

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАФЕДРА МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЭНЕРГЕТИЧЕСКИХ СИСТЕМ

Выпускная квалификационная работа бакалавра

Каплун Игорь Борисович

**СОВРЕМЕННЫЕ МЕТОДЫ АНАЛИЗА ДАННЫХ
В ИНВЕСТИЦИОННОМ РИСК МЕНЕДЖМЕНТЕ**

Прикладная математика и информатика 01.03.02

Научный руководитель:
К. ф. м. наук,
доцент
Балыкина Ю. Е.

Санкт-Петербург
2018

Оглавление

| | |
|---------------------------------------------|-----------|
| Введение | 3 |
| Обзор Литературы | 4 |
| Список использованных обозначений | 14 |
| Постановка задачи | 14 |
| Оценка риска портфеля | 16 |
| Модель оценки рисков | 16 |
| Модели волатильности | 17 |
| Подбор распределения доходностей | 21 |
| Оценка ожидаемой доходности | 26 |
| Нейросети прямого распространения | 26 |
| Рекуррентные нейронные сети | 28 |
| Результаты | 31 |
| Диверсификация портфеля | 31 |
| Оценка ожидаемой доходности | 32 |
| Прогнозирование рискованности | 34 |
| Бэктестинг | 37 |
| Заключение | 39 |
| Приложения | 40 |
| Литература | 45 |

Введение

Математические основы для современных методов, используемых в инвестиционном риск менеджменте, берут истоки в середине XX века. Однако они получили развитие в области численных финансов сравнительно недавно. Толчком к росту необходимости в подобных методах стала повсеместная компьютеризация биржевой деятельности в 1970-х годах и открытие доступа к интернет торгам в 1990-х, что привело к нескольким серьезным изменениям на фондовых рынках: значительно увеличилось количество сделок; резко возросло количество информации, генерируемой биржами. На исследования и разработку современных подходов к риск менеджменту также оказало сильное влияние развитие компьютерных технологий и достижения в области анализа данных, которые позволили обрабатывать большие объемы информации и дали доступ к компьютеризированному решению задач из области численных финансов большому количеству людей. Актуальность подобных задач объясняется большим практическим потенциалом и научной сложностью решений.

В настоящей работе рассматривается задача управления риском инвестиционного портфеля. Основной подход к управлению риском – перенос риска. Он располагает тремя схемами: страхованием, хеджированием и диверсификацией. Первые два способа управления риском заключаются в создании производных финансовых инструментов и реализации торговых стратегий, основанных на их использовании, что выходит за рамки выбранной темы исследования, сконцентрированной на процессе формирования диверсифицированного инвестиционного портфеля. Классический подход к диверсификации модифицирован при помощи современных методов анализа данных, которые позволили получить более точные оценки рисков и прогнозов доходностей финансовых инструментов. Основной целью работы является демонстрация возможностей современных методов анализа данных к увеличению точности работы классических моделей портфельной оптимизации.

Обзор Литературы

1. Диверсификация инвестиционного портфеля

Инвестиционный портфель (investment portfolio) представляет собой набор из нескольких финансовых инструментов, доли которых определяются набором весов. *Диверсификация* инвестиционного портфеля предполагает настройку весов входящих в него активов таким образом, чтобы полученный портфель удовлетворял заранее установленным ограничениям. Инвестиционные портфели появились в то же время, как и само понятие инвестиции, однако в самостоятельную теорию составление инвестиционных портфелей отделилось лишь в середине XX века. Американский экономист Г. Марковиц в 1952 году представил свою "Современную теорию портфелей" (Modern Portfolio Theory (MPT)). В ней постулировались две основные идеи [1]:

1. для того, чтобы рассчитывать на больший доход, необходимо мириться с большим уровнем риска;
2. при инвестировании в несколько ценных бумаг возможно уменьшить уровень риска при заданном уровне дохода путем составления диверсифицированного инвестиционного портфеля.

Эти утверждения легли в основу Mean-Variance asset allocation model (M-V), которая стала одной из первых математических моделей формирования инвестиционного портфеля. Таким образом, основными параметрами моделей портфельной оптимизации являются ожидаемая доходность и рискованность входящих в состав портфеля финансовых инструментов. Модель Г. Марковица базируется на формировании эффективного множества портфелей и определения кривой Efficient Frontier (EF) – геометрического места точек, соответствующих всевозможным портфелям, составленным из определенного набора активов и имеющих наивысший уровень доходности для фиксированного уровня риска. Р. Мертон позднее развил идею EF, привел несколько теорем и свойств, раскрывающих возможности использования метода и описывающих ограничения, накладываемые на него [2].

В то же время Д. Трейнор и У. Шарп разработали Capital Asset Pricing Model (CAPM), позволяющую рассматривать стоимость активов

вместе с их рыночными рисками [3]. Модель предполагает формирование Capital Market Line, которая использует понятие безрискового актива для поиска оптимального инвестиционного портфеля, тем самым дополняя EF. В работе [4] рассматривается современное применение моделей EF и CAPM в оценке решения задачи портфельной оптимизации.

Указанные подходы к формированию портфеля обладают несколькими недостатками, которые делают их популярными лишь в научных кругах: ряд ограничений и предположений, таких как запрет на короткие позиции и предположение о соответствии доходностей определенному распределению; неинтуитивность определения весов; низкая точность оценок параметров моделей. Модель Блэка-Литтермана, описанная в [5], считается менее подверженной перечисленным недостаткам, во многом благодаря тому, что она позволяет учитывать видение и предпочтения инвесторов в своей математической формулировке. Также существует множество современных методов решения задачи диверсификации. В работе [6] предлагается усовершенствованная динамическая модель Блэка-Литтермана. В [7] рассматривается возможность оценки меры риска портфельной инвестиции при помощи авторегрессионных моделей, а задача портфельной оптимизации представляется в виде задачи многокритериальной оптимизации, для решения которой необходимо найти такой портфель, параметры которого доставляли бы минимум мере риска и максимум ожидаемой доходности. Авторы статьи решают обозначенную задачу путем нахождения множества оптимальных решения по Парето (Парето фронта) при помощи эволюционного алгоритма. Авторы [8] предлагают производить оценку ожидаемой доходности элементов портфеля при помощи нейронной сети, тем самым обходя проблему неточных предположений о характере распределения доходностей финансовых инструментов.

2. Оценка рисков инвестиционного портфеля

Для оценки рискованности портфельной инвестиции Г. Марковиц одним из первых определил *меру риска* как меру неопределенности, в качестве которой стал использовать стандартное отклонение [3]:

$$\sigma = \sqrt{\frac{1}{T} \sum_{i=1}^T (r_i - \hat{r})^2},$$

где T – количество рассматриваемых периодов времени; r_i – доходность финансового инструмента в момент времени i ; \hat{r} – выборочное среднее доходностей финансового инструмента за время T .

Стандартное отклонение является наиболее простым и широко применимым инструментом оценки риска, однако оно не является универсальным, поскольку несет в себе небольшое количество информации: лишь максимально возможное изменение доходности финансового инструмента в следующий период времени с некоторой вероятностью. Стандартное отклонение не способно сообщить численную оценку величины возможных потерь, а также включает в себя как положительную, так и отрицательную динамику доходностей.

В связи с низкой информативностью стандартного отклонения для оценки рискованности портфелей был разработан целый ряд альтернативных метрик. Так, в модели CAPM применяется коэффициент β , который показывает, насколько ценная бумага или инвестиционный портфель зависят от колебаний рынка [3]:

$$\beta = \frac{cor_{A,B}\sigma_A}{\sigma_B},$$

где $cor_{A,B}$ – коэффициент корреляции между активами A и B ; σ_A и σ_B – стандартные отклонения доходностей активов A и B .

Коэффициент Шарпа (Sharpe Ratio) показывает, насколько оптимально сформирован портфель. Чем выше его значение, тем больший прирост ожидаемой доходности ожидает инвестора при малом приросте рискованности [3]:

$$SharpeRatio = \frac{E_r - R_f}{\sigma},$$

где E_r и R_f – ожидаемые доходности портфеля и безрискового актива соответственно.

Для того, чтобы избежать проблемы объединения положительной и отрицательной волатильности (изменчивости) в одном показателе, можно использовать отрицательную волатильность (downside volatility) вместо стандартного отклонения. Она основана на отклонении реальной доходности от установленного инвестором уровня желаемой доходности инвестиций. В [9] описан коэффициент Сортино, в основе которого лежит приме-

нение отрицательной волатильности, а также приводится сравнение его с коэффициентом Шарпа:

$$S = \frac{R - E_r}{DV},$$
$$DV = \sqrt{\sum_{i=1}^N \text{Min}(0, R_i - E_r)},$$

где R – средний уровень доходности за N периодов; E_r – желаемый уровень доходности; DV – отрицательная волатильность; R_i – доходность за каждый период i , $i = 1, \dots, N$.

В 1980-х годах финансистами компании J. P. Morgan был разработан Value-at-Risk метод, который стал основой для их собственной технологии RiskMetrics [10]. Лишь в 1994 году идея VaR была впервые представлена общественности, а двумя годами позднее – опубликован RiskMetrics Technical Document, описывающий вначале лишь параметрический метод вычисления оценки VaR, а затем, уже дополненный, и непараметрические методы оценки.

В качестве меры риска VaR определяет максимальный уровень потерь, которые может понести инвестор в заданном доверительном интервале, т. о. VaR предоставляет не только оценку вероятности потерь, как это делает стандартное отклонение, но и их численное выражение:

$$VaR = P_t \sigma q_\alpha,$$

где P_t – стоимость актива в момента времени t ; q_α – квантиль распределения доходностей уровня α .

С момента публикации RiskMetrics было представлено много научных трудов на тему применения и измерения VaR. Так, например, в 1996 году была опубликована работа [11], описывающая идею VaR, а также подходы к его измерению. В ней рассматриваются два основных подхода к измерению VaR, которые и сегодня являются наиболее актуальными: параметрический (вариационно-корреляционный метод) и непараметрический (историческая симуляция и метод Монте-Карло). В упомянутой работе, а также во многих последующих трудах, делаются выводы о том, что указанные методы измерения VaR далеко не универсальны и не способны дать

точную оценку рискованности инвестиции. Основными недостатками VaR как правило называют слишком низкую точность для высокой заявленной надежности и информативности данной меры, вызванные низкой точностью оценок параметров, а также различия в значениях VaR, полученных по результатам разных вычислений.

