

УДК 330.47

Прогнозирование направления движения курса акций компаний на основе корреляционного анализа

Малимонов Д.А., Козлов Д.Ю.

Алтайский государственный университет
dacool731@mail.ru, dyk.barnaul@gmail.com

Аннотация

В статье предлагается идея использования обратной корреляции курсов акций различных компаний для получения сигналов на покупку или продажу акций на фондовом рынке. Поставлена и решена задача автоматизированного отыскания пар акций, для которых коэффициент корреляции имеет не просто отрицательное значение ($-0,6$ до -1), но и сохраняет его на протяжении относительно длительного периода. В ходе работы использовался функционал портала impactoria.com, предоставляющий доступ к курсам 7 тысяч акций. Для автоматизации работы с этим ресурсом был написан скрипт-парсер, который получал на входе символ акции, вписывал его в текстовое поле на сайте "Impactoria", находил для нее пару с наилучшим значением коэффициента корреляции и сохранял результат. Для визуализации результатов построены графики изменения коэффициента корреляции на протяжении 6 месяцев между акциями с положительной и отрицательной корреляцией. Далее построена модель прогнозирования направления движения курса акций на период до 10 дней. Написан скрипт на языке программирования R для мониторинга и выявления длительного периода отрицательной корреляции между парами курсов акций, а также с помощью коэффициента детерминации осуществлена численная оценка качества прогноза.

Часто опытный инвестор на фондовом рынке для большей сбалансированности включает в свой инвестиционный портфель акции, которые «стоят против рынка». Это позволяет компенсировать потери в случае движения фондового рынка вниз. Один из способов выявить такие акции – провести корреляционный анализ с целью найти пары акций с отрицательным значением коэффициента корреляции их курсов.

При поиске пар акций, чьи коэффициенты корреляции в данный момент имеют положительное значение, мы, как правило, ожидаем увидеть акции компаний одинаковой или близкой отрасли. Это объясняется влиянием различных факторов (политических, социальных и т.д.) на отрасль в целом, что приводит к сходному изменению направления курса акций таких компаний. Высокие коэффициенты корреляции иногда устанавливаются на весьма продолжительный срок (Рисунок 1).



Рисунок 1. Курс акций компаний "Сбербанк" и "ВТБ24", и их коэффициент корреляции на период с 01.04.18 по 01.10.18 ("ВТБ24" - простая линия, "Сбербанк" - линия с точками, Коэффициент корреляции – серая область)

В случае с отрицательным коэффициентом корреляции появляется возможность найти отличные по отрасли пары акционерных компаний. Для построения прогностической модели необходимо иметь представление не только о величине коэффициента, но и о длительности сохранения этого значения.

В первую очередь требуется найти пары акций, чьи абсолютные показатели коэффициента отрицательной корреляции имеют максимальные значения. На данный момент общее количество компаний, акции которых торгуются на рынках всех стран, насчитывает более 15 тыс. Из них всего 7 тысяч (где порядка 300 российских) тех, чьи данные курса находятся в открытом доступе [1]. Для нахождения коэффициента корреляции между данными всех доступных нам пар необходимо потратить большие вычислительные ресурсы, а также стоит учитывать, что доступ к некоторым курсам предоставляется на ограниченное количество запросов в день, а это значит, что для проверки всех возможных пар потребуется минимум 140 дней. В интернете имеется несколько ресурсов, которые упрощают этот процесс, но, тем не менее, они имеют большие недостатки. Различные крупные финансовые порталы дают возможность нахождения коэффициента корреляции лишь между выбранными парами, но не ведут мониторинг в глобальных масштабах. На некоторых сайтах брокеров имеется подобный инструмент, но он выдает результат только среди валютных пар. Тем не менее нам удалось найти ресурс, который не только показывает коэффициент корреляции между данными курса акций, но и выдает для нее ту пару, с которой это значение максимальное по модулю (как положительное, так и отрицательное) (Рисунок 2) [2].

