

## Застосування та ефективність методів пошуку розладок при аналізі нестационарних процесів спекулятивного валютного ринку

ЛУЦЕНКО О.П., БАЙБУЗ О.Г.

Дніпропетровський національний університет ім. О. Гончара

Запропоноване використання методів пошуку розладок для визначення точок розвороту тренду спекулятивного валютного ринку. Проведено порівняльний аналіз ефективності використання запропонованих методів.

Предложено использование методов поиска разладок для определения точек разворота тренда спекулятивного валютного рынка. Проведен сравнительный анализ эффективности использования методов.

Proposed using disorder detection algorithms to detect points of trend change on the speculative foreign exchange market. A comparative analysis of the methods' efficiency has been carried out.

Протягом 20 років з моменту приходу в Україну ринок валют Forex набрав величезну популярність завдяки своїй доступності і великій потенційній прибутковості. Висока прибутковість торгівлі на Forex поєднується з ризиками, для мінімізації яких застосовується фундаментальний і технічний аналіз ринку. Фундаментальний аналіз досліджує причини зміни цін, у той час як технічний аналіз керується твердженням «курс враховує все» і ґрунтується лише на поточній поведінці курсу валют, без врахування причин зміни. В силу надто великої кількості взаємозалежних факторів, що впливають на курс валют, технічний аналіз відіграє домінуючу роль.

Оскільки ринок Forex є наддинамічним, точне прогнозування значення валютного курсу в майбутньому на певний момент часу є неможливим. Завданням технічного аналізу є виявлення напрямку і сили діючої ринкової тенденції (тренду), а також моментів зміни напряму тенденції (розвороту тренду) шляхом аналізу історії зміни цін на ринку. Виділяють три види тренду: висхідний (бичачий), спадний (ведмежий) і бічний (флет).

Серед методів, що найчастіше застосовуються у технічному аналізі, можна виділити наступні: допоміжні побудови на графіках курсів валют (лінії опору та підтримки, лінії тренду тощо), індикатори і осцилятори. Існує велика кількість різновидів і модифікацій цих методів, але для більшості з них характерні спільні недоліки:

1) суб'єктивність допоміжних побудов, що унеможливує їхню автоматизацію;

2) значна затримка у часі, викликана використанням у розрахунках індикаторів і осциляторів ковзних середніх. Затримка стає тим більшою, чим більший порядок ковзної середньої. Враховуючи, що для збільшення достовірності сигналу про розворот тренду потрібне збільшення порядку ковзної середньої, можна стверджувати, що затримка у часі різко зростає зі збільшенням достовірності визначення;

3) велика кількість хибних сигналів розвороту тренду при використанні осциляторів.

У зв'язку з зазначеними недоліками економічних методів була поставлена задача запропонувати такий метод технічного аналізу валютних котировань, який давав би мінімальну затримку у часі виявлення точок розвороту тренду, будучи надійним в умовах як стійкої бичачої або

ведмежою тенденції, так і флету.

Слід відмітити, що нестабільність досліджуваного часового ряду, викликана високою динамічністю ринку Forex, унеможливує ефективний аналіз даних в їхньому вихідному вигляді, тому перед дослідженням ряду необхідна додаткова первинна обробка — згладжування. Часовий ряд валютних котировань було згладжено шляхом апроксимації В-сплайнами, після чого решта операцій проводилася зі згладженою кривою.

Вибір В-сплайн апроксимації у якості методу згладжування пояснюється наступними причинами [1]:

1) даний метод не дає лага за часом, на відміну від традиційно застосовуваних в економічному аналізі методів експоненційного згладжування і ковзних середніх;

2) крива, отримувана на виході, має прийнятну точність апроксимації, і, що важливо для задачі пошуку розладок, не має різких перегинів;

3) метод дозволяє розрахувати значення проміжних точок у випадку, якщо для розрахунку використовуються не всі точки вибірки;

4) існують відносно швидкі способи обчислення точок В-сплайну, що критично в даному випадку, так як на тикових і хвилинних графіках обсяги вибірки досягають розмірності порядку тисяч;

5) ступінь згладжування не залежить від кількості точок у виборці.

