Федеральное государственное автономное образовательное учреждение высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет экономических наук

КУРСОВАЯ РАБОТА

«Методы обнаружения структурных сдвигов в GARCH-моделях»

по направлению подготовки Экономика

образовательная программа «Статистическое моделирование и актуарные расчеты»

Выполнил:

Студент группы МСМ181

Новиков Лев Ильич

Руководитель:

Старший преподаватель

Борзых Дмитрий Александрович

Москва 2019

Введение

При построении эконометрической модели, в процессе оценивания её параметров, если мы хотим, чтобы оценки параметров были как можно более точными, то, как нам известно, зачастую нужно чтобы выполнялись некоторые базовые требования (например, $\varepsilon \sim N(0, \sigma^2)$ или независимость ошибок регрессии между собой). Также для состоятельности оценок обычно требуется сравнительно большое количество наблюдений. И если с первой проблемой в большинстве случаев можно бороться, то вторая проблема является камнем преткновения при анализе временных рядов: например, если нужно построить прогноз для некоторого временного ряда, то нам будут доступны только наблюдения до текущего момента T включительно, что значительно ухудшает оценки коэффициентов.

Более того, даже если нехватки в наблюдениях нет, со временем параметры модели, описывающей данные, могут случайным образом меняться. Данная ситуация характерна для моделирования стоимости ценных бумаг, поскольку состояние финансового рынка в значительной степени зависит от новостей: так, при смене новостного фона, параметры модели могут измениться скачкообразно (например, вызывая резкий рост волатильности). Подобные моменты времени, в которые наша модель перестает объяснять имеющиеся закономерности, называются разладками случайного процесса или структурными сдвигами. Основная задача этой работы – научиться их выявлять, чтобы тем самым иметь возможность в подобной ситуации улучшить оценки параметров (оценивая их отдельно для каждого интервала без разладок).

В данной работе рассматривается применение модифицированного CUSUM-теста для GARCH(1,1)-моделей, имеющих несколько моментов разладки.

Данный подход, исторически впервые предложенный в работе [1] (Kokoszka & Leipus) для ARCH(∞)-процессов с одним структурным сдвигом, также может быть использован в качестве основы для построения итеративной процедуры, выявляющей все имеющиеся структурные сдвиги.

В ходе работы мы для примера смоделируем GARCH(1,1)-процессы, имеющие один и два момента структурных сдвигов соответственно, и посмотрим, какие результаты мы получим при применении указанного подхода.

Затем для реальных данных – логарифмических доходностей акций за последние несколько лет, мы ищем моменты, которые с достаточно большой вероятностью (p=0.99 или p=0.95) являются структурными сдвигами, после чего пытаемся проверить, какая новость или факт могли в этот момент времени вызвать такую разладку процесса.

Напоминание

GARCH(1,1)-процесс задается уравнениями $Y_t = \varepsilon_t$, $\varepsilon_t = \sigma_t \xi_t$, $\sigma_t^2 = \omega + \delta\sigma_{t-1}^2 + \gamma\varepsilon_{t-1}^2$, где (ω, δ, γ) – параметры процесса, $\xi_t \sim WN(0, 1)$, σ_t^2 – дисперсия процесса Y_t.

В качестве CUSUM-теста в этом случае оптимальнее рассматривать не кумулятивную сумму рекурсивных остатков, а кумулятивную сумму *квадратов* рекурсивных остатков (CUSUM-sq). Это связано с тем, что Y_t между собой не коррелируют, но их квадраты Y_t^2 коррелируют (поскольку $E(\xi_t) = 0$, но $E(\xi_t^2) = 1$).

То есть мы рассматриваем $X_k = Y_k^2$, после чего происходит тестирование гипотез:

H_0: выборка $X_1, \dots X_T$ описывается уравнениями $Y_t = \varepsilon_t$, $\varepsilon_t = \sigma_t \xi_t$, $\sigma_t^2 = \omega + \delta\sigma_{t-1}^2 + \gamma\varepsilon_{t-1}^2$, для некоторого фиксированного вектора параметров $b = (\omega, \delta, \gamma)$. Альтернативная гипотеза H_1 подозревает наличие одного структурного сдвига в момент τ^*, то есть подвыборка $X_1, \dots, X_{\tau^*}$ описывается уравнениями $Y_t = \varepsilon_t$, $\varepsilon_t = \sigma_t \xi_t$, $\sigma_t^2 = \omega_1 + \delta_1\sigma_{t-1}^2 + \gamma_1\varepsilon_{t-1}^2$, а подвыборка $X_{\tau^*+1}, \dots X_T$ описывается уравнениями $Y_t = \varepsilon_t$, $\varepsilon_t = \sigma_t \xi_t$, $\sigma_t^2 = \omega_2 + \delta_2\sigma_{t-1}^2 + \gamma_2\varepsilon_{t-1}^2$, для некоторых фиксированных $b_1 = (\omega_1, \delta_1, \gamma_1)$ и $b_2 = (\omega_2, \delta_2, \gamma_2)$.

P. Kokoszka и R. Leipus в [1] показали, что при выполнении нулевой гипотезы статистика $KL(k) = \frac{1}{\sqrt{T}}\left(\sum_{i=1}^{k} X_i - \frac{k}{T}\sum_{i=1}^{T} X_i\right) \to \sigma W^0(k)$, где σ – истинное стандартное отклонение модели, а $W^0(k)$ – Броуновский мост.

Более того, это утверждение также верно для оценки σ^2, $\hat{s}_{T,q}^2 = \sum_{j=-q}^{q} \omega_j(q)\hat{\gamma}_j$, где $\hat{\gamma}_j = \frac{1}{T}\sum_{i=1}^{T-|j|}(X_i - \bar{X})\left(X_{i+|j|} - \bar{X}\right)$ – выборочные ковариации, $\omega_j(q) = \left(1 - \frac{|j|}{q+1}\right) \cdot I\{|j| \leq q\}$ – веса Бартлетта (коэффициенты для треугольного ядра), а $q \in N$ – константа, зависящая от T. Единственное обязательное условие – сходимость $\frac{q}{T}$ при T, стремящемся к бесконечности: $\lim_{T\to\infty}\left(\frac{q}{T}\right) = 0$. В нашем исследовании мы положили $q(T) = \sqrt{T}$ (другие возможные варианты, такие как q = ln(T) или log(T) дают гораздо меньшую величину q, из-за чего носитель ядра получается слишком маленьким, и точность наших оценок ощутимо падает)

Итак, $\frac{KL(t)}{\hat{s}_{N,q}} \to W^0(t)$. Таким образом, у нас есть возможность свести нашу задачу к исследованию статистики, поведение которой нам хорошо известно – супремуму модуля броуновского моста: нулевая гипотеза об отсутствии структурного сдвига отвергается на уровне значимости 1-p при $\frac{|KL(\tau^*)|}{\hat{s}_{N,q}} \geq q_p$, где q_p – квантиль уровня p для супремума модуля броуновского моста, а τ^* – момент структурного сдвига для альтернативной гипотезы. Поскольку очевидно, что таких τ^* может быть много, логично положить $\tau^* = \min\{k\colon |KL(k)| = \max|KL(k)|\}$.