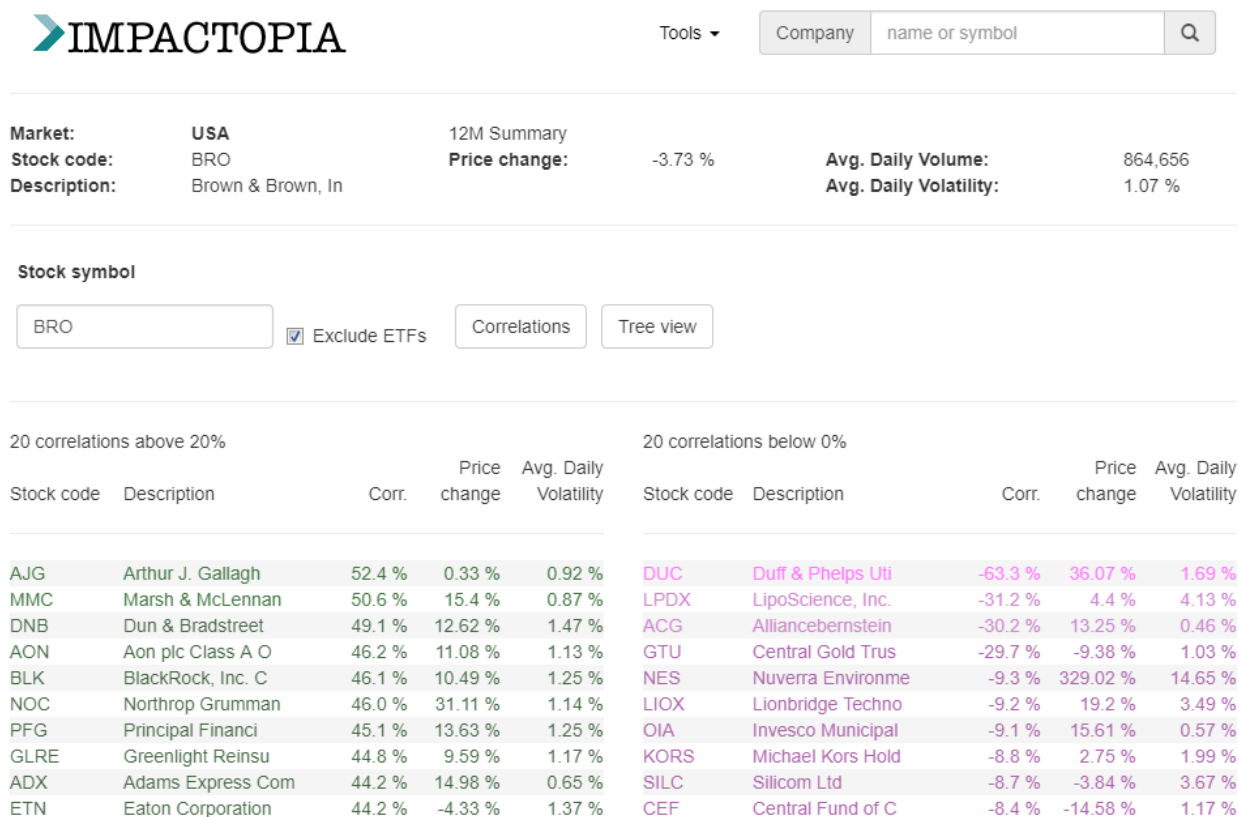


Рисунок 2. Внешний вид ресурса "Impactoria"

Но, как и в предыдущих случаях, у этого ресурса есть недостаток: для нахождения нужной нам пары необходимо вбивать вручную символы каждой из 7 тысяч акций. Для ускорения этого процесса был написан скрипт-парсер, который получал на входе символ акции, вписывал его в текстовое поле на сайте "Impactoria", находил для нее пару с наилучшим результатом (коэффициент корреляции от -0,6 до -1) и сохранял результат. В качестве порогового значения решено было выбрать показатель, равный -0.6, что является значительным и в большей мере хорошо характеризующим связь временных рядов [3]. Те

пары акции, чья связь была слабее, к дальнейшему анализу не допускались. По завершению работы скрипта было получено 64 пары, имеющие коэффициент корреляции в целевом интервале.