При рішенні задачі були використані однорідні В-сплайни 4-го порядку з алгоритмом побудови вузлового вектору, що найчастіше використовується при побудові В-сплайнів. Базисна функція в такому випадку однакова для кожного сегменту і в матричній формі має наступний вигляд:

$$S_i(t) = [t^3 t^2 t 1] \begin{pmatrix} \frac{1}{6} \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix} \quad (1)$$

де  $t \in [0,1]$

Звідси випливає, що для кожної точки кривої справедлива наступна систем рівнянь:

$$x(t) = ((a_3 * t + a_2)t + a_1)t + a_0, \\ \text{для } 0 \leq t \leq 1, \text{ де:}$$

$$\begin{aligned} a3 &= (-x_{i-1} + 3x_i - 3x_{i+1} + x_{i+2})/6 \\ a2 &= (x_{i-1} - 2x_i + x_{i+1})/2 \\ a1 &= (-x_{i-1} + x_{i+1})/2 \\ a0 &= (x_{i-1} + 4x_i + x_{i+1})/6 \end{aligned} \quad (2)$$

Керування ступенем наближення згладженої кривої до вихідної здійснюється шляхом вибору кратності точок кривої спостережень, що використовуються при побудові В-сплайну.

Для рішення задачі застосовані наступні методи пошуку розладок: алгоритм кумулятивних сум (АКС) для зміни середнього, АКС для зміни дисперсії, метод заключення в коридор і визначення розладок виміром кутів між апроксимаційними прямими.

Алгоритм кумулятивних сум для зміни середнього ґрунтується на послідовному обчисленні значень статистики  $R$  згідно значенням  $x$  в реальному часі. Підраховані значення  $R$  зіставляють з порогом  $h$ . Якщо  $R < h$ , то процес контролю триває, інакше — дається сигнал про присутність розладки [2].

Послідовні значення функції  $R_n$  розраховуються рекурентно, виходячи зі значень статистики на попередньому кроці  $R_{n-1}$  і наступного значення  $x_i$ .

При розрахунках використовується наступна інформація:

$m_1$  — значення математичного очікування до розладки;

$m_2$  — значення математичного очікування після розладки;

$\sigma_2$  — вважаємо, що дисперсія не змінювалася.

Формула розрахунку кумулятивної суми:

$$\begin{aligned} R_i &= \begin{cases} \max(0, R_{n-1} + g(i)), & \text{якщо } i > 0 \\ 0, & \text{якщо } i = 0 \end{cases} \\ g(i) &= \begin{cases} \frac{x_i - m}{\sigma}, & m_1 < m_2 \\ \frac{m - x_i}{\sigma}, & m_1 > m_2 \end{cases} \quad m = \frac{m_1 + m_2}{2} \end{aligned} \quad (3)$$

Момент розладки  $N$  визначається наступним чином:  $N = \inf(i: R_i > h)$ .

Метод заснований на переборі всіх моментів часу  $t$  (передбачуваних моментів розладки). На інтервалах  $[t - w; t - l]$ ,  $[t; t + w]$  здійснюється підрахунок значень статистики  $m_1$ ,  $m_2$  відповідно. Кумулятивна сума накопичується до досягнення заданої межі в тих точках, значення курсу валют в яких лежить вище середнього на інтервалі при зростанні середнього (висхідному тренді) і нижче середнього при спадному середньому (на спадному тренді), інакше значення суми зменшується.

Величина порогу  $h$  впливає на чутливість методу і затримку при виявленні розладки. При високому значенні  $h$  кумулятивна сума довше буде накопичуватися до заданої межі і в певних випадках може почати зменшуватися, так і не досягнувши цієї межі. У той же час, при дуже низькому значенні порогу збільшиться кількість хибних сигналів.

Недоліком методу є те, що алгоритм буде видавати сигнали про розладку у всіх точках часового ряду, де тренд буде мати постійну тенденцію до зростання або спаду. Тому точки розладок потрібно видалити на таких ділянках, де можна провести апроксимуючу пряму, і відхилення часового ряду від неї не буде перевищувати встановленого бар'єру.

