Федеральное государственное автономное образовательное учреждение высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет экономических наук

КУРСОВАЯ РАБОТА

Методы обнаружения структурных сдвигов в моделях временных рядов

Выполнил:
Студент группы МСМ-181
Кафтанов И. А.

Руководитель:
с.с.
Борзых Д. А.

Москва 2019

Содержание

Введение

Задача анализа и прогнозирования временных рядов является одной из самых востребованных и актуальных: практически во всех областях науки и прикладных отраслей существуют процессы, которые упорядочены во времени. Развитие теорий, методов и технологий в области решения этой задачи в первую очередь подкреплено огромным спросом на анализ подобных процессов.

Временной ряд представляет собой последовательность наблюдений, упорядоченных по какому-либо параметру (обычно, реальному времени или номеру события) для которых порядок наблюдения, играет существенную роль. Существует масса примеров процессов, где значения упорядочены во времени и скоррелированны между собой, и в зависимости от области, в которой данный ряд порождается, перед исследователями ставятся различные цели и задачи по его анализу и прогнозированию. С помощью статистического анализа находятся аномальные значения во временных рядах, прогнозируют его поведение или изучают механизм порождения.

Основным средством для этого служат так называемые модели. Понятие модели включает в себя два ключевых компонента – как модель временного ряда и как прогнозная модель. Модель временного ряда описывает правило генерации членов ряда, а прогнозирование даёт оценку будущих значений членов ряда. Во многих случаях модель можно определить с точностью до конечного числа параметров. Задача получения статистических выводов базируется на данных параметрах. При этом важно понимать, что для нахождения наиболее точных оценок параметров модели, как правило, необходимо большое количество наблюдений.

Такое ограничение подразумевает, что для нахождения оптимальных оценок коэффициентов для различного рода моделей, например эконометрических, требуется большой объем данных (количество наблюдений). Ограничение влечёт за собой логичный вывод, что существуют процессы, структура которых, с определённой вероятностью, будет изменяться с течением времени, а при увеличении количества наблюдений это вероятность лишь увеличивается. Конечно, это не относиться к временным рядам, которые являются стационарными и не имеют внешних факторов воздействия. Но в силу специфики предметной области эконометрических моделей (которые будут рассмотрены ниже) мы, как правило, сфокусированы на временных рядах, в которых влияние внешних факторов существенно.

Таким образом исследование такого рода временных рядов затруднено наличием структурных сдвигов (разладки случайного процесса) игнорирование которых в большинстве случаев приводит к некорректных результатам. Поэтому задача обнаружения структурных сдвигов является актуальной и востребованной при анализе временных рядов.

В работе рассматривается задача детектирования структурных сдвигов для кусочно-заданной $GARCH(1,1)$ модели. Пусть имеется временной ряд $(Y_t)_{t=1}^T$ для которого существует $k \geq 0$ структурных изменений (сдвигов), тогда обозначив моменты разладок как $\tau_1\tau_2 \ldots \tau_k$ уравнение для j-того фрагмента будет выглядеть следующим образом:

$$Y_t = \varepsilon_t, \quad \sigma_t^2 = w_j + \delta_j \sigma_{t-1}^2 + \gamma_j \varepsilon_{t-1}^2 \qquad \text{где } t \in [\tau_{j-1}; \tau_j - 1] \tag{1}$$

а $(w_j, \delta_j, \gamma_j)$ тройка неизвестных параметров модели которые принадлежат множеству значений $\Theta = \{(w, \delta, \gamma) : w > 0, \delta \geq 0, \gamma \geq 0, \delta + \gamma \leq 0\}$

В результате задача заключается в нахождении алгоритмов поиска всех возможных сдвигов во временном ряде по двум параметрам:

1. Точность нахождения моментов сдвига

2. Временная и пространственная сложность алгоритмов

В работе рассматривается $ICSS$ процедура с применением $KL{-}CUSUM$ для детектирования изменений garch модели. Для проверки выдвигаемых гипотез методы апробируются на данных о цене акций следующих компаний:

- AAPL:NasdaqGS валюта USD

- Что-то:TSX валюта CAD

1 Метод детектирования одного сдвига

Метод обнаружения одного структурного сдвига $KL - CUSUM$, предложенный Парнем основан на разнице двух взвешенных сумм по временным промежуткам разной длинны. Вводятся статистика KL и $\hat{\upsilon}$ на основании которых принимается решение о наличии структурного сдвига или его отсутствия. Критерий был сформулирован (ссылка на книжку) и опирается на теорему:

Пусть существует кусочно-заданный GARCH(1,1)-процесс $(Y_t)_{t=1}^{T}$ допускающий наличие одного структурного сдвига в момент времени $\tau \in [2;T]$ и описывающийся следующей системой уравнений:

$$\begin{cases} Y_t = \epsilon_t,\ \epsilon_t = \sigma_t \xi,\ \sigma_t^2 = w_1 + \delta_1 \sigma_{t-1}^2 + \gamma_1 \epsilon_{t-1}^2, & t \in [1; \tau - 1] \\ Y_t = \epsilon_t,\ \epsilon_t = \sigma_t \xi,\ \sigma_t^2 = w_2 + \delta_2 \sigma_{t-1}^2 + \gamma_2 \epsilon_{t-1}^2, & t \in [\tau; T] \end{cases} \tag{2}$$