Для прогнозирования направления курса выбранных акций необходимо иметь высокие показатели не только на текущий момент, но при условии сдвига данных одной из акций на несколько дней относительно другой. При таком условии количество пар сократилось до 6. Нужно помнить, что показатель коэффициента корреляции не постоянен [4]. Более вероятно, что спустя несколько дней может появиться еще несколько пар с хорошим результатом, но на данном этапе, в качестве теста, были выбраны эти 6 пар: DUC/BRO, CLX/VTBR, DUC/SBER, CLX/BRO, DUC/VTBR, CLX/SBER (Рисунок 3). Как можно заметить, среди них есть 2 акции российских компаний “Сбербанк” и “ВТБ24”, и одна американская “Brown & Brown Inc” (компания, занимающаяся представлением услуг посредника в сфере страхования). На них скрипт в дальнейшем и должен давать прогноз, основываясь на анализе опорных акций DUC (Duff & Phelps Utility and Corporate Bond Trust Inc. - закрытая диверсифицированная компания по управлению инвестициями) и CLX (Clogox Company - транснациональный производитель и распространитель потребительских и профессиональных товаров).

Пара	Уровень корреляции										
	Сдвиг первой акции на n дней относительно второй										
	0	1	2	3	4	5	6	7	8	9	10
DUC/BRO	-0,79	-0,88	-0,9	-0,94	-0,92	-0,93	-0,88	-0,83	-0,77	-0,76	-0,73
CLX/SBER	-0,62	-0,66	-0,67	-0,69	-0,71	-0,76	-0,8	-0,83	-0,8	-0,79	-0,76
DUC/SBER	-0,69	-0,78	-0,81	-0,82	-0,83	-0,82	-0,82	-0,84	-0,81	-0,77	-0,75
CLX/VTB	-0,76	-0,72	-0,71	-0,69	-0,68	-0,66	-0,69	-0,62	-0,47	-0,45	-0,44
DUC/VTB	-0,82	-0,8	-0,78	-0,74	-0,73	-0,74	-0,69	-0,57	-0,44	-0,43	-0,45
CLX/BRO	-0,6	-0,71	-0,83	-0,9	-0,92	-0,89	-0,84	-0,78	-0,72	-0,7	-0,69

Рисунок 3. Показатель коэффициента корреляции при сдвиге данных на период до 10 дней

Как видно из рисунка 3, никаких резких изменений между разными интервалами сдвигов не прослеживается. В то же время этот результат не дает представлений о длительности сохранения этих показателей. Для ее определения обратимся к историческим данным и посмотрим, как долго держится высокое значение отрицательного коэффициента корреляции и случаются ли (и если да, то, как часто) резкие скачки в положительную сторону (Рисунок 4-5).

