В.А. Даниленко

АЛЬТЕРНАТИВНІ МЕТОДИКИ ПРОВЕДЕННЯ ФРАКТАЛЬНОГО АНАЛІЗУ

Останнім часом набувають популярності дослідження, спрямовані на врахування ефектів нелінійності у функціонуванні систем різноманітної природи, зокрема соціально-економічних систем. У результаті синтезу наукових знань у цьому напрямі досліджень почала формуватися міждисциплінарна наукова течія – синергетика. Один із напрямів дослідження в рамках синергетики спрямований на вивчення ефектів нелінійності в динаміці часових рядів, зокрема таких, що пов'язані з відхиленням характеру розподілу випадкових величин від нормального. Цей напрям називається фрактальним аналізом.

Фрактальний аналіз містить проведення кількох метричних тестів, які виділяються в роботі Л.Н. Сергєєвої: «…обчислення кореляційної розмірності, яка є нижньою оцінкою фрактальної розмірності; обчислення максимального показника Ляпунова; оцінка К-ентропії Колмогорова; обчислення показника Херста; тест залишків Брока; BDS-тест …» [1, *6*].

У рамках даного дослідження акцентовано увагу саме на обчисленні показника Херста, який визначається в результаті проведення R/S-аналізу. Процедура R/S-аналізу розглядається в роботах Є. Федера [2, *152-154*]; Е. Петерса [3, *107-109*; 4, *69-70*]; Н.К. Максишко, В.О. Перепелиці [5, *99-100*]; О.В. Михайловської [6, *220*] та інших дослідників. При цьому застосовуються різні системи позначень, що формалізують різні алгоритми проведення цього аналізу. До того ж застосування різних алгоритмів може призводити до істотно відмінних результатів обчислення числових значень показника Херста. Усе це ускладнює сприйняття й інтерпретацію значень показника Херста, а можливі розбіжності значень цього показника, обчисленого за різними методиками, до того ж викликають сумніви в надійності результатів такого аналізу.

Метою статті є дослідження альтернативних методик проведення фрактального аналізу при визначенні показника Херста, виявлення спільних та відмінних рис окремих методик, сфери їх застосування.

Розглянемо методики проведення R/S-аналізу, який дозволяє визначити фрактальні характеристики часового ряду, такі як показник Херста та показник фрактальної розмірності. Найбільш відомі модифікації вказаних методик наведені в роботах Е. Петерса [3, *107-109*; 4, *69-70*], де пропонуються дуже схожі методики, що відрізняються способом поділу загального часового ряду на досліджувані підінтервали. Детальніше сутність цієї відмінності буде розглянута пізніше. Задля ідентифікації базової методики проведення R/S-аналізу домовимося називати її методикою Петерса. Для забезпечення кращого порівняння сутності даної методики з альтернативними підходами при її розгляді використаємо систему позначень, наведену в роботі Н.К. Максишко, В.О. Перепелиці [5, *99-100*].

Отже, нехай розглядається часовий ряд Z, що складається з елементів z_i:

$$Z = \{z_i\},\ i = 1, 2, \ldots, m, \quad (1)$$

де m – кількість спостережень, що формують ряд.

Першим кроком процедури R/S-аналізу, за методикою Е. Петерса, є перехід від вихідного ряду Z до ряду Y, що складається з m-1 елементів y_i, кожен з яких отримується за формулою

$$y_i = \lg(z_i) / \lg(z_{i-1}),\ i = 1, 2, \ldots, m-1. \quad (2)$$

Проте в роботі Н.К. Максишко, В.О. Перепелиці, де наводиться ця ж методика, перший крок пропускається. Дійсно, доцільність переходу від вихідних елементів ряду до відношення логарифмів наступного та попереднього рівнів ряду не є незаперечною і залежить від специфіки досліджуваних даних. Ураховуючи

© Даниленко В'ячеслав Анатолійович – аспірант.
Київський національний університет імені Тараса Шевченка.

зазначене, теж пропустимо перший крок, маючи на увазі, що вказана процедура є різновидом додаткової обробки вихідних даних, і в разі необхідності до вихідного ряду може бути застосована різного роду попередня обробка, а не лише така, як наведена у формулі (2).