В дальнейшем предпринимались неоднократные попытки более точного вычисления VaR. Так, в работах [12, 13] описываются возможности параметрического дельта-нормального метода вычисления VaR и усовершенствованные методы исторической симуляции. Иной подход к улучшению точности оценки VaR предлагается к рассмотрению в исследованиях [14, 15, 16, 17]: в указанных работах демонстрируется возможность прогнозирования рискованности посредством моделей волатильности на разных финансовых рынках: ARCH, GARCH и их производных. Подобные модели позволяют учитывать гетероскедастичность данных, т. е. непостоянство дисперсии. Авторы указанных статей приходят к схожим выводам касательно наилучшей модели волатильности для оценки VaR: в [14] были выбраны модели T-GARCH(1,1) и EGARCH(1,1) с ошибками, распределенными по Стьюденту и Johnson SU соответственно; в [15, 16] – модели ARCH(1) и GARCH(1,1) с ошибками, распределенными по Стьюденту (в [15]) и нормальному распределению (в [16]); в [17] – модели EGARCH(1,1) и GARCH(1,1) с ошибками, распределенными по Стьюденту. В исследовании [17] также заключается, что более сложные модели представляют более точные оценки, а вот повышение порядка модели не всегда приводит к повышению точности результатов.

В статье [18] детально описаны модели Markov Switching GARCH (MS-GARCH), объединяющие в себе концепцию марковских моделей смены режима и моделей GARCH, делая последние более восприимчивыми к разнице между волатильностью в период роста и падения рынка. Основной проблемой, с которой сталкиваются исследователи при использовании моделей класса MS-GARCH, является проблема зависимости значений временного ряда от последовательности предыдущих (path dependence problem), что делает оценку параметров модели чересчур трудоемкой задачей. Авторы статьи рассматривают два подхода к решению обозначенной проблемы: аппроксимация, предполагающая переход к более простым мо-

делям, и симуляция, позволяющая свести задачу оценки параметров к более простой. Они заключают, что второй подход является более надежным и эффективным. Одной из первых и самых известных работ, в которых применяется первый подход, является [19]. В ней авторы предлагают перейти к модели MS-ARCH, значительно сокращая количество оцениваемых параметров. В исследовании [20] применяется алгоритм сэмплинга, сочетающий в себе элементы алгоритма Monte Carlo Markov Chain (MCMC) и алгоритма Гиббса. Несмотря на хорошие результаты работы обозначенного алгоритма, скорость его работы и сложность реализации делает его мало применимым на практике.

Предположение о характере распределения доходностей, лежащего в основе VaR, заставляет исследователей обратиться к проблеме оценки хвостов этих распределений (expected tail loss estimation), т. е. оценке тех событий, которые лежат вне доверительного интервала VaR. Одним из самых популярных методов оценки хвостов, является Expected Shortfall (или Conditional VaR (CVaR)), описание и применение которой, демонстрируются, в частности, в работах [21, 22]. Авторы статей [23, 24] проводят сравнение целого ряда мер рисков. В [23] утверждается, что VaR не может корректно измерять риск, и делается попытка найти более подходящую меру риска, которой, по результатам исследования, становится CVaR. В [24] исследуется наиболее подходящая риск метрика для портфельной оптимизации. Автор статьи заключает, что, несмотря на то, что некоторые метрики в определенных случаях демонстрируют выдающиеся результаты, в общем случае ни одна из них, в том числе и VaR, не может быть названа доминирующей. Ниже (см. рис. 1) представлена сводная таблица вышеперечисленных мер риска.

3. Оценка ожидаемой доходности

В качестве оценки ожидаемой доходности зачастую используют выборочное среднее доходностей финансового инструмента за взятые периоды времени или математическое ожидание распределения, по которому, предположительно, распределены доходности. Подобные подходы к оценке ожидаемой доходности не являются точными и приводят к низкому качеству работы моделей портфельной оптимизации. Альтернативой является подход, согласно которому оценка ожидаемой доходности является прогно-

| Мера риска | Назначение |
|------------------------|------------------------------------------------------------------------------------|
| Стандартное отклонение | основа для других мер риска |
| Коэффициент β | оценка степени зависимости от рынка финансового инструмента |
| Коэффициент Шарпа | мера оптимальности сформированного портфеля |
| Коэффициент Сортино | выявление отрицательной волатильности |
| VaR | оценка величины максимально возможных потерь |
| CVaR | оценка средней величины возможных потерь за пределами доверительного интервала VaR |

Рис. 1: Метрики риска

зом некоторой модели, описывающей поведение временного ряда.

Сложность анализа финансовых данных обусловлена нелинейным характером движения цен финансовых инструментов, высокой волатильностью и зашумленностью данных, однако их ключевой особенностью является зависимость последующих элементов данных от предыдущих, что и заставляет исследователей зачастую обращаться к временным рядам. Временной ряд – это множество данных, получаемых последовательно во времени.

Авторегрессионные модели – самые распространенные и широко применимые модели прогнозирования финансовых временных рядов. Они используют линейную регрессию на предыдущие члены временного ряда для предсказания будущих значений ряда. Чаще всего в научных работах прибегают к использованию моделей класса ARIMA. С целью оценить эффективность предсказательной способности моделей класса ARIMA авторы статьи [25] протестировали различные модели ARIMA на акциях 53 индийских компаний. Они выявили, что модели способны прогнозировать стоимость акций со средней точностью 85%.

Тем не менее, ARIMA имеет серьезные ограничения, связанные со своей линейной природой. Финансовые данные зачастую не могут быть полностью описаны линейно, поэтому многие исследователи прибегают к использованию гибридных моделей. Так, к примеру, авторы [26] примени-

ли метод опорных векторов для анализа нелинейной части временного ряда, где в качестве регрессоров использовали текстовые новостные данные. Впрочем, наиболее распространенные гибридные модели на основе ARIMA основаны на использовании нейронных сетей для предсказания нелинейной части данных, о чем будет сказано ниже.

Финансовые рынки склонны периодически менять свое поведение довольно резко. Смена поведения финансового рынка зачастую называется “сменой режима”. Описанные выше модели прогнозирования финансовых рынков не могут предсказать такого резкого изменения. Для отслеживания подобных случаев используются *скрытые марковские модели* (СММ) (Hidden Markov Models (НММ)), суть которых состоит в том, чтобы определить скрытое состояние (режим), в котором в данный момент находится объект (рынок), используя лишь внешние наблюдения (цены). Изначально скрытые марковские модели использовались преимущественно в задачах распознавания речи. Одними из первых, кто представил работу по применению НММ для предсказания режимов финансовых рынков были А. Вигенд и С. Ши. В своей работе [27] 1998 г. они использовали НММ для определения трех режимов рынка с целью дальнейшего обучения нейросетей (“экспертов”) на классифицированных данных.

В работах [28, 29] развили их идею, протестировав НММ на разных типах данных и сильно увеличив точность результатов: они перешли к смеси Гауссовских распределений на каждом режиме, ввели переоценку модели на каждом шаге и добавили экспоненциальную взвешенность в основной алгоритм обучения модели с тем, чтобы присваивать большие веса новым данным.

В последние годы рядом специалистов по анализу численных финансов были предложены проекты по предсказанию режимов рынка при помощи НММ. Так, М. Холлс-Мур [30] построил торговую стратегию на основе обнаружения периодов разной волатильности при помощи НММ, а Б. Кристофер [31] изучил возможность применения НММ для предсказания “рыночного дна”. Однако, в приведенных работах авторы признают, что их результаты сложно проверить, и они точны настолько, насколько точны их предположения.

Благодаря своей гибкости и способности к прогнозированию дан-

ных с нелинейными зависимостями в прогнозировании финансовых рынков с конца 20-го века стали активно применяться *нейронные сети*, демонстрирующие стабильно высокие результаты. В научной литературе представлено много трудов, сравнивающих работу нейронных сетей различной архитектуры. В работе [32] рассматривается широкий спектр возможных нейросетей, использованных для прогнозирования финансовых данных: нейросеть прямого распространения (Feed Forward Neural Network (FNN)), обучаемая при помощи метода обратного распространения ошибки (Back Propagation (BP)), сеть с радиально-базисными функциями (Radial Basis Network (RBN)), рекуррентная нейросеть (Recurrent Neural Network (RNN)) и нейросеть обобщенной регрессии (Generalized Regression Network (GRNN)). Авторы статьи приходят к нескольким значимым выводам: прогнозы RNN дают наименьшую RMSE

$$RMSE = \sqrt{\frac{\sum_{i=1}^{T_1} (r_i - \bar{r}_i)^2}{T_1}},$$

где T_1 – размер исследуемой выборки; r_i – истинные доходности, а \bar{r}_i – предсказанные. С ростом числа скрытых нейронов и число нейронов, подаваемых на вход сети, RMSE вначале падает, а затем вновь начинает расти. Авторы [33] демонстрируют работу улучшенной RNN – нейросети с долгой краткосрочной памятью (Long – Short Term Memory Network (LSTM)). Они приходят к заключению о том, что LSTM сеть наилучшим образом подходит для анализа финансовых данных и демонстрируют ее эффективность на примере торговой стратегии на основе LSTM сети.

Как было обозначено ранее, зачастую авторегрессионные модели дополняют нейронными сетями. Подобные подходы основаны на предположении о том, что исходные данные имеют линейную и нелинейную части. Г. Джанг одним из первых предложил использовать комбинацию из ARIMA и нейронной сети. В его модели, на которой основаны или ссылаются многие дальнейшие работы, использующие нейронные сети в коллаборации с ARIMA, предполагается моделирование нелинейной части - остатков ARIMA [34]. В модели применяется FNN с одним скрытым слоем и Q скрытыми нейронами. На вход нейросеть принимает несколько элементов ряда с лагом p . Таким образом, модель задается двумя параметрами - p и Q . В

последствии было предложено несколько путей улучшения предложенной модели, выдающих более точные прогнозы, нежели модель Джанга. В работе [35] была представлена гибридная модель, в которой был использован иной подход к выявлению нелинейной составляющей финансовых данных: дискретная вейвлетная трансформация. В исследовании [36] предложили еще один подход к разделению исходного временного ряда на линейную и нелинейную компоненты: считать нелинейными наиболее волатильные части данных, производя декомпозицию исходного временного ряда при помощи авторегрессионной модели скользящего среднего.

