

АНАЛИЗ ФОНДОВОГО РЫНКА S&P-500 И ЦЕЛЕСООБРАЗНЫЙ АЛГОРИТМ ТОРГОВЛИ

А.А. Рубчинский

Высшая школа экономики

Россия, 101000, г. Москва, ул. Мясницкая, д. 20

E-mail: arubchinsky@yahoo.com

Ключевые слова: фондовый рынок, рыночный граф, стоимость акций, кластеризация, алгоритм торговли, правила остановки.

Аннотация: В докладе предлагается новый алгоритм торговли на фондовом рынке S&P-500, который обеспечивает положительные результаты на достаточно длительном промежутке времени. Никаких предположений о поведении этого рынка (в том числе вероятностных) не делается. Предлагаемый общий алгоритм содержит как один из этапов одновременный прогон одного и того же подалгоритма с совпадающими входными данными. Включение стандартного случайного датчика обеспечивает вероятностный характер результатов. Общий алгоритм использует только известные фиксированные данные о стоимости акций при закрытии торговли в предыдущие дни. Он включает в себя специальный подалгоритм принятия решения о продолжении или прекращении торговли до конца определённого периода, основанное на достижении накопленным доходом заданного максимального или минимального уровня.

1. Введение

Основной подход к анализу фондовых рынков состоит в поиске стохастических закономерностей, определяющих их поведение, и в прогнозировании будущих значений на основе этих закономерностей (см., например, [1 – 4] и имеющиеся в них ссылки). Однако во многих теоретически и практически значимых случаях поведение временных рядов можно рассматривать как стохастическое только с большой натяжкой. Поэтому и прогнозируемые члены временных рядов также могут рассматриваться как случайные величины только в рамках таких же ограничений. Само по себе это не означает отсутствие достаточно точных прогнозов для многих конкретных ситуаций. Однако формальная оценка надёжности прогнозов, не опирающаяся на предположения о стохастическом характере рассматриваемых временных рядов, является трудной и почти не исследованной задачей.

В настоящем докладе намечен некоторый общий подход не столько к преодолению, сколько к обходу этих трудностей. Суть этого подхода состоит в следующих двух пунктах:

А. Негативный пункт. Отказ от каких бы то ни было предположений о поведении рассматриваемого временного ряда. Значения всех известных до очередного момента времени величин считаются просто фиксированными данными.

В. Позитивный пункт. В разрабатываемых алгоритмах прогнозирования очередного значения временного ряда предполагается использовать стандартный случайный датчик. Алгоритм запускается K раз с одними и теми же фиксированными исходными данными. Тогда результаты работы этого алгоритма (т.е. K значений одного прогнозируемого параметра) являются реализациями одной и той же «законной»

случайной величины, для которой применимы основные утверждения традиционной теории вероятности, включая ЦПТ и её следствия.

Качество работы упомянутого в пункте В алгоритма естественно оценивается по близости предсказанного результата к реальному, т.е. его оценка является в конечном счёте экспериментальной в рамках любого подхода. Однако именно потому, что сравнивается среднее значение случайной величины с фиксированным реальным значением, можно говорить о вероятности тех или иных отклонений и всех остальных выводах, основанных на традиционной теории вероятности. Можно упомянуть также, что все упомянутые величины могут быть векторными.

Разумеется, для построения вышеупомянутых алгоритмов и других деталей описанной общей схемы для тех или иных конкретных ситуаций, требуется работа по моделированию и анализу соответствующих временных рядов и, возможно, моделированию и анализу порождающих эти временные ряды систем, явлений и алгоритмов. Весь последующий материал настоящего доклада посвящён одной хорошо известной реальной системе – **фондовому рынку S&P-500**, связанному с торговлей акциями 500 компаний США с наибольшей (на данный день) капитализацией.

Первые материалы по этой теме уже опубликованы [5]. В данном докладе приводятся результаты дальнейших исследований. В разделе 2 дано краткое описание фондового рынка S&P-500 и приводится блок-схема предложенного алгоритма торговли. В разделе 3 приводятся основные экспериментальные результаты. В заключении указываются некоторые выводы и направления дальнейших исследований.

2. Фондовый рынок S&P-500 и блок-схема общего алгоритма торговли

В докладе рассмотрен фрагмент фондового рынка США – S&P-500 (500 крупнейших компаний в США). Использовались ежедневные данные о курсах акций на момент закрытия торговли с 01.12.1989 по 31.12.2010. В качестве основных периодов рассматривались отдельные кварталы, а также более короткие периоды – месяцы, называемые далее под-периодами. Число дней в 1-ом и 2-ом под-периодах равно 21, а в 3-ем под-периоде оно равно числу оставшихся в квартале дней (от 16 до 21).

Набор акций, торгуемые на рынке S&P-500 перед днём t , как и их курсы на момент закрытия торговли, известны в течение l ($l = 15$) рабочих дней перед днём t . Курсы акций в день t и во все последующие дни неизвестны.

Укрупненная блок-схема алгоритма показана на рис. 1. Ниже приводятся краткие пояснения к отдельным блокам.