Етап 1. Отже, першим етапом буде визначення величини кроку $\Delta \geq 1$ та формування послідовності довжин відрізків, на які розбивається вихідний ряд:

$$n_1, n_2, \ldots, n_k, \ldots, n_l, \tag{3}$$

де $n_{k+1} = n_k + \Delta$, для всіх k від 1 до l-1. Максимальне значення l визначається за формулою: $n_l \leq [m/2]$. Ця вимога означає, максимальна довжина відрізка n_l має дозволяти розподілити загальний ряд довжини m хоча б на два відрізки однакової довжини (з однаковою кількістю елементів).

Такий підхід до формування послідовності довжин відрізків (3) пропонується в роботі Е. Петерса [3, *107-108*]. Проте пізніше він вводить додаткові обмеження на значення n_k., а саме пропонує емпіричним шляхом отримане правило, що $n_k \geq 10$, тобто мінімальна довжина відрізка має бути не менше 10. Крім того, пропонується обирати лише ті значення n_k з набору (3), які є дільниками числа m, тобто m має бути кратне n_k [4, *70*]. Саме такі відмінності відрізняють більш ранню й більш пізню модифікації методик R/S-аналізу, що пропонуються в роботах Е. Петерса.

Надалі домовимося під методикою Петерса мати на увазі більш пізній варіант, який передбачає врахування додаткових обмежень на значення n_k (що $n_k \geq 10$, а m має бути кратне n_k).

Подальші етапи реалізуються для всіх значень k від 1 до l.

Етап 2. Для чергового значення індексу k ряд Z розбивається на відрізки довжини n_k, кількість яких r_k визначається за співвідношенням: $r_k = m/n_k$. Позначимо такі відрізки Z_k^t, їх елементи $\{ z_j^t \}$, $j = 1, 2, \ldots, n_k$; t = 1, 2, … r_k. Відрізки Z_k^t мають таку властивість, що вони не перетинаються (тобто містять унікальні елементи вихідного ряду). Для кожного відрізка Z_k^t визначається середнє значення z^t за формулою

$$z^t = \frac{1}{n_k} \sum_{j=1}^{n_k} z_j^t, \; t = \overline{1, r_k}. \tag{4}$$

Етап 3. Для кожного відрізка Z_k^t обчислюється ряд накопичених відхилень, елементи якого x_j^t визначаються за формулою

$$x_j^t = \sum_{i=1}^{j} (z_i^t - z^t), \; j = \overline{1, n_k}. \tag{5}$$

Формула для визначення накопичених відхилень (5) у роботі Н.К. Максишко, В.О. Перепелиці наводиться дещо в іншому вигляді. Ними задається додатковий індекс, який виконує такі ж функції, що й індекс j у формулі (5). Проте нижнє значення діапазону зміни цього індексу починається не з 1, а з 3 [5, *100*]. Тобто перші два елементи кожного відрізка мають пропускатися. Це не відповідає алгоритму, наведеному в роботі Е. Петерса, і, оскільки це окремо не пояснюється, то нами залишається індексація, яка запропонована Е. Петерсом [4, *69-70*]. На основі значень накопичених відхилень для кожного відрізка Z_k^t визначається величина розмаху за формулою

$$R_k^t = \max_{1 \leq j \leq n_k} x_j^t - \min_{1 \leq j \leq n_k} x_j^t, \; \forall\, t = \overline{1, r_k}. \tag{6}$$

Етап 4. Для кожного відрізка Z_k^t обчислюється його середньоквадратичне відхилення за формулою

$$S_k^t = \sqrt{\frac{1}{n_k} \sum_{j=1}^{n_k} (z_j^t - z^t)^2}, \; \forall\, t = \overline{1, r_k}. \tag{7}$$

Після цього для кожного відрізка визначається нормоване (на середньоквадратичне відхилення) значення розмаху:

$$\left(\frac{R}{S} \right)_k^t = \frac{R_k^t}{S_k^t}, \; \forall\, t = \overline{1, r_k}. \tag{8}$$