Многие исследования в области применения нейронных сетей в задачах предсказания финансовых инструментов применяют улучшенную модель FNN, поскольку она проста в применении. Автор статьи [37] исследовал возможности улучшить результаты работы простой FNN путем применения Байесовского регуляризатора к входным данным. Результатом явилась модель, устойчивая к переобучению, а также выдающая столь же хорошие результаты как и более сложные модели. В работе [38] была представлена FNN, составленная из двух скрытых слоев, улучшенная за счет замены метода оптимизации на этапе подбора весов с градиентного спуска на алгоритм Левенберга-Маквардта.

Список использованных обозначений

Ниже представлен список обозначений глобальных переменных, использованных в настоящей работе. В некоторых пунктах некоторые переменные могут иметь другое значение, согласно общепринятым обозначениям, что будет оговорено отдельно.

1. T – количество рассматриваемых периодов времени;
2. N – количество рассматриваемых финансовых инструментов в инвестиционном портфеле;
3. $p_{i,t}, t = 1, \dots, T$ – цена i -го элемента портфеля в момент времени t ;
4. $W = \sum_{i=1}^N w_i = 1$ – вектор весов инструментов в портфеле;
5. $R_t = \sum_{i=1}^N r_{i,t}, t = 1, \dots, T$ – взвешенная доходность портфеля в момент времени t ;
6. $r_{i,t} = \ln \frac{p_{i,t}}{p_{i,t-1}}, t = 1, \dots, T$ – логарифмическая доходность i -го элемента портфеля в момент времени t ;
7. x^T – знак транспонирования;
8. MLE – метод максимального правдоподобия
(Maximum Likelihood Estimation);
9. LSE – метод наименьших квадратов (Least Squares Error).

Постановка задачи

Пусть имеется множество ценных бумаг, состоящее из N элементов, поведение которых за T периодов времени описано массивом данных размерности $N \times T$. Необходимо решить задачу портфельной оптимизации: сформировать путем определения весов $W = \hat{W}$ диверсифицированный портфель из имеющегося множества ценных бумаг таким образом, чтобы он имел максимальный уровень ожидаемой доходности $E(R)$ при заданном инвестором уровне риска $V = \hat{V}$:

$$\max_w E(r) = \sum_{i=1}^N \hat{w}_i r_i,$$

$$V = \hat{V},$$

$$\hat{W} = \sum_{i=1}^N w_i = 1, 0 \leq w_i \leq 1.$$

Данные представлены в виде финансового временного ряда, включающего в себя стоимости каждой ценной бумаги на конец каждого из взятых периодов.

Оценка рисков портфеля

Модель оценки рисков

Как было описано в пункте 2 обзора литературы, существует большое количество мер риска и подходов к их вычислению. Несмотря на то, что многие исследователи считают VaR неточной оценкой риска, VaR используется в большом количестве работ по портфельной оптимизации, что объясняется простой и наглядностью вычислений. Поскольку задача настоящей работы состоит в том, чтобы продемонстрировать возможности современных методов анализа к улучшению оценок, вычисленных классическими методами, поэтому без потери точности была выбрана наглядная и популярная оценка VaR.

В качестве меры риска VaR определяет максимальный уровень потерь, которые может понести инвестор в заданном доверительном интервале, т. е. VaR предоставляет не только оценку вероятности потерь, как это делает стандартное отклонение, но и их численное выражение. Аналитически VaR вычисляется по следующей формуле [12]:

$$VaR = P_0 \sigma q_\alpha,$$

где P_0 – стоимость портфеля на момент вычисления; σ – стандартное отклонение доходности портфеля; q_α – квантиль уровня α выбранного распределения доходности.

Традиционно параметры модели, σ и q_α , вычисляются на основе выборочных оценок и предположении о распределении доходностей. Подобный подход не способен привести к точной оценке VaR. В следующих пунктах этой части рассматриваются модели, при помощи которых возможно уточнить компоненты уравнения VaR.

Проверку качества оценки VaR была проведена при помощи теста Купика (см. приложения 1.5), описанного и предложенного к применению, в [14, 15, 16, 17]. Он основан на вычислении метрики, равной отношению количества периодов времени, в которых потери инвестора превышали значения VaR.

Модели волатильности

Рассмотрим три модели волатильности, описанные в [39] и предложенные к использованию для оценки VaR в исследованиях [14, 15, 16, 17]: ARCH, GARCH и EGARCH. Модель T-GARCH, рекомендованная к использованию в [14] имеет схожую с EGARCH структуру и выдает очень близкие результаты, поэтому в настоящей работе рассмотрена не была.

Условная Гетероскедастичность

Преимущество используемых в данной работе моделей волатильности состоит в том, что они, имея авторегрессионную природу, способны выявлять и учитывать влияние прошлых элементов временного ряда на последующие, а также специфические особенности финансового временного ряда: гетероскедастичность и кластеризацию волатильности. Гетероскедастичность – это свойство временного ряда, заключающееся в непостоянстве дисперсии. Кластеризация волатильности, (volatility clustering) проявляется в наличии периодов высокой и низкой волатильности и является частым явлением в финансовых временных рядах. Гетероскедастичность временного ряда не является достаточным условием для наличия кластеризации волатильности, однако наличие условной гетероскедастичности, т. е. зависимости будущих значений дисперсии от предыдущих, является.

Модель ожидаемой доходности

Предполагается, что доходности состоят из двух частей: предсказываемой некоторой *моделью ожидаемой доходности* (mean model (mm)) и непредсказуемой (остатков):

$$r_t = E(r_t|\psi_{t-1}) + \epsilon_t,$$

где ψ_t представляет всю информацию, доступную на момент времени t . Для инициализации модели волатильности необходимо задать условное (зависящее от ψ_t) математическое ожидание $\mu(t) = E(r_t|\psi_{t-1})$ при помощи mm, остатки которой будут описываться моделью волатильности.

Автор [17] приводит в своей работе сравнение трех самых популярных моделей mm: нулевого среднего (zero mean), постоянного среднего (constant mean) и среднего, заданного авторегрессионной моделью (AR). В исследо-

вании делается вывод о том, что модель нулевого среднего демонстрирует несостоятельные результаты, тогда как две другие стабильно хорошо подходят для оценки ожидаемой доходности. Итак, в данной работе были использованы и сравнены обе обозначенные модели:

$$\mu = \phi_0 \text{ - constant mm;}$$

$$\mu = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + \epsilon_t \text{ - AR mm порядка } p.$$

Проверка на гетероскедастичность

Как было упомянуто ранее, в исходном временном ряде может не наблюдаться кластеризации волатильности. Следовательно, следующим предварительным этапом инициализации модели волатильности является проверка временного ряда на указанное свойство. Для проверки этого необходимо рассмотреть квадраты остатков mm на наличие в них автокорреляции (ARCH effect). Проверка была осуществлена при помощи двух статистических тестов: теста Льюинга-Бокса и теста множителей Лагранжа (см. приложения 1.3 и 1.4).

ARCH(p)

Определим модель ARCH(1) (AutoRegressive Conditional Heteroscedastic Model) следующим образом:

$$a_t = r_t - \mu_t = \sigma_{a,t} \epsilon_t,$$

$$\sigma_{a,t}^2 = \alpha_0 + \alpha_1 a_{t-1}^2,$$

$$\alpha_0 > 0, \alpha_1 \geq 0, \alpha_0 + \alpha_1 < 1,$$

$$\epsilon_t \sim \mathcal{N}(0, Var(\epsilon_t)),$$

где взаимосвязь между которыми может также быть представлена путем подобного перехода:

$$\begin{aligned} \sigma_{a,t}^2 &= E(a_t^2) - (E(a_t))^2 = E(a_t^2) = E(\epsilon_t^2) E(\alpha_0 + \alpha_1 a_{t-1}^2) = \\ &= Var(\epsilon_t) (E(\alpha_0) + E(\alpha_1 a_{t-1}^2)) = \alpha_0 + \alpha_1 \sigma_{a,t-1}^2. \end{aligned}$$

Таким образом была представлена дисперсия шумовой компоненты (остатков) модели исходного ряда в виде авторегрессионного процесса. Мы мо-

жем обобщить второе уравнение порядка p для задания ARCH(p):

$$\sigma_{a,t}^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_p a_{t-p}^2,$$

$$\alpha_0 > 0, \forall i \alpha_i \geq 0, \sum_{i=0}^p \alpha_i < 1$$

или

$$\sigma_{a,t} = \alpha_0 + A'_{p,t-1} \alpha,$$

где $A_{p,t-1} = (a_{t-1}, \dots, a_{t-p})$, а α - неотрицательная диагональная матрица параметров α .

Для оценки параметров используем самые распространенные и предложенные к применению в [39] методы: для определения порядка модели p - последовательный перебор моделей с целью определения той, АИС которой является наименьшим (см. приложение 1.1); для оценки коэффициентов уравнения для $\sigma_{a,t}^2$ - LSE.

Для прогноза на один шаг вперед используем следующую формулу:

$$\sigma_{a,t+1} = \alpha_0 + \alpha_1 a_t^2 + \dots + \alpha_p a_{t+1-p}^2.$$

Модели класса ARCH хорошо описывают кластеризацию волатильности, поскольку, исходя из их структуры, чем больше (или меньше) предыдущее значение волатильности, тем больше (меньше) его текущее значение. Однако у них есть и существенные недостатки: они медленно реагируют на "выбросы" если они все же происходят, а также демонстрируют одинаковые результаты как для периодов падения, так и для периодов роста. Помимо всего, они, требуют большого порядка для описания процесса, которому следует волатильность, и, как следствие, определения большого количества параметров.

GARCH(p,q)

Для решения проблемы слишком большого количества параметров модели ARCH перейдем к обобщенной модели ARCH или к модели GARCH (General AutoRegressive Conditional Heteroscedastic Model). Определим модель GARCH(p,q):

$$a_t = \sigma_{a,t} \epsilon_t,$$

$$\sigma_{a,t}^2 = \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{a,t-j}^2,$$

$$\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1.$$

Параметры модели GARCH, прогнозирование и оценка результатов определяется также как и в случае модели ARCH.