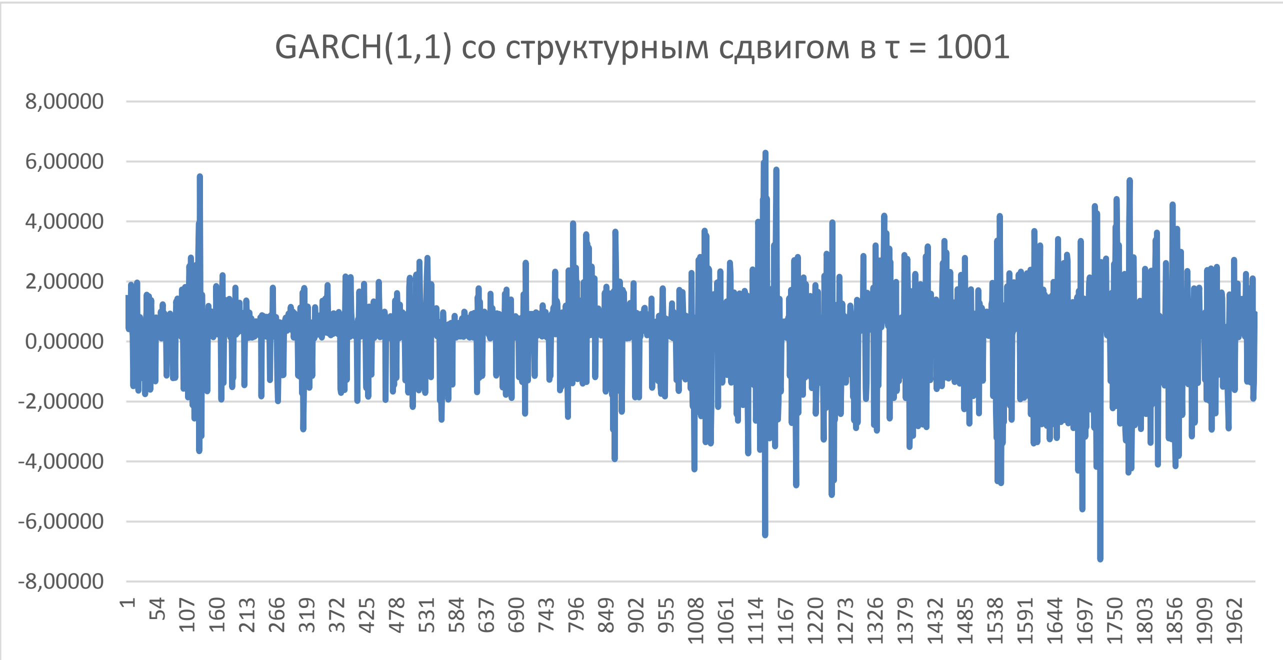
В дальнейшем будем обозначать за $\boldsymbol{KL}(t)$ нормированную величину $\frac{|KL(t)|}{\hat{s}_{N,q}}$.

Из таблиц квантили для $sup|W^0(t)|$ равны $q_{0.95} = 1{,}358$ и $q_{0.99} = 1{,}628$.

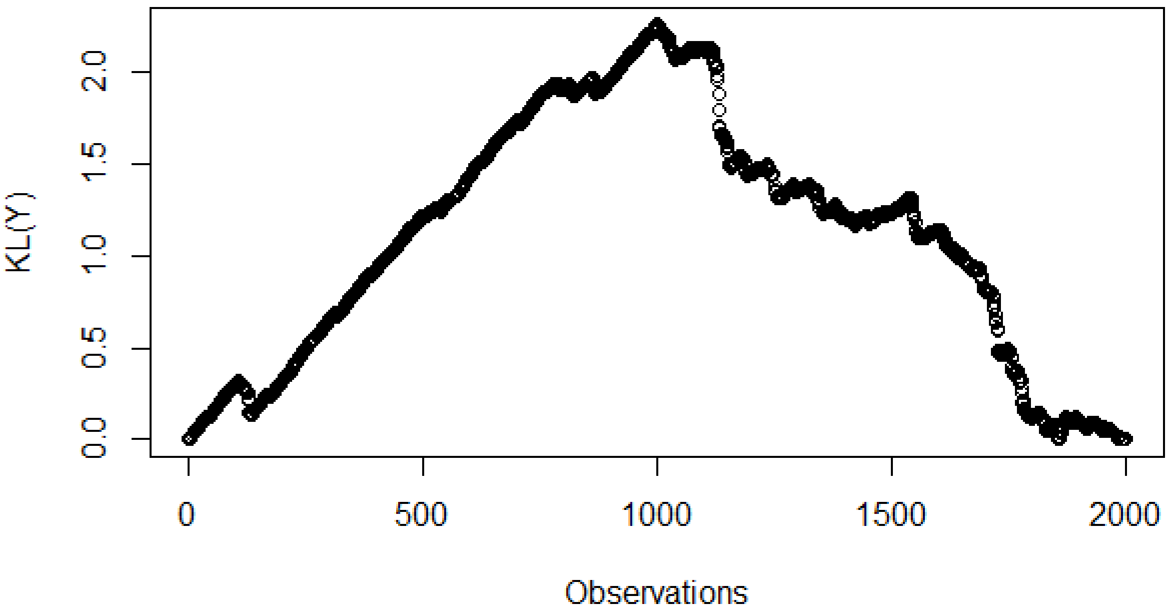
Симуляция GARCH(1,1)-процесса с одним структурным сдвигом

В качестве примера рассмотрим результат применения указанного теста к симулированным 2000 наблюдениям, соответствующим GARCH(1,1) со структурным сдвигом на 1001 наблюдении (данные и параметры указаны в таблице ниже, код на языке R для генерации всех необходимых случайных величин и функции, осуществляющие тест CUSUM приведены в приложении)

Параметры:	w	δ	γ
$t \leq 1000$	0,1	0,7	0,2
$t \geq 1001$	0,3	0,7	0,2



Статистика супремума модуля Броуновского моста для этого GARCH-процесса изображена на графике ниже:



Как мы видим, максимум статистики достигается примерно на 1000-ном наблюдении, что означает наличие там структурного сдвига (в данном случае максимальное значение статистики на промежутке [1; 2000] превышает 2, что больше 1.628 = $q_{0.99}$, и гипотеза об отсутствии структурного сдвига отвергается на 1%-ном уровне значимости).

```
> print(paste0('Для данных ',length(Y), ' наблюдений с помощью KL-статистики
 выявлен структурный сдвиг в момент времени ', 't_break - ', BB_stat(Y)[1]))
[1] "Для данных 2000 наблюдений с помощью KL-статистики выявлен структурный
сдвиг в момент времени t_break - 1001"
```

В нашем случае получилось, что время сдвига угадано с идеальной точностью. Однако сейчас мы знали точно, что сдвиг в модели в точности один, а в реальной жизни так почти никогда не бывает.

