

Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра Математической экономики

**СИНК-АППРОКСИМАЦИЯ КОТИРОВОК ЦЕННЫХ БУМАГ
С NOSQL ХРАНИЛИЩЕМ ДАННЫХ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

Студента 2 курса 247 группы
направления 09.04.03 – «Прикладная информатика»

Механико-математического факультета

Соболева Василия Михайловича

Научный руководитель
профессор, д.ф.-м.н., доцент А.Ю. Трынин

Зав. кафедрой
д.ф.-м.н., профессор С.И. Дудов

Саратов 2018

Введение. Актуальность исследования. Актуальность исследования обусловлена тем, что в настоящее время рынок ценных бумаг в Российской Федерации набирает обороты, и изучение методов моделирования их поведения становится отнюдь не последним по значимости. Используя методы моделирования поведения котировок любой интересующийся человек сможет принять решение как, в конечном счёте, о совершении манипуляций с ценными бумагами, так и в целом о желании войти на данный рынок. Зная и понимая его состояние и возможности, держатель ценных бумаг имеет возможность планировать свои расходы и доходы, лучше ориентироваться в сфере для прогнозирования своего финансового будущего и принятия важнейших рыночных решений. Моделирование поведения котировок финансовых инструментов способно помочь участникам рынка в определении его перспектив, динамики, а также более прибыльного направления деятельности. Помимо всего прочего, создание модели способно предупредить о ряде возможных кризисных явлений на различных видах рынков. Внедрение моделирования способно дать колоссальный эффект для экономики и финансовую выгоду для держателей ценных бумаг.

Целью исследования является разработка программного обеспечения для синк-аппроксимации котировок ценных бумаг.

В соответствии с данной целью поставлены и решены следующие **задачи**:

1. Сделать описание базовых методов аппроксимации, а также изучить аппроксимацию синками;
2. Дать основные понятия о ценных бумагах, котировках и проанализировать брокеров, выбрав самого удобного;
3. Предоставить описание NoSQL-хранилищ и провести их сравнительный анализ, остановившись на наиболее подходящем;
4. Создать программный продукт, используя данные, полученные на предыдущих этапах.

Объектом исследования является программный продукт, который даст возможность аппроксимировать неизвестные значения котировок на основе некоторых базовых значений.

Предметом исследования является синк-аппроксимация и её применение для анализа данных о котировках.

В первой части «Теория аппроксимации и интерполяции функций» описываются основные понятия, подходы и проблемы, присущие задаче интерполяции функций. От более простых и распространённых методов происходит постепенный переход к синк-аппроксимации.

Интерполяция функций занимает одну из ведущих ролей в вычислительной математике, имея при этом и практическое, и теоретическое значение. Данное понятие можно интерпретировать как построение по некой заданной функции какой-либо другой, обычно более простого вида, имеющая с исходной одинаковые значения в определённых точках. Нередкой практической задачей является восстановление по табличным значениям непрерывной функции. Для вычисления широкого спектра функций нередко применяются приближения дробно-рациональными функциями или полиномами. Сама по себе теория интерполирования имеет большую сферу применения, в которую входит, например, отыскание способов решения интегральных и дифференциальных уравнений, исследование и построение квадратурных формул численного интегрирования.

Пусть функция $f(x)$ задана на отрезке $[a, b]$. Задача интерполяции состоит в построении функции $g(x)$, совпадающей с заданной $f(x)$ в некотором множестве точек (узлы интерполяции) $\{x_1, x_2, \dots, x_{n+1}\}$ из отрезка $[a, b]$, т.е. должны выполняться условия: $g(x_k) = y_k$, где $k = \overline{1, n+1}$.

В случае, когда $f(x)$ представляет из себя табличную функцию, которая была получена, например, в результате эксперимента, и известны только её значения y_k в точках x_k , то, в целом, о качестве итогового приближения судить довольно непросто. Но если значения $f(x)$ могут быть вычислены в любой точке отрезка $[a, b]$, то возможно исследование качества получающегося

приближения, к примеру, с помощью нахождения отклонения функции $g(x)$ от $f(x)$. Также на качество интерполяции в немалой степени влияет гладкость функции $f(x)$, расположение и число узлов, а также некоторые другие факторы.

