

СРАВНЕНИЕ МЕТОДОВ ПРОГНОЗИРОВАНИЯ ФИНАНСОВЫХ ВРЕМЕННЫХ РЯДОВ

Андрианова Е.Г., Чукалина Е.Р.

*МИРЭА - Российский технологический университет, 119454, Россия, г. Москва, проспект Вернадского, 78,
e-mail: dtghmflys@gmail.com, ekaterina.chukalina@yandex.ru*

Методы прогнозирования служат для исследования системных связей и закономерностей функционирования объектов и процессов с использованием современных методов обработки информации и представляют собой важное средство в анализе сложных прикладных систем, в работе с информацией, в целенаправленном воздействии человека на объекты исследования с целью повышения эффективности их функционирования. В современной экономике прогнозирование занимает особое место. В статье приведено сравнение моделей и методов прогнозирования финансовых временных рядов.

Ключевые слова: финансовые временные ряды, прогнозирование, регрессионные модели, авторегрессионные модели, модели экспоненциального сглаживания, нейросетевые модели.

COMPARISON OF FINANCIAL TIME SERIES FORECASTING METHODS

Andrianova E.G., Chukalina E.R.

*MIREA - Russian Technological University, 119454, Moscow, 78 Vernadskogo Avenue, Russia
e-mail: dtghmflys@gmail.com, ekaterina.chukalina@yandex.ru*

Forecasting methods serve to study the system relationships and patterns of operation of objects and processes using modern methods of information processing and are an important tool in the analysis of complex application systems, in working with information, in the targeted human impact on research objects in order to increase their efficiency. In the modern economy, forecasting occupies a special place. This paper provides a comparison of models and methods of forecasting financial time series.

Key words: financial time series, forecasting, regression models, autoregression models, exponential smoothing models, neural network models.

Введение

Первостепенной задачей естественных наук является описание некоторого процесса или явления на основе экспериментальных исследований и наблюдений. Под описанием подразумевается построение математической модели с некоторыми допущениями. В свою очередь такая модель должна как можно полнее описывать наблюдаемое явление или процесс, а также подразумевать решение задачи прогнозирования.

В современной экономике прогнозирование занимает особое место. Фондовый рынок в настоящее время остается одной из самых сложных систем для моделирования. Он быстро меняется, так как на него влияют различные факторы. Способность прогнозировать рынок и его ожидаемый исход - актуальная проблема, с которой акционеры сталкиваются сегодня [1].

Показателями состояния и динамики рынка ценных бумаг являются фондовые индексы. Поведение рынка можно изучить, сопоставляя их значения с предыдущими показаниями. С развитием экономики увеличиваются фондовые индексы. Однако этот рост наблюдается не линейно. Такая неустойчивость вызывает сложности в проведении анализа и прогнозирования.

Любая задача, связанная с инвестиционной деятельностью и финансовыми инструментами, сопровождается риском и требует тщательного расчета. Предсказание поведения будущих процессов подразумевает грамотный и качественный анализ данных, на основе которых строится прогноз. Также немаловажно – выбрать метод прогнозирования. На выявление скрытых закономерностей в накопленных данных и должны опираться

прогнозы таких систем. Поэтому предсказание финансовых временных рядов – необходимый элемент любой инвестиционной деятельности [2].

Фундаментальный и технический анализ

Существуют два основных подхода к анализу рынка: технический и фундаментальный.

Фундаментальный анализ – метод прогнозирования рыночной стоимости компании, основанный на анализе финансовых и производственных показателей деятельности. Истинные стоимости ценных бумаг компании сравниваются с текущими ценами с целью определения величин отклонения. Однако, в случае определения оптимального момента для извлечения прибыли более оптимален технический анализ.

При таком виде анализа финансового рынка не берутся в расчет возможные связи с некоторыми ценными бумагами. На основе анализа изменений конкретного финансового инструмента в прошлом прогнозируется его динамика. Этот подход обоснован сосредоточенностью участников именно на том инструменте, с которым они в определенный момент работают. Наиболее часто методы технического анализа используются для анализа цен, изменяющихся свободно, например, на биржах.