Рассмотрим итеративную процедуру ICSS (Iterated Cumulative Sums of Squares), использованную в работе [2] (Inglan, Tiao), применительно к нашей ситуации она позволит выявить все имеющиеся сдвиги:

Краткое описание ICSS-процедуры:

После нахождения времени τ^* структурного сдвига для выборки $Y_1, \ldots, Y_T$ (если его нет, значит сдвигов нет совсем), мы рассматриваем левую подвыборку $Y_1, \ldots, Y_{\tau^*}$ и ищем сдвиг в ней, после чего повторяем эти действия с новым τ^*, и так до тех пор, пока не окажется, что левая подвыборка не содержит сдвигов (для того, чтобы алгоритм работал корректно, необходимо ввести дополнительное условие и не проверять выборки единичного размера, в противном случае мы получим оценку дисперсии $\hat{s}_{N,q} = 0$)

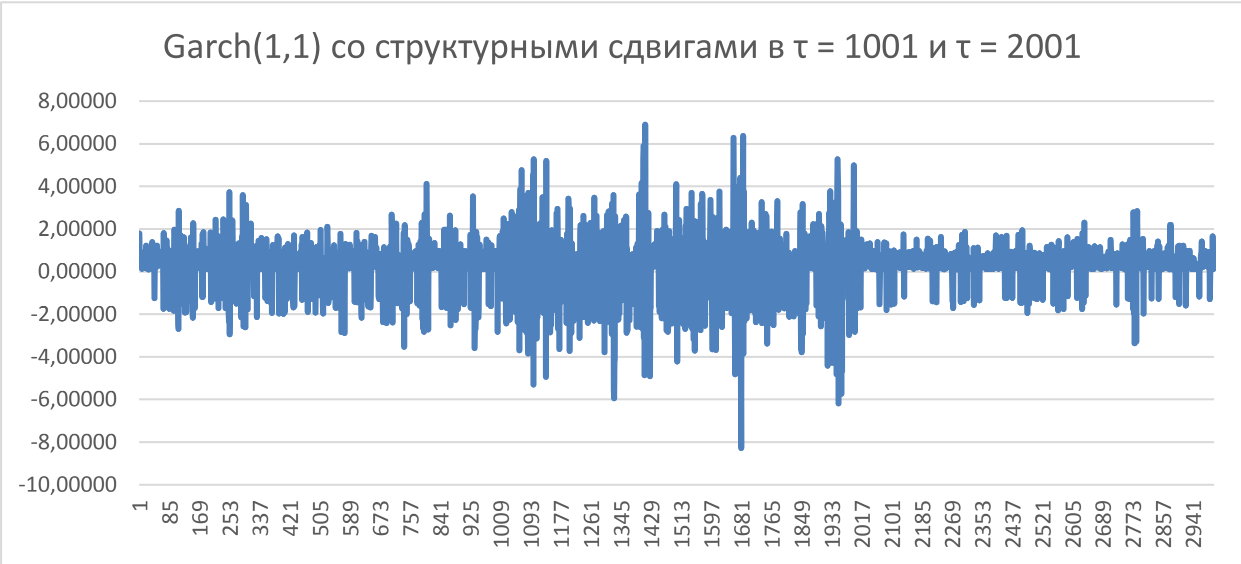
Найденный самый левый структурный сдвиг мы запоминаем, и после повторения аналогичной процедуры с правыми подвыборками получаем также самый правый структурный сдвиг. Итеративно повторяя указанные выше шаги для выборки, лежащей между самым левым и самым правым сдвигами, полученными на предыдущем шаге ICSS-процедуры, находим все левые и правые сдвиги по очереди. После этого остается только перепроверить, что все найденные τ_i действительно разделяют подвыборки с разными параметрами, то есть для каждой выборки $[\tau_{i-1}; \tau_{i+1}]$ провести тест, и, если τ_i не является структурным сдвигом, удалить его. Повторяя эту процедуру до тех пор, пока количество структурных сдвигов не перестанет меняться, получим то, что требовалось.

Симуляция GARCH(1,1)-процесса с двумя структурными сдвигами

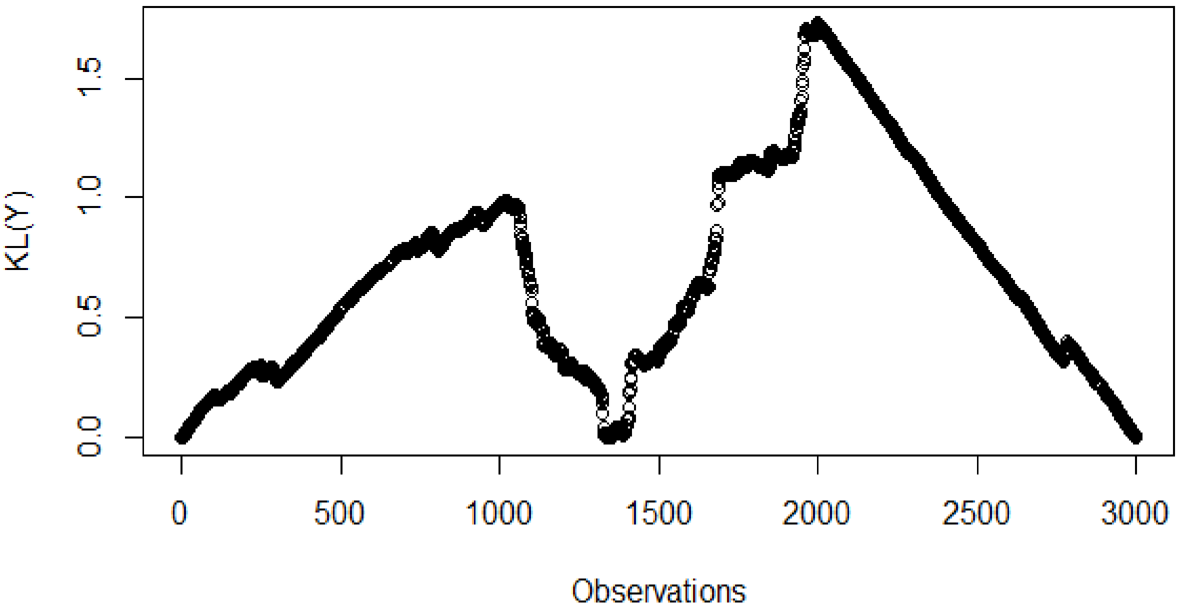
Опять же, будем проверять работоспособность данного алгоритма на симулированных данных, теперь симулируем GARCH(1,1)-процесс, состоящий из 3000 наблюдений, со структурными сдвигами в моменты $\tau_1^* = 1001$ и $\tau_2^* = 2001$ и заданными параметрами.

Параметры:	w	δ	γ
$T \leq 1000$	0,1	0,7	0,2
$1001 \leq T \leq 2000$	0,3	0,7	0,2

$2001 \leq T \leq 3000$	0,1	0,6	0,2



Статистика супремума модуля Броуновского моста:




```
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.99)))
[1] "Для данных 3000 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 1056"
[2] "Для данных 3000 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 2000"
```

В итоге процедура ICSS выдает два структурных сдвига: $\tau_1 = 1056$ и $\tau_2 = 2000$ (два локальных пика на графике), что почти совпадает с истинными значениями.

Значит, алгоритм хорошо выявляет структурные сдвиги (несложно проверить, смоделировав 100 аналогичных рядов с двумя структурными сдвигами), в том случае, если данные *действительно* соответствуют модели GARCH(1,1).

Структурные сдвиги в реальных данных

Попробуем найти структурные сдвиги в реальных данных и объяснить их. Возьмем в качестве данных цену закрытия акций крупных компаний США (мы рассматриваем именно крупные компании, во-первых, для того, чтобы полученную информацию о времени структурных сдвигов можно было бы проверить с помощью открытых и легкодоступных источников информации, а во-вторых, крупные компании более стабильны, и наша модель будет более приближена к реальности).