EGARCH(p,o,q)

Модель Exponential GARCH (EGARCH) отличается от обычной модели GARCH тем, что учитывает асимметричность дисперсии в зависимости от позитивных и негативных отклонений (предполагается, что негативные отклонения вызывают более высокую волатильность, нежели позитивные). Определим модель EGARCH(p,q):

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^q \alpha_i (|z_{t-i} - E(|z|)|) + \sum_{i=1}^q \gamma_i z_{t-i} \sum_{j=1}^o \beta_j \ln(\sigma_{t-j}^2),$$

где $z_{t-1} = \frac{\epsilon_{t-1}}{\sigma_{t-1}}$.

Знак $(|z_{t-i} - E(|z|)|)$ определяет знак отклонения. Если $\gamma_i < 0$, то волатильность является асимметричной. Параметры модели GARCH, прогнозирование и оценка результатов определяется также как и в случае модели ARCH.

Проверка качества построенных моделей

Для проверки качества модели волатильности необходимо проверить остатки моделей на соответствие стандартному нормальному распределению, на стационарность и на наличие в них зависимостей. В [39] для проверки остатков на автокорреляцию предложено рассмотреть квадраты стандартизированных остатков \hat{a}_t :

$$\hat{a}_t = \frac{a_t}{\sigma_t}.$$

Проверку на соответствие стандартному нормальному распределению осуществим при помощи критерия согласия Колмогорова, описанного и предложенного к использованию в [40]; на стационарность – при помощи Рас-

ширенного теста Дики-Фуллера (см. приложение 1.2) двух статистических тестов: теста Льюинга-Бокса и теста множителей Лагранжа (см. приложения 1.3 и 1.4).

Подбор распределения доходностей

Одной из самых серьезных проблем в анализе финансовых временных рядов является частая и порой резкая смена поведения временного ряда, которая влечет за собой изменение его параметров и, как следствие, доходностей финансового инструмента. Периоды такого одинакового поведения финансовых рынков в научной литературе принято называть режимами рынка (market regimes), а модели, описывающие смену режимов – regime switching models. Подобный подход к прогнозированию финансовых временных рядов позволяет учитывать зависимость элементов ряда от предыдущих и разбивать исходный временной ряд на множества периодов того или иного состояния, что позволяет сделать более точным подбор распределения доходностей путем оценки параметров распределения, с которым распределены доходности инструментов, отдельно для каждого режима. Построение подобной модели требует использования нескольких предположений: что доходности финансовых инструментов независимо и хотя бы частично одинаково распределены. Также необходимо допущение, что одинаково распределены доходности, полученные в периоды одинакового поведения рынка.

Скрытая марковская модель

Задачу моделирования смены режимов рынка принято решать при помощи реализации *скрытых марковских моделей* (СММ, Hidden Markov Models (НММ)). НММ задается набором скрытых состояний, за которые примем режимы рынка, последовательностью наблюдений, за которые примем финансовый временной ряд, а также параметром λ :

1. Пусть имеется Цепь Маркова с дискретным временем и числом состояний N , которая представляет из себя случайный процесс $Q : q_t = s_j : S = \{s_1, \dots, s_N\}$ – набор состояний, в которые система переходит в ходе процесса Q ;

2. Пусть также имеется T различных наблюдений:

$O = \{o_1, \dots, o_T\}$ – последовательность наблюдений;

3. Зададим матрицу переходных вероятностей:

$$A = \{a_{i,j}\}, a_{i,j} = P\{q_{t+1} = s_j | q_t = s_i\};$$

4. Зададим распределение вероятностей наблюдений:

$$b_j(o_t) = P\{o_t | q_t = s_j\};$$

5. Зададим вектор начальных распределений:

$$\pi = (\pi_i), \pi_i = P(q_0 = s_i);$$

$\lambda = (A, B, \pi)$ – параметр НММ.

Подбор распределения

Зачастую при инициализации НММ используется дискретное распределение вероятностей наблюдений. Дискретный случай определенно является самым простым, однако его использование приводит к значительному снижению точности результатов. Это вынуждает нас перейти на этапе обучения модели к непрерывному распределению. Поскольку мы имеем дело с финансовыми данными, предположение о нормальности распределения исходного массива данных также приведет к ухудшению точности. В работах [28, 29] для разрешения данной проблемы предлагается перейти к смеси гауссовских распределений. Смесь гауссовских распределений – это модель, основанная на определении количества распределений M , условных вероятностей принадлежности \forall элементов массива данных к тому или иному распределению и определения, исходя из этого, вектора весов w_j , где m определяет компоненту распределения из смеси:

$$b_j(o_t) = \sum_{m=1}^M w_{j,m} \mathcal{N}(o_t, \mu_{j,m}, \sigma_{j,m}),$$

$$\mathcal{N}(o_t, \mu_{j,m}, \sigma_{j,m}) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(o_t - \mu_{j,m})^2}{2\sigma_{j,m}^2}}$$

$$\sum_{m=1}^M w_{j,m} = 1$$

$$j = 1, \dots, N.$$

Для повышения точности результатов определим для каждого мно-

жества периодов, относящегося к тому или иному режиму, новую смесь гауссовских распределений. От режима к режиму могут меняться не только параметры Гауссовских распределений в смеси, но и их количество. Задачу определения параметров каждой смеси распределений решим при помощи применения алгоритма k -средних.

Уже на этапе инициализации алгоритма нам понадобится количество распределений, участвующих в смеси. Рассмотрим подробнее задачу отыскания этого числа. Введем логарифмическую вероятностную функцию правдоподобия (log-likelihood function), которая бы оценивала вероятность генерации наблюдения o_i Гауссовским распределением m :

$$L = \log P(o_1, o_2, \dots, o_T) = \sum_{i=1}^T \log \sum_{m=1}^M P(o_i|m)P(m),$$

где M – число Гауссовских распределений в смеси.

Самым очевидным подходом к решению проблемы отыскания количества распределений, формирующих смесь, было бы максимизировать значение функции L , тем самым отыскав наиболее подходящее для исходных данных число M . Однако на практике сходимость такой задачи оптимизации достигается в том случае, когда число Гауссовских распределений соответствует количеству наблюдений, что едва ли можно применить в дальнейших вычислениях. Другой подход заключается в использовании информационных критериев, при помощи которых можно оценить два параметра: функцию правдоподобия L и p – количество параметров смеси Гауссовских распределений. Для этих целей был использован АИС (см. приложение 1.1).

Постановка задачи

В теории скрытых марковских процессов данная задача называется задачей декодирования: пусть имеется модель λ и последовательность наблюдений O . Необходимо определить последовательность скрытых состояний модели λ , которая лучше всего “объясняет” последовательность наблюдений O . Обозначенная задача решается при помощи динамического алгоритма декодирования *Витерби*. Он предлагается к использованию и детально описывается в работах [28, 29], поэтому в настоящей работе

описан не будет.

Общий алгоритм

Рассмотрим общий алгоритм решения задачи прогнозирования режима рынка.

1. Инициализация:

- 1.1. возьмем последовательность доходностей O , длины T , и определим количество N и набор S скрытых состояний: $N = 2$, $S = \{\text{падение}, \text{рост}\}$;
- 1.2. преобразуем входные данные: определим периоды роста и падения стоимости актива; произведем сглаживание ряда с окном шириной в n_1 периодов, чтобы нивелировать влияние скачков в цене. Получим временной ряд той же длины T , в котором каждое значение будет заменено на среднее между n_1 соседями; разделим полученный временной ряд на две части: тренировочную (длины T_1), на которой обучим модель, и тестовую (длины T_2), на которой оценим работу модели (далее будем работать с тренировочной выборкой); повторим процедуру сглаживания еще два раза: с окном n_2 , а затем n_3 ;
- 1.3. для полученного временного ряда вычислим булев вектор R , отвечающий за участки падения и роста цен, а затем на основе полученных значений вектора R вычислим количество смен режимов рынка (с роста на падение и с падения на рост), а также общее число элементов, относящихся к каждому из режимов рынка;
- 1.4. на основе значений, полученных на предыдущем шаге, инициализируем вектор начальных распределений π и матрицу переходных вероятностей A ;
- 1.5. на основе значений вектора R разобьем тренировочную выборку на два множества: значения ряда, относящихся к падению, и значения, относящихся к росту;
- 1.6. для получения матрицы B – последнего компонента параметра λ , перейдем к смеси Гауссовских распределений для каждого

из состояний. Этот шаг позволит нам без существенной потери точности воспользоваться предположением о нормальности распределения исходных данных. Для подбора параметров распределений воспользуемся алгоритмом К-средних.

2. Прогнозирование

На данном этапе необходимо определить наиболее подходящую последовательность скрытых состояний, которая бы как можно точнее описывала последовательность наблюдений тестовой выборки. Последним периодом, для которого будет найдено подходящее состояние, будет являться следующий, $T_1 + 1$ период времени. Для решения этой задачи применим алгоритм декодирования Витерби.

3. Оценка результатов

Разобьем тестовые данные на подпоследовательности длины l , тем самым получив множество последовательностей в количестве $T_1 - l - 1$. Для каждой последовательности при помощи алгоритма Витерби получим последовательность состояний и сравним значение последнего элемента с реальным. Для получения оценки работы модели будет служить метрика, равная отношению числа правильно определенных состояний к общему числу предсказанных состояний.

Оценка ожидаемой доходности

Поскольку модели прогнозирования финансовых рынков, основанные на предположении о линейности взаимосвязей между элементами финансовых данных зачастую не способны продемонстрировать высокую точность предсказаний, обратимся к для решения задачи прогнозирования к *нейронным сетям*. В настоящей работе рассмотрены несколько архитектур нейросетей: многослойный персептрон, а также две архитектуры рекуррентных нейросетей. В связи с особенностями ряда доходностей (наличие отрицательных и близких к нулю значений) нейросети были обучены на ценах закрытия элементов портфеля. Предсказания затем были представлены в виде логарифмических доходностей.