В техническом анализе все инструменты и методы основаны на анализе временных рядов. Можно спрогнозировать поведение цен, выделяя тренды.

Модели прогнозирования временных рядов

Временным рядом $\{x_i\}_{i=1}^n$ называется массив из n чисел, представляющих собой значения некоторой наблюдаемой динамической переменной $x(t)$ с некоторым постоянным шагом τ по времени, $t_i = t_0 + (i-1)\tau$: $x_i = x(t_i), i=1, \dots, n$.

Таким образом, любой временной ряд имеет два обязательных элемента: время и уровень ряда (конкретное значение показателя). Их можно различать по следующим признакам:

- время (моментные, интервальные);
- форма представления уровней (временные ряды абсолютных, относительных и средних величин);
- расстояние между интервалами времени (полные и неполные временные ряды);
- содержание показателей (временные ряды частных и агрегированных показателей).

Следует отметить, что для корректного формирования временных рядов важнее всего является сопоставимость уровней, образующих ряд.

В теории временных рядов выделяют две основные задачи: идентификации и прогноза. Ответ на вопрос как был порожден данный временной ряд можно получить, решая задачу идентификации. Задача прогноза встает при попытке на основе данных временного ряда в прошлом предсказать будущие значения ряда. К сожалению, вместо полного решения уравнений наличие только лишь временного ряда значительно ограничивает знания об изучаемой системе.

Всякий временной ряд имеет свои индивидуальные характеристики такие как волатильность, стационарность и другое. Следовательно, для разных временных рядов подходят разные методы и модели прогнозирования.

Методы прогнозирования разделяют на две группы [3]: интуитивные и формализованные. Первая группа методов не предполагает разработку моделей прогнозирования и отражает суждения экспертов. Они применяются, когда объект слишком прост или настолько сложен, что невозможно аналитически учесть влияние внешних факторов. Вторая группа методов основывается на построении модели прогнозирования. В обзоре [4] модели разделяются на статистические модели и структурные модели. К таким моделям временных рядов относят: регрессионные модели, авторегрессионные модели, модели экспоненциального сглаживания, марковские модели, модели, основанные на деревьях решений и другие.

Архитектура IoT систем

IoT системы в настоящее время применяются в разных сферах. Наиболее частыми решения по IoT системам предлагаются в следующих сферах:

- умный дом;
- умный город;
- предсказание погоды/катаклизмов;
- отслеживание транспорта;
- офисные корпоративные системы;
- индустрия 4.0;
- автомобильная промышленность;

Для каждой из этих областей нет единой, стандартизированной архитектуры IoT систем, так как каждая отрасль по - своему уникальна. Для каждой области предоставляется свой набор требований и правил к функционированию и безопасности системы. Это является одной из главных проблем в области IoT систем – отсутствие общепринятой эталонной архитектуры [5].

Несмотря на то, что все IoT системы имеют общие компоненты, каждый производитель предлагает свои решения в отдельных элементах архитектуры IoT систем. В таблице 1 приведены предложения по эталонной архитектуре IoT систем.

Наличие большого количества подходов к стандартизации означает то, что данная проблема не будет решена в ближайшее время. Соответственно, под вопросом стоит безопасность таких систем как со стороны информационной, так и функциональной безопасности.

Регрессионные модели прогнозирования (простая линейная регрессия, множественная регрессия, нелинейная регрессия)

В задачах, где требуется изучить отношение между двумя и более переменными, используется регрессионный анализ.

Модель в регрессионном анализе представляет собой функцию независимой переменной и параметров с добавленной случайной переменной. Параметры модели настраиваются таким образом, что модель наилучшим образом приближает данные: неизвестные параметры представлены в виде скаляра или вектора; ошибка представляет собой случайную величину; невозможно выразить какой-либо предиктор в виде линейной комбинации других и т.д.

Модель линейной регрессии предполагает, что связь между зависимой и независимой переменной является линейной. Это соотношение моделируется с помощью случайных возмущений или ошибки ε_t – ненаблюдаемой случайной величины. Простая линейная регрессия описывается следующим уравнением:

$$Z(t) = \alpha_0 + \alpha_1 X(t) + \varepsilon_t,$$

где, α_0 , α_1 - коэффициенты регрессии, ε_t - ошибка модели.