При инициализации $t = 0$. В блоке 1 осуществляется подготовка данных (матрицы расстояний D и рыночного графа G), по которым в блоке 2 определяется кластеризация A множества 500 акций. Кластеры состоят из акций с близким поведением в течение 15 последних перед днём t рабочих дней. В блоке 3 в результате операций купли-продажи указанных алгоритмом акций (сразу после открытия торговли и перед её окончанием) вычисляется доход (возможно, отрицательный) за день t . В блоке 4 определяется накопленный с первого дня периода доход. В последнем блоке 5 происходит принятие решений о продолжении или прекращении торговли в зависимости от накопленного дохода и текущего дня. Возможны три варианта: продолжение торговли в текущем месяце, прекращение торговли до начала следующего месяца и прекращение торговли до конца текущего квартала. В последнем случае накопленный доход является результатом работы общего алгоритма в рассматриваемом квартале.

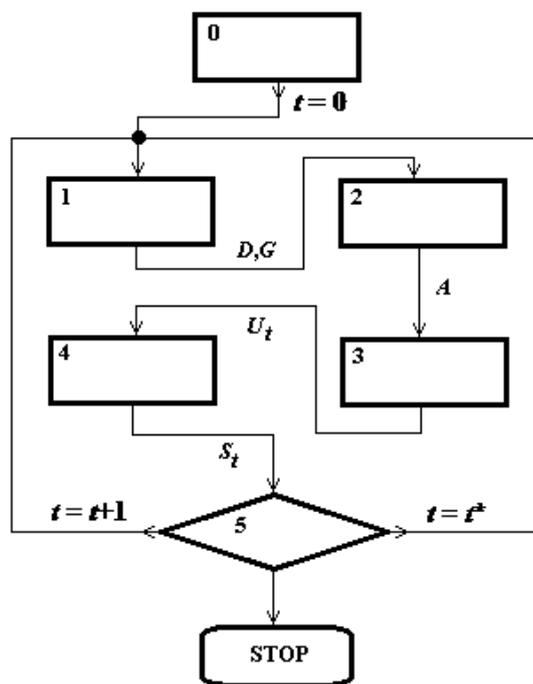


Рис. 1. Укрупненная блок-схема общего алгоритма торговли на фондовом рынке.

При инициализации положим $t = 0$. В блоке 1 осуществляется подготовка данных (матрицы расстояний D и рыночного графа G), по которым в блоке 2 определяется кластеризация A множества 500 акций. Кластеры состоят из акций с близким поведением в течение 15 последних перед днем t рабочих дней. В блоке 3 в результате операций купли-продажи указанных алгоритмом акций вычисляется доход (возможно, отрицательный) за день t . В блоке 4 определяется накопленный с первого дня периода доход. В последнем блоке 5 происходит принятие решений о продолжении или прекращении торговли в зависимости от накопленного дохода и текущего дня. Возможны три варианта: продолжение торговли в текущем месяце, прекращение торговли до начала следующего месяца и прекращение торговли до конца текущего квартала. В последнем случае накопленный доход является результатом работы общего алгоритма в рассматриваемом квартале.

3. Экспериментальные результаты

В блоке 2 производилось 5 параллельных прогонов при совпадающих входах по одному и тому же алгоритму. В результате определялось 5 векторов, компоненты которых представляют собой реальные ежедневные доходы. Для определения устойчивости предложенных алгоритмов анализировалась близость найденных векторов. Считались коэффициенты корреляции для всех 10 пар векторов, после чего определились средние значения 10 чисел. Численные результаты для всех рассмотренных 84-х кварталов приведены в таблице 1. Они говорят о значительной устойчивости квартальных ежедневных доходов на протяжении 21 года. В частности, кризисы 2008-го и 2001-го годов не сказались на их устойчивости. В силу детерминированной зависимости между этими последовательностями и квартальными доходами по предложенному алгоритму торговли, можно говорить и о стабильности квартальных доходов. Сразу подчеркнем, что само по себе это не говорит о знаке доходов. Действительно, в некоторые периоды квартальные и даже годовые доходы являются отрицательными. Подробнее это обсуждается ниже.

Таблица 1. Средние коэффициенты корреляции по кварталам.