Вказаного недоліку позбавлений алгоритм кумулятивних сум для зміни дисперсії.

Для використання цього алгоритму нам необхідна наступна інформація:

$\sigma_1$  — значення дисперсії до розладки;

$\sigma_2$  — значення дисперсії після розладки;

$m$  — вважаємо, що математичне очікування не змінювалося.

Значення  $R_n$  обчислюються за формулою, подібною до попереднього випадку:

$$\begin{aligned} R_i &= \begin{cases} \max(0, R_{n-1} + g(i)), & \text{якщо } i > 0 \\ 0, & \text{якщо } i = 0 \end{cases} \\ g(i) &= \begin{cases} \frac{(x_i - m)^2}{\sigma}, & \sigma_1 < \sigma_2 \\ -\frac{(m - x_i)^2}{\sigma}, & \sigma_1 > \sigma_2 \end{cases} \quad \sigma = \frac{\sigma_1 + \sigma_2}{2} \end{aligned} \quad (4)$$

Як і в попередньому випадку, момент розладки  $N$  визначається за принципом  $N = \inf(i: R_i > h)$ .

У даному методі кумулятивна сума накопичується в тих точках, дисперсія в яких вище усередненого значення, (тобто збільшується відхилення від середнього), у іншому випадку значення кумулятивної суми зменшується (так як при зменшенні відхилення від середнього розладки немає).

Алгоритм кумулятивних сум належить до послідовних методів пошуку розладки. Він може бути застосований для визначення розладки у реальному часі, по мірі надходження даних про нові точки часового ряду. На відміну від послідовних методів, апостеріорні методи пошуку розладок, такі як метод заключення в коридор і метод вимірювання кутів між апроксимаційними прямими, потребують деякої вже отриманої сукупності даних. Ці методи дозволяють більш точно визначити положення і кількість точок розладки. В розрізі даної задачі, апостеріорні методи були застосовані для отримання еталонної картини, яка була використана при складанні порівняльної характеристики результатів.

Метод вимірювання кутів між апроксимаційними прямими полягає в наступному: для кожної точки часового ряду знаходяться рівняння двох апроксимаційних прямих, перша з яких найкраще апроксимує задану кількість точок  $l$  часового ряду зліва від заданої точки, а інша — таку ж кількість точок справа від неї. Позначимо:  $x_t$  — значення точки часового ряду, яку спостерігаємо;  $t$  — відповідний момент часу;  $k_1$ ,  $b_1$ ,  $k_2$ ,  $b_2$  — коефіцієнти рівнянь апроксимаційних прямих.

Досягнемо найкращої апроксимації за допомогою рішення наступної системи рівнянь:

$$\begin{aligned} x_t &= k_1 t - b_1, \\ x_t &= k_2 t - b_2, \\ \sum_{i=t-l}^{i=t} (x_i - k_1 t - b_1)^2 &\rightarrow \min \\ \sum_{i=t}^{i=t+l} (x_i - k_2 t - b_2)^2 &\rightarrow \min \end{aligned} \quad (5)$$

Перетворюючи третє рівняння системи, отримуємо:

$$\sum_{i=1}^t x_i t_i = k_1 \sum_{i=1}^t t_i^2 + b_1 \sum_{i=1}^t t_i$$

$$\sum_{i=1}^t x_i = k_1 \sum_{i=1}^t t_i + nb_1$$
(6)

Аналогічно для четвертого рівняння

$$\sum_{i=1}^{t+1} x_i t_i = k_2 \sum_{i=1}^{t+1} t_i^2 + b_2 \sum_{i=1}^{t+1} t_i$$

$$\sum_{i=1}^{t+1} x_i = k_2 \sum_{i=1}^{t+1} t_i + nb_2$$
(7)

З отриманих систем рівнянь (6) і (7) знайдемо невідомі  $k_1, k_2$ , що є тангенціальними коефіцієнтами рівнянь прямих. Знаючи їх, можемо знайти кут  $\varphi_i$  між двома відрізками. Нанісши знайдені кути для всіх точок часового ряду на графік, отримаємо криву, що має вигляд хвилі, вершина кожної з яких відповідає моменту розладки у досліджуваному процесі (рис.1). Чим більше значення піку хвилі, тим більше кут на графіку процесу, отже, тим сильніше змінювався стан системи. Чутливість методу визначається заданням нижнього бар'єру значення піку, вище якого піковий момент  $t$  вважатиметься моментом розладки.