Етап 5. Для кожного k обчислюється середньоарифметичне значення нормованих розмахів, визначених за формулою (8):

$$\left(\frac{R}{S} \right)_k = \frac{1}{r_k} \sum_{t=1}^{r_k} \left(\frac{R}{S} \right)_k^t, \; \forall\, k = \overline{1, l}. \tag{9}$$

Етап 6. Для кожного k обчислюється логарифм довжини відрізка n_k, що становитиме абсцису точки при подальшій побудові графіка (x_k), а також логарифм усередненого значення нормованих розмахів $(R/S)_k$, обчисленого за формулою (9). Останній становитиме ординату зазначеної точки (y_k). У символьному вигляді це такі формули:

$$x_k = \lg n_k,\ y_k = \lg\left(\frac{R}{S}\right)_k,\ k = \overline{1,l}. \quad (10)$$

Етап 7. Використовуючи отримані за формулою (10) значення x_k та y_k, методом найменших квадратів будується рівняння лінійної регресії:

$$y = a \cdot x + b. \quad (11)$$

Коефіцієнт a в побудованому на основі фактичних даних x_k та y_k рівнянні (11) становить усереднену оцінку показника Херста H для часового ряду Z.

Н.К. Максишко та В.О. Перепелиця, коментуючи отриману за наведеною методикою оцінку показника Херста, підкреслюють, що ця оцінка являє собою саме усереднене для цього ряду значення показника Херста.

У роботі О.В. Михайловської наводиться більш проста методика визначення показника Херста, яку вона називає «швидкий алгоритм» [6, *220*]. Розглянемо її докладніше і при цьому будемо використовувати вже задану нами раніше систему позначень, згідно з якою аналізується часовий ряд Z, що складається з елементів z_i, кількість яких дорівнює m.

Етап 1. Визначаються середні значення для інтервалів часового ряду, що послідовно збільшуються від 1 елементу до m, за формулою

$$z^i = \frac{1}{i}\sum_{j=1}^{i} z_j,\ i = \overline{1,m}.$$

Етап 2. Обчислюються накопичені відхилення для кожного інтервалу i за формулою

$$x^i = \sum_{j=1}^{i}(z_j - z^i), i = \overline{1,m}.$$

Етап 3. Визначається максимальне та мінімальне відхилення x^i для кожного з m можливих інтервалів, позначимо їх $\max_i(x_1, x_2, ..., x_i)$ та $\min_i(x_1, x_2, ..., x_i)$ відповідно. Також визначається розмах накопиченого відхилення при різних значеннях довжини інтервалу i за формулою:

$$R_i = \max_i(x_1, x_2, ..., x_i) - \min_i(x_1, x_2, ..., x_i), i = \overline{1,m}.$$

Етап 4. Для кожного інтервалу знаходиться середньоквадратичне відхилення:

$$S_i = \sqrt{\frac{1}{i}\sum_{j=1}^{i}(z_j - z^i)^2},\ i = \overline{1,m}.$$

Етап 5. Розмах накопиченого відхилення (R_i) для кожного інтервалу i нормалізується шляхом ділення на S_i. У результаті отримуємо ряд значень (R_i/S_i) для кожного інтервалу i.

Етап 6. Значення (R_i/S_i) та i логарифмуються й будується графік залежності $\lg(R_i/S_i)$ від $\lg(i)$.

Етап 7. Методом найменших квадратів будується рівняння лінійної регресії, коефіцієнт при незалежній змінній якого й буде показником Херста.

Слід зазначити, що етапи 6 і 7 обох розглянутих методик є цілком тотожними. Якщо ж порівнювати складність цих двох алгоритмів визначення показника Херста, то варто погодитися, що другий із них значно простіший. Хоча, можливо, ця простота не очевидна, оскільки формально ми отримали також сім етапів. Проте в дійсності другу методику можна реалізувати взагалі без допомоги програмування, просто скориставшись формулами електронних таблиць. Тоді як розглянута раніше методика Петерса доволі трудомістка, а без програмування її реалізувати практично неможливо.