Будем анализировать на наличие структурных сдвигов логарифмическую доходность ($y_t = ln\frac{P_t}{P_{t-1}}$), где P_t – цена акции в конце дня t (цена закрытия)

Отдельно отметим, что GARCH(1,1)-модели обычно неплохо показывают себя при моделировании доходностей финансовых активов, поскольку они умеют объяснять кластеризацию волатильности в данных, что постоянно наблюдается на фондовом рынке.

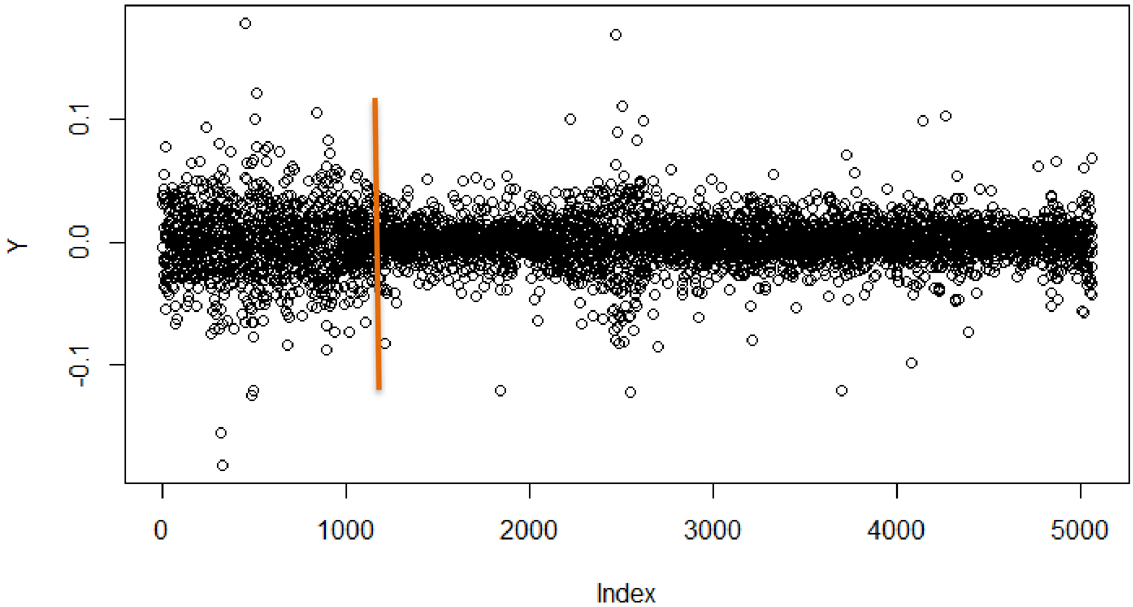
Пример 1. Microsoft (MSFT), данные с 1999 по 2018 год:

```
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.99)))
[1] "Для данных 5064 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 1118"
> print(dates[1118])
[1] "2003-06-17"
```

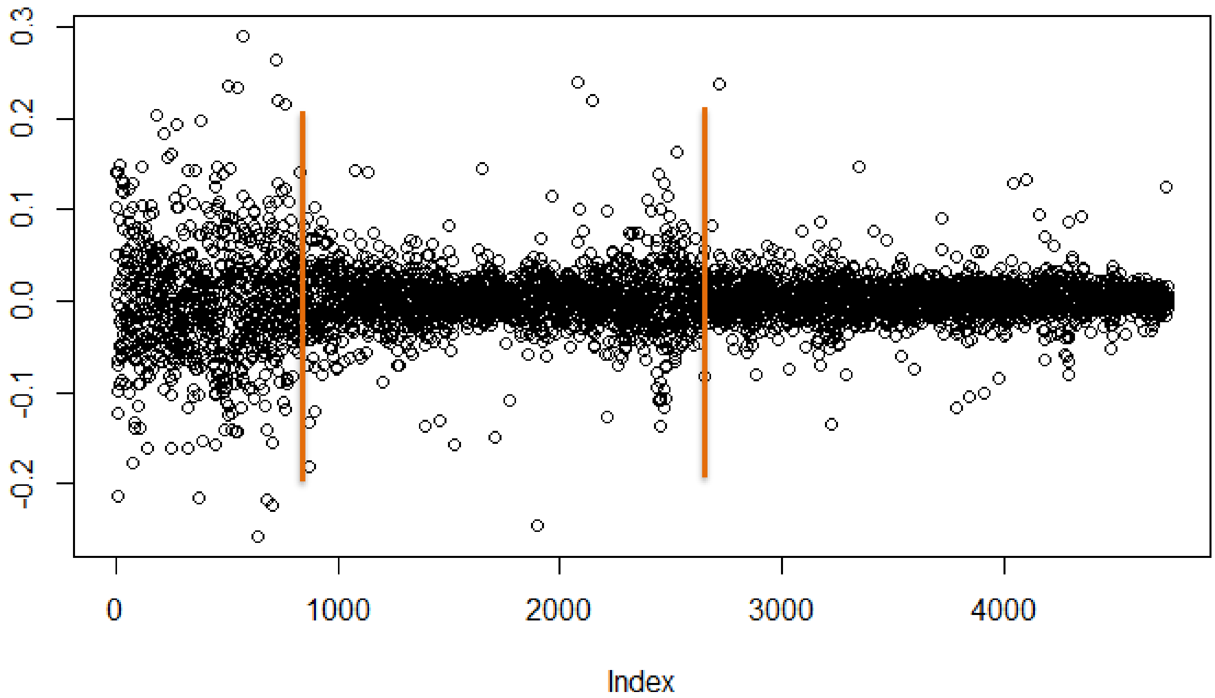
Из всех 5064 наблюдений сдвиг есть только 17 июня 2003 года. После ознакомления с источниками [3] становится понятно, чему он соответствует: в начале июля 2003 года Microsoft объявил о том, что поменяет свою систему бонусных вознаграждений для топ-менеджмента и будет вместо опционов выдавать пакеты акций. Это дало сигнал инвесторам, что Microsoft из молодого стартапа превратился в зрелую компанию, что очевидно увеличило количество долгосрочных инвесторов и уменьшило число спекулянтов, в результате чего волатильность уменьшилась.

Как мы и можем видеть на графике, до 2003 года волатильность была стабильно высокой: после кризиса доткомов не было очевидным, как пойдут дела у компании. После этого же волатильность практически не менялась, даже во время кризиса 2007-2008 года

(примерно 2500-е – 2600-е наблюдения).