Множественной называют линейную регрессию, в модели которой число независимых переменных две или более. Почти во всех реальных регрессионных моделях используется несколько предикторов, а базовые описания линейной регрессии часто формулируются в терминах модели множественной регрессии.

Нелинейная регрессия - это форма регрессионного анализа, в которой данные наблюдений моделируются функцией, которая представляет собой нелинейную комбинацию параметров модели и зависит от одной или нескольких независимых переменных. Примеры нелинейных функций включают экспоненциальные функции, логарифмические функции, тригонометрические функции, степенные функции и другие.

Целью регрессионного анализа является выявление зависимости между исходной переменной и внешними факторами. Для определения коэффициентов регрессии используются метод наименьших квадратов и метод максимального правдоподобия.

Авторегрессионные модели прогнозирования (ARIMAX, GARCH, ARDLM)

Модель временных рядов, в которой значения временного ряда в данный момент линейно зависят от предыдущих значений этого же ряда. Наиболее популярными являются модель авторегрессии (1) и модель скользящего среднего (2).

$$Z(t) = C + \varphi_1 Z(t-1) + \varphi_2 Z(t-2) + \dots + \varphi_n Z(t-p) + \varepsilon_t, (1)$$

где C - вещественная константа, $\varphi_1, \dots, \varphi_n$ - коэффициенты, p - порядок авторегрессии, ε_t - ошибка модели.

$$Z(t) = 1/q (Z(t-1) + Z(t-2) + \dots + Z(t-q)) + \varepsilon_t, (2)$$

где q - порядок скользящего среднего, ε_t - ошибка прогнозирования.

Часто в одной модели объединяют авторегрессию и скользящее среднее. Такая модель представляет собой фильтр в виде скользящего среднего порядка q и авторегрессию фильтрованных значений процесса порядка p .

Чаще всего модель используется для моделирования волатильности.

Модели экспоненциального сглаживания (взвешенное скользящее среднее, экспоненциальное сглаживание, модель Хольта, модель Хольта-Винтерса)

Выявление и анализ тенденции временного ряда часто производится с помощью его выравнивания или сглаживания. В основу экспоненциального сглаживания заложена идея постоянного пересмотра прогнозных значений по мере поступления фактических. Функция модели имеет вид:

$$Z(t) = S(t) + \varepsilon_t,$$

$$S(t) = \alpha \cdot Z(t-1) + (1-\alpha) \cdot S(t-1),$$

где α - коэффициент сглаживания, $0 < \alpha < 1$, $S(1) = Z(0)$ - начальные условия. В данной модели каждое последующее сглаженное значение $S(t)$ является взвешенным средним между предыдущим значением временного ряда $Z(t)$ и предыдущего сглаженного значения $S(t-1)$.

Модели экспоненциального сглаживания наиболее популярны для долгосрочного прогнозирования.

Модель на нейронных сетях

Модели на основе нейронных сетей являются наиболее популярными среди структурных методов прогнозирования временных рядов. Принцип работы искусственной нейронной сети основан на некоей математической интерпретации работы биологической нейронной сети. За счет этого они способны видеть нелинейные связи.

Нейросети активно используют пенсионные фонды и страховые компании, работающие с большими портфелями, для которых особенно важны корреляции между различными рынками.

Они способны строить оптимальную модель прогнозирования. Такая модель адаптивна и меняется вместе с рынком, что имеет огромное значение для современных высокочастотных финансовых рынков.

При помощи этой модели возможно моделирование нелинейной зависимости будущего значения временного ряда от его фактических значений и от значений внешних факторов. Модель на основе искусственных нейронных сетей, состоящих из нейронов, описывающихся уравнениями:

$$U(t) = \sum_{i=1}^m \omega_i \cdot Z(t-i) + b,$$

$$Z(t) = \varphi(U(t)),$$

где $Z(t-1), \dots, Z(t-m)$ — входные сигналы, $\omega_1, \dots, \omega_m$ — синаптические веса нейрона, p — порог, $\varphi(U(t))$ — функция активации (функция единичного скачка, кусочно-линейная функция, сигмоидальная функция).