1990_01 0.67	1997_01 0.76	2004_01 0.88
1990_02 0.83	1997_02 0.86	2004_02 0.85
1990_03 0.89	1997_03 0.86	2004_03 0.78
1990_04 0.77	1997_04 0.80	2004_04 0.75
1991_01 0.72	1998_01 0.73	2005_01 0.79
1991_02 0.78	1998_02 0.77	2005_02 0.98
1991_03 0.71	1998_03 0.91	2005_03 0.78
1991_04 0.81	1998_04 0.79	2005_04 0.61
1992_01 0.65	1999_01 0.69	2006_01 0.73
1992_02 0.79	1999_02 0.78	2006_02 0.87
1992_03 0.74	1999_03 0.76	2006_03 0.87
1992_04 0.72	1999_04 0.74	2006_04 0.65
1993_01 0.74	2000_01 0.78	2007_01 0.85
1993_02 0.61	2000_02 0.52	2007_02 0.85
1993_03 0.70	2000_03 0.51	2007_03 0.87
1993_04 0.74	2000_04 0.62	2007_04 0.79
1994_01 0.73	2001_01 0.68	2008_01 0.83
1994_02 0.82	2001_02 0.74	2008_02 0.74
1994_03 0.61	2001_03 0.92	2008_03 0.80
1994_04 0.77	2001_04 0.85	2008_04 0.83
1995_01 0.59	2002_01 0.45	2009_01 0.90
1995_02 0.55	2002_02 0.82	2009_02 0.77
1995_03 0.76	2002_03 0.94	2009_03 0.84
1995_04 0.71	2002_04 0.88	2009_04 0.84
1996_01 0.75	2003_01 0.96	2010_01 0.76
1996_02 0.70	2003_02 0.89	2010_02 0.87
1996_03 0.84	2003_03 0.85	2010_03 0.86
1996_04 0.67	2003_04 0.86	2010_04 0.84

В таблице 2 указаны доходы во всех 84 кварталах, суммы за год, накопленные суммы, а также накопленные суммы при тех же самых ежедневных доходах без деления кварталов на месячные под-периоды.

Таблица 2. Ежеквартальные доходы по предложенному алгоритму.

Годы	1-ый квартал	2-ой квартал	3-ий квартал	4-ый квартал	Суммы за год	Накопления с помесячной схемой	Накопление с поквартальной схемой
1990	204	322	291	222	1234	1234	554
1991	341	-83	256	214	728	1962	666
1992	235	-75	-27	287	420	2382	945
1993	-13	159	127	-13	260	2642	1083
1994	-128	-78	-10	111	-105	2537	924
1995	159	110	63	215	547	3084	1126
1996	240	216	146	363	965	4049	1654
1997	345	344	-146	488	1031	5080	2129
1998	254	213	392	226	1085	6165	1721
1999	196	339	-106	398	827	6992	2266
2000	271	321	80	-380	292	7284	2492
2001	545	-509	286	-52	270	7554	2914
2002	242	361	381	-399	585	8139	2990
2003	-396	236	137	297	274	8413	2822

2004	129	417	292	134	972	9385	2862
2005	-89	-132	252	-183	-152	9233	3078
2006	-444	317	182	-153	-98	9135	3139
2007	266	68	-199	-154	-19	9116	3092
2008	461	-340	-1039	446	-472	8644	2790
2009	179	485	-271	122	515	9159	2974
2010	49	363	-147	-143	122	9281	2657

4. Заключение

Данные из таблиц 1 и 2 позволяют сделать некоторые выводы.

1. Накопленный доход растёт по близкому к линейному закону в 1990-2004 гг. и становится близким к постоянному в 2005-2010 гг. Это относится как к новой схеме с разделением квартала на месяцы, так и к старой схеме без этого деления.
2. Указанное деление квартала на месяцы позволяет увеличить накопленные доходы примерно в 3 раза при тех же самых ежедневных доходах на обоих участках. Это говорит об эффективности и целесообразности такого деления.
3. Простой взгляд на числа в таблице 2 позволяет увидеть, что минимальный квартальный доход (или максимальная потеря) в \$ 1039 приходится на 3-ий квартал 2008 года. Вторая по величине (хотя в два раза меньшая) квартальная потеря в \$ 509 приходится на 2-ой квартал 2001 года. Первый из этих кварталов непосредственно предшествовал ипотечному кризису 2008 года, второй совпал с началом доткомовского кризиса 2001 года. Оба кризиса были самыми большими кризисами в рассматриваемом периоде. Оба числа в таблице выделены.
4. Аналогичное исследование рынка S&P-500 предполагается довести до 2024 года.
5. Предполагается уделить специальное внимание прогностическим возможностям предложенного подхода к анализу других временных рядов.

Список литературы

1. Jallo D., Budai D., Boginski V., Goldengorin B., Pardalos P. M. Network-Based Representation of Stock Market Dynamics: An Application to American and Swedish Stock Markets // Springer in Mathematics & Statistics. 2013. Vol. 32. P. 91-108.
2. Lohrmann C., Luukka P. Classification of intraday S&P500 returns with a Random Forest // International Journal of Forecasting. 2019. Vol. 35, No. 1. P. 28-39.
3. Yanru Guo. Stock Trading Based on Principal Component Analysis and Clustering Analysis // IOP Conf. Series: Materials Science and Engineering. 2020. P. 740-748.
4. Bruni R. Stock Market Index Data and indicators for Day Trading as a Binary Classification // Data in Brief. 2017. Vol. 10. P. 569-575.
5. Fung P.Y. Online two-way trading: Randomization and advice // Theoretical Computer Science. 2021. Vol. 856, No. 8, P. 41-50.
6. Rubchinsky A., Baikova K. Algorithm of Trading on the Stock Market, Providing Satisfactory Results // In: Goldengorin B., Kuznetsov S. (Eds.) Data Analysis and Optimization. Springer Optimization and Its Applications. Cham: Springer, 2023. Vol 202. P. 331-347.