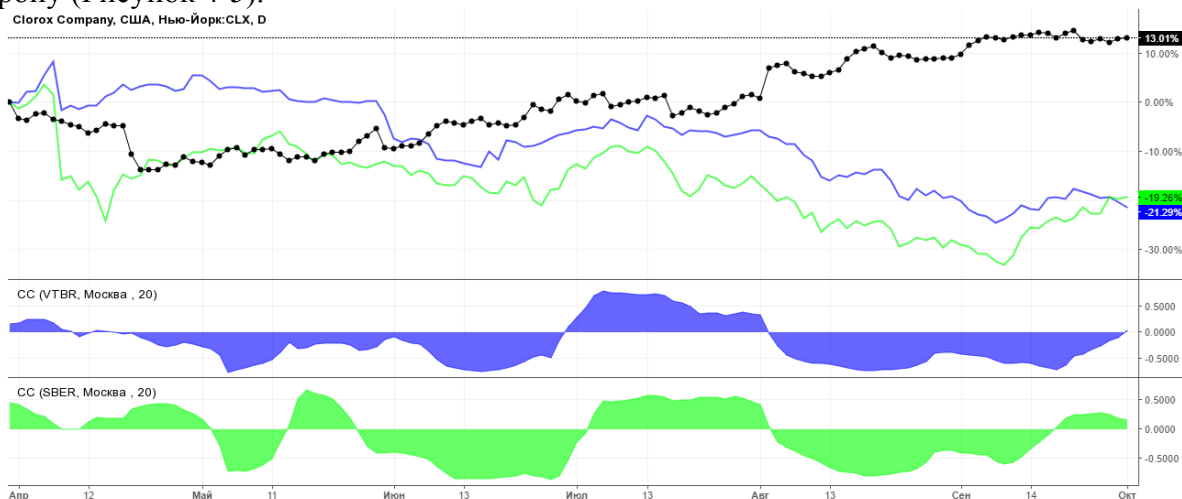


Рисунок 4. Курс акций SBER и VTBR относительно акций CLX, и их коэффициент корреляции на период с 01.04.18 по 01.10.18 (VTBR – простая синяя линия, SBER – простая зеленая линия, CLX - черная линия с точками. Коэффициент корреляции: CLX/SBER – зеленая область, CLX/VTBR – синяя область)

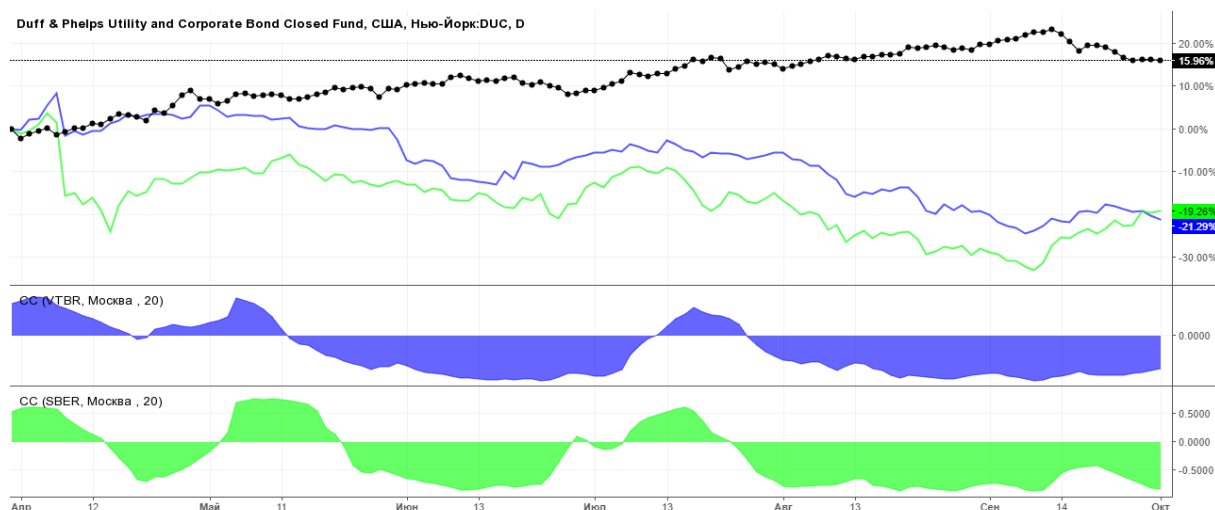


Рисунок 5. Курс акций SBER и VTBR относительно акций DUC, и их коэффициент корреляции на период с 01.04.18 по 01.10.18 (VTBR – простая синяя линия, SBER – простая зеленая линия, DUC - черная линия с точками. Коэффициент корреляции: DUC /SBER – зеленая область, DUC /VTBR – синяя область)

Из рисунков 4-5 видно, что коэффициент отрицательной корреляции сохраняется в заданном интервале (от -0.6 до -1.0) меньшее время, нежели коэффициент положительной, однако этот показатель удерживается минимум в течении 10-20 дней, что дает возможность проанализировать и спрогнозировать направление курса акции за этот период.