Нейронные сети делятся на несколько типов в зависимости от способа связи нейронов [3]: однослойные нейронные сети прямого распространения, многослойные нейронные сети прямого распространения, рекуррентные нейронные сети.

Модель на цепях Маркова

Данная модель предполагает, что будущее состояние системы не зависит от предыдущих состояний, а зависит только от текущего [4].

Модель на классификационно-регрессионных деревьях

Прогнозное значение модели временного ряда зависит от предыдущих значений, а также некоторых независимых переменных. Классификационно-регрессионные деревья разработаны для моделирования процессов, на которые оказывают влияние как непрерывные внешние факторы, так и категориальные.

Кроме моделей, перечисленных выше, существуют модель на основе генетического алгоритма, модель на опорных векторах, модель на основе передаточных функций, модель на нечеткой логике и т.д. Также отметим, что на сегодняшний день существует множество модификаций моделей.

Сравнение моделей прогнозирования

Регрессионные модели отличаются простотой, гибкостью, единообразием их анализа и проектирования, не смотря на сложность определения вида функциональной зависимости, параметров модели и отсутствие способности моделирования нелинейных процессов. Если использовать линейные регрессионные модели, результат прогноза может быть получен быстрее, нежели при использовании остальных моделей. Также нельзя не отметить прозрачность моделирования, то есть все промежуточные вычисления доступны для анализа. Определение функциональной зависимости переменных можно отнести к недостаткам в силу своей сложности. Кроме того, вычисление параметров модели является трудоемким процессом. Наибольшую эффективность регрессионные модели и методы показывают на стационарных рядах.

Авторегрессионные модели содержат большое число параметров, что влияет на адаптивность моделей. Главным достоинством является простота моделирования, единообразие анализа и проектирования. Но модель линейна. Кроме того, авторегрессионные модели достаточно популярны при решении задач прогнозирования временных рядов различных областей. Недостатки: определение большого числа параметров (также недостаток) – процесс ресурсоемкий; модели недостаточно адаптивны.

Методы экспоненциального сглаживания чаще всего используются для долгосрочного прогнозирования из-за простоты и единообразия их анализа и проектирования. Недостатком данного класса моделей является отсутствие гибкости. Однако, если во временных рядах присутствует сезонная составляющая, то модели экспоненциального сглаживания дают наилучший результат

Модели на нейронных сетях являются нелинейными, адаптивными и масштабируемыми. Также присутствует единообразие их анализа и проектирования. При этом отсутствует прозрачность моделирования. Следующими недостатками являются сложность выбора архитектуры, высокие требования к непротиворечивости обучающей выборки, сложность выбора алгоритма обучения и ресурсоемкость процесса их обучения. Но для временных рядов, которые подвержены кризисным процессам лучшую эффективность дают методы на основе нейронных сетей.

Отсутствие моделирования процессов с длинной памятью является главным недостатком моделей на базе цепей Маркова. В то время как простота и единообразие анализа и проектирования являются достоинствами.

Модели на базе классификационно-регрессионных деревьев отличает возможность использовать категориальные внешние факторы. Достоинствами также являются масштабируемость, быстрота и однозначность процесса обучения дерева. Однако, присутствует неоднозначность алгоритма построения структуры дерева, сложность вопроса останова, отсутствие единообразия их анализа и проектирования.

Заключение

Для многих предметных областей задача прогнозирования имеет огромную важность и является неотъемлемой частью повседневной работы многих компаний. Теоретические исследования, в которых используется анализ временных рядов, могут дать мощный инструмент для понимания многих явлений. Прогноз будущих значений на основе прошлых наблюдений позволяет наиболее эффективно принимать решения в настоящем.

Можно сделать вывод о том, что не существует универсального метода прогнозирования временных рядов, каждый метод находит свое применение для разных типов временного ряда. Например, на рядах, которые имеют ярко выраженную тенденцию, хорошую прогностическую способность проявляют регрессионные модели; на рядах, в которых присутствует сезонная составляющая – модели экспоненциального сглаживания. Для временных рядов, которые подвержены кризисным процессам лучшую эффективность дают методы на основе нейронных сетей.