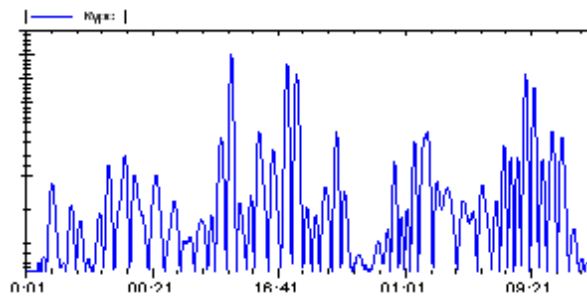


Рис.1. Графік зміни кутів між апроксимаційними прямими

Алгоритм роботи методу закінчення в коридор наступний:

1. Зафіксуємо початковий і кінцевий моменти в процесі, всередині якого будемо шукати точку розладки:  $m$  - початковий момент часу,  $n$  - кінцевий момент часу. На цьому проміжку будемо шукати лише одну точку розладки.

2. Для цього проміжку побудуємо масив

$$\{d_i, i = \overline{n, m}\}$$

$$d_i = \frac{|kx_i - y_i + b|}{\sqrt{k^2 + 1}}$$
(8)

де  $k, b$  - знайдені коефіцієнти рівняння прямої ( $y=kx+b$ ), що проходить через точки  $(m, x_m)$  і  $(n, x_n)$ .

3. Точку розладки знаходимо в такий момент часу  $r$ , що:

$$x_r = \max_{i=m, n} d_i, r \in (m, n), x_r > bar$$
(9)

де  $bar$  - бар'єр чутливості методу.

4. Якщо точка розладки була знайдена, то момент часу  $r$  розбиває проміжок  $[m, n]$  на два:  $[m, r]$  та  $[r, n]$ , які розглядаємо аналогічним способом, переходячи на пункт 1.

5. Процес закінчується, коли ні одне значення  $d_i$  на проміжку не перевищує встановленого бар'єру.

Практично,  $x_r$  - максимальна відстань від

прямої до точки часового ряду на кожному заданому проміжку.

Для оцінки результатів роботи використаних методів створений інформаційно-програмний продукт, який дозволяє застосувати вказані методи пошуку розладки на реальних значеннях часового ряду валютних котирувань і дати чисельну оцінку результатам їх роботи. Програмний продукт реалізований на мові програмування C# з використанням СКБД MS SQL Server. Схема роботи програмного продукту з даними наведена на рис. 2.



Рис. 2. Структура програмного продукту

Використані методи пошуку розладок були порівняні між собою та з методом пошуку точок розвороту тренду за допомогою ковзних середніх за наступними критеріями:

**1. Величина відхилення.** Припустимо, що на часовому ряду  $x_i, i = \overline{0, N-1}$  були знайдені точки розладки у таких моментах часу:  $r_i, i = \overline{0, N-1}$ . Тоді величину відхилення  $D$  обчислюємо наступним чином:

$$D = \sum_{i=0}^{Nr-2} \sum_{j=r(i)}^{r(i+1)} dist_i(x_j)$$

$$dist_i(x_j) = \frac{k_i j - x_j + b_i}{\sqrt{k_i^2 + 1}}$$
(10)

де  $k_i, b_i$  - коефіцієнти рівняння прямої, що проходить через дві точки:  $(r_i, x_{r_i}), (r_{i+1}, x_{r_{i+1}})$ .

Знайдена сума відстаней від точок кривої до прямої між сусідніми точками розладок є мірою апроксимації кривої за допомогою набору прямих. Чим менша сума відхилень, тим точніша апроксимація.