Точность работы нейронных сетей была оценена при помощи среднеквадратичной ошибки MSE:

$$MSE = \frac{\sum_{i=1}^{T_1} (r_i - \bar{r}_i)^2}{T_1},$$

где T_1 – размер исследуемой выборки; r_i – истинные доходности, а \bar{r}_i – предсказанные.

Нейросети прямого распространения

Нейросеть прямого распространения (Feedforward Neural Network (FNN)) - широкий класс простых и распространенных архитектур нейронных сетей. FNN предложены к использованию в финансовой отрасли во многих научных работах: [32, 34, 35, 37, 38]. Исследователи зачастую заявляют об определенных недостатках подобной архитектуры нейронной сети, таких как низкая скорость сходимости и склонность к переобучению. Несмотря на указанные изъяны, в настоящей работе реализованы несколько вариантов FNN для сравнения с более продвинутыми архитектурами.

Реализованные нейросети имеют архитектуру многослойного персептрона Румельхальта (см. рис. 2). Они состоят из одного выходного слоя размерности n ; Q скрытых полносвязных слоев, включающих в себя H нейронов; а также одного выходного слоя размерности M . Нейроны представляют из себя персептроны с нелинейными функциями активации. Обу-

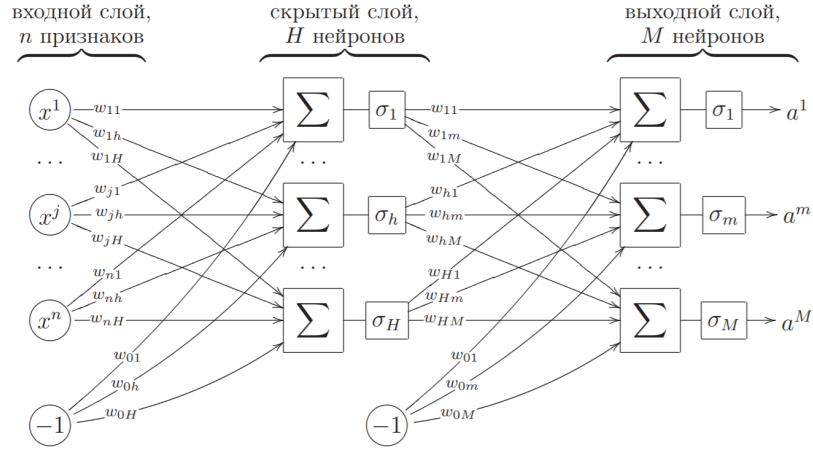


Рис. 2: Многослойный перцептрон [41]

чение происходит путем последовательной подачи на вход сети рядов значений доходности размерности n , которые проходят по указанной на рис. 2 схеме слева направо (прямое распространение). Затем вычисленная ошибка подается на выход сети и последовательно вычисляются изменения для весов сети справа налево. Тем самым реализуется метод обратного распространения ошибки (Back Propagation (BP)). Выходные значения поданного элемента ряда x_i вычисляются следующим образом:

$$a^m = \sigma_m\left(\sum_{h=0}^H w_{hm} u^h(x_i)\right);$$

$$u^h(x_i) = \sigma_h\left(\sum_{j=0}^J w_{ij} v^j(x_i)\right),$$

где w_{hm} - веса синоптических связей между нейронами, а v^j - выходное значение j -го слоя.

В качестве функции активации была использована функция ReLU (Rectified Linear Unit):

$$f(x) = \max(0, x).$$

В [45] авторы предлагают использовать данную функцию в целях увеличения скорости сходимости, а также для повышения стабильности процесса оптимизации, при условии, что исходные данные больше нуля.

Алгоритм построения и обучения, представленный выше, детально описан в [41]. Для обеспечения стабильной и быстрой работы нейросети в

процессе обучения используем Adam – модификацию метода стохастического градиента (Stochastic Gradient Descent (SGD)). Выбранный метод использует сразу два улучшения базового алгоритма: масштабирование градиента и инерции. Описание SGD и Adam, сравнение со всеми популярными методами оптимизации, а также рекомендация к использованию приведены в [43]. Для решения возможной проблемы переобучения используем dropout слои, суть применения которых состоит в том, чтобы выборочно исключать из работы несколько нейронов каждого слоя с некоторой вероятностью, а затем усреднять полученные по разным комбинациям нейронов результаты. Подобный подход описан и предложен к использованию в [42].

Рекуррентные нейронные сети

Также были построены *рекуррентные нейросети* (Recurrent Neural Network(RNN)). Зависимость последующих элементов временного ряда от предыдущих, обуславливает преимущество рекуррентных нейросетей. Однако, как было упомянуто в [33], обычные RNN не подходят для случая финансовых временных рядов, поскольку не способны находить зависимости, проявляющиеся со значительным лагом (тренды, сезонности и т. д.). Следовательно, были рассмотрены две архитектуры указанного класса: управляемые рекуррентные нейроны (Gated Recurrent Units (GRU)) и рекуррентная нейросеть с долгой краткосрочной памятью (Long–Short Term Memory Network (LSTM)). Нейросети, имеющие одну из представленных архитектур способны накапливать и использовать информацию, переданную прошлыми элементами ряда для предсказания последующих, даже при условии того, что между этими элементами были расположены многие другие. LSTM и GRU сети отличаются главным образом структурой нейронов: первая имеет 3 информационных фильтра, а вторая – 2. Количество фильтров способно существенным образом сказаться на скорости и точности работы сети. Представленный ниже алгоритм построения и обучения LSTM подробно описан в [33, 42]. Алгоритм построения и обучения GRU сети детально описан и предложен к сравнению в [44]. Он представляет из себя небольшую модификацию алгоритма для LSTM и не будет описан в данной работе.

LSTM сеть состоит из входного слоя, H внутренних скрытых слоев и выходного слоя. В начале разобьем исходный временной ряд на тренировочную часть длины $T1$ и тестовую длины $T2$. Далее тренировочные данные разделим на $T1 - l - 1$ последовательностей, где l – размер окна.

На вход сети последовательно подаются полученные ряды размерности l . Скрытые слои состоят из N нейронов, общее устройство которых для каждого момента времени t проиллюстрировано схемой (см. рис. 3).

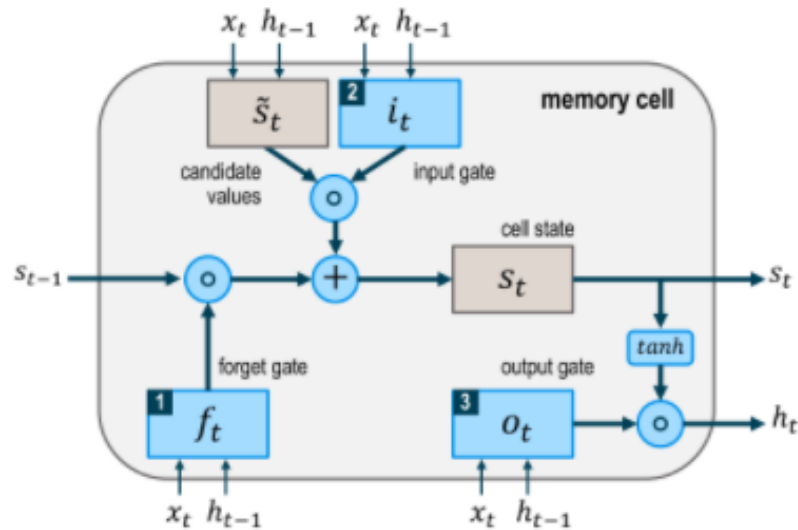


Рис. 3: Схема работы нейрона в момент времени t [33]

Нейрон принимает на вход элемент вектора $s_{t-1}(1 \times N)$ — кумулятивное значение состояния нейрона, измененное последний раз в момент $t - 1$ и несущее некоторую информацию о подаваемых на вход данных; элемент вектора $h_{t-1}(1 \times N)$ – состояние нейрона в момент времени $t - 1$; параметр x_t , который является соответствующим элементом временного ряда. Такая процедура повторяется в каждый момент времени для $i = 1, \dots, N$ -го нейрона каждого из скрытых слоев.

Алгоритм действий нейрона в момент времени t :

1. Параметры $s_{t-1,i}$, x_t и $h_{t-1,i}$ подаются на вход в «фильтр забвения» (forget gate), который возвращает число от 0 до 1 для каждого числа из состояния s_{t-1} , где 1 означает «полностью сохранить», а 0 – «полностью выбросить»;
2. Параметры x_t и $h_{t-1,i}$ подаются на вход в функцию отбора значений, которые будут добавлены в вектор s , и «входной фильтр» (input

gate), который добавит отобранные данные в вектор s , формируя i -й элемент вектора s_t ;

3. Параметры $s_{t,i}$, x_t и $h_{t-1,i}$ подаются на вход в «фильтр вывода» (output gate), который формирует i -й элемент вектора h_t .

В процессе обучения также использована функция активации ReLU [45] и метод оптимизации Adam – модификация алгоритма SGD [43].

Результаты

Продемонстрируем решение сформулированной задачи на реальных рыночных данных. Вопрос выбора финансовых инструментов для включения в инвестиционный портфель не входит в поставленную задачу настоящей работы и, как следствие, не будет рассмотрен. В качестве исследуемых ценных бумаг были выбраны $N = 3$ эмитента разных секторов американского рынка, входящих в число крупнейших по рыночной капитализации в каждом из секторов и котирующиеся более 10 лет: BP, GE, Intel (INTC). Данный срок был выбран с целью обеспечения одинаково большого объема данных для обучения моделей и, вместе с тем, включил в себя кризис 2008 года, вследствие которого произошло беспрецедентное падение стоимости перечисленных инструментов.

За рассматриваемый период времени был выбран один день. Данные о ценах закрытия (в долларах США) были взяты за $T = 2912$ дней из открытого финансового портала Yahoo Finance [48]: с 01.01.2007 по 01.03.2018. Для обучения разных моделей требовалось разное количество исходных данных: промежуток с 01.01.2007 по 01.07.2014 (65%) были использованы для тренировки НММ, с 11.07.2014 по 01.01.2018 (35%) — для их тестирования; промежуток с 01.01.2014 по 09.08.2016 и с 10.08.2016 по 01.01.2018 были выбраны для тренировки и тестирования моделей волатильности и нейронных сетей. Последние 300 дней (с 06.01.2017) тестовой части данных были использованы для бэктестинга портфельной оптимизации при помощи реализации торговой стратегии. Для реализации методов была написана программа на языках Python(3.6.4) с использованием библиотек statsmodels(0.8.0), arch(4.3.1), scikit-learn (0.19.1), keras (2.1.5) и tensorflow (1.4.0) и Matlab(r2018a) с использованием библиотеки Optimization Toolbox (см. приложение 2).