Пример 2. Amazon (AMZN), данные с 1999 по 2018 год:



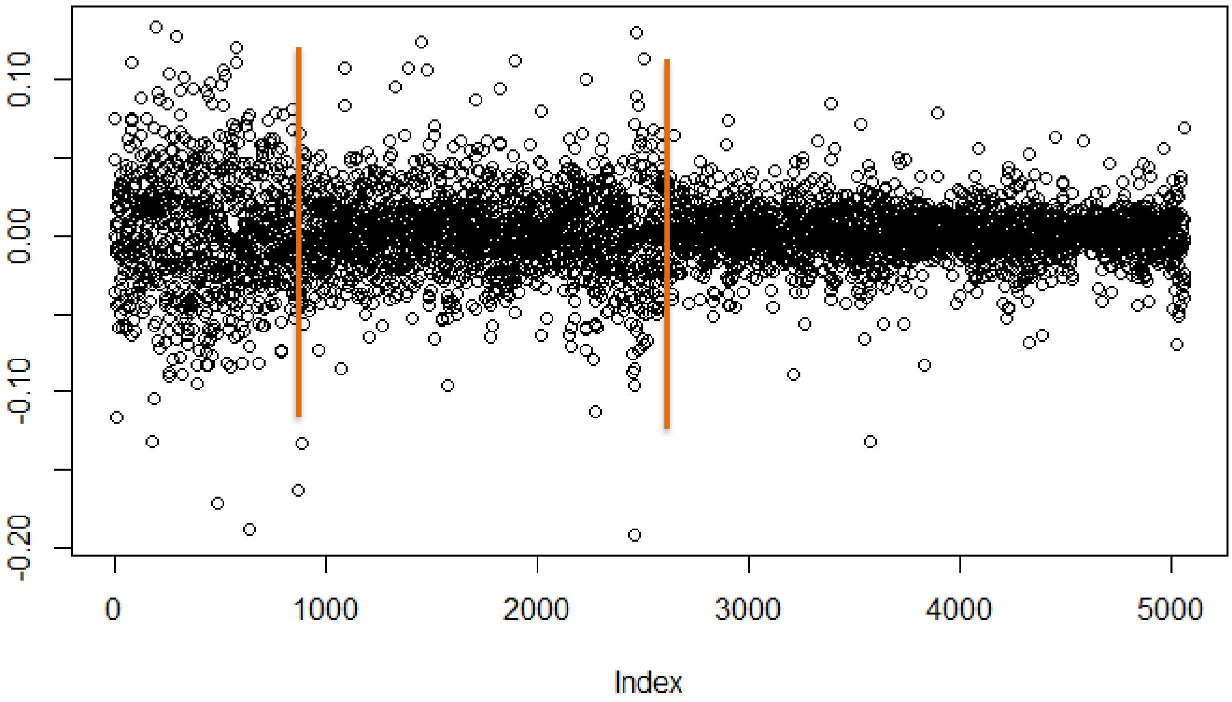

```
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.99)))
[1] "Для данных 4741 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 893"
[2] "Для данных 4741 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 2715"
> print(dates[893])
[1] "2002-07-29"
> print(dates[2715])
[1] "2009-10-26"
```

На этот раз сдвига два – в конце июля 2002 года и в конце октября 2009.

Сдвиг в конце июля понятен: после конца 90-х, когда в бум доткомов компания сильно закредитовалась, даже с работающим бизнесом и понятной бизнес-моделью компания вызывала сомнения у многих инвесторов, поэтому бумага была сильно волатильной. Но в отчетах за конец 2001 года и апрель 2002 года компания показала чистую прибыль и сокращение расходов, одновременно с ростом продаж на 20%. [4] Вследствие этого доверие к компании начало восстанавливаться, волатильность немного уменьшилась.

Со вторым сдвигом картина тоже ясна: после кризиса 2007-2008 года и ценовой войны с Walmart, на фоне новости [5] в конце октября о том, что продажи Amazon по некоторым разделам выросли на 40%, а прибыль – на 70%, параметры модели, описывающей данные, резко поменялись.

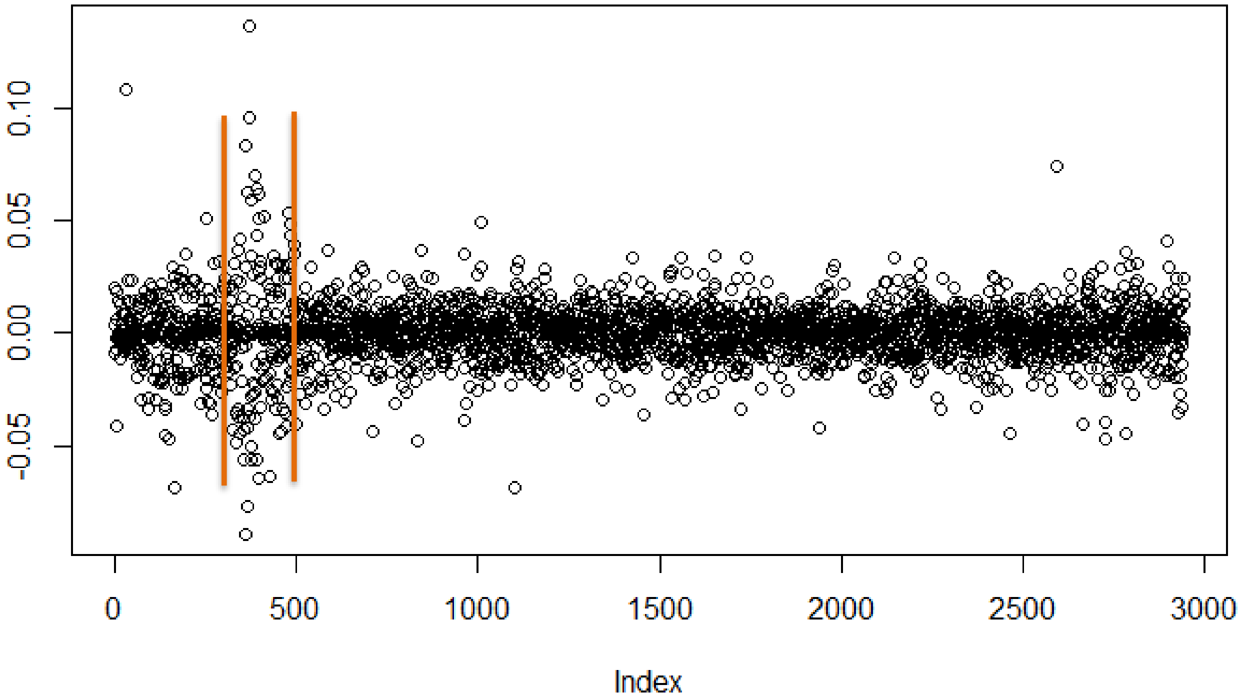
Пример 3. Apple (AAPL), данные с 1999 по 2018 год:




```
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.99)))
[1] "Для данных 5064 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 901"
[2] "Для данных 5064 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 2610"
> print(dates[901])
[1] "2002-08-06"
> print(dates[2610])
[1] "2009-04-09"
```

Аналогично акциям Amazon, структурный сдвиг в 2002 году – выход Apple из зоны высокой волатильности после кризиса доткомов в связи с возвратом доверия инвесторов, сдвиг в 2009 году – аналогичный возврат к стабильности после кризиса 2008 года (компания показала прибыль выше рыночных ожиданий) [6]