Рассчитывается статистика KL такая что

$$KL(k) = \sqrt{T} \frac{k(T-k)}{T^2} \left(\frac{1}{k} \sum_{t=1}^{k} Y_t^2 - \frac{1}{T-k} \sum_{t=k+1}^{T} Y_t^2 \right) = \\ = \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{k} - \frac{k}{T} \sum_{t=1}^{T} Y_t^2 \right) \qquad k \in \{1, \dots, T\} \tag{3}$$

Далее среди всех возможных значений статистики KL выбирается k по следующему правилу:

$$\tau^* = min\{k : |KL(k)| = max_{j \in \{1,\dots,T\}} |KL(j)|\} \tag{4}$$

Для проверки критерия наличия структурного сдвига, введём следующую статистику:

$$\hat{\upsilon^2_{r,T}} = \sum_{|j| \le r} w_j \hat{c_j},\ \text{где} \tag{5}$$

$$w_j = 1 - \frac{|j|}{r+1}, \qquad r \in N \tag{6}$$

$$\hat{c_j} = \frac{1}{T} \sum_{i=1}^{T-|j|} (Y_i^2 - \bar{Y^2})(Y_{i+|j|}^2 - \bar{Y^2}) \tag{7}$$

Из теоремы №2 следует, что про $T \to \infty$ и $\frac{r}{T} \to 0$ в случае отсутствия структурного сдвига имеет место сходимость по распределению:

$$\frac{|KL(\tau^*)|}{\hat{\upsilon_{r,T}}} \xrightarrow{d} \sup_{0 \le u \le 1} |B^0(u)| \tag{8}$$

где $B^0(u)$, $u \in [0;1]$, - процесс броуновского моста. При этом эмпирически r может быть выбрана двумя способами (ссылка на статьи):

1. $r = \lfloor lnT \rfloor$
2. $r = \lfloor \sqrt{T} \rfloor$

Тогда критерий наличия структурного сдвига можно определить следующим образом: если $\frac{|KL(\tau^*)|}{\hat{\upsilon_{r,T}}} \ge q_p$, то гипотеза об отсутствии структурного сдвига в момент времени τ^* отвергается на уровне значимости $1-p$, где $q_{0.95} = 1.358$ и $q_{0.99} = 1.628$ - значение квантилей 0.95 и 0.99 соответственно для $\sup_{0 \le u \le 1} |B^0(u)|$ (супремум броуновского моста)

2 Итеративные алгоритмы для поиска всех возможных сдвигов

Идея процедуры поиска всех возможных структурных сдвигов во временном ряду заключается в последовательном применении алгоритма KL-CUSUM (KL) к временным рядам полученных в результате деления исходного ряда по предполагаемым сдвигам τ_j.

Соответственно, задача поиска формируется стандартным способом: рассматриваются подмножется исходного временного ряда границами которых являеются предполагаемые структурные сдвиги. Алгоритм можно представить в виде последовательных шагов:

1. Находим первый сдвиг τ_0

2. Движемся влево меняя правую границу постоянно (т.е. рассматриваем новый ряд $Y\left[0; KL(Y\left[0;T\right])\right]$, продолжая операцию до тех пор пака предполагаемые сдвиги существуют слева)

$$\tau_{first} = KL(Y\left[0; KL(Y\left[0; \ldots KL(Y\left[0;T\right])\right])\right]) \tag{9}$$

3. Аналогично, находим новую правую границу

$$\tau_{last} = KL(Y\left[KL(Y\left[\ldots KL(Y\left[0;T\right])\right]); T\right]) \tag{10}$$

4. Теперь рассматриваем новый ряд $Y\left[\tau_{first}+1; \tau_{last}\right]$

5. Далее в полученный вектор сдвигов временного ряда добавляются точки начала и конца исходного ряда (т.е. $\tau_{vec} = (0, \ldots, T)$)

6. После формируются интервалы вида $\left[\tau_{j-1}+1; \tau_{j+1}\right]$ И применяется алгоритм KL для проверки является ли точка сдвигом или нет.

На языке описания алгоритмов процесс поиска всех возможных структурных сдвигов выглядит следующим образом [ICSS алгоритм 1]:

Algorithm 1: Алгоритм распространения близости

```
Function ICSS (Y)
    t_1 = 1
    // применяем метод для определения первого сдвига
    KL-CUSUM(Y[t_1 : T])
    if KL-CUSUM then
        // если сдвиг был найден считаем его правой границей
          нового ряда, который вложен в Y
        t2 = τ*(Y[t1 : T])
        while not KL-CUSUM(Y[t1 : t2]) do
            t2 = τ*(Y[t1 : t2])
        end
        τ_first = t_2
        while not KL-CUSUM(Y[t1 : t2]) do
            t1 = τ*(Y[t1 : t2])
        end
        τ_last = t_1 − 1
        if τ_first = τ_last then
            return τ_first // считается что сдвиг только один
        end
        if τ_first < τ_last then
            τ_vec ⟵ (τ_first, τ_last)
            ICSS(Y[τ_first + 1; τ_last])
        else
            τ_vec ⟵ (0, T)
            sort(τ_vec, type="ascending") // Сортируем всевозможные
              моменты сдвигов по возрастанию
            foreach τ_j do
                if KL-CUSUM(Y[τ_{j−1} + 1 : τ_{j+1}]) then
                    ts_breakpoints ⟵ τ_j
                end
            end
            return ts_breakpoints // Массив уже не будет содражать
              начальную и конечные точки
        end
    else
        return no change-points found
    end
```

3 Апробация методов на реальных данных