Диверсификация портфеля

Решение поставленной задачи портфельной оптимизации реализовано двумя разными способами. В обоих случаях оценка параметров производилась согласно модели M-V Г. Марковица [3], однако оценки параметров модели проводились разными способами. В первом (первая вариация оп-

тимизации) использовались классические оценки ожидаемой доходности и риска: выборочное среднее и VaR, вычисленный при помощи вариационно-корреляционного метода [12]:

$$E(R)_t = w^T \mu_t, VaR_t = P_t \sqrt{\Sigma_t} q_\alpha,$$

где μ – выборочное среднее доходностей, посчитанное на основе данных, доступных к моменту времени t ; P_t – стоимость портфеля на момент времени t ; $\Sigma = w^T C w$ – взвешенная ковариационная матрица доходностей портфеля, посчитанная на основе данных, доступных к моменту времени t ; q_α – квантиль уровня α распределения Стьюдента с 10 степенями свободы, математическим ожиданием, равным 0, и дисперсией, равной 1.

Во втором случае (вторая вариация оптимизации), оценки выполнены при помощи модификации и объединения вышеописанных методов, предложенных в работах [7, 8]:

$$E(R)_t = w^T \hat{R}_{t+1}, VaR_t = P_t \sqrt{\hat{\Sigma}_t} \hat{q}_\alpha,$$

$$\hat{\Sigma}_t = \sum_{i=1}^N w_i^2 \hat{\sigma}_i^2 + \sum_{i=1}^N \sum_{j=1, j \neq i}^N w_i w_j \gamma_{i,j},$$

$$\gamma_{i,j} = \frac{1}{n-1} \sum_{t=1}^T r_{i,t} r_{j,t},$$

где $\hat{R}_{t+1} = \sum_{i=1}^N w_i \hat{r}_i$ – вектор прогноза доходностей по каждому элементу портфеля на $t+1$ -й момент времени; P_t – стоимость портфеля на момент времени t ; $\hat{\Sigma}_t$ – взвешенная матрица ковариаций, представленная в виде суммы дисперсий элементов портфеля, оцененных при помощи моделей волатильности, и ковариаций, вычисленных на основе элементов вектора доходностей R ; \hat{q}_α – квантиль распределения Стьюдента, параметры которого были оценены при помощи НММ и MLE.

Оценка ожидаемой доходности

Выбор нейронной сети осуществлялся на основе сравнения точности и скорости работы нескольких нейронных сетей: FNN сети с двумя, тремя

и шестью полносвязными скрытыми слоями; а также сети GRU и LSTM. На вход были поданы цен на финансовые инструменты длины $l = 100$. Начальные параметры нейросетей (длина последовательностей для входа, количество слоев и нейронов в каждом слое) были взяты согласно рекомендациям, в частности, авторов статей [33, 41], далее они последовательно увеличивались с целью достижения наибольшей точности. Ниже (см. рис. 4) представлена таблица с параметрами лучших по точности нейросетей и оценками их работы на примере акций компании BP.

BP

| Архитектура | Количество слоев | Количество нейронов | Количество эпох | MSE | Время обучения |
|-------------|------------------|---------------------|-----------------|------|----------------|
| Перцептрон | 2 | 200 | 200 | 2.35 | 0.05 |
| Перцептрон | 4 | 200 | 100 | 1.81 | 0.05 |
| Перцептрон | 6 | 100 | 3000 | 0.47 | 0.5 |
| Перцептрон | 6 | 500 | 6000 | 1.12 | 0.75 |
| GRU | 3 | 100 | 200 | 0.25 | 0.6 |
| LSTM | 3 | 100 | 200 | 0.18 | 1.15 |

Рис. 4: Характеристики архитектуры и результатов работы нейросетей на примере акций компании BP

По результатам подбора нейросетей были сделаны несколько выводов:

1. рекуррентные нейросети демонстрируют более точные предсказания, нежели обыкновенные перцептроны;
2. значения MSE на тренировочной и тестовой выборках не различались значительно, проблема переобучения не была обнаружена;
3. за увеличением числа нейронов и скрытых слоев не обязательно следует увеличение точности предсказаний;
4. нейронные сети не достаточно быстро реагируют на резкие изменения цены;

5. LSTM сети дают более точные предсказания, нежели GRU сети, однако они обладают рядом недостатков по сравнению с последними, такими как сложная настройка, длительное время обучения, не универсальность.

Поскольку проблема переобучения не возникла, dropout слои [42], призванные бороться с проблемами переобучения, значительно ухудшали предсказательные возможности нейронных сетей, и не приведены в сравнительной таблице. В следствие перечисленных характеристик работы нейросетей GRU сети с двумя скрытыми слоями, содержащими 100 нейронов каждый, были применены для прогнозирования доходностей элементов портфеля. Графики прогноза, а также 95%—й доверительный интервал для них представлены ниже (см. рис. 5). На тестовых частях данных по ценам акций компаний BP, GE и INTC были получены следующие значения MSE: 0.279, 0.255 и 0.375 соответственно.



Рис. 5: Графики прогнозов нейронной сети

Прогнозирование рискованности

Прогноз волатильности

Выбор модели волатильности осуществлялся отдельно для каждого рассматриваемого финансового инструмента, основываясь на величине значения AIC (см. приложение 1.1). Помимо нескольких возможных значений параметров (p, o, q) для каждой модели были также рассмотрены два распределения ошибок [14, 15, 17]: нормальное и Стьюдента, параметры которых были оценены при помощи MLE; а также две модели ожидаемой доходности [17]: постоянного среднего и AR порядка \hat{p} . В результате подбора моделей было выявлено, что для всех рассматриваемых инструментов

наилучшим образом подходит модель EGARCH с моделью ожидаемой доходности постоянного среднего. Для двух моделей было выбрано распределение Стьюдента (для акций компании BP и INTLC), а для одной (для акций компании GE) – нормальное. Ниже (см. рис. 6) представлена таблица с параметрами лучших моделей для каждого из элементов портфеля.

| Модель | BP | GE | INTLC |
|--------|-----------|-----------|-----------|
| EGARCH | (3, 1, 2) | (2, 0, 1) | (4, 0, 1) |

Рис. 6: Лучшие модели волатильности для соответствующих инструментов

При помощи выбранных моделей волатильности были произведены прогнозы на следующий период времени для всей тестовой выборки. Остатки моделей были проверены на соответствие стандартному нормальному распределению при помощи критерия Колмогорова [40] и на стационарность при помощи расширенного теста Дики-Фуллера (см. приложение 1.2), а также проверены на наличие автокорреляции при помощи тестов Льюинга-Бокса и множителей Лагранжа (см. приложения 1.3 и 1.4). В результате проверки распределение остатков по нормальному закону было подтверждено, они были признаны стационарными, а наличие автокорреляций не выявлено, что может свидетельствовать о хорошем качестве работы подобранных моделей.

Подбор распределения

В связи с тем, что распределение доходностей инвестиционного портфеля склонно иметь более тяжелые хвосты, нежели нормальное, ряд исследователей, в частности, [12, 13], рекомендует использовать распределение Стьюдента в качестве оценки распределения доходностей при оценке VaR. При помощи НММ исходный финансовый временной ряд был разбит на множество периодов роста и падения доходностей портфеля, для каждого из которых посредством MLE были подобраны параметры своего распределения Стьюдента.

Ниже (см. рис. 7) представлены графики стоимости элементов портфеля за тестовый период времени, окрашенные в цвета, согласно предсказаниям НММ. Зеленый цвет отображает рост, а синий – падение. Модель

была протестирована на нескольких комбинациях размеров окон, однако наилучшим образом подошли следующие значения: $n_1 = 7, n_2 = 25, n_3 = 15, l = 20$. Видно, что модель довольно хорошо предсказывает моменты резкой смены тренда, в отличие от нейросетей, однако плохо реагирует на отдельные движения против тренда. Такой характер работы модели обуславливает значение описанной ранее метрики: 65%, 62% и 66% успешных предсказаний соответственно.

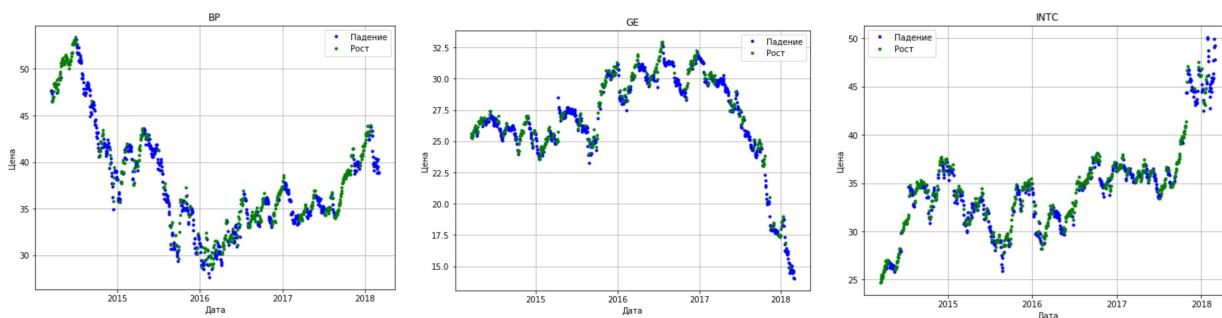


Рис. 7: Графики прогнозов НММ

Оценка VaR

Оценка VaR производилась для каждого из 300 периодов тестовой выборки тремя разными способами: классическим (с использованием выборочных оценок и предположением о распределении доходностей по Стьюденту с 10 степенями свободы, математическим ожиданием, равным 0, и дисперсией, равной 1); с использованием прогнозов модели волатильности и распределением Стьюдента, параметры которого были оценены при помощи MLE; с использованием прогнозов волатильности и с распределением Стьюдента, параметры которого были оценены при помощи MLE отдельно для периодов роста и падения, согласно прогнозам НММ. По результатам подобной оценки была вычислена статистика теста Купика (см. приложение 1.5), результаты которой представлены ниже (см. рис. 8).

| | 1 | 2 | 3 | Доверительный интервал |
|---------------------|--------|-------|-------|------------------------|
| Значение статистики | 35.119 | 4.360 | 2.754 | (0 ; 3.841) |

Рис. 8: Результаты теста Купика

По значениям, указанным в таблице, можно сделать вывод о том,

что классическая оценка VaR дает оценки низкого качества, что приводит к слишком большому значению статистики; использование моделей волатильности заметно приближает значение статистики к критическому; а применение раздельной оценки параметров распределения на основе прогнозов НММ позволяет значению статистики попасть в 95%-й доверительный интервал. Следовательно, примененные модели способны существенно улучшить оценку VaR для имеющихся исходных данных.