Далее, для построения прогностической модели, требуется сделать прототип скрипта, который бы следил за изменениями коэффициента корреляции и при удовлетворительной ситуации выдавал сигнал на повышение или понижение курса [5]. Так как значение коэффициента корреляции не постоянно, и плавно меняется с течением времени, необходимо было выставить максимальную дальность прогнозирования, такую, чтобы за это время результат (коэффициент корреляции) не выходил за пределы выше обозначенного интервала [6]. Как было замечено ранее, в среднем в течение 5-10 дней уровень корреляции у выбранных пар может разнонаправленно меняться относительно допустимой границы, поэтому было решено установить максимальный срок прогноза в 10 дней.

Охарактеризуем работу написанного на языке программирования R скрипта, реализующего Shiny web-приложение [5]. При запуске первым делом выгружается актуальный курс выбранных акций (Рисунок 6).

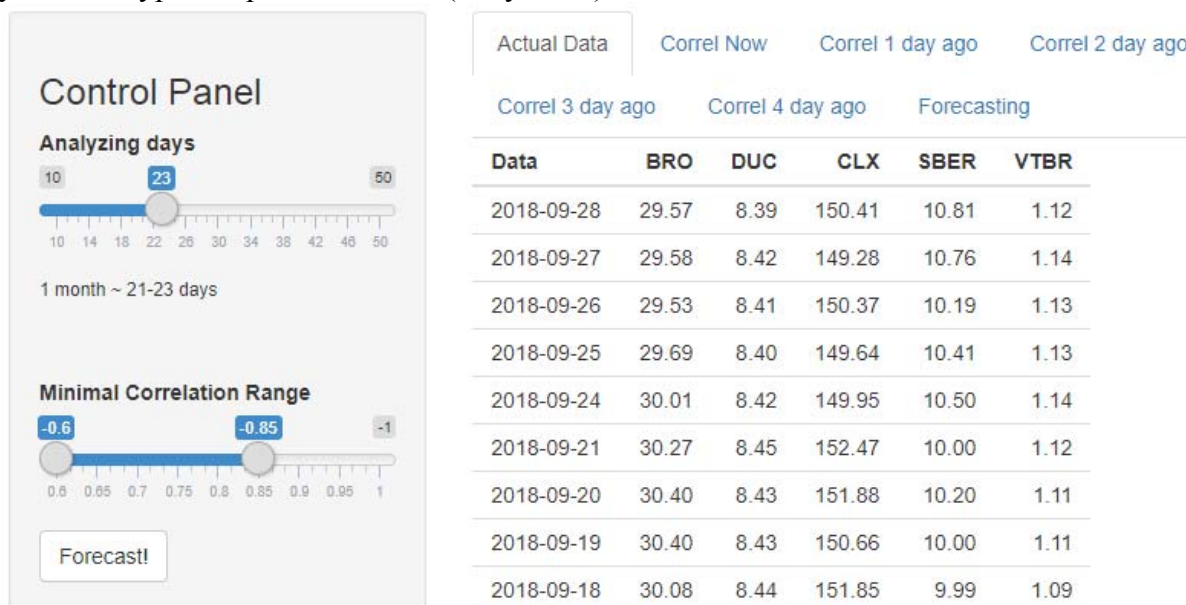


Рисунок 6. Стартовое окно web-приложения

После того, как коэффициент корреляции определен для всех пар, при различных вариантах сдвига (от 1 до 10 дней) скрипт проделывает тоже самое для 4 предыдущих дней, чтобы оценить в какой стадии находится (растущий или падающий по модулю) уровень корреляции. Если наблюдается рост коэффициента, значит, соответствующие акции в паре показывают разные направления курса, а значит, по акции можно сделать прогноз (Рисунок 7).