Первые упоминания о применении синк-аппроксимаций можно найти ещё в работах Джованни Плейна в 1820 году. Спустя полвека, в рамках рассмотрения данной темы Эмилем Борелем (1897 год) и Эдмундом Уиттекером (1914-1915 годы) на свет появились понятия обычной и усечённой кардинальной функции. Их сужения на отрезок $[0, \pi]$ выглядят следующим образом:

$$\begin{aligned} L_n(f, x) &= \sum_{k=0}^n \frac{\sin(nx - k\pi)}{nx - k\pi} f\left(\frac{k\pi}{n}\right) = \\ &= \sum_{k=0}^n \frac{(-1)^k \sin(nx)}{nx - k\pi} f\left(\frac{k\pi}{n}\right) = \sum_{k=0}^n l_{k,n}(x) f\left(\frac{k\pi}{n}\right) \end{aligned}$$

На текущий момент использование приближения синками функций, которые убывают по экспоненте на бесконечности, но при этом являются аналитическими на действительной оси, немало освещено в научном сообществе. Помимо всего прочего, уже существуют исследования, базирующиеся на теореме Котельникова, которые позволили получить различные варианты представления целых функций рядами по синкам с узлами интерполирования, а также появились работы, изучающие взаимосвязь теоремы дискретизации и интерполяции Лагранжа по узлам для спектра задачи Штурма–Лиувилля.

До недавнего времени аппроксимация операторами на ограниченном интервале или отрезке могла быть применена с использованием отображений лишь для некоторых классов аналитических функций. В 2007 году А.Ю. Трыниным было проведено исследование, результатом которого стала возможность определения наличия-отсутствия аппроксимативной сходимости в точке конкретно значений операторов. При этом для самой приближаемой функции f нет необходимости ни в чём, кроме непрерывности на $[0, \pi]$, более

того, информация о ней может быть ограничена лишь её значениями в узлах $\frac{k\pi}{n}$, принадлежащих окрестности точки, для которой и исследуются аппроксимативные свойства. Помимо всего прочего, был выведен критерий равномерной сходимости операторов для непрерывных функций внутри интервала $(0, \pi)$, который аналогичен критерию Привалова сходимости интерполяционных многочленов Лагранжа–Чебышёва и критерию равномерной сходимости интерполяционных процессов Лагранжа–Штурма–Лиувилля внутри интервала $(0, \pi)$.

Во второй части «Ценные бумаги» даётся развёрнутое понятие самой ценной бумаги, приводятся её виды, классификация, а также рассказывается о котировках и брокерах. Проводится анализ представленных на российской территории брокеров с целью выбора подходящего для решения задач исследования .

Ценная бумага (ЦБ) представляет из себя особый товар - обязательство её эмитента (юрлица, выпустившего от своего имени ЦБ) выплачивать доход на вложенный капитал её держателю и/или вернуть его через заданный срок. ЦБ возможно также определить и как документ, который удостоверяет с соблюдением установленной формы и/или обязательных реквизитов права на имущество, передача или осуществление которых возможны лишь при его предъявлении. ЦБ выступает как титул собственности в системе экономических отношений. Это означает, что она, будучи бумажным дубликатом капитала, может представлять производительный капитал (имущество предприятия, которое приходится на одну акцию) или быть отдельно стоящей сферой вложения ссудного капитала (эмиссия облигаций под конкретный инвестиционный проект). Различие имущественных прав, которые связаны с ЦБ, а также существование особых договорных отношений и разных систем расчетов объясняют многообразие их видов/подвидов.

В качестве основных свойств ЦБ можно выделить срочность - протяженность во времени связанных ими экономических отношений. Для большей части ЦБ характерны высокая ликвидность и получение каких-либо

дополнительных прав/льгот её владельцами. При помощи ЦБ возможно достаточно гибко перераспределяются капиталы и денежные средства между субъектами рынка, отраслями экономики, регионами и странами. Также при их помощи капитал концентрируется в самых выгодных сферах для вложения. Использование чеков и векселей позволяет сократить издержки обращения. Как инструмент платежа ЦБ занимают одну из ведущих позиций в платёжном обороте государства.