Бэктестинг

Сравним работу двух вышеописанных методов диверсификации путем реализации простой торговой стратегии на последних $T = 300$ днях тестовой части данных. Стратегия заключается в том, чтобы в конце каждого периода времени t определять оптимальный портфель одним из вышеописанных вариаций оптимизации, т. е. такой, который при заданном максимальном желаемом для инвестора уровне риска согласно прогнозу должен обеспечить максимальную доходность в следующий, $t + 1$ -й период времени; переформировывать согласно новым весам портфель в начале следующего периода и держать его до окончания $t + 1$ -го периода.



Рис. 9: Результаты реализации торговой стратегии

Реализация стратегии ограничена запретом на короткие позиции и отсутствием комиссий за сделки. За желаемый максимальный уровень по-

терь (VaR) была взята сумма, равная 10% от текущей стоимости портфеля. Оптимизация в обоих случаях осуществлялась при помощи квадратичной оптимизации, предложенной к использованию, в частности, в [8].

Выше (см. рис. 9) представлены графики стоимости портфеля в случае реализации первой вариации оптимизации (обычный портфель) и в случае реализации второй вариации оптимизации (улучшенный портфель). Можно видеть, что к последнему периоду времени разрыв между улучшенным и обычным портфелями равен 10.34 пунктам, что составляет 31% от стоимости обычного портфеля в тот момент времени и может говорить о значительном улучшении процесса формирования портфеля за счет вышеописанных методов.

Заключение

В настоящей работе была рассмотрена задача портфельной оптимизации, описаны и реализованы методы, позволяющие улучшить классические подходы к ее решению: авторегрессионные модели, скрытые марковские модели и нейронные сети. Были использованы несколько типов каждого из указанных методов, выбраны лучшие из них, и произведена оценка качества их работы согласно введенным метрикам. Результаты работы методов были использованы в реализации торговой стратегии, и произведено сравнение с ее реализацией без применения перечисленных методов. Таким образом, в данной работе продемонстрирована возможность увеличения точности работы классической модели портфельной оптимизации при помощи современных методов анализа данных.

Приложения

1. Статистические тесты

1.1 Информационный критерий Акаике

Существует несколько часто применимых информационных критериев, одним из которых является *информационный критерий Акаике* (Akaike Information Criterion (AIC)), описанный и предложенный к использованию в [39]:

$$AIC = \frac{-2}{T} \ln(P) + \frac{2}{T}L,$$

где T – размер исследуемой выборки; L – количество параметров модели; P – вероятность, оцениваемая при помощи MLE.

Метод использования AIC, также как и других подобных критериев, основан на принципе Бритвы Оккама, согласно которому, из нескольких непротиворечащих друг другу, одинаково хороших объяснений некоторого явления стоит выбрать самое простое из них. В данном случае P отвечает за качество построенной модели, а L – за ее простоту. По результатам вычисления значений критерия для каждой модели из множества допустимых моделей выбор делается в пользу той, чье значение AIC является наименьшим.

В настоящей работе используется именно этот критерий, поскольку, несмотря на отсутствие подтверждений в научной литературе, эмпирически было получено, что он дает наиболее адекватные оценки параметров моделей, построенных в данной работе, в наибольшем числе случаев.

1.2 Расширенный тест Дики-Фуллера

Проверка временного ряда r_t на стационарность предполагает решение проблемы существования единичного корня (unit-root problem). Она заключается в том, что временной ряд является стационарным тогда и только тогда, когда принадлежат единичному кругу корни характеристического уравнения

$$1 - \phi_1x - \phi_2x^2 - \dots - \phi_px^p = 0$$

для разностного уравнения автокорреляционной функции (ACF)

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_l = 0,$$

где $\phi_i, i = 1, \dots, p$ – коэффициенты авторегрессионной модели порядка p (AR(p)); $\rho_l, l > 0$ – автокорреляционные коэффициенты лага l ; $B\rho_l = \rho_{l-1}$ – оператор обратного сдвига (back-shift operator).

В настоящей работе используется один из самых распространенных тестов на наличие единичного корня (unit root tests) – *расширенный тест Дики-Фуллера* (Augmented Dickey-Fuller (ADF) test), который описывается и предлагается к использованию в [39, 46, 47]. Данный тест является модифицированной версией теста Дики-Фуллера (Dickey-Fuller (DF) test), главным недостатком которого является его предположение о том, что исходный временной ряд представлен моделями класса AR(p)

$$r_t = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + \epsilon_t,$$

и, что случайная компонента ϵ_t является белым шумом. В отличие от DF теста, ADF учитывает возможные зависимости последовательности случайных компонент временного ряда, предполагая, что он может быть представлен моделью класса ARMA и описан уравнением вида:

$$r_t = c_t + \beta_t x_{t-1} + \sum_{i=1}^{p-1} \phi_i \nabla x_{t-i} + \epsilon_t,$$

где c_t – некоторая константа, а β_t, ϕ_i – коэффициенты, оцененные при помощи LSE. Статистикой ADF является величина, равная:

$$ADF_\tau = \frac{\beta - 1}{std(\beta)}.$$

Для инициализации подобной модели необходимо подобрать порядок, т. е. лаг, p . Будем делать это при помощи информационного критерия AIC (см. приложение 1.1).

Гипотезы теста: $H_0 : \beta = 1$; $H_1 : \beta < 1$. Нулевая гипотеза отвергается при условии, что $ADF_\tau < critical\ value$. Соответствующее критическое значение (critical value) может быть найдено в таблице значений критерия

DF.

Если критерий ADF показал, что временной ряд не является стационарным, необходимо продифференцировать его с лагом 1, тем самым перейдя от модели класса ARMA к ARIMA, и проверить на стационарность новый ряд.

1.3 Тест Льюинга-Бокса

Тест Льюинга-Бокса (Ljung-Box (LjB) test) позволяет исследовать остатки временного ряда на наличие автокорреляции сразу для m первых значений. Он описан и предложен к использованию в [39, 47]. Тест основан на Q -статистике Бокса-Пирса:

$$Q(m) = T \sum_{l=1}^m \bar{\rho}_l^2,$$

где T – размер исследуемой выборки; m – число лагов, для которых исследуется зависимость; $\bar{\rho}_l$ – выборочная автокорреляция k -го порядка. Однако, статистика LjB

$$Q_L(m) = T(T+2) \sum_{l=1}^m \frac{\bar{\rho}^2}{T-l}$$

гораздо лучше соответствует асимптотически распределению χ^2 на конечных выборках.

Гипотезы теста: $H_0 : \rho_i = 0, i = 1, \dots, m$ (остатки являются белым шумом); $H_1 : \exists i \in (1, \dots, m) : \rho_i \neq 0$ (остатки коррелированы между собой). Нулевая гипотеза отвергается, если $Q_L > \chi_{(\alpha, m)}^2$, где $\chi_{(\alpha, m)}^2$ – квантиль распределения χ^2 с m степенями свободы и $100(1 - \alpha)\%$ -м доверительным интервалом. В данной работе используются $\alpha = 0.05$, $m = \ln(T)$ [39].

1.4 Тест множителей Лагранжа

Тест множителей Лагранжа (Lagrange Multiplier (LM) test) также как и тест Льюинга-Бокса (см. приложение 1.3) производит множественную проверку m первых элементов ряда на наличие у остатков a_t временного ряда автокорреляции. Однако, в отличие от него, LM тест был сконструирован для проверки остатков на ARCH-эффект, и, как следствие, он исследует не сами остатки временного ряда, а их квадраты, представленные

в виде уравнения линейной регрессии:

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_m a_{t-m}^2 + \epsilon_t, \quad t = m + 1, \dots, T,$$

где m – число лагов, для которых исследуется зависимость.

Тест основан на F –статистике [39]:

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - 2m - 1)},$$

где $SSR_0 = \sum_{t=m+1}^T (a_t^2 - \bar{\omega})^2$, $\bar{\omega}$ – выборочное среднее a_t^2 ; $SSR_1 = \sum_{t=m+1}^T \bar{\epsilon}_t^2$, $\bar{\epsilon}_t$ – ошибки оценки регрессионной модели MSE.

F –статистика асимптотически распределена с распределением $\chi_{(\alpha, m)}^2$. Гипотезы теста: $H_0 : \alpha_i = 0, i = 1, \dots, m$ (остатки являются белым шумом); $H_1 : \exists i \in (1, \dots, m) : \alpha_i \neq 0$ (остатки имеют зависимости между собой). Нулевая гипотеза отвергается, если $F > \chi_{(\alpha, m)}^2$, где $\chi_{(\alpha, m)}^2$ – квантиль распределения χ^2 с m степенями свободы и $100(1 - \alpha)\%$ -м доверительным интервалом. В данной работе используются $\alpha = 0.05, m = \ln(T)$ [39].

1.5 Тест Купика

Тест Купика (Kupiec Test (КТ)), как описано, в частности, в [15] основан на определении пропорции ошибок VaR среди общего числа прогнозов (периодов). Его основной задачей является проверка того, насколько отклоняется процент правильных прогнозов от выбранного уровня доверия $1 - \alpha$. Для вычисления статистики теста вначале необходимо задать индикатор I :

$$I_t = \begin{cases} 1, & \text{loss}_t \leq -VaR_t, \\ 0, & \text{loss}_t > -VaR_t, \end{cases}$$

где loss_t – величина убытков за период времени t .