Також є роботи, у яких наводиться дуже схожа на «швидкий алгоритм» методика визначення показника Херста. Так, у роботі Є. Федера розглядається методика, яка охоплює етапи 1-5 «швидкого алгоритму» [2, *152-154*]. Після цього пропонується формула для обчислення показника Херста, що має такий вигляд:

$$\frac{R_i}{S_i} = \left(\frac{i}{2}\right)^H.$$

Хоча це не вказано явно, для того щоб отримати оцінку показника Херста для всього ряду в цілому, слід прийняти $i=m$. Також у роботі Л.Н. Сергєєвої наводиться формула для визначення показника Херста за цією методикою [1, *55*]:

$$H = \frac{\ln(R/S)}{\ln(m/2)}.$$

Те, що в даній формулі фігурують натуральні логарифми, а не десяткові, як у попередніх методиках, не є принциповим і на значення показника Херста не впливає.

Крім простоти з точки зору технічної реалізації, «швидкий алгоритм» має ще й таку перевагу, що він не висуває жорстких вимог до кількості спостережень в аналізованому ряді даних. Проте перш ніж аналізувати відносні переваги та недоліки наведених вище методик, варто зазначити, що в роботі Н.К. Максишко та В.О. Перепелиці пропонується вдосконалення методики Петерса, головною перевагою якого є кращі результати, які дає вдосконалена методика при визначенні тривалості циклів. Розглянемо зміст цієї методики, домовившись для зручності називати її «методикою Максишко» [5, *105*]. Самі автори називають цю методику «алгоритмом послідовного R/S-аналізу».

Як і в попередньому випадку, при розгляді цієї методики будемо використовувати вже задану нами раніше систему позначень, згідно з якою аналізується часовий ряд Z, що складається з елементів z_i, загальна кількість яких дорівнює m.

Етап 1. Для заданого часового ряду Z розглядаються його початкові відрізки $Z_\tau = z_1, z_2, \ldots, z_\tau$, $\tau = 3, 4, \ldots, m$. Для кожного з них обчислюється їх середнє значення за формулою

$$\overline{z_\tau} = \frac{1}{\tau}\sum_{i=1}^{\tau} z_i, \ \tau = 3,4,\ldots,m,$$

а також знаходиться накопичене відхилення:

$$x_{\tau\, t} = \sum_{i=1}^{t}(z_i - \overline{z_\tau}), \forall t = \overline{1,\tau}.$$

Етап 2. Для кожного відрізка Z_τ відповідно до формули (6) обчислюється розмах. Із врахуванням специфіки індексів, більш точно ця формула матиме такий вигляд:

$$R_\tau = \max_{1\le t\le \tau} x_{\tau\, t} - \min_{1\le t\le \tau} x_{\tau\, t}, \ \forall \tau = \overline{3,m}.$$

Також для цього ж відрізка Z_τ визначається середньоквадратичне відхилення за формулою

$$S_\tau = \sqrt{\frac{1}{\tau}\sum_{i=1}^{\tau}\left(z_i - \overline{z_\tau}\right)^2}, \ \forall \tau = \overline{3,m}.$$

Після цього знаходиться нормований розмах шляхом ділення розмаху R_τ на середньоквадратичне відхилення S_τ.

Як певне пояснення до перших двох кроків слід зазначити, що вони в цілому відповідають першим шести етапам «швидкого алгоритму», наведеного в роботі О.В. Михайловської. Відмінність у кількості етапів пояснюється різною глибиною деталізації по суті тієї самої процедури. Єдина змістовна відмінність полягає у виборі мінімальної довжини початкових відрізків Z_τ. За методикою Максишко пропонується мінімальне значення τ обрати на рівні 3, тоді як у роботі Михайловської мінімальна довжина відрізка становить 1. Коментуючи цю розбіжність, слід зазначити, що з обчислювальної точки зору обирати мінімальну довжину відрізка на рівні 1 недоцільно. Оскільки в такому випадку розмах R_τ та середньоквадратичне відхилення S_τ гарантовано будуть нульовими. Разом із тим, якщо обирається значення більше 2, то такий підхід теж потребує пояснень. Узагалі у методиці фрактального аналізу є ряд правил, які носять емпіричний характер і не мають відповідного теоретичного обґрунтування. Проте в даному випадку вибір авторами мінімального значення τ на рівні 3 має пояснення, яке буде наведене при розгляді етапу 3.