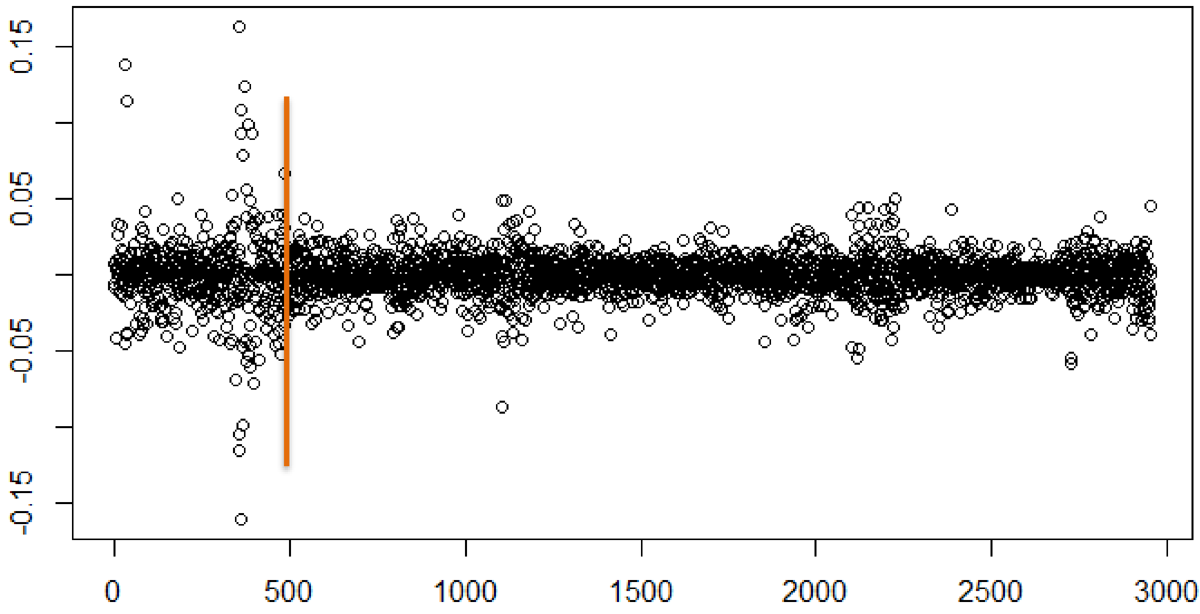
Пример 4. Verizon (VZ), данные с 2007 по 2018 год:




```
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.99)))
[1] "Для данных 2944 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - отсутствуют"
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.95)))
[1] "Для данных 2944 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 333"
[2] "Для данных 2944 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 498"
> print(dates[333])
[1] "2008-09-13"
> print(dates[498])
[1] "2009-04-07"
```

Здесь можно наглядно видеть структурные сдвиги во время начала и конца кризиса 2008 года. Отметим, что эти структурные сдвиги есть лишь на уровне значимости 5%, при проведении теста на 1%-ном уровне значимости выясняется, что структурных сдвигов нет.

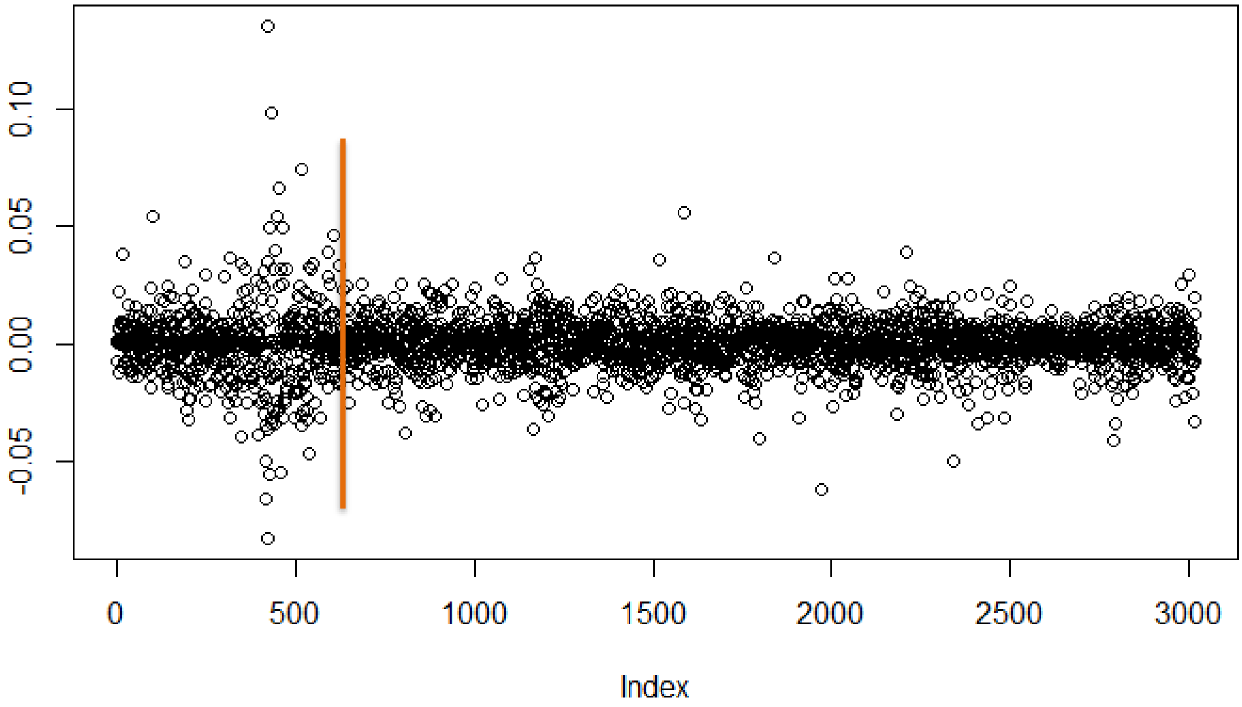
Пример 5. Exxon Mobil (XOM), данные с 2007 по 2018 год:




```
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.99)))
[1] "Для данных 2954 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - отсутствуют"
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.95)))
[1] "Для данных 2954 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 492"
> print(dates[492])
[1] "2009-03-30"
```

Судя по графику, это выход доходности из кластера с повышенной волатильностью в момент кризиса 2008 года. Структурный сдвиг тут, как и в предыдущем случае, есть только на 5%-ном уровне значимости.

Пример 6. Coca-Cola (KO), данные с 2007 по 2018 год:




```
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.99)))
[1] "Для данных 3016 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - отсутствуют"
> print(paste0('Для данных ', length(Y), ' наблюдений с помощью
KL-статистики выявлены следующие моменты структурных сдвигов: ',
 't_break - ', ICSS_refinement(Y, p = 0.95)))
[1] "Для данных 3016 наблюдений с помощью KL-статистики выявлены
 следующие моменты структурных сдвигов: t_break - 624"
> print(dates[624])
[1] "2009-06-22"
```

Точной информации по наличию/отсутствию структурного сдвига в этом случае найти не удалось, но судя по графику, это также успешный выход компании из кризиса 2008 года. Опять-таки, отметим, что структурный сдвиг тут есть только на 5%-ном уровне значимости.

Заключение и выводы

В работе был рассмотрен модифицированный алгоритм, соединяющий в себе KL-статистику и процедуру ICSS. Мы убедились, что данный алгоритм очень хорошо работает для обнаружения структурных сдвигов в GARCH(1,1) моделях, однако, на реальных данных он срабатывает не всегда.

Мы смогли интерпретировать большинство структурных сдвигов, так:

- для компании Microsoft структурный сдвиг происходит 17 июня 2003 года – одновременно с изменением системы бонусных вознаграждений (замена опционов на акции), давая сигнал инвесторам, что из молодого стартапа Microsoft превратился в зрелую компанию.

- для компании Amazon два сдвига: 29 июля 2002 и 26 октября 2009. Первый – компания показала чистую прибыль и сокращение расходов, одновременно с ростом продаж на 20%, из-за чего доверие к компании начало восстанавливаться, а волатильность уменьшилась. Второй – после кризиса 2007-2008 года, на фоне новости в конце октября о том, что продажи Amazon по некоторым разделам выросли на 40%, а прибыль – на 70%, параметры модели, описывающей доходность акции, резко поменялись.

