierrez, C. Noè et al. – *Liuc Papers n. 200 – Serie Tecnologia 11*, marzo. – 2007. – 51 p.

7. Fabretti A. Recurrence plot and recurrence quantification analysis techniques for detecting a critical regime. Examples from financial market indices / A. Fabretti, M. Ausloos // *Int. J. Mod.Phys. C.* – 2005. – V. 16. – № 5. – P. 671–706.

8. Bastos J. A. Recurrence quantification analysis of global stock markets / J. A. Bastos, J. Caiado // *Physica A.* – 2011. – V. 390. – № 7. – P. 1315 – 1325.

9. Recurrence plot based measures of complexity and its application to heart rate variability data / N. Marwan, N. Wessel, U. Meyerfeldt et al. // *Phys. Rev. E.* – V. 66. – № 2. – 026702.

10. Marwan N. Encounters With Neighbours – Current Developments Of Concepts Based On Recurrence Plots And Their Applications: Ph.D. Thesis / Marwan Norbert. – University of Potsdam, 2003. – 159 p.
11. Sauer T. Embedology / Sauer, J. A. Yorke, M. Casdagli // *J. Statist. Phys.* – V. 65. – №3–4. – 1991. – P. 579–616.
12. Kantz H. Nonlinear Time Series Analysis / H. Kantz. T. Schreiber. – Cambridge: University Press, 1997. – 306 p.
13. Eckmann J. P. Recurrence Plot of dynamical system / J. P. Eckmann, S. O. Kamphorst, D. Ruelle // *Europhys. Lett.* – 1987. – № 5. – P. 973–977.
14. Webber C. L. Dynamical assessment of physiological systems and states using recurrence plot strategies / C. L. Webber, J. P. Zbilut // *J. Appl. Physiol.* – 1994. – V. 76. – № 2. – P. 965 – 973.
15. Zbilut J. P. Embeddings and delays as derived from quantification of recurrence plots / J. P. Zbilut. C. L. Webber // *Physics Letters A.* – 1992. – № 171. – P. 199 – 203.
16. Gao J. B. Recurrence time statistics for chaotic systems and their applications / J. B. Gao // *Phys. Rev. Lett.* – 1999. – V. 83. – № 16. – P. 3178 – 3181.
17. Estimation of dynamical invariants without embedding by recurrence plots / M. Thiel, M. C. Romano, P. L. Read, J. Kurths // *Chaos.* – № 14. – 2004. – P. 234–243.
18. Ambiguities in recurrence-based complex network representations of time series / R. V. Donner, Y. Zou, J. F. Donges et al. // *Physical Review E.* – 2010. – V. 81. – № 1. – 015101(R).
19. Schinkel S. Selection of recurrence threshold for signal detection / S. Schinkel, O. Dimigen, N. Marwan // *European Physical Journal, Special Topics.* – 2008. – № 164. – P. 45–53.
20. Marwan N. How to avoid potential pitfalls in recurrence plot based data analysis / N. Marwan // *International Journal of Bifurcation and Chaos.* – 2011. – V. 21. – № 4. – P. 1003–1017.
21. Mhalsekar S. D. Investigation on feasibility of recurrence quantification analysis for detecting flank wear in face milling / S. D. Mhalsekar, S. S. Rao, K. V. Gangadharan // *International Journal of Engineering, Science and Technology.* – 2010. – V. 2. – № 5. – P. 23–38.
22. Piskun O. Recurrence Quantification Analysis of Financial Market Crises and Crashes [Електронний ресурс] / O. Piskun, S. Piskun. – Режим доступу: <http://arxiv.org/pdf/1107.5420.pdf>