На сегодняшний день наиболее распространенными моделями прогнозирования являются авторегрессионные модели и модели на нейронных сетях [6,7]. В то же время, создание комбинированных моделей и методов является наиболее перспективным направлением в прогнозировании временных рядов. Такой подход дает возможность компенсировать недостатки одних моделей с помощью других. Также делается упор на повышение точности прогнозирования, в качестве главного критерия эффективности модели.

Список литературы

1. Андрианова Е.Г., Новикова О.А. Роль методов интеллектуального анализа текста в автоматизации прогнозирования рынка ценных бумаг. *Cloud of Science*. 2018. Т. 5. № 1. с.202-217.
2. Андрианова Е.Г., Головин С.А., Зыков С.В., Лесько С.А., Чукалина Е.Р. Обзор современных моделей и методов анализа временных рядов динамики процессов в социальных, экономических и социотехнических системах. *Российский технологический журнал*. 2020. 8(4):7-45. DOI: 10.32362/2500-316X-2020-8-4-7-45
3. Alfares H.K., Nazeeruddin M. Electric load forecasting: literature survey and classification of methods // *International Journal of Systems Science*. 2002, Vol 33. P. 23 – 34.
4. Jingfei Yang M. Sc. Power System Short-term Load Forecasting: Thesis for Ph.d degree. Germany, Darmstadt, *Elektrotechnik und Informationstechnik der Technischen Universitat*, 2006.
5. Zhu J., Hong J., Hughes J.G. Using Markov Chains for Link Prediction in Adaptive Web Sites // 1st *International Conference on Computing in an Imperfect World*, UK, London, 2002. P. 60 – 73.
6. Егошин А.В. Анализ и прогнозирование сложных стохастических сигналов на основе методов ведения границ реализаций динамических систем: Автореферат диссертации ... канд. техн. наук. Санкт-Петербург, 2009. 19 с.
7. Тихонов Э.Е. Прогнозирование в условиях рынка. Невинномысск, 2006. 221 с.
8. Хайкин С. Нейронные сети: полный курс. М.: ООО «И. Д. Вильямс», 2006. 1104 с.

References

1. Andrianova E.G., Novikova O.A. Rol' metodov intellektual'nogo analiza teksta v avtomatizacii prognozirovaniya rynka cennyh bumag. Cloud of Science. 2018. T. 5. № 1. s.202-217.
2. Andrianova E.G., Golovin S.A., Zыkov S.V., Les'ko S.A., CHukalina E.R. Obzor sovremennyh modelej i metodov analiza vremennyh ryadov dinamiki processov v social'nyh, ekonomicheskikh i sociotekhnicheskikh sistemah. Rossijskij tekhnologicheskij zhurnal. 2020. 8(4):7-45. DOI: 10.32362/2500-316X-2020-8-4-7-45
3. Alfares H.K., Nazeeruddin M. Electric load forecasting: literature survey and classification of methods // International Journal of Systems Science. 2002, Vol 33. P. 23 – 34.
4. Jingfei Yang M. Sc. Power System Short-term Load Forecasting: Thesis for Ph.d degree. Germany, Darmstadt, Elektrotechnik und Informationstechnik der Technischen Universitat, 2006.
5. Zhu J., Hong J., Hughes J.G. Using Markov Chains for Link Prediction in Adaptive Web Sites // 1st International Conference on Computing in an Imperfect World, UK, London, 2002. P. 60 – 73.
6. Egozhin A.V. Analiz i prognozirovanie slozhnyh stohasticheskikh signalov na osnove metodov vedeniya granic realizacij dinamicheskikh sistem: Avtoreferat dissertacii ... kand. tekhn. nauk. Sankt-Peterburg, 2009. 19 s.
7. Tihonov E.E. Prognozirovanie v usloviyah rynka. Nevinnomyssk, 2006. 221 s.
8. Hajkin S. Nejronnye seti: polnyj kurs. M.: OOO «I. D. Vil'yams», 2006. 1104 s.