Котировка представляет собой цену товара, объявленную продавцом/покупателем, по которой они готовы совершить сделку - покупку/продажу. Чаще всего имеется в виду относительно быстро меняющаяся цена, к примеру, биржевая. В некоторых случаях котировкой считается цена лишь заключенной сделки, но не цена оферты покупки/продажи.

На биржах цены регистрируются с помощью специальной котировальной комиссии. Чаще всего публикуются цены открытия и закрытия биржевой сессии, максимальная и минимальная цена дня. Данный вид публикации называется «официальная котировка». Котировки сами по себе отображают складывающуюся на торгах конъюнктуру рынка, баланс спроса и предложения.

Обычное физлицо не может просто так пойти на биржу с целью подачи заявки на покупку/продажу ценных бумаг, в котором оно заинтересовано. Все торги на биржах проводятся профессиональными участниками рынка, например, брокерскими компаниями. Они уже и дают возможность своим клиентам открывать торговые счета, подавать заявки на куплю/продажу, причём за это брокер получает от клиента комиссионные. Ключевая роль брокера – выводить на биржу заявки клиентов. Брокером может выступать только тот субъект, который получил соответствующую лицензию на деятельность от Центрального банка РФ.

В России ведущие брокеры предоставляют не только посредническую, но и также аналитическую функцию, помогая своим клиентам с обучением, что положительно сказывается на продолжительном сотрудничестве.

В третьей части «NoSQL базы данных» описываются основные черты нереляционных хранилищ, их классификация и основные понятия, проводится сравнение и выбор наиболее подходящего для задач исследования.

Термин NoSQL объединяет в себе несколько подходов, которые направлены на реализацию хранилищ БД, имеющих существенные отличия от тех моделей, что используются в реляционных СУБД с доступом к данным с помощью SQL. Он применим к БД, в которых делается попытка решить ряд проблем: масштабируемость и доступность за счёт атомарности и согласованности данных.

Реляционные СУБД ориентируются на ряд требований (ACID) к системе транзакций:

- атомарность (atomicity);
- согласованность (consistency);
- изолированность (isolation);
- надёжность (durability).

NoSQL несколько расширяет данный список, дополняя его следующими (BASE):

- базовая доступность (basic availability) - всякий запрос завершается гарантированно;
- гибкое состояние (soft state) - состояние системы способно изменяться с течением времени с целью достижения согласования данных;
- согласованность в конечном счёте (eventual consistency) - данные рассогласованы какое-то время, но в любом случае приходят к согласованию через некоторый промежуток.

Системы, построенные на основе BASE, не могут быть использованы в любых приложениях. Например, для нормального функционирования банковских и биржевых систем применение транзакций является жизненной необходимостью, но, с другой стороны, свойства ACID практически нереально поддерживать в системах с многомиллионной web-аудиторией. Из этого следует, что архитекторы систем, построенных на NoSQL, должны жертвовать

согласованностью данных для достижения пары других свойств. Существует ряд СУБД (к примеру, Riak), позволяющих изменять под себя требуемые характеристики доступности-согласованности в том числе и для отдельных запросов с помощью задания числа узлов, которые необходимы для подтверждения успешности транзакции.

Для описания схемы данных NoSQL-решения можно использовать разные структуры данных: деревья, хеш-таблицы и другие. Выделяют 4 типа хранилищ в зависимости подходов к распределённости, репликации и модели данных:

- Хранилища «ключ-значение» представляют собой простейшее хранилищем данных, в котором для доступа к значению используется ключ. Данный тип хранилищ часто применяется в качестве кэшей для объектов, создания специализированных файловых систем, для хранения изображений, а также в системах, нацеленных на масштабируемость. Примеры: Redis, MemcacheDB, Berkeley DB.
- В хранилище семейств колонок данные находятся в виде разреженной матрицы, чьи строки и столбцы используются как ключи. Распространённым примером использования данного типа СУБД может быть web-индексирование, а также задачи, которые связаны с BigData, с более низкими требованиями к согласованности данных. Примеры: Cassandra, HBase, Accumulo.
- Сферой применения графовых БД являются задачи, в которых данные имеют большое кол-во связей, к примеру, выявление мошенничества или социальные сети. Обход графа не требует каких-либо дополнительных вычислений, так как его рёбра материализованы, но для нахождения начальной вершины обхода необходимо наличие индексов. Примеры: Blazegraph, OrientDB, Neo4j.
- Документо-ориентированные СУБД предназначены для хранения иерархических структур данных. Они находят себе применение в

издательском деле, системах управления контентом, документальном поиске и т.п. Примеры: MongoDB, CouchDB, MarkLogic.