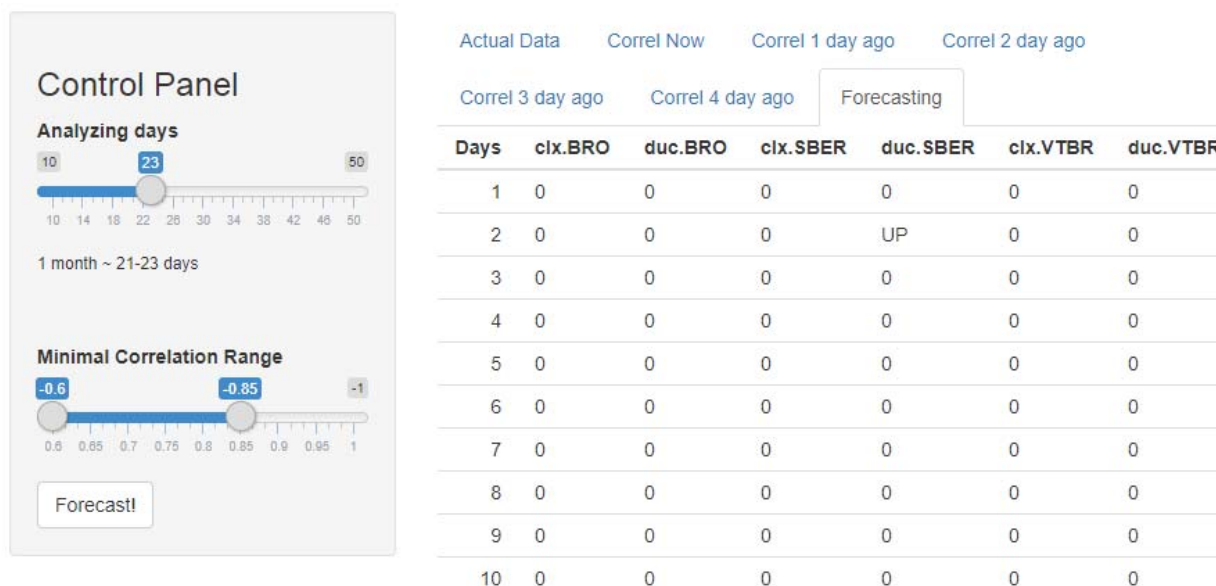


Рисунок 7. Внешний вид матрицы прогнозов (показатель UP обозначает рост курса, DOWN – падение, строки представляют из себя дальность прогноза в днях, а столбцы – пары курсов акций)

Прежде чем выдать прогноз, скрипт находит коэффициент детерминации курса опорной акции, тем самым проверяя наличие и плавность изменения показателей отрицательного коэффициента корреляции на протяжении последних 5 дней [7]. Для нашей модели было выбрано значение коэффициента детерминации не ниже 0,5, что является достаточным и в большей мере хорошо характеризующим ежедневное усиление связи между курсами акций показателем. Если его модуль ниже, то прогноз не выдается. На рисунке 7 показан пример прогноза, который дает скрипт для мониторинга уровня корреляции. Он интерпретируется как рост курса акций SBER в течение ближайших 2 дней на основе анализа курса акции DUC. Если обратиться к рисунку 5, то можно заметить, что тенденции роста у акций “Сбербанка” прослеживаются. Это в некоторой степени подтверждает возможность прогнозирования направления движения курса акций компаний с помощью отрицательной корреляции. Ее выявление позволяет открывать новые пары, имеющие сильную связь, а значит увеличивать количество обнаружений возможных спрогнозированных ситуаций.

Библиографический список

1. Investing.com – котировки и финансовые новости [Электронный ресурс] URL: <https://ru.investing.com>
2. Impactoria [Электронный ресурс] URL: <http://impactopia.com/correlation/>
3. Дженкинс Г.М., Кулаков В.И., Анализ временных рядов, прогноз и уравнение. - Великобритания, 1969. – 405 с.
4. Айвазян С.А. Прикладная статистика. Основы эконометрики. – М., 2001. – 432 с.
5. Share Price Forecasting – мониторинг уровня корреляции [Электронный ресурс] URL: <https://asutest1.shinyapps.io/forecasting/>
6. Кабаков Р.И., Анализ и визуализация данных в программе R. – Нью-Йорк, 2014. – 589 с.
7. Мاستицкий С.Э. Статистический анализ и визуализация данных с помощью R. – Тольятти, 2014. – 401 с.