Етап 3. Спираючись на наведений у роботі Е. Петерса «емпіричний закон Херста», будується H-траєкторія шляхом розрахунку значень $H(\tau)$ за формулою [3, *99*]:

$$H(\tau) = \left(\lg(R(\tau)/S(\tau))\right)/\lg(\tau/2).$$

При побудові H-траєкторії обчислюється ордината точки $y_\tau = H(\tau)$ за наведеною вище формулою, а також абсциса точки $x_\tau = \lg(\tau/2)$ для всіх $\tau = 3, 4, \ldots, m$.

Як видно з наведеної формули для обчислення $H(\tau)$, у її знаменнику знаходиться вираз $\lg(\tau/2)$. При $\tau = 2$ знаменник перетворюється на нуль і виникає відповідна арифметична помилка ділення на нуль. Для уникнення цього за методикою Максишко і пропонується обирати $\tau \ge 3$.

Етап 4. Будується R/S-траєкторія шляхом формування послідовності точок (x_τ, y^o_τ). Координата x_τ для кожної точки визначається за формулою $x^o_\tau = \lg(\tau/2)$, а координата за формулою $y_\tau = \lg(R(\tau)/S(\tau))$.

До переваг методики Максишко, або алгоритму послідовного R/S-аналізу, слід віднести те, що вона, як і швидкий алгоритм, допускає реалізацію з використанням лише стандартних можливостей електронних таблиць. Крім того, ця методика може бути ефективно використана для визначення тривалості циклів і квазіциклів.

Висновки. Розглянуто три альтернативних підходи до проведення фрактального аналізу шляхом обчислення показника Херста. Перша з розглянутих методик, яку умовно названо методикою Петерса, попри свою обчислювальну трудомісткість і вибагливість до обсягів інформаційних масивів, є найкращою з точки зору отримання абсолютного значення показника Херста для ряду даних у цілому.

Друга методика, що названа «швидким алгоритмом», має перевагу, яка полягає у простоті реалізації та технічній можливості обчислення показника для обмежених за кількістю спостережень інформаційних масивів. Проте точність таких обчислень є недостатньо високою.

Перевагою третьої методики, умовно названої «методикою Максишко», є також відносна простота обчислювальних процедур, а також можливість ідентифікації з її допомогою тривалості циклів та квазіциклів.

Перспективи подальших досліджень у цьому напрямі полягають у розробці методів фрактального аналізу, які дозволяють ідентифікувати атрактори, потрапляння в зону дії яких призводить до ефекту «втрати пам'яті», наявність якої діагностується за допомогою обчислення показника Херста. Це важливо для прогнозування катастрофічних змін у динаміці систем, що характеризуються радикальною зміною минулих тенденцій і втратою ними стійкості.

Література

1. Сергеева Л.Н. Моделирование поведения экономических систем методами нелинейной динамики (теории хаоса) / Л.Н. Сергеева. – Запорожье: Запорожский гос. ун-т, 2002. – 227 с.

2. Федер Е. Фракталы: пер. с англ. / Е. Федер. – М.: Мир, 1991. – 254 с.

3. Петерс Э. Хаос и порядок на рынках капитала. Новый аналитический взгляд на циклы, цены и изменчивость рынка: пер. с англ. / Э. Петерс. – М.: Мир, 2000. – 333 с.

4. Петерс Э. Фрактальный анализ финансовых рынков: применение теории хаоса в инвестициях и экономике / Э. Петерс. – М.: Интернет-трейдинг, 2004. – 304 с.

5. Максишко Н.К. Анализ и прогнозирование эволюции экономических систем / Н.К. Максишко, В.А. Перепелица; Запорожский национальный ун-т. – Запорожье: Полиграф, 2006. – 236 с.

6. Михайловська О.В. Самоорганізація світового інвестиційного процесу в умовах глобалізації: можливості фрактального аналізу / О.В. Михайловська // Актуальні проблеми економіки. – 2009. – № 1 (91). – С. 218-228.