Тогда значение статистики вычисляется по формуле:

$$K = 2 \ln\left(\left(1 - \frac{F}{N}\right)^{N-F} \left(\frac{F}{N}\right)^F\right) - 2 \ln\left((1 - \alpha)^{N-F} (\alpha)^F\right),$$

где N – число рассматриваемых периодов, а $F = \sum_{i=1}^N I_t$.

Статистика K распределена с распределением $\chi_{\alpha, 1}^2$. Гипотезы теста: $H_0 : \frac{F}{N} = \alpha$ (оценка VaR соответствует выбранному уровню доверия);

$H1 : \frac{F}{N} \neq \alpha$ (оценка VaR не соответствует выбранному уровню доверия). Нулевая гипотеза отвергается, если $K > \chi^2_{(\alpha,1)}$, где $\chi^2_{(\alpha,1)}$ – квантиль распределения χ^2 с одной степенью свободы и $100(1 - \alpha)\%$ -м доверительным интервалом. В данной работе используются $\alpha = 0.05$.

2. Программная реализация

1. Полный код программы на языке Python доступен по ссылке:

URL: <https://pastebin.com/DHtrkL78> (дата обращения: 01.05.18).

2. Полный код программы на Matlab доступен по ссылке:

URL: <https://pastebin.com/JpHnitEW> (дата обращения: 01.05.18).

Литература

- [1] Markowitz H. M. "The early history of portfolio theory: 1600–1960". Financial Analysts Journal, 1999. Vol. 55. No 4. P. 5 - 16.
- [2] Merton R. C. An analytical derivation of Efficient Portfolio Frontier // Journal of Financial and Quantitative Analysis, 1972. P. 1851 - 1872.
- [3] Шарп У. Ф., Александер Г. Д., Бэйли Д. В. Инвестиции. Изд. 8. М: ИНФРА-М, 2015. 1028 с.
- [4] Lee M. C., Su L. E. Capital Market Line Based on Efficient Frontier of Portfolio with Borrowing and Lending Rate // Universal Journal of Accounting and Finance, 2014. Vol. 2, No 4, P. 69 - 76.
- [5] Брошюра банка Goldman Sachs: The Intuition Behind Black-Litterman Model Portfolios. URL: <http://www.cis.upenn.edu/~mkearns/finread/intuition.pdf> (дата обращения: 31.05.18).
- [6] Harris R. D. F., Stoja E., Tan L. The dynamic Black–Litterman approach to asset allocation // European Journal of Operational Research, 2017. Vol. 259, No 3. 16, P. 1085 - 1096.
- [7] Ranković V., Drenovak M, Urosevic B., Jelic R. Mean-univariate GARCH VaR portfolio optimization: Actual portfolio approach // Computers Operations Research, 2016. Vol. 72, P. 83 - 92.
- [8] Freitas F. D., De Souza A. F., De Almeida A. R. Prediction-based portfolio optimization model using neural networks // Neurocomputing, 2009. Vol. 72, P. 2155 – 2170.
- [9] Брошюра инвестиционной компании Sunrise Capital. Tom Rollinger, Scott Hoffman. Sortino ratio: a better measure of risk. URL:

https://www.sunrisecapital.com/wp-content/uploads/2014/06/Futures_Mag_Sortino_0213.pdf (дата обращения: 31.05.18).

- [10] Технический документ банка J. P. Morgan. RiskMetrics — Technical Document. Fourth Edition. URL: <https://www.msci.com/documents/10199/5915b101-4206-4ba0-aea2-3449d5c7e95a> (дата обращения: 31.05.18).
- [11] Linsmeier T. J., Pearson N. D. Risk management: An Introduction to Value at Risk. URL: <https://ageconsearch.umn.edu/bitstream/14796/1/aceo9604.pdf> (дата обращения: 31.05.18).
- [12] Kondapaneni R. A study of the Delta Normal Method of Measuring VaR. URL: https://web.wpi.edu/Pubs/ETD/Available/etd-050905-104553/unrestricted/Rajesh_Thesis.pdf (дата обращения: 31.05.18).
- [13] Gueant O. Computing the Value at Risk of a Portfolio: Academic literature and Practitioners' response. URL: http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/738bb4496121216ac1257a7c006702ba/\protect\T2A\textdollarFILE/computing_the_value_at_risk_of_a_portfolio.pdf (дата обращения: 31.05.18).
- [14] Smolović J. C., Lipovina-Božović M., Vujošević S. GARCH models in value at risk estimation: empirical evidence from the Montenegrin stock exchange // Journal Economic Research-Ekonomska Istraživanja, 2017. Vol. 30, No 1, P. 477 - 498.
- [15] Orhan M., Köksal B. A comparison of GARCH models for VaR estimation // Expert Systems with Applications 2012. Vol. 39, No 3, P. 3582 - 3592.
- [16] Johansson A., Sowa V. A comparison of GARCH models for VaR estimation in three different markets. URL: <http://www.diva-portal.org/smash/get/diva2:630568/FULLTEXT01.pdf> (дата обращения: 31.05.18).
- [17] Wennström A. Volatility Forecasting Performance: Evaluation of GARCH type volatility models on Nordic equity indices. URL: <https://www.math.>

kth.se/matstat/seminarier/reports/M-exjobb14/140616b.pdf (дата обращения: 31.05.18).

- [18] Billio M., Casarin R., Osuntuyi A. Efficient Gibbs Sampling for Markov Switching GARCH Models. URL: <https://arxiv.org/pdf/1212.5397v1.pdf> (дата обращения: 31.05.18).
- [19] Hamilton J. D., Susmel R. Autoregressive conditional heteroskedasticity and changes in regime // Journal of Econometrics, 1994. Vol. 64, P. 307 - 333.
- [20] Bauwens L., Dufays A., Rombouts J. V. K. Marginal likelihood for Markov-switching and change-point GARCH models // Journal of Econometrics, 2014. Vol. 178, P. 508 - 522.
- [21] Kisiala J. Conditional VaR: theory and applications. URL: http://www.maths.ed.ac.uk/~prichtar/docs/Kisiala_Dissertation.pdf (дата обращения: 31.05.18).
- [22] Васильев В. А., Летчиков А. В., Лялин В. Е. Математические модели оценки и управления финансовыми рисками хозяйствующих субъектов // Финансы и аудит, 2006. Вып. 4, С. 200 - 237.
- [23] Szegö G. Measures of risk // Journal of Banking Finance, 2002. Vol. 26, P. 1253 - 1272.
- [24] Righi M. B., Borenstein D. A simulation comparison of risk measures for portfolio optimization // Finance Research Letters, 2018. Vol. 24, P. 105 - 112.
- [25] Mondal L., Shit L., Goswami S. Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices // International Journal of Computer Science, Engineering and Applications (IJCSEA), 2014. Vol. 4, No 2, P. 13 - 29.
- [26] Wang B., Huang H., Wang X. A novel text mining approach to financial time series forecasting // Neurocomputing, 2012. Vol. 83, P. 136 - 145.

- [27] Weigend A. S., Shi S. Predicting Daily Probability Distributions of SP500 Returns // Density Forecasting in Economics and Finance, 2000. Vol. 19, No 4, P. 375 - 392.
- [28] Zhang Y. Prediction of financial time series with hidden Markov models. URL: <http://summit.sfu.ca/item/7728> (дата обращения: 31.05.18).
- [29] Idvall P., Jonsson N. Algorithmic trading: hidden Markov models on foreign exchange data. URL: <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-10719> (дата обращения: 31.05.18).
- [30] Личный сайт инвестиционного аналитика Холлс-Мура М. URL: <https://www.quantstart.com> (дата обращения: 31.05.18).
- [31] Личный сайт инвестиционного аналитика Кристофера Б. URL: <http://www.blackarbs.com> (дата обращения: 31.05.18).
- [32] Tarsauliya A., Kant S., Kala R., Tiwari R., Shukla A. Analysis of Artificial Neural Network for Financial Time Series Forecasting // International Journal of Computer Applications, 2010. Vol. 9, No 5, P. 16 - 22.
- [33] Fischer T., Krauss C. Deep learning with long short-term memory networks for financial market predictions // 2018. European Journal of Operational Research. P. 1 - 16.
- [34] Jang G. P. Time series forecasting using a hybrid ARIMA and neural network model // 2003. Neurocomputing. Vol. 50, P. 159 – 175.
- [35] Khandelwal I., Adhikari R., Verma G. Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition // 2015. Procedia Computer Science. Vol. 48, P. 173 - 179.
- [36] Narendra Babu C. N., Reddy B. E. A moving-average-filter-based hybrid ARIMA–ANN model for forecasting time series data Applied // 2014. Soft Computing. Vol. 23, P. 27 - 38.
- [37] Ticknor J. L. A Bayesian regularized artificial neural network for stock market forecasting // 2013. Expert Systems with Applications. Vol. 40, P. 5501 - 5506.

- [38] Mammadia S. Financial time series prediction using artificial neural network based on Levenberg-Marquardt algorithm // 2017. Procedia Computer Science. Vol. 120, P. 602 - 607.
- [39] Tsay R. S. Analysis of financial time series. 3-rd ed. John Wiley Sons Inc., 2010, 714 p.
- [40] Буре В. М., Парилина Е. М., Седаков А. А. Методы прикладной статистики в R и Excel. Изд. 1. Лань, 2016. 152 с.
- [41] Воронцов К. В. Математические методы обучения по прецедентам (теория обучения машин). URL: <http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf> (дата обращения: 31.05.18).
- [42] Блог инвестиционной компании WunderFund. URL: https://habr.com/users/wunder_editor/posts/ (дата обращения: 31.05.18).
- [43] Ruder S. An overview of gradient descent optimization algorithms. URL: <https://arxiv.org/pdf/1609.04747.pdf> (дата обращения: 31.05.18).
- [44] Chung J., Gulcehre C., Cho K., Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. URL: <https://arxiv.org/pdf/1412.3555v1.pdf> (дата обращения: 31.05.18).
- [45] Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks. URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf> (дата обращения: 31.05.18).
- [46] Greene W. H. Econometric Analysis. 7-th International ed. Pearson Education Ltd., 2012. 1241 p.
- [47] Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс. Изд. 6-е. М.: Дело, 2004. 576 с.
- [48] Yahoo Finance. URL: <https://finance.yahoo.com/> (дата обращения: 31.05.18).