**2. Рівномірність.** Виражене у відсотковому вигляді

число, яке показує, наскільки точки розладки розташовані рівномірно по часовому ряду. Точки розладки можуть бути розташовані дуже нерівномірно, тому рівномірність не обов'язково характеризує якість їх виявлення. Швидше навпаки: надто високе значення рівномірності дає підставу припустити, що метод видає сигнали про розладки там, де розладок немає.

Рівномірність — сума відхилень довжин відрізків часу між розладками від середнього значення відрізка часу між розладками, отже:

$$R = 100\% \left( 1 - \frac{1}{N_r - 1} \sum_{i=0}^{N_r-2} \left| \frac{N}{N_r} - (r_{j+1} - r_j) \right| \right) \quad (11)$$

**3. Середній час затримки сигналу про розладку.** У даному методі знайдений вектор точок розладки  $r_i, i = 0, N_r - 1$  зіставляється з еталонним значенням вектора точок розладки  $r_{ki}, i = 0, N_{rk} - 1$ . В якості еталонного методу застосовується метод заклучення в коридор, який знаходить точки розладки апостеріорно, з

високою точністю і нульовою затримкою. Для кожної точки  $r$  знаходиться відхилення від найближчої зліва точки  $r_k$ . Після чого знаходиться середнє значення відхилення. Щоб помилкові сигнали, якщо такі присутні, не впливали на результати, вводиться поріг чутливості  $p$  методу. Якщо на відрізку  $[i-p, i]$  не знайдено жодної точки контрольного вектора, сигнал вважається хибним.

Необхідно відзначити, що для забезпечення можливості коректного порівняння чутливості методів пошуку розладок повинна бути попередньо налаштована таким чином, щоб всі вони видавали приблизно однакову кількість сигналів.

Ще одним важливим зауваженням є те, що для послідовних методів (АКС для дисперсії та середнього) для визначення середнього часу затримки більш коректно брати не самі точки розладки, а моменти їх виявлення, так як на практиці користувач отримує сигнал заднім числом, з затримкою, що дорівнює ширині класу  $w$ .

Таблиця 1. Результати порівняльного аналізу методів пошуку розладок

Метод	Знайдено точок	Рівномірність, %	Відхилення	Затримка	Хибних сигналів
АКС для дисперсії з перевіркою приналежності точок до прямої	86	53,25	314,09	14,04	14
АКС для дисперсії	102	39,82	394,48	15,91	21
АКС для зміни середнього	88	56,56	236,97	13,69	13
Вимірювання кутів	87	47,21	280,19	—	—
Заклучення в коридор	91	47,39	158,5	—	—
Ковзна середня	89	70,29	321,27	49	13

В результаті порівняння, результати якого приведені в табл. 1, встановлено, що:

1) АКС для зміни середнього дає невеликі значення відхилення при малих значеннях затримки визначення розладок. За умови підключення алгоритму видалення точок розладок, що лежать на одній прямій, він може бути ефективно використаний для визначення точок розвороту тренду. Найбільш ефективним даний метод стає при пошуку точок розвороту тренду в умовах флету;

2) АКС для зміни дисперсії, який є менш точним, але більш універсальним, може використовуватись при будь-яких напрямі і силі діючого тренду. На відміну від економічних методів, що ґрунтуються на ковзних середніх, АКС для зміни дисперсії дає значно меншу величину затримки визначення точок розвороту тренду, яка є менш залежною від параметрів чутливості метода.

Запропоновані методи є більш гнучкими в налаштуванні, ніж більшість індикаторно-осциляторних методів, можуть бути налаштовані на роботу з будь-яким масштабом часового ряду і здатні знаходити точки розладки з малою затримкою у часі.

Для максимізації ефективності технічного аналізу валютних котирувань запропоновано використовувати методи пошуку розладок наряду з традиційними економічними методами при визначенні точок розвороту тренду на фінансовому ринку.

#### ЛІТЕРАТУРА

1. Carl de Boor A Practical Guide to Splines — Springer-Verlag, 1978. — 368 p.
2. Никифоров И.В. Последовательное обнаружение изменения свойств временных рядов. — М.: Наука, 1983.- 198 с.