В четвёртой части *«Реализация программного продукта» описывается архитектура решения на языке Java и платформе Spring Framework.*

Для решения задачи аппроксимации котировок ценных бумаг был реализован программный продукт с использованием ряда современных технологий. В качестве языка программирования выбрана Java версии 1.8. Выбор на него пал не случайно, а в виду широкой распространённости, гибкости в использовании, большого числа сторонних библиотек и личного опыта. Для упрощения реализации промышленного приложения применялся широко распространённый фреймворк Spring, о котором стоит рассказать далее, поскольку он является основополагающей частью всего проекта в целом.

Spring Framework предоставляет комплексный подход к разработке и конфигурированию современных бизнес-приложений на Java. Его ключевой элемент - это забота об инфраструктуре на уровне приложения. Другими словами: разработчики могут сосредоточиться на реализации бизнес-логики без излишних настроек среды исполнения. Некоторые возможности:

- Внедрение зависимости (Context and Dependency Injection, CDI);
- Аспектно-ориентированное программирование (Aspect Oriented Programming, AOP);
- Создание MVC (паттерн Model View Controller) web-приложений, RESTful/SOAP web-сервисов (в том числе Hypermedia API);

Встроенная поддержка JDBC, JPA, JMS.

На сегодняшний день под термином «Spring» имеется в виду целый ряд взаимосвязанных проектов, которые развиваются и курируются компанией Pivotal, а также силами сообщества. К ключевым проектам можно отнести:

- Spring Core - основа всей платформы, которая даёт возможность создавать приложения: менеджмент компонентов, CDI, MVC, базовый доступ к БД и транзакции;

- Spring MVC - управляет контроллерами, маппингами запросов, а также различными HTTP-абстракциями;
- Spring Data - доступ к данным: реляционные и нереляционные БД, Key-Value хранилища;
- Spring Cloud - используется для микросервисной архитектуры: service discovery, трассировка и диагностика, балансеры запросов, circuit breaker-ы, роутеры;
- Spring Security - авторизация и аутентификация, доступ к данным, методам OAuth, LDAP, Kerberos и т.д.;
- Spring Integration - обработка данных из разных источников в pipeline-стиле.

Отдельно стоит сказать о Spring Boot - связующее звено, которое позволяет быстро создать и сконфигурировать приложение, упаковать его в исполняемый независимый артефакт:

- не использует конфигурирование посредством XML, вместо этого используются Java-аннотации;
- максимально возможный уровень автоконфигурирования;
- используется convention over configuration - для большинства конфигураций не нужно ничего настраивать;
- конфигурацию по умолчанию легко перекрывается вручную при необходимости.

Разработка велась с помощью IDE IntelliJ IDEA, поскольку она предоставляет разработчику широкий спектр возможностей для удобной реализации проекта, начиная с автозавершения ввода, анализа кода на опечатки и заканчивая мощным декомпилятором, отладчиком и облачной синхронизацией.

Для хранения информации о котировках была локально развёрнута последняя версия документо-ориентированной NoSQL базы данных MongoDB. Для её администрирования применялось поставляемое вместе БД свободное ПО Compass, представляющее собой пользовательский интерфейс (GUI) к базе.

В приложениях доступны исходные коды программной реализации, а также анализ результатов работы.

Заключение. Проведённое исследование показало актуальность темы моделирования поведения котировок ценных бумаг. В особенности это важно для разрастающегося рынка ценных бумаг Российской Федерации. Аппроксимация массива исторических данных позволяет проследить за тем, как себя ведёт цена финансового инструмента, и сделать для себя соответствующие выводы, принять решение о дальнейших операциях с ним, сделать планирование своих расходов и доходов, а также оценить перспективность рынка и указать на возможные критические изменения.

Для аппроксимации данных можно использовать различные методы. Целью же данной работы был анализ возможности применения синков к приближению значений котировок. В результате были получены данные, на основании которых можно судить о том, что метод ведёт себя ничуть не хуже широко известных, устойчив и вполне может быть применим к такой задаче.