- для компании Apple также есть два сдвига: 6 августа 2002 и 9 апреля 2009. Причины смены поведения в точности такие же, как у Amazon (да еще и даты почти такие же)

- для компании Verizon два сдвига (но с другим уровнем значимости 0.05): 13 сентября 2008 и 7 апреля 2009. Эти структурные сдвиги – в точности начало и конец кризиса 2008 года (банкротство Lehman Brothers – 15 сентября). Уровень значимости структурных сдвигов 5%, на 1%-ном уровне значимости сдвигов нет.

- для компании Exxon Mobil один сдвиг (с уровнем значимости 0.05), 30 марта 2009. Судя по графику, это выход доходности из кластера с повышенной волатильностью в момент конца кризиса 2008 года.

- для компании Coca-Cola также есть один структурный сдвиг, (также с уровнем значимости 0.05) – 22 июня 2009 (видимо, конец кризиса 2008 года).

Возможные дальнейшие направления работы:

Попробовать реализовать новый метод, основанный на тесте Колмогорова-Смирнова (предложен в статье Д. А. Борзых и А. А. Языкова [8])

Применить KL-ICSS метод к реальным данным, которые лучше описываются GARCH-моделями (предварительно продифференцировать данные, вычитая тренд), для того чтобы предпосылка о нулевом среднем не нарушалась.

Список литературы:

1. *Kokoszka P., Leipus R.* Testing for parameter changes in ARCH models. // Lithuanian Mathematical Journal, 1999. Vol. 39, No. 2. pp. 182-195
2. *Inglan C., Tiao G.* Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. // Journal of the American Statistical Association, 1994. Vol. 89, No. 427. pp. 913-923.
3. *New York Times.* Microsoft to Give Its Employees Stock Instead of Options. July 8, 2003. – URL – https://www.nytimes.com/2003/07/08/business/microsoft-to-give-its-employees-stock-instead-of-options-2003070894060647345.html
4. *New York Times.* Amazon II: Will This Smile Last? May 19, 2002. – URL – https://www.nytimes.com/2002/05/19/business/amazon-ii-will-this-smile-last.html
5. *The Wall Street Journal.* Amazon Lights Up E-Commerce. October 23, 2009. – URL – https://www.wsj.com/articles/SB10001424052748703816204574489750561367182
6. *New York Times.* Apple sales rise despite recession. February 2, 2009. – URL – https://www.nytimes.com/2009/01/22/technology/22iht-apple.4.19605999.html
7. *Д. А. Борзых, М. А. Хасыков*. Процедура уточнения ICSS алгоритма обнаружения структурных сдвигов в GARCH-моделях. // Прикладная эконометрика, 2018, т. 51, с. 126–139.
8. *Д. А. Борзых, А. А. Языков.* The new KS method for a structural break detection in GARCH(1,1) models.

Приложение

Использованный код на языке R:

```
library(readxl)

# GARCH(1,1) modelling
GARCH_1break <- function(w = 0.1, b = 0.7, g = 0.2, dw = 0.2, db = 0, dg = 0) {
	ksi <- rnorm(1999, mean = 0, sd = 1)
	sigma <- c(0)
	Y <- c(rnorm(1, mean = 0, sd = 1))

	for (i in 2:1000){
		sigma[i] <- sqrt(w + b*sigma[i-1]^2 + g*Y[i-1]^2)
		Y[i] <- ksi[i-1]*sigma[i]
	}

	w <- w + dw # external shock w_2 = w_1 + dw (0.2 by default)
	b <- b + db # external shock b_2 = b_1 + db (0 by default)
	g <- g + dg # external shock g_2 = g_1 + dg (0 by default)

	for (i in 1001:2000){
		sigma[i] <- sqrt(w + b*sigma[i-1]^2 + g*Y[i-1]^2)
		Y[i] <- ksi[i-1]*sigma[i]
	}
	return(Y)
}
GARCH_2breaks <- function(w = 0.1, b = 0.7, g = 0.2, dw1 = 0.2, db1 = 0, dg1 = 0, dw2 = 0.1, db2 = -0.1, dg2 = 0) {
	ksi <- rnorm(2999, mean = 0, sd = 1)
	sigma <- c(0)
	Y <- c(rnorm(1, mean = 0, sd = 1))

	for (i in 2:1000){
		sigma[i] <- sqrt(w + b*sigma[i-1]^2 + g*Y[i-1]^2)
		Y[i] <- ksi[i-1]*sigma[i]
	}

	w <- w + dw1 # external shock w_2 = w_1 + dw1 (0.2 by default)
	b <- b + db1 # external shock b_2 = b_1 + db1 (0 by default)
	g <- g + dg1 # external shock g_2 = g_1 + dg1 (0 by default)
```

```
	for (i in 1001:2000){
		sigma[i] <- sqrt(w + b*sigma[i-1]^2 + g*Y[i-1]^2)
		Y[i] <- ksi[i-1]*sigma[i]
	}

	w <- w + dw2 # external shock w_3 = w_2 + dw2 (0.1 by default)
	b <- b + db2 # external shock b_3 = b_2 + db2 (-0.1 by default)
	g <- g + dg2 # external shock g_3 = g_2 + dg2 (0 by default)

	for (i in 2001:3000){
		sigma[i] <- sqrt(w + b*sigma[i-1]^2 + g*Y[i-1]^2)
		Y[i] <- ksi[i-1]*sigma[i]
	}
	return(Y)
}

Y <- GARCH_2breaks(dw2 = -0.2)
plot(Y, xlab = "Observations")
plot(KL(Y), xlab = "Observations")

print(paste0('Для данных ', length(Y), ' наблюдений с помощью KL-статистики выявлен структурный сдвиг в момент времени ', 't_break - ', BB_stat(Y)[1]))
print(paste0('Для данных ', length(Y), ' наблюдений с помощью KL-статистики выявлены следующие моменты структурных сдвигов: ', 't_break - ', ICSS_refinement(Y, p = 0.99)))

# CUSUM test realization

KL <- function(Y) {	# calculates KL statistic for each observation from the sample
	X <- Y*Y
	T <- length(Y) # number of observations in our sample
	Xmean <- sum(X)/T # mean value of Y^2 series

	KL <- 0
	for (k in 1:T){
		KL <- c(KL, (sum(X[1:k]) - k*Xmean)/sqrt(T)) # calculating KL statistic
	}
	KL <- abs(KL[-1])
	q <- floor(sqrt(T)) # using square root function (logarithm finds more breaks, but also gives more mistakes)
	w <- c((1:q)/(q+1), ((q+1):1)/(q+1)) # Bartlett weights (triangular kernel with window [-q; q])
	C <- c()
	for (j in 0:q){
		Cj <- 0
```

```
		for (i in 1:(T-j)){
			Cj <- Cj + (X[i]-Xmean)*(X[i+j]-Xmean) # sample covariances in [-q; q] window
		}
		C <- c(C, Cj/T)
	}
	C <- c(rev(C[-1]), C) # here we use symmetry of covariance to simplify the code
	s <- sqrt(sum(w*C)) # triangular kernel
	return(KL/s)
}

BB_crit <- function(p = 0.99) {	# asymptotic quantiles for supremum of Brownian bridge's absolute value (KL)
	if (p == 0.05) # significance level 0.05
		BB_cr <- 0.520
	if (p == 0.10) # significance level 0.10
		BB_cr <- 0.571
	if (p == 0.25) # significance level 0.25
		BB_cr <- 0.677
	if (p == 0.50) # significance level 0.50
		BB_cr <- 0.828
	if (p == 0.75) # significance level 0.75
		BB_cr <- 1.019
	if (p == 0.90) # significance level 0.90
		BB_cr <- 1.224
	if (p == 0.95) # significance level 0.95
		BB_cr <- 1.358
	if (p == 0.99) # significance level 0.99
		BB_cr <- 1.628
	return(BB_cr)
}

BB_stat <- function(Y) {	# returns suspicious time moment and value of KL statistic
	X <- Y*Y
	T <- length(Y) # number of observations in our sample
	Xmean <- sum(X)/T # mean value of Y^2 series

	KL <- 0
	for (k in 1:T){
		KL <- c(KL, (sum(X[1:k]) - k*Xmean)/sqrt(T)) # calculating KL statistic
	}
	KL <- abs(KL[-1])
	tau <- min(which(KL == max(KL))) # suspicious time moment
```

```
    q <- floor(sqrt(T)) # using square root function (logarithm finds more breaks, but also gives more mistakes)
    w <- c((1:q)/(q+1), ((q+1):1)/(q+1)) # Bartlett weights (triangular kernel with window [-q; q])

    C <- c()
    for (j in 0:q){
        Cj <- 0
        for (i in 1:(T-j)){
            Cj <- Cj + (X[i]-Xmean)*(X[i+j]-Xmean) # sample covariances in [-q; q] window
        }
        C <- c(C, Cj/T)
    }
    C <- c(rev(C[-1]), C) # here we use symmetry of covariance to simplify the code
    s <- sqrt(sum(w*C)) # triangular kernel
    return(c(tau, KL[tau]/s))
}

BB_test <- function(Y, p = 0.99) { # returns most suspicious time moment and value of statistic
    BB <- BB_stat(Y)
    if (BB[2] > BB_crit(p)){ # check whether statistic is larger than asymptotic quantile or not
        return(c('Break', BB))}
    else
        return(c('No break', BB))
}

# ICSS procedure

ICSS_iter <- function(Y, p = 0.99) {     # returns series of suspicious moments
    T <- length(Y) # number of observations in our sample
    t1 <- 1
    t2 <- T # starting endpoints
    tfirst <- 0
    tlast <- T
    tau <- c(tfirst, tlast) # all possible points of structural break (including the borders)
    tcus <- 0

    while (tfirst < tlast - 1) {
        if (BB_test(Y[t1:t2], p)[1] == 'No break'){
            tau <- sort(unique(tau))
            return(tau) # stop the procedure, if there are no more breaks
        }
        else {
```

```
                        tcus <- as.numeric(BB_test(Y[t1:t2], p)[2]) + t1-1 # taking into account that BB_test
returns relative position
                        tau <- c(tau, tcus) # remember first break-like point
                        t2 <- tcus # looking at left interval

                        repeat {
                                if (BB_test(Y[t1:t2], p)[1] == 'No break'){
                                        break      # go further, if there are no more breaks to the left
                                }
                                else {
                                        t2 <- as.numeric(BB_test(Y[t1:t2], p)[2])+t1-1
                                }
                        }

                        tfirst <- t2
                        tau <- c(tau, tfirst) # remember break-like point
                        t1 <- tcus + 1
                        t2 <- T # new endpoints for next part of an algorithm

                        repeat {
                                if (BB_test(Y[t1:t2], p)[1] == 'No break'){
                                        break # go further, if there are no more breaks to the right
                                }
                                else {
                                        t1 <- as.numeric(BB_test(Y[t1:t2], p)[2])+1 + t1-1
                                }
                        }

                        tlast <- t1 - 1
                        tau <- c(tau, tlast) # remember break-like point
                        t1 <- tfirst + 1
                        t2 <- tlast # new endpoints for next iteration of an algorithm
                }
        }
        tau <- sort(unique(tau))
        return(tau) # stop the procedure, if the remaining interval is too small
}

ICSS_refinement <- function(Y, p = 0.99) {     # refining the moments of structural breaks
        tau <- ICSS_iter(Y, p) # getting series of possible structural breaks
        if ((length(tau) <= 2)) { # in case there are no structural breaks at all
                #tau <- BB_stat(Y)[1]
```

```
		#D <- BB_stat(Y)[2] # if info about p-value of structural break is needed
		return('отсутствуют')
	}

	tau_ref <- tau[2:(length(tau)-1)] # series of possible structural breaks without 0 and T
	iteration <- 1
	while (iteration < 20) { # setting limit to number of iterations; though, it coincides quickly
		tau <- c(0, tau_ref, length(Y))
		tau_ref <- c()
		K <- length(tau)

		for (n in 2:(K-1)){
			tprev <- tau[n-1]+1
			tnext <- tau[n+1]
			if (BB_test(Y[tprev:tnext], p)[1] == 'Break') # check if the moment is really a structural break for interval from two adjacent moments
				tau_ref <- c(tau_ref, tau[n-1] + as.numeric(BB_test(Y[tprev:tnext], p)[2]))
		}

		iteration <- iteration + 1

		if ((length(tau_ref) == length(tau[2:(length(tau)-1)])) & ((is.null(tau_ref)) | (max(abs(tau_ref-tau[2:(length(tau)-1)]))<3)))
			break # end refinement process if there is no change in tau or if change is very small
	}
	return(tau_ref)
}

# Data preprocessing

#Y <- read_excel("GARCH_with_breaks.xlsx", col_names = FALSE, range = "CUSUM (1 break)!A1:A2000")
#Y <- read_excel("C:/Users/Лев/Documents/R/coursework/GARCH_with_breaks.xlsx", col_names = FALSE, range = "ICSS (2 breaks)!A1:A3000")

# 1. Microsoft
#Y <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.1999-2018!D3:D5066")
#dates <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.1999-2018!B3:B5066")

# 2. Amazon
```

```
#Y <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.1999-2018!I3:I4743")
#dates <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.1999-2018!G3:G4743")

# 3. Apple
#Y <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.1999-2018!N3:N5066")
#dates <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.1999-2018!L3:L5066")

# 4. Verizon
#Y <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.2007-2018!D3:D2946")
#dates <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.2007-2018!B3:B2946")

# 5. Exxon Mobil
#Y <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.2007-2018!N3:N2956")
#dates <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.2007-2018!L3:L2956")

# 6. Coca Cola
#Y <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.2007-2018!AH3:AH3018")
#dates <- read_excel("C:/Users/Лев/Documents/R/coursework/US - 2007-2018.xlsx", col_names = FALSE, range = "US.2007-2018!AF3:AF3018")

Y <- as.matrix(Y)
storage.mode(Y) <- "numeric"
Y <- as.vector(Y)

dates <- as.matrix(dates)
storage.mode(dates) <- "character"
dates <- as.vector(dates)

plot(Y) # our data
plot(KL(Y)) ) # statistic
print(paste0('Для данных ', length(Y), ' наблюдений с помощью KL-статистики выявлены следующие моменты структурных сдвигов: ', 't_break - ', ICSS_refinement(Y, p = 0.99))) # searching with p=0.99
```

```
print(paste0('Для данных ', length(Y), ' наблюдений с помощью KL-статистики выявлены следующие моменты структурных сдвигов: ', 't_break - ', ICSS_refinement(Y, p = 0.95))) # searching with p=0.95

# for printing date from number of observation
# print(dates[1118])
```